HHS Public Access

Author manuscript

Epilepsy Behav. Author manuscript; available in PMC 2022 March 01.

Published in final edited form as:

Epilepsy Behav. 2021 March ; 116: 107714. doi:10.1016/j.yebeh.2020.107714.

Quantitative [18]FDG PET asymmetry features predict long-term seizure recurrence in refractory epilepsy

Lohith G. Kini[#a,b], **Ashesh A. Thaker**[#c], **Peter N. Hadar**[d], **Russell T. Shinohara**[e], **Mesha-Gay Brown**[f], **Jacob G. Dubroff**[g,2], **Kathryn A. Davis**[b,d,*,2]

[a]Department of Bioengineering, University of Pennsylvania, 240 Skirkanich Hall, 210 S 33rd St, Philadelphia, PA 19104, United States

[b]Center for Neuroengineering and Therapeutics, University of Pennsylvania, 240 Skirkanich Hall, 210 S 33rd St, Philadelphia, PA 19104, United States

[c]Department of Radiology, Division of Neuroradiology, University of Colorado School of Medicine, 12401 E. 17th Ave, Aurora, CO 80045, United States

[d]Department of Neurology, Hospital of the University of Pennsylvania, 3400 Spruce St, 3 West Gates Bldg, Philadelphia, PA 19104, United States

[e]Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, 217 Blockley Hall, 423 Guardian Dr, Philadelphia, PA 19104, United States

[f]Department of Neurology, University of Colorado School of Medicine, 1635 Aurora Ct Ste 4200, Aurora, CO 80045, United States

[g]Department of Radiology, Hospital of the University of Pennsylvania, 3400 Spruce Street, 1 Silverstein Pavilion, Philadelphia, PA 19104, United States

[#] These authors contributed equally to this work.

Abstract

Objective: Fluorodeoxyglucose-positron emission tomography (FDG-PET) is an established, independent, strong predictor of surgical outcome in refractory epilepsy. In this study, we explored the added value of quantitative [^{18}F]FDG-PET features combined with clinical variables, including electroencephalography (EEG), [^{18}F]FDG-PET, and magnetic resonance imaging (MRI) qualitative interpretations, to predict long-term seizure recurrence (mean post-op follow-up of 5.85 ± 3.77 years).

Methods: Machine learning predictive models of surgical outcome were created using a random forest classifier trained on quantitative features in 89 patients with drug-refractory temporal lobe epilepsy evaluated at the Hospital of the University of Pennsylvania epilepsy surgery program

[*]Corresponding author at: Department of Neurology, Hospital of the University of Pennsylvania, 3400 Spruce St, 3 West Gates Bldg, Philadelphia, PA 19104, United States. kathryn.davis@pennmedicine.upenn.edu (K.A. Davis).

[2]These authors are joint last authors.

Ethical publication statement

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

(2003–2016). Quantitative features were calculated from asymmetry features derived from image processing using Advanced Normalization Tools (ANTs).

Results: The best-performing model used quantification and had an out-of-bag accuracy of 0.71 in identifying patients with seizure recurrence (Engel IB or worse) which outperformed that using qualitative clinical data by 10%. This model is shared through open-source software for research use. In addition, several asymmetry features in temporal and extratemporal regions that were significantly associated with seizure freedom are identified for future study.

Significance: Complex quantitative [^{18}F]FDG-PET imaging features can predict seizure recurrence in patients with refractory temporal lobe epilepsy. These initial retrospective results in a cohort with long-term follow-up suggest that using quantitative imaging features from regions in the epileptogenic network can inform the clinical decision-making process.

Keywords

3D-SSP; Asymmetry; PET; Drug-resistant epilepsy; Surgical outcome; Imaging features; Machine learning; Open source

1. Introduction

Epilepsy affects 65 million people in the world of which 20–40% are refractory to medical therapy [1]. Surgical resection has been the mainstay of therapy for these patients and identifying patients at risk of suboptimal surgical outcome can allow clinicians to provide them with alternative therapies that continue to be improved, such as responsive neurostimulation, vagal nerve stimulation, and deep brain stimulation. Despite significant advances in imaging techniques, the use of neuroimaging in epilepsy as a predictive biomarker of both medical and surgical interventions has yet to be fully realized [2].

