Individualized epidemic spreading models predict epilepsy surgery outcomes: a pseudo-prospective study

Ana. P. Millán[1,2], **Elisabeth C.W. van Straaten**[1,5,6], **Cornelis J. Stam,** [1,4,6], **Ida A. Nissen**[1], **Sander Idema**[3,5,7], **Piet Van Mieghem**[8],

and Arjan Hillebrand[1,4,5]

[1]Amsterdam UMC location Vrije Universiteit Amsterdam, Department of Clinical Neurophysiology and MEG Center, De Boelelaan 1117, Amsterdam, The Netherlands

[2]Institute "Carlos I" for Theoretical and Computational Physics, and Electromagnetism and Matter Physics Department, University of Granada, E-18071 Granada, Spain

[3]Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Neurosurgery, De Boelelaan 1117, Amsterdam, The Netherlands

[4]Amsterdam Neuroscience, Brain Imaging, Amsterdam, The Netherlands

[5]Amsterdam Neuroscience, Systems & Network Neurosciences, Amsterdam, The Netherlands

[6]Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands

[7]Amsterdam Neuroscience, Cancer Biology and Immonology, Amsterdam, The Netherlands

[8]Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands

Keywords: Epilepsy, Epilepsy Surgery, Large-scale brain network, Magnetoencephalography, Epidemic Spreading Model, Personalized Medicine, Seizure Modeling, Whole-brain Modeling

ABSTRACT

Epilepsy surgery is the treatment of choice for drug-resistant epilepsy patients, but up to 50% of patients continue to have seizures one year after the resection. In order to aid presurgical planning and predict postsurgical outcome on a patient-by-patient basis, we developed a framework of individualized computational models that combines epidemic spreading with patient-specific connectivity and epileptogeneity maps: the Epidemic Spreading Seizure and Epilepsy Surgery framework (ESSES). ESSES parameters were fitted in a retrospective study ($N = 15$) to reproduce invasive

Corresponding author: Ana P Millán, apmillan@ugr.es

electroencephalography (iEEG)-recorded seizures. ESSES reproduced the iEEG-recorded seizures, and significantly better so for patients with good (seizure-free, SF) than bad (non-seizure-free, NSF) outcome. We illustrate here the clinical applicability of ESSES with a *pseudo-prospective study* ($N = 34$) with a blind setting (to the resection strategy and surgical outcome) that emulated presurgical conditions. By setting the model parameters in the retrospective study, ESSES could be applied also to patients without iEEG data. ESSES could predict the chances of good outcome after *any* resection by finding patient-specific model-based optimal resection strategies, which we found to be smaller for SF than NSF patients, suggesting an intrinsic difference in the network organization or presurgical evaluation results of NSF patients. The actual surgical plan overlapped more with the model-based optimal resection, and had a larger effect in decreasing modeled seizure propagation, for SF patients than for NSF patients. Overall, ESSES could correctly predict 75% of NSF and 80.8% of SF cases pseudo-prospectively. Our results show that individualised computational models may inform surgical planning by suggesting alternative resections and providing information on the likelihood of a good outcome after a proposed resection. This is the first time that such a model is validated with a fully independent cohort and without the need for iEEG recordings.

AUTHOR SUMMARY

Individualized computational models of epilepsy surgery capture some of the key aspects of seizure propagation and the resective surgery. It is to be established whether this information can be integrated during the presurgical evaluation of the patient to improve surgical planning and the changes of a good surgical outcome. Here we address this question with a pseudo-prospective study that applies a computational framework of seizure propagation and epilepsy surgery – the *ESSES* framework– in a pseudo-prospective study mimicking the presurgical conditions. We found that, within this pseudo-prospective setting, ESSES could correctly predict 75% of NSF and 80.8% of SF cases. This finding suggests the potential of individualised computational models to inform surgical planning by suggesting alternative resections and providing information on the likelihood of a good outcome after a proposed resection.

INTRODUCTION

Surgical resection is often the most effective treatment to achieve seizure control for patients with drug-resistant focal epilepsy. The surgery requires the generation of an hypothesis of the epileptogenic zone (EZ) by means of extensive presurgical evaluations, and its subsequent removal or disconnection during surgery (Lüders, Najm, Nair, Widdess-Walsh, and Bingman (2006)). Despite extensive investigations, there has only been a slight improvement in prognosis over the past two decades (Baxendale et al. (2019); Jehi et al. (2015)), and between 30 to 50% of the patients who undergo surgery continue to have seizures one year later, depending on etiology and location of the EZ (Englot et al. (2015)). A key conceptual change in recent years is the notion of *epileptogenic networks*, which takes into account the complex interplay between different brain regions in promoting and inhibiting seizure generation and propagation (Bartolomei et al. (2017); Kramer and Cash (2012); van Diessen, Diederen, Braun, Jansen, and Stam (2013)). As a consequence, the effect of a given surgery is to be measured against the whole epileptogenic network: a small resection involving heavily connected regions may have widespread effects, but it may also be compensated for by the remaining network (Hebbink, Meijer, Huiskamp, van Gils, and Leijten (2017); Nissen et al. (2018)). This perspective aligns with the commonly accepted view that large-scale brain organization can be regarded as an emerging phenomenon taking place on a complex network, which has spurred numerous data- and model-based studies (Seguin, Jedynak, et al. (2023); Seguin, Sporns, and Zalesky (2023)). Several network-based studies have found group-level differences between seizure-free and non-seizure-free patients (da Silva et al. (2020); Nissen et al. (2018); Taylor et al. (2018)), with removal of pathological hub (i.e. central) regions typically associated with seizure-freedom (Nissen et al. (2017)). These results highlight the need to consider *patient-specific connectivity* (van den Heuvel and Sporns (2019)) in order to tailor the surgery to each patient (Gerster et al. (2021)).

A data-driven manner to study the relation between individual brain networks and surgical outcomes involves *computational models of seizure dynamics*, which allow us to simulate seizure propagation *in silico*. Different resection strategies can be tested on the computational model before the actual surgery (Goodfellow et al. (2016); Hutchings et al. (2015); V. Jirsa et al. (2017); Laiou et al. (2019); Lopes et al. (2017); Nissen et al. (2021); Olmi, Petkoski, Guye, Bartolomei, and Jirsa (2019); Proix, Bartolomei, Chauvel, Bernard, and Jirsa (2014); Sinha et al. (2017); Taylor, Kaiser, and Dauwels (2014)). The models

can be fitted to patient-specific data of brain structure and seizure dynamics, allowing us to tailor the resection strategy for each patient. Within this perspective, previous studies have obtained remarkable success at a group level: Sinha et al. (2017) found that the removal of regions identified as epileptogenic according to an EEG-brain network dynamical model predicted surgical outcome with 81.3% accuracy. Proix, Bartolomei, Guye, and Jirsa (2017), using a seizure model known as the *epileptor* (V. K. Jirsa, Stacey, Quilichini, Ivanov, and Bernard (2014)) based on MRI (magnetic resonance imaging) connectivity, found significant differences in the overlap between the model-based propagation zone and the area sampled by iEEG between patients with good (Engel class I) and bad (Engel class III) outcomes at the group level. Subsequent studies also found a better match between the modeled and clinically observed epileptogenic regions for seizure-free than non-seizure-free patients (Makhalova et al. (2022); Vattikonda et al. (2021)). Similarly, Sip et al. (2021) simulated patient-specific resection strategies by means of *virtual resections*, and found that virtual resections in their model correlated with surgical outcome, with larger effects found for patients with good outcome (Engel classes I and II). In an independent study, Goodfellow et al. (2016) also found significant differences in the model prediction between Engel class I and class IV patients, using an electrocorticogram modeling framework.

Following the same rationale, we developed a computational model of seizure propagation and epilepsy surgery based on epidemic spreading dynamics and patient-specific MEG brain connectivity (Millán et al. (2022)), to which we refer here as the *Epidemic Spreading Seizure and Epilepsy Surgery model (ESSES)*. Epidemic models describe the spread of an infectious agent through a network. Epidemic processes on fixed networks have a rich mathematical history (Pastor-Satorras, Castellano, Van Mieghem, and Vespignani (2015)) with a plethora of models that can be exploited for epilepsy surgery optimization (Millán et al. (2022); Nissen et al. (2021)). Although such models ignore the underlying bio-physical processes that lead to seizure generation and propagation, they describe the basic rules that govern spreading processes. In previous studies (Millán et al. (2022, 2023)), we found that epidemic spreading models could reproduce stereotypical patterns of seizure propagation as recorded via invasive electroencephalography (iEEG) recordings. Moreover, once fitted with patient-specific data, ESSES could identify alternative resection strategies, either of smaller size or at a different location than the actual surgery (Millán et al. (2022); Nissen et al. (2021)). In a more recent study Millán et al. (2023), we showed that the goodness-of-fit of ESSES seizures to those recorded via iEEG predicted surgical

outcome –with an area under the curve of 88.6% – indicating that ESSES not only reproduces the basic aspects of seizure propagation, but it also captures the differences, either in the location of the resection area relative to the EZ, or intrinsically in the iEEG or MEG data, between patients with good and bad outcome. Importantly, ESSES's global parameters were defined at the population level, and the model was individualized for each patient via patient-specific MEG networks, which characterized the local spreading probabilities. As a consequence, ESSES can be extended to patients without iEEG recordings, in contrast to previous modeling studies, which typically required the existence of patient-specific iEEG data to individualize the model for each patient (Bernabei et al. (2023); Gunnarsdottir et al. (2022); Makhalova et al. (2022); Proix et al. (2017); Runfola, Sheheitli, Bartolomei, Wang, and Jirsa (2023); Sinha et al. (2017); Y. Wang et al. (2023)). IEEG allows for a highly resolved description of seizure dynamics, but its spatial sampling is sparse and it is highly invasive. Consequently, it is only part of the presurgical evaluation in a selection of patients.

Here we performed a pseudo-prospective blind study (34-patient validation cohort) to validate the clinical applicability of ESSES to a) identify model-based optimal resection strategies and b) predict the likelihood of a good outcome after a proposed resection strategy, on a patient-by-patient basis. In order to emulate the clinical presurgical conditions, the research team was blind to the patients' postsurgical data, namely the resection area and surgical outcome, during ESSES's analyses, and the multimodal presurgical information available for each patient was integrated into ESSES. ESSES can identify resection strategies that perform optimally in the model, i.e. by minimizing modeled seizure propagation, for a given resection size. We refer to these resections as *optimal resections*, in agreement with previous works (An, Bartolomei, Guye, and Jirsa (2019); Millán et al. (2022); Nissen et al. (2021); Sinha et al. (2017)). ESSES can also simulate the effect of a given resection *in silico*. Within this set-up, we tested three hypotheses: a) seizure-free (SF) patients would have smaller optimal resections than non-seizure-free (NSF) patients, b) SF patients would have a larger overlap between optimal and planned (clinical) resections, and c) the planned resection would have a larger effect (in ESSES) for SF than for NSF patients. We found that these three ESSES biomarkers, namely the size of the optimal resection, their overlap with the planned resection, and the effect of the planned resection on ESSES seizures, provided estimates of the likelihood of a good outcome after the surgery, as well as suggesting alternative resection strategies that performed optimally in the model. We envisage that the implementation of a

modeling scheme such as ESSES in clinical practice may inform the planning of epilepsy surgery. Different surgical plans can be tested with ESSES for each patient, such that strategies that lead to a large decrease of propagation in the model are more likely to lead to seizure freedom. ESSES may also suggest optimal (alternative) resection strategies, for cases where ESSES predicts a bad outcome with the planned resection. Optimal strategies can then lead to new surgical plans, the effect of which can then be tested in ESSES again.

RESULTS

Here we validated the clinical applicability of ESSES to **A)** identify optimal resection strategies that may improve surgical outcomes and **B)** provide estimates of the probability of postsurgical seizure freedom, given a surgical plan. The key goal of ESSES is to identify surgical candidates who would have a bad outcome (NSF patients) so that the surgical plan can be adjusted. This study combined a retrospective analysis on a *modeling cohort* ($N = 15$) that was used to set the model hyperparameters (following our previous retrospective study (Millán et al. (2023)) on this same cohort), and a pseudo-prospective study on a *validation cohort* ($N = 34$) to validate ESSES findings and to emulate its clinical application in a blind set-up that mimics the clinical presurgical conditions. The researchers were blind to the performed surgery and surgical outcome during the application of ESSES to the validation cohort.

The study was performed as follows:

1. **Seizure model: definition and fitting (modeling cohort).** An SIR-type of epidemic spreading process modeled seizure propagation over patient-specific brain connectivity. IEEG data from the modeling cohort was used to fit the global parameters of the spreading model so that ESSES-modeled seizures matched those recorded via iEEG, as shown in figure 1A.
2. **Individualized ESSES framework: patient-specific models.** ESSES was individualized for each patient: patient-specific MEG brain connectivity defined the network on which ESSES computed seizure propagation. Multi-modal patient-specific data, available from presurgical evaluations, defined the seed regions (i.e. the seizure onset regions) based on *epileptogenicity* or *seed-probability maps*.

3. **Alternative resection strategies (aim A).** ESSES incorporates an optimization algorithm to determine model-based optimal resection strategies for each patient. These acted as a benchmark against which the planned resection for each patient could be tested. These resections were optimal in the model in the sense that they minimized modeled seizure propagation.
4. **Simulation of the planned resection strategy (aim B).** The resection plan for each patient was simulated in ESSES with a virtual resection that emulated the actual surgical resection, and the subsequent decrease in seizure propagation was measured.
5. **Statistical analyses (aim B).** We compared ESSES's predictions (steps 3 and 4) between patients with good and bad outcome. We defined the NSF class as the positive class for classification and prediction testing.

This analysis pipeline was first implemented in the modeling cohort in a retrospective study that served to set all model hyperparameters. Then, steps $2-5$ were applied to the validation cohort in a pseudo-prospective study with a blind set-up. The pipeline for the model implementation, detailing at which step the de-blinding of each data-type took place, is illustrated in figure 2. A detailed pipeline including also the model set-up (modeling cohort) is also included as Supp. figure 7.

Seizure propagation as an epidemic spreading process

We modeled seizure propagation by a Susceptible-Infected-Recovered (SIR) epidemic process, as illustrated in figure 1. The S-I-R states account respectively for the healthy (pre-ictal), ictal and healthy (post-ictal) states, coupled with patient-specific brain connectivity (derived from MEG data) to define the local spreading probabilities. The SIR model describes the spread of an infection from an initial set of infected nodes, the seed regions, to the other nodes in the network, and the recovery of the infected nodes, without re-infections (Barrat, Barthelemy, and Vespignani (2008); Pastor-Satorras et al. (2015)). Here we confined ourselves to one of the simplest compartmental SIR models, using a discrete-time setting where the spreading probability from node i to node j corresponded to the coupling strength w_{ij} on the patient-specific brain network and where the recovery probability γ was set to be equal for all nodes. The brain network was initially thresholded (by setting the weakest links to zero) at different densities ρ indicating the fraction of non-zero links remaining in the network after thresholding (see Methods section and Supp. section 5).

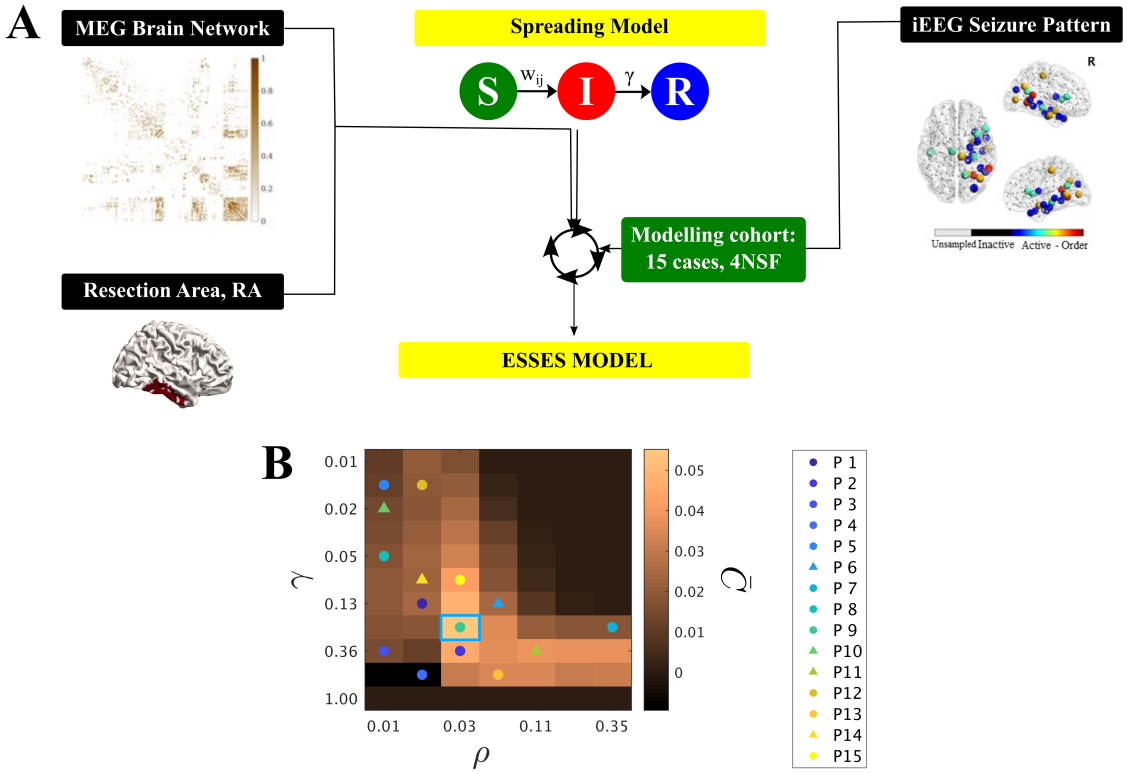

Figure 1. **A** Sketch of ESSES's parameter-fitting scheme. The parameters controlling seizure propagation, namely the density of links in the network ρ and the global recovery probability γ, were set so as to maximize the similarity between ESSES-modeled seizures and iEEG-recorded ones for the modeling cohort (eq. 1). Seizures were simulated via SIR dynamics over MEG patient-specific brain networks, and setting the resection area as the seed of epidemic spreading. **B** $\bar{C}(\rho, \gamma)$ map displaying the average model fit (modeling cohort). The data points indicate the parameters corresponding to the best individual fit for each patient, with circles (triangles) indicating SF (NSF) cases (corresponding C values can be seen in Supp. figure 2). Most individual best fits (data-points) fall within the same region (SIR phase transition) but there is large variability (in fact, we found low signal to noise ratios of approx. $1/2$, see Supp. figure 3A). The blue square marks the maximum of the goodness-of-fit, and the corresponding (ρ, γ) values were used for the subsequent analyses. The y-axis is shown using a logarithmic scale.

The two control parameters of ESSES are thus the global recovery probability γ and the network density ρ. We followed the inference method presented in our previous study (Millán et al. (2023)) to fit the model parameters to iEEG-recorded seizures of the modeling cohort. We note that the modeling framework as presented here differs slightly from the one in Millán et al. (2023), which included an extra parameter to set the global spreading rate. The details of the model fit can be found in the methods

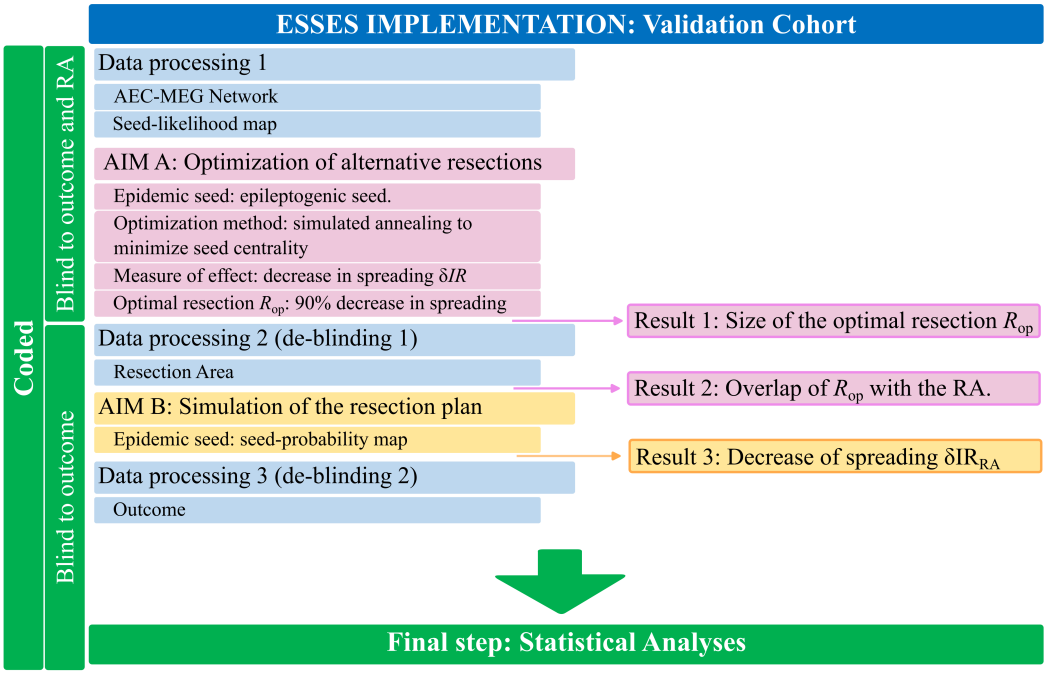

Figure 2. Processing and analysis pipeline. The patient data were processed in three different steps (blue boxes) for the validation cohort. Firstly, ESSES's key ingredients, the patient-specific MEG brain network and the seed-likelihood map, were processed. The research team remained blind to the resection area and outcome of each patient. The first analysis (AIM A: Optimization of alternative resections, pink boxes) then took place and the first result (Result 1: Size of the optimal resection R_{op}) was obtained. Then, the patients's resection areas were processed (de-blinding step 1) and the second result was obtained (Result 2: overlap of R_{op} with the resection area, RA). AIM B (Simulation of the resection plan, yellow boxes) could then take place: the simulation of the resection plan, by performing a virtual resection of the resection area. The third and final result (Result 3: Decrease of spreading $\delta IR(RA)$) was then obtained. Then, the second and final de-blinding took place to recover the outcome of each patient and perform the statistical analyses.

section, and the fit results are reported in the supplementary information (Supp. section 5.2, see also Supp. figures 2 and 3).

The degree of similarity between the ESSES and iEEG seizures was measured with the *goodness-of-fit* $C(\rho, \gamma)$ (eq. 1). The resulting diagram resembled a familiar phase transition (figure 1B), with an interface of high goodness-of-fit (yellow regions) corresponding to a roughly constant spreading-to-recovery ratio $\rho/\gamma = \text{const}$, in agreement with other studies (Moosavi, Jirsa, and Truccolo (2022)). The maximum goodness-of-fit is indicated by a blue square in figure 1B, and sets the working point of ESSES for the

remaining analyses. At this working point, the SF group presented a significantly better fit than the NSF group ($p = 0.04$, see Supp. table 3 and Supp. figure 3B for details).

A ROC classification analysis indicated a good classification (AUC $= 0.79$, see Supp. table 4 and Supp. figure 3C) between the SF and NSF groups. At the optimal classification point (Youden criterion, Supp. figure 3D), all NSF patients were correctly identified. The high sensitivity suggests that all patients identified as SF by ESSES could proceed to surgery with high expectations (100% in this group) of a good outcome. On the contrary, patients identified as NSF should be examined further (e.g. by performing further presurgical evaluations or considering other resection plans) as they had a 57% chance of bad outcome with the proposed surgery (to be compared with a 26% chance of bad outcome expected simply from the relative group sizes).

Presurgical hypothesis of the seed regions

A key ingredient of ESSES is the definition of the epileptogenic or seed regions. Here we defined epileptogenicity or seed-probability maps SP_i, indicating the probability that each brain region i gave rise to a seizure. The *seed-probability* maps integrated patient-specific multimodal presurgical information (encoded in the local patient database (*Castor Electronic Data Capture.* (n.d.))) in a quantitative and systematic manner that was adapted for each patient to include the data from the presurgical evaluations that they had undergone (see Methods section and Supp. section 4 for details). The resulting seed-probability maps for two representative patients (modeling cohort) are illustrated in figure 3B,D together with the corresponding resection areas (panels A, C). The seed-probability maps show wider spatial patterns than the resection areas, and may involve several lobes in both hemispheres. The resection areas for the two cases shown here were contained within the most likely seed regions. In general, the resection areas had a larger seed-probability than expected by chance for all patients. We did not find significant differences in the overlap between the resection areas and the seed-probability maps between SF and NSF patients (see Supp. figure 1).

Optimal resection strategies

ESSES can derive individualized alternative resection strategies –that minimize modeled seizure propagation– via an optimization algorithm based on simulated annealing (Millán et al. (2022); Nissen et

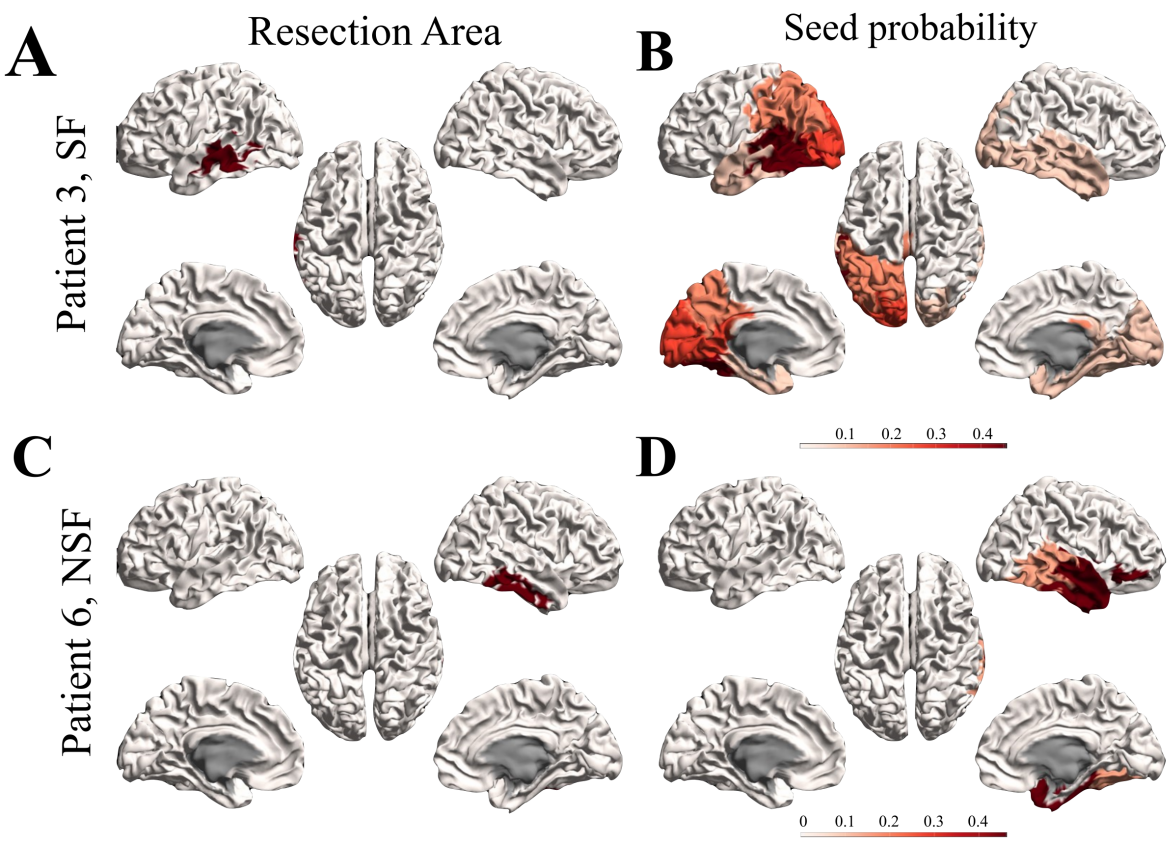

Figure 3. Seed-probability maps. Resection areas (left) and seed-probability maps (right) as derived from the database with presurgical information for two representative cases from the modeling cohort: patient 3 (SF, top) and 6 (NSF, bottom).

al. (2021)). The optimization algorithm parameters were set on the modeling cohort data (see Methods section for the algorithm details, and Supp. section 5.3 and Supp. figures 4 and 5 for the modeling cohort results), and the algorithm was then applied to the validation cohort in a blind setting.