For patients with intractable epilepsy being considered for surgery, testing at tertiary centers can include interictal brain [^{18}F]fluorodeoxyglucose (FDG)-positron emission tomography (PET), structural brain magnetic resonance imaging (MRI), seizure semiology evaluation, neuropsychological testing, and scalp and intracranial electrocorticography. Together, these tests are used to best delineate the seizure onset zone for resection. While qualitative visual [^{18}F]FDG-PET interpretation is important to the clinical workup for surgical candidates, quantitative imaging features derived from these studies can be used when other clinical and imaging tests do not unequivocally localize the seizure focus. In these situations, [^{18}F]FDG-PET has shown promise in being able to better identify patients who may not be ideal candidates for surgery. Several studies have identified important temporal and extratemporal regions on functional and structural imaging that predict prognosis [3–7]. A recent study suggests that lateralized hypometabolism and extent of resection are strong indicators of seizure recurrence with differential outcomes based on seizure lateralization [8]. Quantification has the advantage of precision and objectivity over qualitative visual [^{18}F]FDG-PET interpretation, potentially improving presurgical planning.

Application of robust, machine learning-based, semiautomated, and generalizable models in clinical epilepsy practice remains largely unexplored. Most previous investigations of [^{18}F]FDG-PET in intractable epilepsy have been limited by sample size, minimizing broad

applicability. Two recent studies used larger populations and more quantitative approaches [8,9]. Although six- to twelve-month surgical outcome are Engel 1 in 70–90% of patients following temporal lobectomy, long-term outcomes are highly variable in epilepsy surgery, with better outcomes in patients with focal findings on MRI or PET [10–12]. To evaluate such an application, we identified quantitative [^{18}F]FDG-PET imaging features predictive of surgical recurrence in a cohort of temporal lobe epilepsy patients, relying on asymmetry measurements derived from automated image analysis. We present additional preliminary findings in a small cohort of extratemporal epilepsy patients in the supplementary material. We also measured the importance of quantitative PET features relative to traditional clinical data, such as EEG and qualitative visual PET interpretation, in improving prediction of seizure recurrence. These models were applied and cross-validated to a dataset with patient outcomes measured at least 6 months after epilepsy surgery, most having outcome measures at least 2 years after surgery (mean 5.9 ± 3.8 years follow-up). All models used to compute these features are made publicly available via open-source software for both research and clinical use. Translating these models into routine clinical practice could potentially improve quality of life for many epilepsy patients by enabling physicians and researchers to incorporate quantitative PET information into the decision-making process.

2. Methods

2.1. Patients

This study was approved by the Hospital of the University of Pennsylvania (HUP) Institutional Review Board. Subjects were screened using electronic records and included in the study based on the following criteria: sequential patients at HUP who underwent epilepsy surgery between 2003 and 2016; clinical evaluation by members of the Epilepsy Surgery program; and available histopathologic and neuroimaging assessment by applicable board-certified pathologists, nuclear medicine physicians, and radiologists. Patients were excluded if they lacked preoperative interictal brain [^{18}F]FDG-PET, scalp-EEG, or formal image interpretation (Fig. 1). Of note, [^{18}F]FDG-PET was implemented as a standard of care. As a result, this study used inclusion criteria that delineate a clinically representative cohort and do not bias the cohort toward subpopulations with only discordant findings or bilateral foci. Clinical data were obtained from electronic and paper medical records and included the following: gender, date of birth, date of epilepsy surgery, date of seizure onset, scalp-EEG findings, neuroimaging findings, clinical seizure lateralization and localization, region of surgical resection, and surgical outcome by Engel classification. Extensive presurgical evaluation, as part of the evaluation process for potential surgery at HUP, included detailed clinical examinations, scalp- and invasive- (subdural and depth electrode) EEG, MRI, and interictal [^{18}F]FDG-PET. Resection margins were not extended to remove hypometabolic regions beyond the apparent EEG/MRI “focus” identified by the clinical team. His-tological examination was completed after surgical resection. Seizure outcome for each patient was determined by epileptologists blinded to the results of this study (M-G.B., K.A.D.). For the purposes of this study, the binary seizure recurrence variable was based on surgical outcome classified according to the Engel system. Since the majority of the patient population had seizure onset in the temporal lobe, the main analysis was performed on only temporal onset patients and a separate subanalysis was performed with extratemporal