The optimization algorithm searched for resections R of increasing size $S(R)$ that minimized the *seed efficiency* $E_R(\text{seed})$, i.e. the average distance (on the network) from the seed nodes to the other network nodes. This procedure exploits the link between epidemic spreading dynamics and network structure, such that spreading to a region is strongly influenced by its distance to the seed (Pastor-Satorras et al. (2015)). In figure 4A we show the normalized seed efficiency $e_R(\text{seed})$, which is normalized to the seed efficiency in the unresected network so as to diminish differences due to seed extent and initial efficiency. $e_R(\text{seed})$ decreased with the size of the resection for all patients. At the group level, the SF group showed

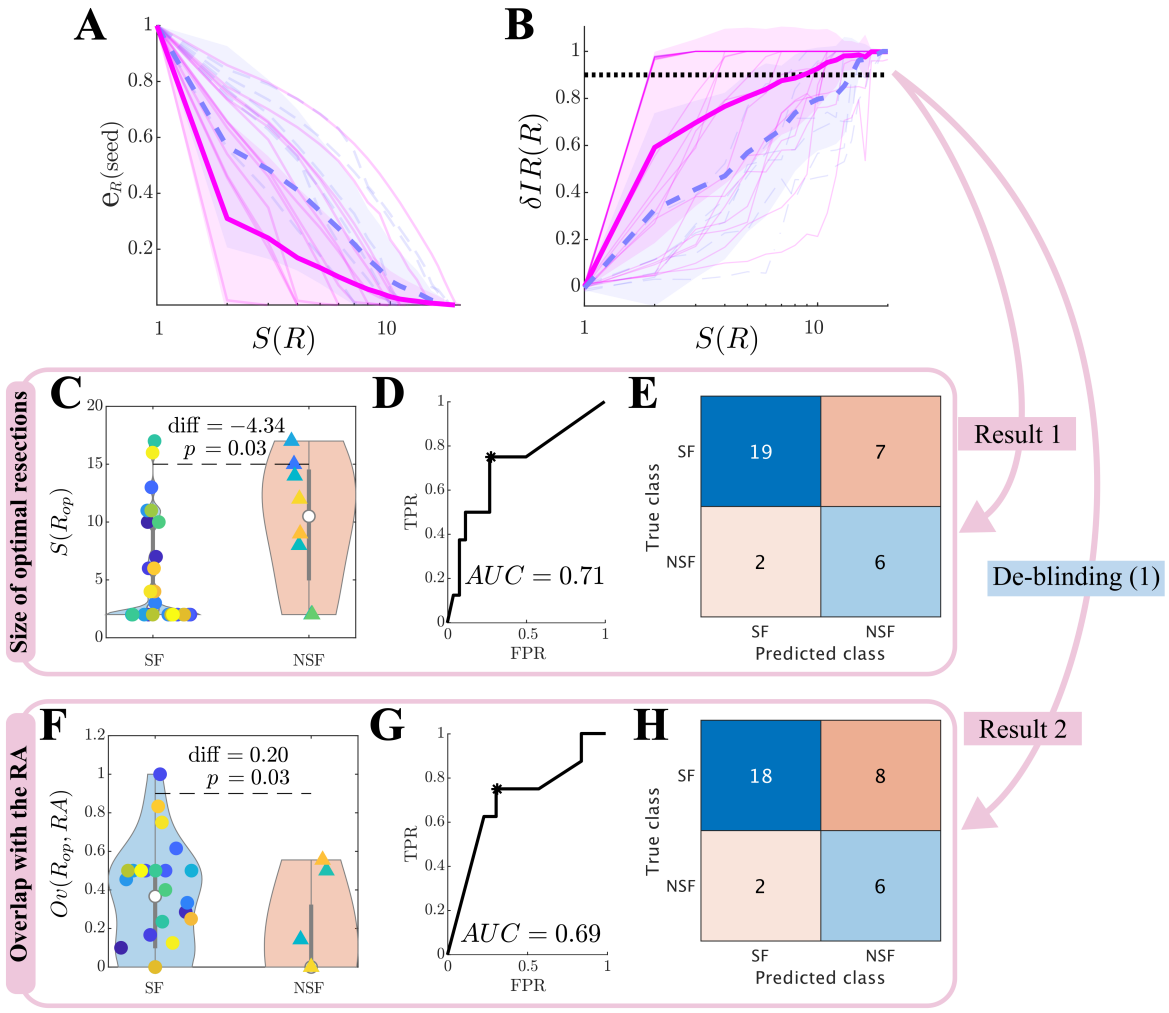

Figure 4. Optimal (alternative) resection strategies (validation cohort). Effect of optimal virtual resections of size $S(R)$ as measured by **A** the normalized seed efficiency $e_R(\text{seed})$, and **B** normalized decrease in seizure propagation $\delta IR(R)$. Blue dashed lines stand for NSF patients, and pink solid lines for SF patients. Thin lines show individual patients, and darker wide lines the group averages, with shaded areas indicating the standard deviations. The apparent darker pink line at the top of the plot arises from overlap of several individual lines. **C-H** Group level comparison of the size of optimal resections $S(R_{op})$ (**C-E**) and their overlap with the resection area $Ov(R_{op}, RA)$ (**F-H**). Panels **C** and **F** show the distribution of values of each patient group, with significance results obtained with exact two-sided Wilcoxon ranksum tests. Panels **D** and **G** show the corresponding ROC classification analyses, where TPR and FPR stand respectively for the true positive (NSF cases classified as NSF) and false positive (SF cases classified as NSF) rates. Finally, panels **E** and **H** show the confusion matrices corresponding to the optimal point (Youden criterion, black asterisks in the middle panels) of the ROC curves.

a significantly smaller $e_R(\text{seed})$ than the NSF group (repeated measures ANOVA test, $F(19) = 37.95$, $p < 10^{-89}$), for all considered seed sizes except $S(R) = 1$. Moreover, the effect of increasing the resection size on $e_R(\text{seed})$ was larger for the SF than for the NSF group ($F(19) = 3.78$, $p < 10^{-6}$).

The actual effect of a resection R on modeled seizure propagation was quantified by measuring the *normalized decrease in seizure propagation* due to the resection, $\delta IR(R)$ (figure 4B), again relative to propagation on the unresected network. Seizure propagation depended heavily on the seed realization such that a bi-stable regime emerged in which ESSES seizures either propagated macroscopically or died locally (an exemplary case is shown in Supp. figure 4). Thus, results reported here were averaged over 300 independent realizations of the seed regions and SIR dynamics. At the group level, the SF group presented a larger decrease in seizure propagation ($F(19) = 25.88$, $p < 10^{-65}$), and a larger effect of increasing the resection size ($F(19) = 2.90$, $p = 4 \cdot 10^{-5}$). There were large differences in the dependence of $\delta IR(R)$ on the resection size between different patients. Whereas in the majority of the cases $\delta IR(R)$ increased roughly exponentially with $S(R)$, for several patients there was an abrupt (discontinuous) jump at a given resection size.

We defined the *optimal resection* R_{op} as the one leading to a 90% decrease in seizure propagation, $\delta IR(R_{op}) = 0.90$. The SF group had significantly smaller optimal resections, and these presented a significantly larger overlap with the actual resection strategy $Ov(R_{op}, RA)$ (see panels C and F of figure 4, and table 1), than the NSF group. We found good classification results using either of these variables to classify between the SF and NSF groups ($AUC = 0.71,\ 0.69$ respectively for $S(R_{op})$ and $Ov(R_{op}, RA)$, see figure 4D,G). Both variables led to very similar classification results at the optimal classification point (Youden criterion), correctly identifying $6/8$ NSF cases (panels E and H). The classification results for the validation cohort are summarized in table 2 (see Supp. table 4 for the modeling cohort results).

In summary, these results indicate that the planned resection strategy (accounted for here by the resection area) presented a larger overlap with the optimal resection for patients with good outcome. In particular, 90.0% of SF and 42.9% of NSF patients were correctly classified by $Ov(RA, R_{op})$. Remarkably, ESSES could also distinguish between SF and NSF patients without taking into account the information of the surgical plan. In fact, up to 90.4% of SF and 46% of NSF patients were correctly identified by $S(R_{op})$ (in relation to only a 76.5% SF-chance and 23.5% NSF-chance according to the group ratios). As this analysis did not depend on the planned resection strategy, a bad prognosis would be indicative of the need

to perform a more exhaustive presurgical evaluation, and potentially imply an unavoidable non-seizure-free outcome after any surgery.

Finally, we note that almost equivalent results may be obtained by considering the disconnecting resection, i.e. the smallest resection leading to disconnection of the seed, instead of the optimal resection (see Supp. section 5.4 and Supp. figure 6). This is due to the strong link between network topology and emergent SIR dynamics, a result that can be used to speed up computations considerably, by using a purely network-based analysis of the effect of different resection strategies.

Simulation of the surgical plan

We simulated the effect of the planned surgery in ESSES for each patient by performing virtual resections of the resection area, which was considered as a proxy for the surgical plan here. We report here on the results for the validation cohort (figure 5), results for the modeling cohort can be found in the supplementary information (Supp. section 7, Supp. figure 8). As in previous sections, all modeling details had already been set during the modeling step. The effect of the resection strategy on (modeled) seizure propagation, $\delta IR(RA)$, was significantly larger for the SF than the NSF group (figure 5B, table 1). A ROC classification analysis revealed a good classification between the two groups ($AUC = 0.78$, figure 5C) and at the optimal point (Youden criterion, black asterisk in panel C) the majority of the patients were correctly identified (figure 5D, table 2). In particular, there was a 91.3% chance that a patient classified as SF had a good outcome, and a 54.5% chance that a patient classified as NSF had a bad outcome, compared to a 76.5% and 23.5% chance based on the relative group sizes.

Prediction of surgical outcome

The classification analyses in the previous sections were informed by each patient's surgical outcome. In a prospective setting the outcome for the patient is not yet known, and thus cannot be used to build the classification model. In order to emulate a true prospective setting, we performed a *prediction analysis* based on leave-one-out crossvalidation. That is, in order to predict the outcome of each patient of the validation cohort, a prediction model was built using data from the remaining 33 cases. Results from this analysis are shown in figure 6, with the statistical details reported in table 3. The prediction results were slightly worse than the classification ones (previous sections), particularly for the NSF class where there

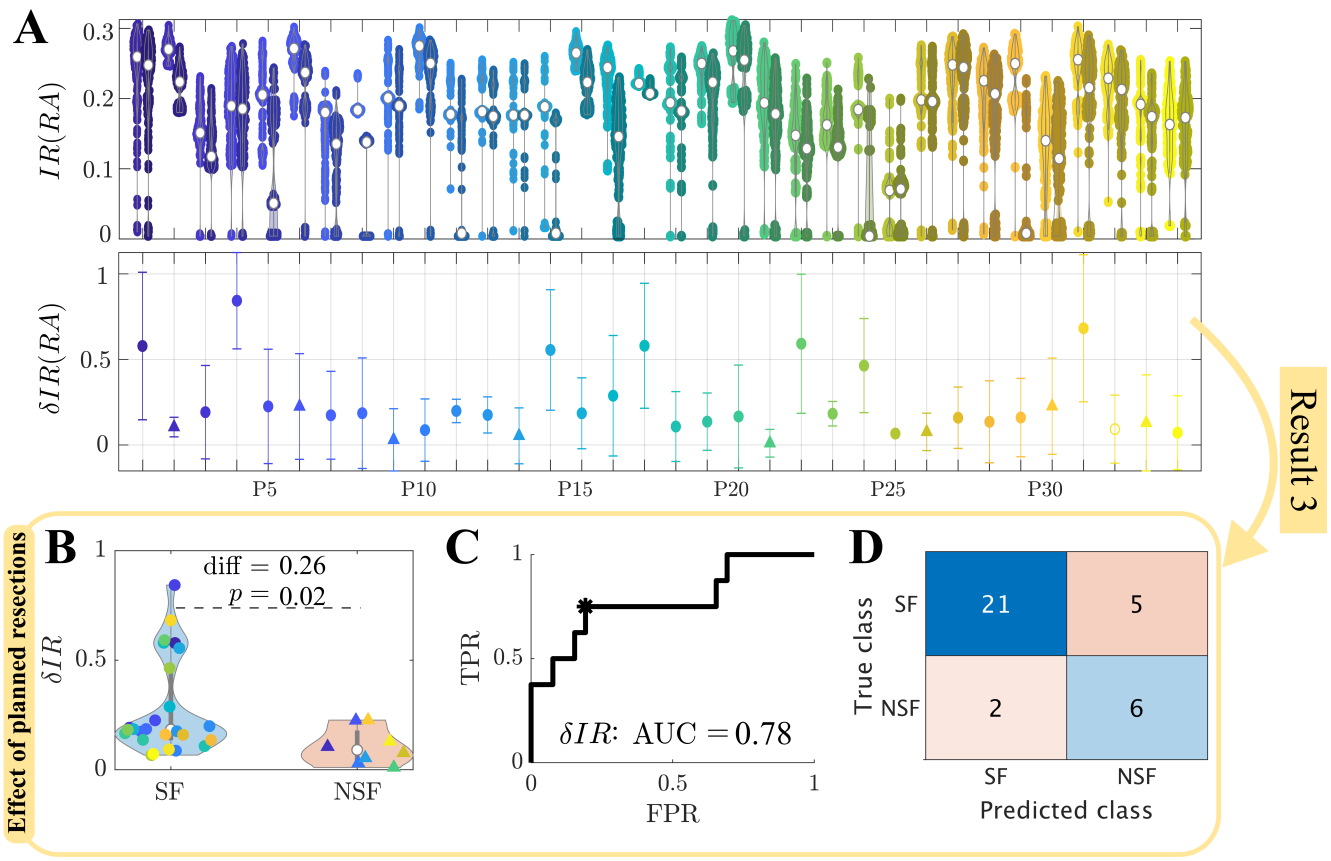

Figure 5. Simulation of the planned resection strategy (validation cohort). **A** The top panel shows seizure propagation IR before (left point cloud for each patient) and after (right point clouds) the resection, for 300 iterations of the seed regions, for each patient. The bottom panel shows the average relative decrease in seizure propagation $\delta IR(RA)$, with errorbars given by the standard deviation over seed iterations. **B** Comparison of the relative decrease in seizure propagation $\delta IR(RA)$ between the SF and NSF groups. Each point corresponds to one patient. **C** ROC curve of the group classification based on $\delta IR(RA)$. TPR and FPR indicate respectively the true positive (NSF cases classified as NSF) and false positive (SF cases classified as NSF) rates. **D** Classification results for the optimal point (black asterisk in panel C) of the ROC curve according to the Youden criterion.

was a 12.5% reduction in the group size. In any case, respectively 4, 5 and 5 NSF cases and 19, 18 and 21 SF cases were correctly identified by each ESSES biomarker (figure 6A). Moreover, 75% of NSF cases $(6/8)$ and only 19.2% $(5/26)$ of SF cases were identified by two or more biomarkers as NSF (figure 6B). For this cohort, if ESSES predicted a good outcome with at least two markers, there was a 80.8% chance

of seizure freedom after the surgery (compared to a 76.5% expectancy of surgery success according to the group rates). Conversely, if the model predicted a bad outcome, then there was a 75% chance that the surgery would fail (compared to a 23.5% expectancy of surgery failure according to the group rates). In clinical practice, a good ESSES prediction could then be interpreted as a large (80.8%) chance of seizure freedom after the surgery and thus support the decision to proceed with surgery. On the contrary, a bad ESSES prediction would indicate a 76.5% chance that the surgery would fail. This may be suggestive of the need of more presurgical evaluations or a different resection strategy, and eventually indicate a low probabily of complete seizure freedom after the surgery.