neocortical epilepsy patients added to the study population. The results of all patients are reported in the supplementary section.

2.2. Image acquisition and processing

Details regarding [^{18}F]FDG-PET acquisition, processing, and the computation of quantitative image preprocessing variables are provided in the supplementary section (Table S1). The PET imaging computational pipeline is outlined in Fig. 2, filtering out any corrupted images or images that could not be processed. A sample feature image is shown in Fig. 3. In addition, the codification of clinical variables, specifically scalp-EEG, MRI findings, and PET findings is also described in the supplementary text.

2.3. Outcome variable

The most recent outpatient visit (5.9 ± 3.8 years post-op) was used to calculate the binary seizure recurrence outcome variable following surgical resection: The Engel outcome variable was coded as 1 for seizure recurrence (Engel IB-D, II, III, IV) or 0 if entirely seizure free (Engel IA). For the purposes of defining outcomes, Engel IA was solely considered as “seizure free” since Engel IB outcome may indicate nondisabling simple partial seizures which could potentially result from a surgical resection or image interpretation failing to completely delineate the boundaries of the seizure focus. Engel IA outcome better corresponds to the ILAE classification scale that separates auras into a separate class (a scale that would have been used if not for the limitations of determining seizure days retrospectively) [13].

2.4. Statistical analysis and machine learning

Pearson’s chi-squared tests were used to compare clinical variables to outcomes for categorical variables: resected region, MRI lesion status, pathology, and qualitative visual [^{18}F]FDG-PET interpretation. Clinical imaging variables were based on qualitative nuclear medicine and neuroradiology interpretations. Extent of [^{18}F]FDG-PET hypometabolism was classified as normal, focal, subtle, multifocal, or diffuse. This was based on expert review of the nuclear medicine report for each patient with attention to specific language including “diffuse”, “multiple foci,” or “subtle/mild asymmetry” in the report findings and impression.

Four models were computed to measure the added benefit of including quantitative imaging features: clinical EEG and MRI lesion status with the addition of qualitative [^{18}F]FDG-PET read which represents a standard of care (Model A), clinical EEG and MRI lesion status with the addition of quantitative PET asymmetry features (Model B), all clinical variables and quantitative PET features together (Model C), and clinical EEG and MRI lesion status alone (Model D). The clinical variables were scalp-EEG localization, lesion status based on MRI lesion status read, and qualitative [^{18}F] FDG-PET read (extent of interictal hypometabolism). All models were random forest classifiers using optimal number of decision tree estimators, splitting at every node based on Gini importance [14]. Random forests (RF) are a combination of decision tree classifiers that minimize generalization error by splitting on randomly selected features. A preanalysis variable selection step was performed in the dataset to create a short list of quantitative imaging biomarkers from the large set of features described in the quantitative imaging preprocessing section in the

supplementary section. One step of feature selection was used with features that were ranked highly by a marginal F test comparing the mean feature value across outcomes, constituting the final shortlist of features. This method of feature selection has been used in a study that applied machine learning models to diagnose Alzheimer's disease to avoid overfitting and identify useful imaging features from a larger set of predictors [15].

Sensitivity/specificity analyses were performed on the cohort using Python and R. By sampling approximately 90% of the cohorts, 1000 bootstraps were conducted and the out-of-bag accuracy was calculated. The feature reduction step described above was performed repeatedly within each bootstrap subsample and resulted in different sets of features each time. Confidence intervals were reported for difference in out-of-bag accuracy (which measures accuracy on samples that were not used in the training of predictor trees in the forest) between models, focusing mostly on comparison to standard of care, Model A, to determine the value of adding quantitative PET imaging features to currently used clinical features: EEG, MRI, and qualitative PET reads. Fig. 1 shows an outline of the pipeline.

2.5. Data availability statement

Retrospective clinical and imaging data collected and analyzed during this study are securely and anonymously stored and will be shared upon request from any qualified investigator. Requests should be made via email to the corresponding author. Furthermore, we share our open-source computational pipeline, as described in Section 3.2 on "Open-Source Shared Tool for Feature Computation."

3. Results

3.1. Patient demographics

In the main study using only temporal lobe patients, there were 89 patients. Table 1 summarizes the demographics of all 96 total patients included in full study (89 temporal and 7 extratemporal cases). Outcome was determined from last clinical encounter (mean follow-up of 5.9 ± 3.8 years after surgery). Eighty-eight of the 96 total patients had Engel outcome measured at least 1 year after surgery, 76 had Engel outcome measured at least 2 years after surgery, and 43 had Engel outcome measured at least 5 years after surgery. Four patients with temporal onset had outcomes measured less than a year after resection, two of which were evaluated >10 months after surgery. Of these 4 patients, 3 were Engel I (IA, IB, and ID) and 1 was Engel IV. The one Engel IV patient was a 49-year-old woman who died with autopsy-confirmed SUDEP (Sudden Unexpected Death in Epilepsy). This patient had left mesiotemporal sclerosis on MRI evaluation and focal left anterior temporal hypometabolism on PET. Fig. 4 shows surgical outcomes for all patients as function of time of Engel outcome evaluation in years, with major timepoints after resection color-coded according to outcome at most recent office visit (used for subsequent analyses).

In the full patient cohort, surgery was performed ipsilateral to the hypometabolic temporal lobe in 75 (78 %) patients (27 were Engel IA at last follow-up). Remaining patients had normal qualitative PET reads (4 were Engel IA at last follow-up). Among patients who had favorable outcomes (Engel IA), 16 had left-temporal lobe resections (LTL) and 16 had right-

temporal lobe (RTL) resections. Among patients who had seizure recurrence (Engel IB-ID, II, III, or IV), 31 had LTL, 26 had RTL, and the rest had extratemporal right-hemispheric resections (full patient cohort results are pre-sented in Supplementary Material).

Of the 61 patients who had focal or subtle hypometabolism PET, 38 had mesial temporal sclerosis (MTS), 11 had gliosis, 3 had focal cortical dysplasia (FCD), 3 had dual pathology (MTS and malformation of cortical development), 3 had low-grade tumors or vascular malformations, and 2 were normal. Of the 35 patients who did not have focal or subtle hypometabolism, 19 had MTS, 6 had gliosis, 3 had FCD, 3 had dual pathology, 1 had a cavernoma, and 3 were normal (pathology not significantly different between outcome groups). In patients with either focal or subtle hypometabolism on PET, 17 of 39 (44%) lesional patients had seizure freedom, and 8 of 22 (36%) nonlesional patients had seizure freedom.

As an exploratory first analysis, we investigated the association between features and surgical outcome based on an F test. The top few features predictive of seizure recurrence are shown in Table S3. Note this is not a feature reduction step and instead serves to simply identify the most associated features to further study with larger well-designed cohorts.

3.2. Model selection for optimal prediction of surgical outcome

These models were measured using out-of-bag bootstrap samples derived from generating trees over the cohort ($n = 89$ with temporal lobe onset). These bootstraps were generated by randomly sampling 90% of the patients to train. Feature reduction was performed to select features that were ranked highly by a marginal F test of the mean feature value to outcomes. Overall, adding quantitative variables (Models B and C) improves performance more than using qualitative PET alone (Model A) (Table 2). Model D, using only clinical EEG and MRI variables (mean out-of-bag accuracy = 0.62), performed similarly to standard of care, Model A (mean out-of-bag accuracy = 0.61), but worse than Model B (mean out-of-bag accuracy = 0.71). We found similar improvement in performance when performing subgroup analyses. We ran the same models when we isolated to only left-temporal and right-temporal lobe patients. We also assessed prediction of seizure recurrence in patients with Engel I outcome only (i.e., predicting Engel IA vs. Engel IB-D). There was statistically significant improvement in seizure recurrence prediction accuracy with Models C and D in most cases (Table 2).