Finally, in order to test whether the information provided by the three biomarkers could be combined to improve the prediction results, we performed a machine learning analysis using an adaptive boosting algorithm with random undersampling and leave-one-out cross-validation (figure 7A,B). The input variables for the classification algorithm were $\delta IR(RA)$, $S(R_{op})$ and $Ov(RA, R_{op})$. We found that, even though the accuracy of the model was good (0.71) the machine learning model was biased towards the majority class (SF), with only 35% of NSF cases correctly identified (precision $= 0.37$, sensitivity $= 0.35$) and a poor result for $F1 = 0.36$, even though the considered algorithm (RUSboost) was designed to correct for class imbalance. However, the minority class in our case contained only 8 cases, likely preventing the model from being able to generalize. In order to address this issue, we created a *combined cohort* ($N = 49$) pooling together the patients from the modeling and validation cohorts (figure 7C,D). The combined cohort had 12 NSF cases (50% increase), and the new model was able to identify the majority of NSF cases correctly (72% of SF cases and 63% of NSF cases). Even though the accuracy of the model (0.70) did not improve, the remaining measures, which are less affected by class imbalance, did (precision $= 0.42$, sensitivity $= 0.63$, $F1 = 0.51$). Overall, the machine learning model was not able to improve upon the results found using the individual variables (see table 3), and indeed the prediction was predominantly based only on one biomarker, namely the effect of the planned resection on the modeled seizures, $\delta IR(RA)$. Due to the small sample size, we could not determine whether this was due to intrinsic model limitations, suboptimal hyperparameters, or simply a too small group size (particularly of the minority class). Our set-up (leave-one-out cross-validation combined with random undersampling) was designed to minimize the effects of the small sample size, but could not avoid them fully.

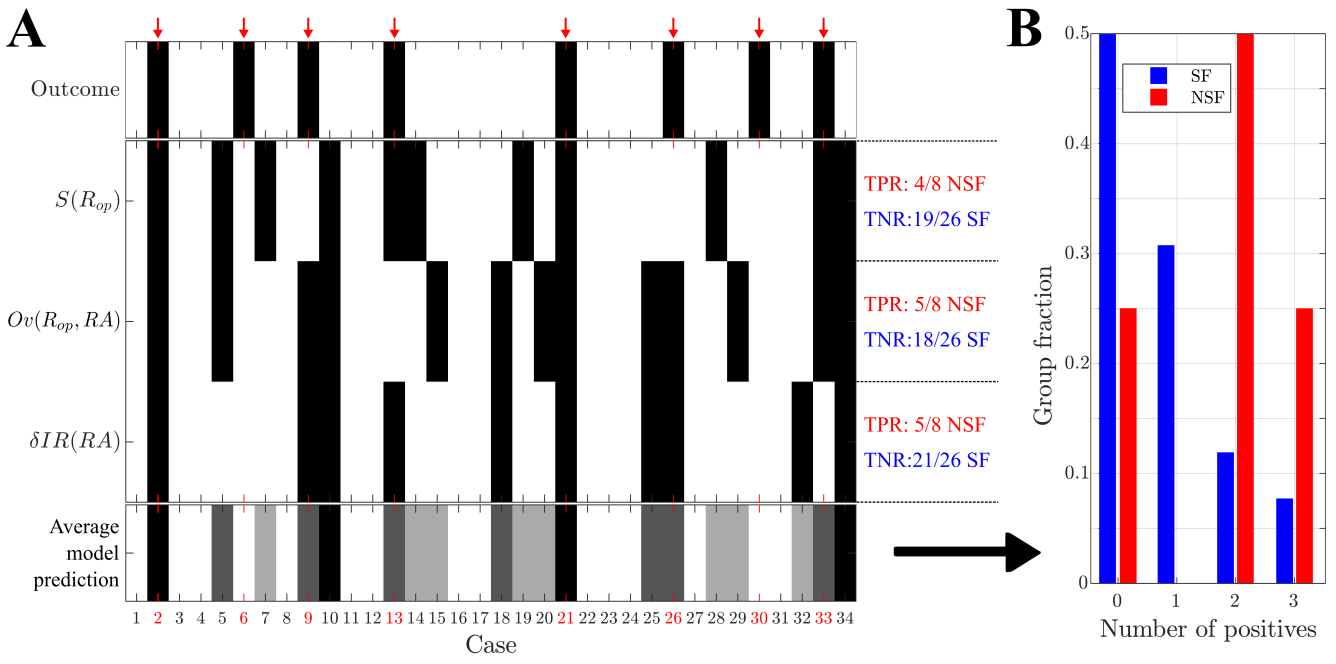

Figure 6. Prediction of surgical outcome: validation cohort. **A** Prediction results using each of the three model-based biomarkers of surgical outcome: the size of optimal resections $S(R_{op})$, the overlap between optimal resections and the resection area, $Ov(R_{op}, RA)$, and the decrease in seizure propagation due to simulation of the planned resection strategy, $\delta IR(RA)$. NSF (SF) cases are shown by black (white) rectangles. The bottom row shows the fraction of biomarkers ($0-3$ out of 3) with a positive (i.e. NSF) classification (refereed to as "Average model prediction" in the figure), for each patient. Surgical outcome is shown in the top row. NSF cases are highlighted by a red arrow and by red labels. **B** Relative number of cases identified as NSF by n biomarkers, $n = 0,\ 1,\ 2,\ 3$, respectively for the SF (blue, left-side bars, $N = 26$) and NSF (red, right-side bars, $N = 8$) groups.

Metric	diff	rks	p
$\delta IR(R_{op})$	-4.34	411.5	0.03
$Ov(R_{op}, RA)$	0.20	495.5	0.03
$\delta IR(RA)$	0.26	513	0.02

Table 1. Summary of statistical comparisons: difference between SF and NSF groups (validation cohort). diff and rks stand respectively for the difference between the SF and NSF groups and the ranksum value.

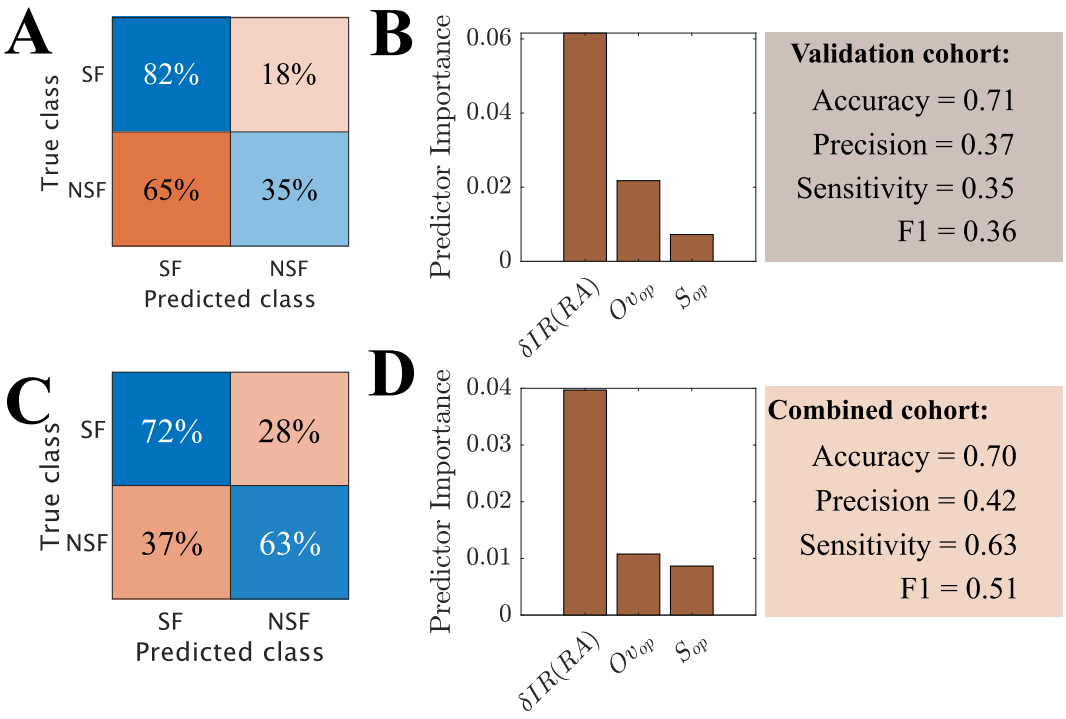

Figure 7. Prediction of surgical outcome using a machine learning algorithm (RUSBoost) with leave-one-out cross validation. As input variables we used the normalized decrease in seizure propagation after virtual resection of the RA, $\delta IR(RA)$, the size of optimal resections $S(R_{op})$ and the overlap of optimal and clinical resections $Ov(R_{op}, RA)$. Panels **A,B** show the confusion matrix and predictor importance for the validation cohort ($N = 34$, 8 NSF), and panels **C,D** are for the combined cohort ($N = 49$, 12 NSF).

Variable	True negatives: SF		True positives: NSF		Acc.	Prec.	Sensitivity	$F1$	AUC
$S(R_{op})$	19	0.73	6	0.75	0.74	0.46	0.75	0.57	0.71
$Ov(RA, R_{op})$	18	0.69	6	0.75	0.71	0.43	0.75	0.55	0.69
$\delta IR(RA)$	21	0.81	6	0.75	0.79	0.55	0.75	0.63	0.78

Table 2. Results of the classification analyses for the validation cohort. Results correspond to the optimal points of the ROC curves according to the Youden criterion to account for class imbalance. For each group (SF, NSF), we show the number of correctly identified cases by absolute number and relative frequency. The remaining columns correspond respectively to the accuracy (Acc.), precision (Prec.), sensitivity, F1 statistic and area under the curve (AUC).

	Variable	True negatives: SF	True positives: NSF	Acc.	Prec.	Sensitivity	$F1$
Validation	$S(R_{op})$	$19/26 \quad (= 0.73)$	$4/8 \quad (= 0.50)$	0.68	0.26	0.50	0.38
	$Ov(RA, R_{op})$	$18/26 \quad (= 0.69)$	$5/8 \quad (= 0.63)$	0.68	0.38	0.63	0.51
	$\delta IR(RA)$	$21/26 \quad (= 0.81)$	$5/8 \quad (= 0.63)$	0.76	0.50	0.63	0.57
	Combined	$21/26 \quad (= 0.81)$	$6/8 \quad (= 0.75)$	0.79	0.55	0.75	0.65
	RUSboost	0.82	0.35	0.71	0.37	0.35	0.36
Combined	RUSboost	0.72	0.63	0.70	0.42	0.63	0.51

Table 3. Results of the prediction analyses for the validation and combined cohorts. For each analysis, we used a leave-one-out crossvalidation such that a predictive model was build to predict the outcome of each patient using the data from the remaining $N-1$ patients. For the individual variables, the results correspond to the optimal points of the ROC curves according to the Youden criterion. For the machine learning analyses, they were derived from an adaptive boosting (AdaBoost1, Matlab 2018) algorithm with leave-one-out crossvalidation, combined with random undersampling (RUSboost) to account for class imbalance. Results were averaged over 10 iterations of the AdaBoost1 algorithm. For the combined method, the results from the three individual analyses were combined, and a NSF classification was assigned to patients with at least two positive (NSF) classifications. For each group (SF, NSF), we show the number of correctly identified cases by absolute number and relative frequency. The remaining columns correspond respectively to the accuracy (Acc.), precision (Prec.), sensitivity and F1 statistic. For machine learning analyses only the average fraction of correctly predicted cases is shown in the true negatives and true positives columns, since absolute results can vary per realization of the prediction algorithm.

DISCUSSION

Personalized models of brain dynamics can aid the treatment of patients with neurological disorders. In this study we presented ESSES (Epidemic Spreading Seizure and Epilepsy Surgery model): a framework to aid epilepsy surgery planning on a patient-by-patient basis. ESSES defines individualized seizure propagation models that integrate multimodal presurgical data, and can propose alternative resection strategies and provide confidence bounds for the probability of success of a given strategy. The implementation of ESSES in clinical practice may thus eventually improve the chances of achieving a good postsurgical outcome.

In this study we proposed a combined setting such that ESSES' parameters were fitted in a retrospective study ($N = 15$) using iEEG data of ictal activity, in analogy with previous studies (Goodfellow et al. (2016); V. Jirsa et al. (2017); Kini et al. (2019); Makhalova et al. (2022); Moosavi et al. (2022); H. E. Wang et al. (2023)). We validated that ESSES captured the main aspects of seizure propagation and was able to reproduce the iEEG-recorded seizures, in agreement with our previous studies (Millán et al. (2022, 2023)). Remarkably, the goodness-of-fit of ESSES-modeled seizures to iEEG data could identify patients with a bad outcome with $AUC = 0.79$, 100% sensitivity and 57% precision. Such information may be integrated in the presurgical evaluation of the patients for whom iEEG data is available: different resection strategies may be tested as the origin of the ESSES-modeled seizures (Millán et al. (2023)), with a low goodness-of-fit being indicative of a low chance of seizure freedom. In particular, a bad prediction by the model would indicate (in this cohort) a 57% chance of a bad outcome (to be compared with only a 26.7% NSF rate in this cohort). Conversely, all patients identified as SF by the model could proceed to surgery with high expectations (100% in this group) of good outcome.

The novel aspect of this study consisted of a subsequent *pseudo-prospective study* with an *independent cohort* and in a *blind setting*. Importantly, we did not require the presence of iEEG data in the pseudo-prospective study, and instead the multimodal presurgical information available for each patient was integrated into seed-probability maps. In this manner ESSES can be adapted to the information available for each patient, in a quantitative and systematic manner. IEEG data is highly invasive and burdensome for the patient, and thus not always part of the presurgical evaluation. For instance, only 19 of the 34 patients of the validation cohort had undergone it. Thus, by not requiring iEEG data ESSES can be applied to a much larger patient population than traditional approaches (Goodfellow et al. (2016); V. Jirsa et al. (2017); Kini et al. (2019)), with the expected wider impact.

ESSES may be applied prospectively as follows. First of all, ESSES may suggest optimal resection strategies, in analogy with previous studies (An et al. (2019); Laiou et al. (2019); Millán et al. (2022); Nissen et al. (2021)), with the advantage that all multi-modal presurgical information available for each patient is integrated into ESSES, instead of considering only one source used for network reconstruction. We note that these resections are optimal within the framework of the model, and this does therefore not guarantee optimal clinical outcome. Nevertheless, we have found that these virtual resections have good predictive value of surgical outcome. The *optimal resection strategy*, defined here as the smallest

resection leading to a 90% decrease in (modeled) seizure propagation, can be used as a first indicator of the chances of seizure-freedom after *any* surgery. In our pseudo-prospective predictive framework (emulating the presurgical conditions) the size of this resection could predict 50% of patients with bad outcome (table 3), whereas the relative NSF rate in this group was 23.5%. This result is independent of the resection strategy and it is completely characterized by the presurgical information available for each patient. Thus, a bad prognosis could indicate that either the presurgical information available is not of sufficient quality, or that the patient is unlikely to be seizure-free with any resection strategy.

ESSES can also provide information about the prognosis after a particular resection by i) comparing it to the optimal ESSES resection strategy and ii) quantifying its effect on seizure propagation in the patient-specific ESSES model. Here we found that resections with a larger overlap with the optimal virtual resection were more likely to lead to seizure freedom, in agreement with previous studies (Goodfellow et al. (2016); Kini et al. (2019); Makhalova et al. (2022)). Similarly, resections leading to a larger decrease in seizure propagation in ESSES were associated with a larger probability of seizure-freedom after the resection, in agreement with other modeling (Goodfellow et al. (2016); Kini et al. (2019)) and network-based (Bartolomei et al. (2017); Lopes et al. (2017); Nissen et al. (2017)) studies. Here we considered only the planned resection strategy, which was approximated here by the resection area, since this information could be derived in a systematic manner, and this set-up allowed us to validate ESSES' findings. In a presurgical setting, different strategies could be tested to measure the probability of seizure freedom after each one. In particular, we found that, when combining the information from the three model-based biomarkers (namely the size of the optimal resection, its overlap with the planned resections, and the effect of the planned resection on modeled seizure propagation) could predict pseudo-prospectively 81% and 75% of SF and NSF cases (see table 3), whereas the relative group ratios were 76.5% and 23.5%, respectively. Clinically, this implies that if a good prognosis is found by at least two biomarkers, then there is a 91.3% (true negative rate, 21 cases were SF of the 23 predicted by the model) chance that the patient will be seizure-free, and the patient can proceed with the surgery with the knowledge that they will likely have a good outcome. Conversely, a bad prognosis by at least two biomarkers indicates a 55% chance of bad outcome, and may be interpreted as an ESSES suggestion to perform more presurgical testing or consider alternative resection strategies. Importantly, epilepsy surgery may still improve the quality of life of the patient even when complete seizure freedom can not

be achieved. Thus, moderate *a priori* chance of a bad outcome is not necessarily a contraindication for surgery, but it is important in the presurgical counseling of the patients.

Our findings here did not depend on the presence of iEEG data, and even when iEEG data were available we only included a low-resolution description of them. IEEG data does provide the most detailed information of epileptogenic activity, and is it often the most valuable tool to identify the epileptogenic zone or predict surgical outcome for patients with complicated ethiology (Bernabei et al. (2023); Gunnarsdottir et al. (2022); Makhalova et al. (2022); Proix et al. (2017); Runfola et al. (2023); Sinha et al. (2017); Y. Wang et al. (2023)). In fact, for the modeling cohort we found the best classification results when using the goodness-of-fit of ESSES-predicted seizure propagation patterns to the iEEG seizures, in agreement with previous studies (Makhalova et al. (2022)). IEEG imaging however is burdensome to the patient, has risk of complications, and has limited spatial coverage. A first prediction of surgical outcome could thus be performed with ESSES when the results of non-invasive testing have been obtained, and an iEEG study might be avoided if the model already predicts a good outcome with the existing data.

In summary, we showed here that ESSES could identify patients with good outcome presurgically based on i) the smaller size of the optimal ESSES resection strategies, ii) a larger overlap of the planned resection strategy with the optimal ESSES resection, and iii) a larger effect of the planned resection strategy on decreasing (modeled) seizure propagation. Our findings here indicate that ESSES could be generalized to other patient populations (as we did with the validation cohort), with the only requirement of a patient-specific brain network, and can incorporate multimodal information from the existing presurgical evaluation, in particular without requiring the presence of iEEG data. The ESSES-based biomarkers identified here could be taken into account during presurgical planning to evaluate the need for more testing, or may lead to the decision to forgo the surgery, if a bad outcome is predicted. This extra information may be particularly valuable for patients with complicated ethiology (e.g. discordant information from different modalities, variable seizure propagation patterns, multiple seizure onset zones), for whom the discussion of whether or not to perform the surgery is challenging.

ESSES modeling framework

ESSES consists of different interconnected elements, namely i) the underlying network structure; ii) the seizure propagation model (and parameter fitting); iii) the seizure onset zone model; iv) the virtual

resection model; and v) the virtual resection optimization algorithm. Each of these different elements was designed to model a particular aspect of epilepsy surgery in a synergistic manner. For instance, the emergent properties of the seizure propagation model (the SIR model) led the design of the virtual resection optimization algorithm. At the same time, the modular organization of the framework allows for the independent improvement or modification of each of the modules. In fact, different modules were developed and analyzed in detail in our previous studies. For instance, the virtual resection algorithm model was initially designed in Nissen et al. (2021) and improved in Millán et al. (2022), whereas the seizure propagation and parameter fitting model as used here was mainly defined in Millán et al. (2023). Below we discuss the main modeling considerations and results for each ESSES module.

As the underlying network structure we considered MEG-derived whole-brain networks as a proxy for structural connectivity, following our previous works (Millán et al. (2022, 2023)), and in contrast with other works (An et al. (2019); V. Jirsa et al. (2017); Nissen et al. (2021); Sip et al. (2021)). MEG provides highly temporally resolved information with good spatial resolution and uniform coverage. Our previous studies showed that MEG networks based on the amplitude envelope correlation (AEC) can integrate information from both short-range structural connections (by not correcting for volume conduction) and long-range functional coupling. Thus, AEC-MEG networks can be used as a cost-effective proxy for structural connectivity (Millán et al. (2022)) with much lower computational cost than DWI (Diffusion-Weighted Imaging) networks, whilst also being more sensitive to long range connections, in particular inter-hemispheric ones, that may often be missed by DWI (Chen et al. (2015)). It would be an interesting question for future studies to discriminate whether structural or functional connections drive seizure propagation, in analogy to recent studies on the spreading of abnormal proteins associated with Alzheimer's disease (Schoonhoven et al. (2023)).

The MEG networks were thresholded at different levels to prune out spurious connections, following previous studies (Millán et al. (2022, 2023); Nissen et al. (2021); Schoonhoven et al. (2023)). This requires the use of an arbitrary threshold, which we fitted to the iEEG data. In all cases we considered sparse networks (the maximum density considered was 0.35), and the operating point of ESSES was set at a very low density (0.03). This small density prevented weak or negative correlations from being included in the thresholded network. The proposed thresholding method can become a limitation if denser networks, including more connections, are considered.

ESSES was based on a simple epidemic spreading model, the SIR model. Epidemic spreading models, such as the SIR or SIS (Susceptible-Infected-Susceptible) models, describe the basic aspects of spreading phenomena on networked systems (Pastor-Satorras et al. (2015)), and have been used to describe other neuro-physiological processes before, such as the spreading of pathological proteins on brain networks (Peraza et al. (2019); Schoonhoven et al. (2023)) or the relation between brain structure and function (Stam et al. (2016)). Epidemic spreading models have been extensively studied on different network substrates (Pastor-Satorras et al. (2015)) and are supported by a well-grounded mathematical and computational framework that we can use to our advantage in the context of epilepsy surgery. For instance, from an epidemic spreading perspective, it is to be expected that hub removal plays a major role in the decrease of seizure propagation, as found experimentally (Lopes et al. (2017); Nissen et al. (2017)), with the spreading threshold heavily influenced by the existence of hubs (Pastor-Satorras et al. (2015)). This theoretical background guided the design of an efficient virtual resection optimization algorithm, such that the decrease in seizure propagation after a virtual resection could be approximated by the decrease of centrality of the seed regions.

As we showed here and in previous works, epidemic spreading models can also reproduce the fundamental aspects of seizure propagation at the whole-brain level in epilepsy patients (Millán et al. (2022, 2023)). As ESSES's working point we chose here the values of the global parameters that led to the maximum average goodness-of-fit of the modeling cohort (figure 1). Importantly, ESSES was still individualized for each patient by means of the patient-specific brain connectivity, setting the local spreading probabilities, and the patient-specific seed regions (based on the seed-probability maps built with multi-modal presurgical information). As we showed in our previous study (Millán et al. (2023)) and in the supplementary information here (Supp. section 5.2), by not individualizing the global model parameters (namely ρ and γ) for each patient we were able to reduce noise effects by integrating together ictal data from different patients. Moreover, this formulation allowed us to generalize ESSES to patients for whom iEEG seizure-propagation patterns were not available.

Our findings in this study indicated that the iEEG seizure propagation patterns were significantly better explained by ESSES for SF patients, and in fact all NSF cases could by identified by a bad ESSES fit, and 73% of the SF cases by a good fit. There are several possible explanations for these findings. Given that the epidemic seed was based on the resection area for each patient in this part of the analyses, a

simple explanation is that the resection strategy might have been better for SF patients given the existing information. However, the difference could also arise from the iEEG data: the sampling may have been inadequate for NSF patients (Sip et al. (2021)), or these may have presented seizure *dynamotypes* (Saggio et al. (2020)) that were not well-explained by the considered epidemic spreading model (SIR model). The fact that the optimization of virtual resections analysis –which did not depend on the clinical resection area– also found differences between the SF and NSF groups points towards an intrinsic difference between the presurgical data of the two groups, and not only to a sub-optimal surgical strategy for the NSF group.

The next ingredient of ESSES was the definition of the seizure onset zone in the model, that is, the set of brain regions from which seizures originate. In this study we presented a method to combine the multimodal presurgical information available for each patient into *seed-probability maps*. This set-up thus emulated the clinical situation prior to the surgery, where a surgical strategy has been devised based on the information that is available from the presurgical evaluation. It would also allow for flexibility in the clinical application of ESSES: if more evaluations become available these could be readily integrated into the seed-probability map to update ESSES's results.

The final key ingredients of ESSES were the simulation and optimization of resection strategies. Here we considered a node-based resection such that the resected nodes were disconnected from the network. This approach however does not take into account possible widespread effects or plasticity mechanisms, which could also be included into the model (Demuru et al. (2020)). The virtual resection optimization algorithm was originally validated in our previous studies (Millán et al. (2022); Nissen et al. (2021)). Given that optimizing virtual resections is highly computationally demanding, the algorithm took advantage of the mathematical link between network structure and SIR dynamics to reduce the dynamics-based optimization problem (i.e. finding the resection leading to a minimum seizure propagation) into a network optimization problem (i.e. finding the resection leading to a minimum seed efficiency). This was also motivated by our previous finding that the effect of a resection on the model depended strongly on the centrality of the seed regions after the resection (Millán et al. (2022); Nissen et al. (2021)). In particular, Nissen et al. (2021) found that removing connections to the network hubs was the most efficient way to decrease seizure propagation, whereas Millán et al. (2022) verified a strong correlation between a decrease in closeness centrality of the seed and a decrease in seizure propagation

following a virtual resection. The effect of a resection on seizure propagation is also influenced by other network and model properties, and as a consequence the optimal network-based and SIR-based resections may differ slightly (Millán et al. (2022)). However, the intrinsic noise in the seed definition, in the seed-probability maps, and in the actual origin and propagation patterns of iEEG-recorded seizures created variability in the clinical data that absorbed the differences between the network-based and SIR-based optimal resections (which we previously found to be small anyway (Millán et al. (2022))).

The virtual resection optimization algorithm considered here imposed no conditions on the location of the resected regions, nor did it force that the resection strategy was made up of only one set of adjacent regions. Conditions on the resection strategies could be imposed, such as preserving eloquent cortex or forbidding bi-hemispheric resections (An et al. (2019); Laiou et al. (2019)). This would limit the dimensionality of the space of possible resection strategies and simplify the computations. However, by not imposing any conditions here we derived an *optimal* ESSES resection against which other, perhaps clinically more realistic, strategies could be tested (by e.g. measuring their overlap as we did here).

Modeling considerations and limitations

There are inherent limitations in the modeling of virtual resections, as the findings cannot be directly tested and we often rely on retrospective data. Here we have attempted to simulate how an epilepsy surgery model could be used in the clinic, i.e. prospectively, by considering only the presurgical information that is typically available to the clinical team. However, the optimal resections suggested by ESSES can still not be tested in practice, and in fact can only be considered optimal within the context of the model. Only long-term testing of the framework in the clinic can truly validate the use of computational models in epilepsy surgery.

ESSES is an abstraction of seizure dynamics that does not aim to reproduce the detailed bio-physiological processes involved in seizure generation and propagation, but aims to focus only on the most relevant features of seizure propagation (Millán et al. (2022, 2023); Nissen et al. (2021); Sip et al. (2021)). In order to validate ESSES as a framework to simulate seizures, we compared the modeled seizures with those recorded via iEEG. This required, however, a simplified representation of the iEEG data. In particular, as there was no intrinsic time-scale in the SIR model, and to avoid introducing an arbitrary one, we reduced the iEEG data to a pattern that describes the activation order of the sampled

ROIs. Furthermore, even if ESSES provides a good representation of the iEEG seizures, extrapolating these results to the simulation of the effect of a resection is not trivial. Moreover, our virtual resection technique assumed that the effect of a surgery could be approximated simply by removing or disconnecting the resected regions, whereas in practice widespread effects and compensation mechanisms are expected (Demuru et al. (2020)). Here we validated ESSES' results against postsurgical outcome, but seizure freedom is not a perfect gold standard either. For instance, in cases with a good outcome a smaller resection could potentially also have led to seizure freedom (Millán et al. (2022); Nissen et al. (2021)).

All modeling frameworks are affected by the need to (sometimes arbitrarily) choose modeling parameters, which go from the data reduction process to the choices of thresholds and metrics for the final analyses. Here we considered well-established data preprocessing techniques (Hillebrand et al. (2016)). ESSES was validated in previous studies (Millán et al. (2022, 2023); Nissen et al. (2021)), and importantly we found that the results held for an independent cohort, and that modeling details (such as the simulation algorithm for the SIR model) did not affect the main results (Millán et al. (2023)). A simple model to simulate seizure propagation (the SIR model), also reduced the number of modeling parameters so that the findings could be more easily generalized. Some arbitrary choices were still needed, such as the definition of the 90% threshold to select the optimal resection strategy. However we validated that similar results were obtained when another resection (the disconnecting resection) was considered.

The seed-probability maps were based on an existing low-resolution database (*Castor Electronic Data Capture.* (n.d.)). Seed regions were consequently widespread over the network. This also led to a large variability in the results of different simulations for each patient (see for instance figures 4A,B and 5A), as these depended strongly on the seed realization. In order to improve the resolution of the model and minimize noise, the data from each modality could be integrated directly into the model, skipping the 34-region description in the database.

Finally, a limitation of this study is the small size of the non-seizure-free group, with only 4 cases in the modeling cohort and 8 in the validation cohort. This small size limited the classification and prediction analyses, and prevented us from building a more sophisticated machine learning model based on our analysis. With the proposed leave-one-out-crossvalidation method, combined with random undersampling and a small input space (only three data-points per patient), we attempted to overcome

these limitations, but we were not able to improve upon the simpler ROC-based prediction results. Future studies involving more that one center have the potential to at least diminish this limitation.

CONCLUSION AND OUTLOOK

Individualized computational models of seizure propagation and epilepsy surgery based on patient-specific brain connectivity can reproduce individual iEEG seizure propagation patterns and aid epilepsy surgery planning by proposing alternative resection strategies and providing estimates on the likelihood of seizure freedom after the surgery. Here we presented the ESSES framework for seizure propagation and epilepsy surgery. ESSES combines SIR epidemic spreading dynamics over patient-specific MEG brain connectivity with a virtual resection framework. We defined a method to derive patient-specific regional epileptogenicity maps from the presurgical evaluations of the patients in a systematic and quantitative manner, and integrated them into ESSES. We performed a pseudo-prospective study emulating the use of ESSES in clinical practice, prior to surgery. In the pseudo-prospective analyses we did not require the presence of iEEG data, demonstrating that the model could be applied to larger patient populations. We found that the goodness-of-fit of ESSES to the iEEG seizures (in a retrospective study), the effect of the planned resection strategy, as well as the size of ESSES optimal resections and their overlap with the planned resection, predicted surgical outcome with $0.68 - 0.76$ AUC and $0.50 - 0.63$ sensitivity to identify non-seizure-free patients. Our results thus prescribe the use of ESSES during the presurgical evaluation to evaluate the need for further presurgical testing on a case-by-case basis or, conversely, support the decision to proceed with surgery in the case of a good-outcome prediction. For cases where a bad outcome is predicted, the surgical plan may be altered to include ESSES's results.

METHODS

The general design of the study is detailed in figure 2 and Supp. figure 7. Namely, we first set the hyperparameters of ESSES using a *modeling cohort* ($N = 15$) for which seizure propagation patterns derived from iEEG recordings were available. Then, ESSES was fitted with multimodal patient-specific data (in the form of seed-probability maps), and it was used to a) identify optimal resection strategies for each patient and b) predict the chance of a good outcome after a given resection. Then, ESSES was

applied to a *validation cohort* ($N = 34$) in a pseudo-prospective analysis with a blind setting to emulate the presurgical conditions. That is, during the application of ESSES to determine optimal resection strategies, the researchers were blind to the actual clinical resection and surgical outcome of each patient. This data was subsequently de-blinded in two stages. First, the resection areas were obtained to be used as a proxy for the surgical plan of each patient to a) compare them with ESSES's optimal resection strategy, and b) simulate the effect of the surgical plan in ESSES. Finally, we de-blinded the one-year surgical outcome to enable a statistical validation of the results.

Patient groups

We included two patient groups in this study, the *modeling cohort* for the model definition (retrospective study) and the *validation cohort* for the pseudo-prospective validation. All patients had undergone resective surgery for epilepsy at the Amsterdam University Medical Center, location VUmc, between 2013 and 2019. All patients had received an MEG recording, and underwent pre- and post-surgical magnetic resonance imaging (MRI). All patients gave written informed consent and the study was performed in accordance with the Declaration of Helsinki and approved by the VUmc Medical Ethics Committee. The excluding criterion was the existence of a prior brain surgery.

Both patient groups were heterogeneous with temporal and extratemporal resection locations and different etiology (see Supp. tables 1 and 2 for details). Surgical outcome was classified according to the Engel classification at least one year after the surgery (Engel Jr (1993)). Patients with Engel class 1A were labelled as seizure-free (SF), and patients with any other class were labelled as non-seizure-free (NSF). The modeling cohort consisted of 15 patients (4 NSF, 11 females) who had also undergone an iEEG (invasive electroencephalography) study, including post-implantation CT-scans. This same cohort was already included in Millán et al. (2023), and partially in Millán et al. (2022). The validation cohort consisted of 34 patients (8 NSF, 13 females). No extra requirements (other than the presence of an MEG recording of sufficient quality) were placed. In order to maintain the pseudo-prospective setting, the research team was blind to the resection area and outcome of the validation cohort patients. In order to perform the final analyses, for which this information was needed, the data was coded to avoid identification. For two cases of the validation cohort (cases 2 and 9) the data of surgical outcome was

de-blinded together with the data of the resection area as the research team became aware of a subsequent resective surgery (indicative of a bad outcome of the first surgery).

Individualized Brain Networks

Seizure propagation was modeled on the patient-specific brain networks, as derived from MEG data, for both cohorts (see Supp. figure 7). For each patient, a 10 to 15 minutes eyes-closed resting-state (supine position) MEG recording was used to derive broadband (0.5 - 48.0 Hz) MEG functional connectivity. All instrumental and methodological details were equal to our previous studies (Millán et al. (2022, 2023)) and are detailed in the supplementary information (Supp. section 2). Functional networks were generated considering each of the 246 ROIs of the Brainnetome (BNA) atlas (Fan et al. (2016)) as nodes. The elements w_{ij} of the connectivity matrix, indicating the strength of the connection between ROIs i and j, were estimated by the AEC (Amplitude Envelop Correlation) (Brookes et al. (2011); Bruns, Eckhorn, Jokeit, and Ebner (2000); Colclough et al. (2016); Hipp, Hawellek, Corbetta, Siegel, and Engel (2012)), without including a correction for volume conduction. The uncorrected AEC maintains information about the structural connections, which are mainly determined by the distance between each ROI pair, by not correcting for volume conduction. We validated the relationship between AEC-MEG and structural networks in a previous study (Millán et al. (2022)) by comparing them with a well-validated model for structural connectivity: the exponential distance rule (EDR) network. Based on animal studies, the EDR specifies that the weights of structural connections in the brain, w_{ij}, decay exponentially with the distance between the ROIs d_{ij} (Ercsey-Ravasz et al. (2013); Gămănuţ et al. (2018); Theodoni et al. (2022)), i.e. $w_{ij} \propto exp(-\alpha d_{ij})$. Recent studies have corroborated this behavior also in human structural connectivity (Deco and Kringelbach (2020); Deco et al. (2021); Roberts, Perry, Roberts, Mitchell, and Breakspear (2017)), although the EDR cannot capture all details of white matter connectivity, as this is not isotropic (Betzel and Bassett (2018); Jbabdi, Sotiropoulos, Haber, Van Essen, and Behrens (2015); Markov et al. (2013)), and includes long-range connections that are missed by the EDR (Roberts et al. (2016)). However, the EDR is enough to capture the overall scaling of structural connections with the distance as observed in the human structural connectome. In Millán et al. (2022) we validated that AEC-MEG networks were strongly correlated ($R^2 = 0.50$) with the corresponding EDR networks, therefore showing that AEC-MEG reproduces at least partially the overall organization of structural connectivity. Moreover, AEC-MEG networks also include long-range connections that may promote

seizure propagation, but that may be missing from structural (i.e. DWI) networks (Jones, Knösche, and Turner (2013); Reveley et al. (2015)). Thus, uncorrected AEC-MEG networks are a convenient way to construct a network that resembles a structural network and includes long-range connections.

AEC values were re-scaled between 0 (perfect anti-correlation) and 1 (perfect correlation), with 0.5 indicating no coupling (Briels et al. (2020)). Functional networks were thresholded at different network densities ρ indicating the fraction of links remaining in the network. We note that the networks were thresholded but not binearized, so that w_{ij} could take values between 0 and 1. The density thresholds were chosen to be logarithmically distributed between 0.01 to 0.35. The weakest non-zero link included in the network had an average weight of 0.54 (range: 0.52 - 0.56) for $\rho = 0.35$. At ESSES's operating point (best model fit) the density was $\rho = 0.03$, and the weakest non-zero weight was 0.71 (range: 0.67 - 0.76).

Resection Area

The resection area (RA) was determined from the three-month post-operative MRI. For the modeling cohort the resection areas were obtained as part of two previous studies (Millán et al. (2022, 2023)). For the validation cohort, to maintain a completely blind setting for the first analysis (*Optimization of alternative resections*), the resection areas were obtained during a second pre-processing step, as described in figure 2. Cases 9 and 20 of the validation cohort underwent the post-operative MRI on a different MRI scanner at their resection center, respectively one day and three weeks after the surgery. Case 9 also lacked a 3-month postoperative MRI, an MRI from 2 years after the surgery was used instead.

The post-resection MRIs were co-registered to the pre-operative MRI using FSP FLIRT (version $4.1.6$) 12 parameter affine transformation. The resection area was then visually identified and assigned to the corresponding BNA ROIs, namely those for which the centroid had been removed during surgery.

iEEG Seizure Propagation Pattern

Patients in the modeling cohort underwent invasive EEG recordings using stereotactic electrode implantation as described in Millán et al. (2023). One characteristic iEEG-recorded seizure from each patient was used to derive a seizure propagation pattern in terms on the BNA ROIs, the *iEEG seizure pattern*, as described in Millán et al. (2023) and in the Supp. section 3.

Seizure Propagation Model

ESSES was based on our previous studies (Millán et al. (2022, 2023); Nissen et al. (2021)) where we showed that simple epidemic spreading models could reproduce the spatio-temporal seizure-propagation patterns derived from invasive EEG recordings, and that they could be used to simulate the effect of different resection strategies *in silico*. ESSES was based on a well-known epidemic spreading model: the Susceptible-Infected-Recovered (SIR) model (Pastor-Satorras et al. (2015)), which was simulated on the patient-specific MEG brain network. The SIR model simulated the propagation of ictal activity from a set of *seed* regions that were set to be infected at the beginning of the simulation to the remaining nodes in the network, and the subsequent recovery of infected nodes. The SIR dynamics were defined by two parameters: the probability β_{ij} that each infected node i propagates the infection to a neighbour j ($S \to I$), and the probability γ_i that each infected node i recovers ($I \to R$). For simplicity, we considered here a global recovery probability $\gamma_i = \gamma$, and spreading probabilities given by the MEG network connectivity: $\beta_{ij} = w_{ij}$. Thus, the spreading rate was determined by the density of connections in network ρ. The two control parameters of ESSES are thus the network density ρ, and the recovery probability γ. Depending on the network structure, the epidemics can show different spatio-temporal spreading profiles described by the probability $p_i(t)$ that each ROI i becomes infected at step t.

ρ and γ were fitted to the iEEG seizure-propagation patterns at the group level. The resection area was set as the seed of epidemic spreading, and an *ESSES seizure propagation pattern* was built that described the set of infected and non-infected ROIs during the SIR-simulated seizures, as well as the order in which infected ROIs became infected. In order to take into account the stochastic nature of the SIR dynamics, the participation of each ROI was weighted by the fraction of realizations in which it was involved in the simulated seizure (since different ROIs became infected in different realizations). The *goodness-of-fit* of the model, $C(\rho, \gamma)$ (Millán et al. (2023)), quantified how similar the ESSES and iEEG patterns were. It took into account two factors: the weighted correlation between activation orders of ROIs that were active in both patterns, C_w, and the overlap between the active and inactive ROI sets of both patterns, P_{overlap}, i.e.

$$C = C_w \cdot P_{\text{overlap}}. \tag{1}$$

The details of this definition can be found in Supp. section 5.2.

We estimated C for a range of values ρ and γ logarithmically distributed (between 0.01 and 0.35 for ρ and between 0.01 and 1.00 for γ), considering $N_R = 10^4$ iterations of the SIR dynamics 10 times in order to determine average C values and their fluctuation for each patient. We then found the parameter set that maximized C for each patient (see Supp. section 5.2 and Supp. figure 2) and at the group level (figure 1A). The model parameters that lead to the best fit at the population level defined the ESSES model and were carried over to the pseudo-prospective analyses. Importantly, even though the SIR global parameters were set equal for all patients, ESSES was individualized for each patient by means of their patient-specific MEG brain connectivity, which defined the spreading probabilities, and their patient-specific seed-probability map, which defined the seed regions.

The SIR dynamics was simulated by an adaptive Monte Carlo method (the BKL algoritihm) in Matlab in discrete time, such that at each time step one new node became infected. $N_R = 10^4$ iterations of the dynamics were run for each model configuration in all analyses.

Presurgical hypothesis of the seed regions

We built seed-probability maps indicating the probability that each ROI started a seizure, for each patient of both cohorts. This is a key difference with our previous studies, where the seed regions were either derived from the resection area (Millán et al. (2022, 2023); Nissen et al. (2021)), which can only be known after the surgery, or from the iEEG data (Millán et al. (2022, 2023)). Here we defined a framework to integrate data from the different presurgical evaluations that were available for each patient, which was encoded in an existing database (Castor EDC, Ciwit B.V., Amsterdam (*Castor Electronic Data Capture.* (n.d.))).

To compute the seed-probability maps, we considered the information available from 6 presurgical modalities: i) presence of ictal activity in EEG, ii) MRI lesions, iii) MEG abnormalities, iv) PET lesions, v) SPECT abnormalities and vi) iEEG recordings of ictal activity. All patients had undergone an EEG, MRI and MEG study, but not all of them presented PET, SPECT or iEEG data. The presence (1) or absence (0) of data of each modality was encoded in a variable $D_m = 0, 1$, $m = 1, 2, ..., 6$, for each patient.

The database included information at the level of 34 regions, consisting of 6 frontal regions (fronto-orbital, frontal-basal, frontal-parasagitaal, frontal-periventricular, frontal-lateral,

frontal-operculum), 6 temporal regions (hippocampus, amygdala, uncus, anterior-neocortical, posterior-neocortical, gyrus-parahippocampalis), 2 insular regions (anterior and posterior insula), 1 central, 1 parietal and 1 occipital region, for each hemisphere. The temporal and frontal lobes are the most often involved in EZ and resection strategies, and thus are described in more detail in the database.

For each region i and modality m, the database indicates the presence (1) or absence (0) of abnormalities, from which we derived binary abnormality maps $a_{i,m} = 0, 1$. The overall abnormality map A_i was obtained by aggregating over all modalities available for each patient. Not all modalities are equally relevant to establish the probability that a region is involved in epileptogenic activity: EEG is the least focal, whereas iEEG provides the most localized information, and its results also integrate information from the other modalities (as these affect where the iEEG electrodes are placed). In order to gauge these differences, we weighted each modality m by a relevance factor ω_m, with $\omega = 1$ for EEG, 2 for MRI, MEG, PET and ISPECT, and 4 for iEEG. Thus, the overall abnormality map was defined as

$$A_i = n^{-1} \sum_{m=1}^{6} D_m \omega_m a_{i,m}, \tag{2}$$

where the normalization factor n is defined as $n = \sum_{m=1}^{6} D_m \omega_m$

A clinician (ECWvS) defined a unique projection of the regions in the database on to the BNA ROIs. In most cases the database regions corresponded to well-defined gyri that are also well-described in the BNA documentation. A table describing the projection is included as supplementary material. We projected the abnormality map A_i from the low-resolution description into the BNA atlas to obtain the seed-probability maps SP_i, with $i = 1, 2, ..., 246$. Given that the description provided by the database was broad and homogeneous (i.e. the considered ROIs are much larger than the BNA ROIs), and that co-occurrence of abnormalities in different modalities is a strong indicator of the epileptogenic zone, we included a re-scaling factor R to produce more focal seed-probability maps: $SP_i = (A_j)^R$, where j is the region in the database corresponding to the BNA ROI i. We found that for $R > 2$ the results did not depend strongly on R, and report here for $R = 3$.

Virtual Resections

We conducted virtual resections of sets of nodes R by disconnecting them from the network, by setting to 0 all their connections. The effect of each resection was characterized by the normalized decrease in

seizure propagation $\delta IR(R)$ in the resected network (R) with respect to the original (0) one:

$$\delta IR(RA) = (IR_0 - IR_R)/IR_0, \tag{3}$$

where IR is the fraction of nodes that became infected at any point during the modeled seizure, namely,

$$IR = I(t \to \infty) + R(t \to \infty). \tag{4}$$

That is, IR takes into account all nodes that became infected during the simulated seizure, regardless of whether they eventually recovered or not. This characterizes the total extent of the simulated seizure.

We performed two virtual resection studies, as detailed in figure 2. Firstly, we performed an *Optimization of alternative resections* analysis. We derived optimal virtual resections R of increasing sizes $S(R)$ (defined as the number of resected nodes) with an optimization algorithm based on simulated annealing (Kirkpatrick, Gelatt, and Vecchi (1983)) and derived in our previous studies (Millán et al. (2022); Nissen et al. (2021)). The optimization method took advantage of the relationship between SIR spreading and network structure to use a structural metric –the seed efficiency– as a proxy for the actual effect of the resection on seizure propagation $\delta IR(R)$. Thus, for each resection size $S(R)$, the simulated annealing algorithm searched for the resection R that minimized the seed efficiency $E_R(\text{seed})$ (Barrat et al. (2008); Brockmann and Helbing (2013); Pinto, Thiran, and Vetterli (2012)). $E_R(\text{seed})$ measures the inverse average distance from the seed nodes to the remaining nodes in the network:

$$E_R(\text{seed}) = \frac{1}{N_{\text{seed}} N_2} \sum_{i \in \text{seed}} \sum_{j \in \mathcal{S}_2} \frac{1}{d_{ij}}, \tag{5}$$

where d_{ij} is the distance (in the network sense) between nodes i and j, $\mathcal{S}_2$ is the set of nodes that do not belong to the seed, N_2 the size of this set, and N_{seed} the number of nodes that belong to the seed. In case of network disconnection, only nodes in the giant component were included in the seed and $\mathcal{S}_2$ sets.

All nodes were considered as possible targets of the resection. To compare between different patients we defined the normalized seed efficiency

$$e_R(\text{seed}) = E_R(\text{seed})/E_0(\text{seed}), \tag{6}$$

where $E_0(\text{seed})$ is the seed efficiency in the original (un-resected) network. The actual effect of each resection was quantified by the seizure propagation level after the resection, $IR(R)$, and the normalized

decrease in seizure propagation $\delta IR(R)$. We defined the *optimal ESSES resection* R_{op}, as the smallest resection leading to (at least) a 90% decrease in (modeled) seizure propagation. This resection was characterized by its size $S(R_{op})$ and overlap with the resection area $Ov(RA, R_{op})$. We also defined the *disconnecting resection* R_D as the smallest resection that lead to seed disconnection (see Supp. section 5.4 and Supp. figure 6).

In the second virtual resection study, we simulated the effect of the planned resection for each patient, to measure its effectiveness in reducing seizure propagation. The resection area was used as a proxy for the resection strategy (figure 2: Simulation of the resection plan), since it could be derived in a systematic manner from the data.

For all virtual resection analyses the seed regions were derived from the patient-specific seed-probability maps, and the underlying network was given from the patient-specific MEG network as before. In order to obtain precise results, the effect of each resection was averaged over 300 independent realizations of the seed regions from the seed-probability maps. As described in figure 2, for the validation cohort we first performed the *Optimization of alternative resections* in a blind setting. Then the resection areas were de-blinded and used as a proxy of the planed resection strategy to i) quantify the overlap of ESSES's optimal resections with the resection strategy and ii) measure the effect of the planed resection in decreasing (modeled) seizure propagation. Finally the one-year postoperative outcome was also de-blinded and used for the statistical analyses.

Statistics

The weighted correlation coefficient was used to determine the correlation between the iEEG and ESSES seizure propagation patterns for the modeling cohort. In all analyses, for comparisons between SF and NSF patients, we used a two-sided Wilcoxon ranksum test. Significance thresholds for statistical comparisons were set at $p < 0.05$.

We performed receiver-operating characteristic (ROC) curve analyses to study the patient classification based on i) the goodness-of-fit of the model (modeling cohort), ii) the size of optimal and disconnecting resections (modeling and validation cohorts), iii) the overlap between optimal resections and the planed resection (modeling and validation cohorts), and iv) the effect of the planed resection on modeled seizure

propagation (modeling and validation cohorts). A positive result was defined as bad outcome (non-seizure-free, NSF) classification.

In order to account for the noise in the SIR model, the spreading dynamics were averaged over 10^4 iterations of the SIR dynamics to derive each ESSES seizure pattern. The model fit analyses were repeated 10 times to obtain averaged values. For the Virtual resection analyses we performed 300 independent realizations of the seed regions and SIR dynamics. Each seed realization was used to measure seizure propagation in the original (before any resections) network and after the selected resection of each size. For the Optimization of resections analysis we also ran the simulated annealing algorithm 10 times for each resection size and selected the iteration that led to the minimal seed efficiency.

For the classification analyses we report the accuracy$= (TP+TN)/(TP+FP+FN+TN)$, precision $= TP/(TP+FP)$, sensitivity$= TP/(TP+FN)$, $F1$ statistic (harmonic mean between precision and sensitivity) $= 2TP/(2TP+FP+FN)$, and area under the curve AUC. For the prediction analyses, we built a predictive model for each patient using the data from the remaining patients, in a leave-one-out crossvalidation-type setting. The predictive model compounded the prediction results from these $N = 34$ models. We measured its accuracy, precision, sensitivity and $F1$ statistic.

In the final analysis of the study we performed a predictive Machine Learning analysis based on the AdaboostM1 algorithm (Matlab 2018) combined with random undersampling. AdaBoost is an adaptive boosting machine learning algorithm in which the weights of mis-classified instances are adjusted iteratively to improve the model. By combining adaptive boosting with random undersampling of the majority class (SF group), the classification algorithm effectively addresses class imbalance and reduces bias to the majority class and overfitting risks (*AdaboostM1 - Matlab 2018.* (n.d.); Friedman, Hastie, and Tibshirani (2000)).

For each patient, three variables were considered as input for the prediction analysis: the size of the optimal resection $S(R_{op})$, its overlap with the resection area $Ov(R_{op}, RA)$, and the effect of the resection strategy on modeled seizure propagation $\delta IR(RA)$). The goal of the machine learning algorithm was to predict surgical outcome. Due to the small cohort size, we performed a leave-one-out-cross-validation procedure, such that N_{pat} different training sets were created, each leaving out one patient, which was

then used to test the prediction model. The training sets were formed by randomly undersampling the majority class (SF) to the size of the minority (NSF) class. The small cohort size also prevented us from including a validation set and performing parameter-tuning. Thus, we used default hyperparameters of AdaboostM1 (see *AdaboostM1 - Matlab 2018.* (n.d.) for details): the number of learners in each model was set equal to the group size minus one, the learning rate was set to 1.0 (default) and results were averaged over 10 iterations of the undersampling and AdaboostM1 procedures for each classification model. The machine learning analysis was performed twice: first considering only the patients in the validation cohort ($N_{pat} = 34$), and secondly considering all patients (combined cohort, $N_{pat} = 49$).

Data availability

The data used for this manuscript are not publicly available because the patients did not consent for the sharing of their clinically obtained data. Requests to access to the data-sets should be directed to the corresponding author. All user-developed codes are publicly available on Github https://github.com/anapmillan/ESSES.

ACKNOWLEDGMENTS

Ana P. Millán and Ida A. Nissen were supported by ZonMw and the Dutch Epilepsy Foundation, project number 95105006. Ana P. Millán acknowledges financial support by the "Ramón y Cajal" program of the Spanish Ministry of Science and Innovation (grant RYC2021-031241-I), and by the Spanish Ministry and Agencia Estatal de investigación (AEI) through Project of I+D+i (PID2020-113681GB-I00), financed by MICIN/AEI/10.13039/501100011033 and FEDER "A way to make Europe", and the *Consejería de Conocimiento, Investigación, Universidad, Junta de Andalucía* and European Regional Development Fund (P20-00173) for financial support. Piet Van Mieghem has been funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101019718). The funding sources had no role in study design, data collection and analysis, interpretation of results, decision to publish, or preparation of the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

Adaboostm1 - matlab 2018. (n.d.). Available at: https://es.mathworks.com/help/stats/ensemble-algorithms.html.

An, S., Bartolomei, F., Guye, M., & Jirsa, V. (2019). Optimization of surgical intervention outside the epileptogenic zone in the virtual epileptic patient (vep). *PLoS Computational Biology*, *15*(6), e1007051. doi: 10.1371/journal.pcbi.1007051

Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks*. Cambridge University Press. doi: 10.1017/CBO9780511791383

Bartolomei, F., Lagarde, S., Wendling, F., McGonigal, A., Jirsa, V., Guye, M., & Bénar, C. (2017). Defining epileptogenic networks: contribution of seeg and signal analysis. *Epilepsia*, *58*(7), 1131–1147. doi: 10.1111/epi.13791

Baxendale, S., Wilson, S. J., Baker, G. A., Barr, W., Helmstaedter, C., Hermann, B. P., ... Smith, M.-L. (2019). Indications and expectations for neuropsychological assessment in epilepsy surgery in children and adults: Report of the ilae neuropsychology task force diagnostic methods commission: 2017–2021 neuropsychological assessment in epilepsy surgery. *Epileptic Disorders*, *21*(3), 221–234. doi: 10.1111/epi.16309

Bernabei, J. M., Li, A., Revell, A. Y., Smith, R. J., Gunnarsdottir, K. M., Ong, I. Z., ... Litt, B. (2023). Quantitative approaches to guide epilepsy surgery from intracranial eeg. *Brain*. doi: 10.1093/brain/awad007

Betzel, R. F., & Bassett, D. S. (2018). Specificity and robustness of long-distance connections in weighted, interareal connectomes. *Proceedings of the National Academy of Sciences*, *115*(21), E4880–E4889. doi: 10.1073/pnas.1720186115

Briels, C. T., Stam, C. J., Scheltens, P., Bruins, S., Lues, I., & Gouw, A. A. (2020). In pursuit of a sensitive eeg functional connectivity outcome measure for clinical trials in alzheimer's disease. *Clinical Neurophysiology*, *131*(1), 88–95. doi: 10.1016/j.clinph.2019.09.014

Brockmann, D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, *342*(6164), 1337–1342. doi: 10.1126/science.124520

Brookes, M. J., Hale, J. R., Zumer, J. M., Stevenson, C. M., Francis, S. T., Barnes, G. R., ... Nagarajan, S. S. (2011). Measuring functional connectivity using meg: methodology and comparison with fcmri. *Neuroimage*, *56*(3), 1082–1104. doi: 10.1016/j.neuroimage.2011.02.054

Bruns, A., Eckhorn, R., Jokeit, H., & Ebner, A. (2000). Amplitude envelope correlation detects coupling among incoherent brain signals. *Neuroreport*, *11*(7), 1509–1514. doi: 10.1097/00001756-200005150-00028

Castor electronic data capture. (n.d.). Available at: https://castoredc.com.

Chen, H., Liu, T., Zhao, Y., Zhang, T., Li, Y., Li, M., … Liu, T. (2015). Optimization of large-scale mouse brain connectome via joint evaluation of dti and neuron tracing data. *Neuroimage*, *115*, 202–213. doi: 10.1016/j.neuroimage.2015.04.050

Colclough, G. L., Woolrich, M. W., Tewarie, P., Brookes, M. J., Quinn, A. J., & Smith, S. M. (2016). How reliable are meg resting-state connectivity metrics? *Neuroimage*, *138*, 284–293. doi: 10.1016/j.neuroimage.2016.05.070

da Silva, N. M., Forsyth, R., McEvoy, A., Miserocchi, A., de Tisi, J., Vos, S. B., … Taylor, P. N. (2020). Network reorganisation following anterior temporal lobe resection and relation with post-surgery seizure relapse: a longitudinal study. *NeuroImage: Clinical*, *27*, 102320. doi: 10.1016/j.nicl.2020.102320

Deco, G., & Kringelbach, M. L. (2020). Turbulent-like dynamics in the human brain. *Cell reports*, *33*(10), 108471. doi: 10.1016/j.celrep.2020.108471

Deco, G., Perl, Y. S., Vuust, P., Tagliazucchi, E., Kennedy, H., & Kringelbach, M. L. (2021). Rare long-range cortical connections enhance human information processing. *Current Biology*, *31*(20), 4436–4448. doi: 10.1016/j.cub.2021.07.064

Demuru, M., Zweiphenning, W., van Blooijs, D., Van Eijsden, P., Leijten, F., Zijlmans, M., & Kalitzin, S. (2020). Validation of virtual resection on intraoperative interictal data acquired during epilepsy surgery. *Journal of Neural Engineering*, *17*(6), 066002. doi: 10.1088/1741-2552/abc3a8

Engel Jr, J. (1993). Outcome with respect to epileptic seizures. *Surgical treatment of the epilepsies*, 609–621.

Englot, D. J., Nagarajan, S. S., Imber, B. S., Raygor, K. P., Honma, S. M., Mizuiri, D., … Chang, E. F. (2015). Epileptogenic zone localization using magnetoencephalography predicts seizure freedom in epilepsy surgery. *Epilepsia*, *56*(6), 949–958. doi: 10.1111/epi.13002

Ercsey-Ravasz, M., Markov, N. T., Lamy, C., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., & Kennedy, H. (2013). A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron*, *80*(1), 184–197. doi: 10.1016/j.neuron.2013.07.036

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., … Jiang, T. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral Cortex*, *26*(8), 3508–3526. doi: 10.1093/cercor/bhw157

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The Annals of Statistics*, *28*(2), 337–407. doi: 10.1214/aos/1016218223

Gerster, M., Taher, H., Škoch, A., Hlinka, J., Guye, M., Bartolomei, F., … Olmi, S. (2021). Patient-specific network connectivity combined with a next generation neural mass model to test clinical hypothesis of seizure propagation. *Frontiers in Systems Neuroscience*, 79. doi: 10.3389/fnsys.2021.675272

Goodfellow, M., Rummel, C., Abela, E., Richardson, M. P., Schindler, K., & Terry, J. R. (2016, July). Estimation of brain network ictogenicity predicts outcome from epilepsy surgery. *Scientific Reports*, *6*(1), 29215. doi: 10.1038/srep29215

Gămănuţ, R., Kennedy, H., Toroczkai, Z., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., & Burkhalter, A. (2018, February). The mouse cortical connectome, characterized by an ultra-dense cortical graph, maintains specificity by distinct connectivity profiles. *Neuron*, *97*(3), 698–715.e10. doi: 10.1016/j.neuron.2017.12.037

Gunnarsdottir, K. M., Li, A., Smith, R. J., Kang, J.-Y., Korzeniewska, A., Crone, N. E., … Sarma, S. V. (2022). Source-sink connectivity: A novel interictal eeg marker for seizure localization. *Brain*, *145*(11), 3901–3915. doi: 10.1093/brain/awac300

Hebbink, J., Meijer, H., Huiskamp, G., van Gils, S., & Leijten, F. (2017). Phenomenological network models: Lessons for epilepsy surgery. *Epilepsia*, *58*(10), e147–e151. doi: 10.1111/epi.13861

Hillebrand, A., Tewarie, P., Van Dellen, E., Yu, M., Carbo, E. W., Douw, L., … Stam, C. J. (2016). Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proceedings of the National Academy of Sciences*, *113*(14), 3867–3872. doi: 10.1073/pnas.1515657113

Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M., & Engel, A. K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, *15*(6), 884–890. doi: 10.1038/nn.3101

Hutchings, F., Han, C. E., Keller, S. S., Weber, B., Taylor, P. N., & Kaiser, M. (2015). Predicting surgery targets in temporal lobe epilepsy through structural connectome based simulations. *PLoS Computational Biology*, *11*(12), e1004642. doi: 10.1371/journal.pcbi.1004642

Jbabdi, S., Sotiropoulos, S. N., Haber, S. N., Van Essen, D. C., & Behrens, T. E. (2015). Measuring macroscopic brain connections in vivo. *Nature neuroscience*, *18*(11), 1546–1555. doi: 10.1038/nn.4134

Jehi, L., Friedman, D., Carlson, C., Cascino, G., Dewar, S., Elger, C., … French, J. (2015). The evolution of epilepsy surgery between 1991 and 2011 in nine major epilepsy centers across the united states, germany, and australia. *Epilepsia*, *56*(10), 1526–1533. doi: 10.1111/epi.13116

Jirsa, V., Proix, T., Perdikis, D., Woodman, M., Wang, H., Gonzalez-Martinez, J., … Bartolomei, F. (2017). The virtual epileptic patient: individualized whole-brain models of epilepsy spread. *Neuroimage*, *145*, 377–388. doi: 10.1016/j.neuroimage.2016.04.049

Jirsa, V. K., Stacey, W. C., Quilichini, P. P., Ivanov, A. I., & Bernard, C. (2014). On the nature of seizure dynamics. *Brain*, *137*(8), 2210–2230. doi: 10.1093/brain/awu133

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion mri. *Neuroimage*, *73*, 239–254. doi: 10.1016/j.neuroimage.2012.06.081

Kini, L. G., Bernabei, J. M., Mikhail, F., Hadar, P., Shah, P., Khambhati, A. N., ... Litt, B. (2019). Virtual resection predicts surgical outcome for drug-resistant epilepsy. *Brain*, *142*(12), 3892–3905. doi: 10.1093/brain/awz303

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, *220*(4598), 671–680. doi: 10.1126/science.220.4598.671

Kramer, M. A., & Cash, S. S. (2012). Epilepsy as a disorder of cortical network organization. *The Neuroscientist*, *18*(4), 360–372.

Laiou, P., Avramidis, E., Lopes, M. A., Abela, E., Müller, M., Akman, O. E., ... Goodfellow, M. (2019). Quantification and selection of ictogenic zones in epilepsy surgery. *Frontiers in Neurology*, *10*, 1045. doi: doi.org/10.3389/fneur.2019.01045

Lopes, M. A., Richardson, M. P., Abela, E., Rummel, C., Schindler, K., Goodfellow, M., & Terry, J. R. (2017). An optimal strategy for epilepsy surgery: Disruption of the rich-club? *PLoS Computational Biology*, *13*(8), e1005637. doi: 10.1371/journal.pcbi.1005637

Lüders, H. O., Najm, I., Nair, D., Widdess-Walsh, P., & Bingman, W. (2006). The epileptogenic zone: general principles. *Epileptic Disorders*, *8*(2), 1–9.

Makhalova, J., Medina Villalon, S., Wang, H., Giusiano, B., Woodman, M., Bénar, C., ... Bartolomei, F. (2022). Virtual epileptic patient brain modeling: relationships with seizure onset and surgical outcome. *Epilepsia*. doi: 10.1111/epi.17310

Markov, N. T., Ercsey-Ravasz, M., Lamy, C., Ribeiro Gomes, A. R., Magrou, L., Misery, P., ... Kennedy, H. (2013). The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences*, *110*(13), 5187–5192. doi: 10.1073/pnas.1218972110

Millán, A. P., van Straaten, E. C., Stam, C. J., Nissen, I. A., Idema, S., Baayen, J. C., ... Hillebrand, A. (2022). Epidemic models characterize seizure propagation and the effects of epilepsy surgery in individualized brain networks based on meg and invasive eeg recordings. *Scientific Reports*, *12*(1), 1–20. doi: 10.1038/s41598-022-07730-2

Millán, A. P., van Straaten, E. C., Stam, C. J., Nissen, I. A., Idema, S., Baayen, J. C., ... Hillebrand, A. (2023). The role of epidemic spreading in seizure dynamics and epilepsy surgery. *Network Neuroscience*, *7*(2), 811–843. doi: 10.1162/netn_a_00305

Moosavi, S. A., Jirsa, V. K., & Truccolo, W. (2022). Critical dynamics in the spread of focal epileptic seizures: Network connectivity, neural excitability and phase transitions. *Plos one*, *17*(8), e0272902. doi: 10.1371/journal.pone.0272902

Nissen, I. A., Millán, A. P., Stam, C. J., van Straaten, E. C., Douw, L., Pouwels, P. J., ... Hillebrand, A. (2021). Optimization of epilepsy surgery through virtual resections on individual structural brain networks. *Scientific Reports*, *11*(1), 1–18.

doi: 10.1038/s41598-021-98046-0

Nissen, I. A., Stam, C. J., Reijneveld, J. C., van Straaten, I. E., Hendriks, E. J., Baayen, J. C., ... Hillebrand, A. (2017). Identifying the epileptogenic zone in interictal resting-state meg source-space networks. *Epilepsia*, *58*(1), 137–148. doi: 10.1111/epi.13622

Nissen, I. A., Stam, C. J., van Straaten, E. C., Wottschel, V., Reijneveld, J. C., Baayen, J. C., ... Hillebrand, A. (2018). Localization of the epileptogenic zone using interictal meg and machine learning in a large cohort of drug-resistant epilepsy patients. *Frontiers in Neurology*, *9*, 647. doi: 10.3389/fneur.2018.00647

Olmi, S., Petkoski, S., Guye, M., Bartolomei, F., & Jirsa, V. (2019). Controlling seizure propagation in large-scale brain networks. *PLoS Computational Biology*, *15*(2), e1006805. doi: 10.1371/journal.pcbi.1006805

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015, August). Epidemic processes in complex networks. *Reviews of Modern Physics*, *87*(3), 925–979. doi: 10.1103/RevModPhys.87.925

Peraza, L. R., Díaz-Parra, A., Kennion, O., Moratal, D., Taylor, J.-P., Kaiser, M., & Bauer, R. (2019). Structural connectivity centrality changes mark the path toward alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *11*, 98–107. doi: 10.1016/j.dadm.2018.12.004

Pinto, P. C., Thiran, P., & Vetterli, M. (2012). Locating the source of diffusion in large-scale networks. *Physical review Letters*, *109*(6), 068702. doi: 10.1103/PhysRevLett.109.068702

Proix, T., Bartolomei, F., Chauvel, P., Bernard, C., & Jirsa, V. K. (2014). Permittivity coupling across brain regions determines seizure recruitment in partial epilepsy. *Journal of Neuroscience*, *34*(45), 15009–15021. doi: 10.1523/JNEUROSCI.1570-14.2014

Proix, T., Bartolomei, F., Guye, M., & Jirsa, V. K. (2017, March). Individual brain structure and modelling predict seizure propagation. *Brain*, *140*(3), 641–654. doi: 10.1093/brain/awx004

Reveley, C., Seth, A. K., Pierpaoli, C., Silva, A. C., Yu, D., Saunders, R. C., ... Ye, F. Q. (2015). Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion mr tractography. *Proceedings of the National Academy of Sciences*, *112*(21), E2820–E2828. doi: 10.1073/pnas.141819811

Roberts, J. A., Perry, A., Lord, A. R., Roberts, G., Mitchell, P. B., Smith, R. E., ... Breakspear, M. (2016). The contribution of geometry to the human connectome. *Neuroimage*, *124*, 379–393. doi: 10.1016/j.neuroimage.2015.09.009

Roberts, J. A., Perry, A., Roberts, G., Mitchell, P. B., & Breakspear, M. (2017). Consistency-based thresholding of the human connectome. *NeuroImage*, *145*, 118–129. doi: 10.1016/j.neuroimage.2016.09.053

Runfola, C., Sheheitli, H., Bartolomei, F., Wang, H., & Jirsa, V. (2023). In pursuit of the epileptogenic zone in focal epilepsy: A dynamical network biomarker approach. *Communications in Nonlinear Science and Numerical Simulation*, *117*,

106973. doi: 10.1016/j.cnsns.2022.106973

Saggio, M. L., Crisp, D., Scott, J. M., Karoly, P., Kuhlmann, L., Nakatani, M., … Stacey, W. C. (2020). A taxonomy of seizure dynamotypes. *Elife*, *9*, e55632. doi: 10.7554/eLife.55632

Schoonhoven, D. N., Coomans, E. M., Millán, A. P., van Nifterick, A. M., Visser, D., Ossenkoppele, R., … Gouw, A. A. (2023). Tau protein spreads through functionally connected neurons in alzheimer's disease: a combined meg/pet study. *Brain*, awad189. doi: 10.1093/brain/awad189

Seguin, C., Jedynak, M., David, O., Mansour, S., Sporns, O., & Zalesky, A. (2023). Communication dynamics in the human connectome shape the cortex-wide propagation of direct electrical stimulation. *Neuron*, *111*(9), 1391–1401. doi: 10.1016/j.neuron.2023.01.027

Seguin, C., Sporns, O., & Zalesky, A. (2023). Brain network communication: concepts, models and applications. *Nature Reviews Neuroscience*, *24*(9), 557–574. doi: 10.1038/s41583-023-00718-5

Sinha, N., Dauwels, J., Kaiser, M., Cash, S. S., Brandon Westover, M., Wang, Y., & Taylor, P. N. (2017, February). Predicting neurosurgical outcomes in focal epilepsy patients using computational modelling. *Brain*, *140*(2), 319–332. doi: 10.1093/brain/aww299

Sip, V., Hashemi, M., Vattikonda, A. N., Woodman, M. M., Wang, H., Scholly, J., … Jirsa, V. K. (2021). Data-driven method to infer the seizure propagation patterns in an epileptic brain from intracranial electroencephalography. *PLoS Computational Biology*, *17*(2), e1008689. doi: 10.1371/journal.pcbi.1008689

Stam, C. J., van Straaten, E. C. W., Van Dellen, E., Tewarie, P., Gong, G., Hillebrand, A., … Van Mieghem, P. (2016, May). The relation between structural and functional connectivity patterns in complex brain networks. *International Journal of Psychophysiology*, *103*, 149–160. doi: 10.1016/j.ijpsycho.2015.02.011

Taylor, P. N., Kaiser, M., & Dauwels, J. (2014). Structural connectivity based whole brain modelling in epilepsy. *Journal of Neuroscience Methods*, *236*, 51–57. doi: 10.1016/j.jneumeth.2014.08.010

Taylor, P. N., Sinha, N., Wang, Y., Vos, S. B., de Tisi, J., Miserocchi, A., … Duncan, J. S. (2018, January). The impact of epilepsy surgery on the structural connectome and its relation to outcome. *NeuroImage: Clinical*, *18*, 202–214. doi: 10.1016/j.nicl.2018.01.028

Theodoni, P., Majka, P., Reser, D. H., Wójcik, D. K., Rosa, M. G., & Wang, X.-J. (2022). Structural attributes and principles of the neocortical connectome in the marmoset monkey. *Cerebral Cortex*, *32*(1), 15–28. doi: 10.1093/cercor/bhab191

van den Heuvel, M. P., & Sporns, O. (2019). A cross-disorder connectome landscape of brain dysconnectivity. *Nature reviews neuroscience*, *20*(7), 435–446. doi: 10.1038/s41583-019-0177-6

van Diessen, E., Diederen, S. J., Braun, K. P., Jansen, F. E., & Stam, C. J. (2013). Functional and structural brain networks in epilepsy: what have we learned? *Epilepsia*, *54*(11), 1855–1865. doi: 10.1111/epi.12350

Vattikonda, A. N., Hashemi, M., Sip, V., Woodman, M. M., Bartolomei, F., & Jirsa, V. K. (2021). Identifying spatio-temporal seizure propagation patterns in epilepsy using bayesian inference. *Communications biology*, *4*(1), 1–13. doi: 10.1038/s42003-021-02751-5

Wang, H. E., Woodman, M., Triebkorn, P., Lemarechal, J.-D., Jha, J., Dollomaja, B., ... Jirsa, V. (2023). Delineating epileptogenic networks using brain imaging data and personalized modeling in drug-resistant epilepsy. *Science Translational Medicine*, *15*(680), eabp8982. doi: 10.1126/scitranslmed.abp8982

Wang, Y., Schroeder, G. M., Horsley, J. J., Panagiotopoulou, M., Chowdhury, F. A., Diehl, B., ... Taylor, P. N. (2023). Temporal stability of intracranial eeg abnormality maps for localising epileptogenic tissue. *arXiv:2302.05734*. doi: 10.48550/arXiv.2302.05734