3.3. Open-source shared tool for feature computation

In order to make the computational pipeline easily shareable and viewable, the authors of this study have created a free open-source software program to allow others to reproduce and build on our work. The code used for this paper is freely available at https://github.com/ieeg-portal/qneuropet.

4. Discussion

In this study, we identified multiple quantitative [^{18}F]FDG-PET features predictive of long-term seizure recurrence in drug-resistant epilepsy patients through a machine learning approach. In the age of precision medicine, incorporation of quantitative PET metrics into

standard clinical decision pathways could lay the groundwork for computer-assisted decision making to identify patients likely to have seizure recurrence after surgery. These features, tures, computed throughout the epileptic network and combined with clinical variables, are predictive of long-term outcome in terms of seizure recurrence, and their application resulted in a higher accuracy as compared to a model that relies solely on clinical variables. We followed multiple steps to build a generalizable model for all patients with drug-resistant epilepsy (focused on temporal onset epilepsy, but easily extendable to extratemporal neocortical onset epilepsy).

There is evidence that focal PET hypometabolism is a predictor of both seizure onset lateralization [16,17] and seizure outcome [3,18–21]. In this study, patients with focal hypometabolism were more likely to be seizure free after resection than those with other PET findings (Table 1). In addition, patients with focal hypometabolism were correctly lateralized to site of surgery in 100% of cases, based on clinical PET interpretations (Table S2). According to prior literature, normal MRI and [^{18}F]FDG-PET are suggestive of suboptimal outcomes, especially when both are normal [5,6,22–27]. This was confirmed in our cohort where nonlesional MRI patients were less likely to have focal hypometabolism and more likely to have suboptimal outcomes. The combination of [^{18}F]FDG-PET and MRI lesion status were stronger predictors of outcome than using either alone in our machine learning models. Therefore, all machine learning models included these clinical variables: lesion status, qualitative extent of PET hypometabolism, and localization of ictal EEG abnormalities.

The best predictive models in our study incorporate quantitative imaging features (Models B and C) compared to clinical variables alone (Model D) or standard of care (Model A). Model B had a 71% out-of-bag accuracy for identifying patients with seizure recurrence (Engel Class IB or greater) which was 10% better than standard of care. Models B and C outperformed clinical only models in all subgroup analyses, including seizure recurrence within Engel I (Engel IA vs. IB–ID) as well as left-only and right-only temporal lobe patients. Our results suggest additive value to using quantitative PET features to augment clinical PET reads in patients with unclear lateralization or subtle imaging findings. In particular, our findings suggest higher surgical outcome prediction accuracy using quantitative PET measures in right-temporal lobe patients alone as compared with prediction accuracy in left-temporal lobe patients alone. Right-temporal lobe cases had an accuracy of 72% predicting seizure recurrence when using quantitative PET variables compared to left-temporal cases which had an accuracy of 64%. This discrepancy between right-only and left-only cohorts is concordant with other recent studies [8,9]. One possibility is that patients with right MTL epilepsy have greater rates of bitemporal hypometabolic changes (especially contralateral temporal lobe hypometabolism portending poorer prognosis) whereas those with left MTL epilepsy have more extensive ipsilateral temporal lobe hypometabolism (as found in Cahill et al., 2019). These findings may be better identified by quantitative features computed across all regions of the brain and fed into the model. Future studies could validate these findings by incorporating extent of the resection zone as determined by postresection imaging to validate these findings. Quantitative asymmetry features in the middle frontal gyri, supramarginal gyri, and other extratemporal regions were important variables in ensuring high accuracy, as shown by features listed in Table S3. A recent study

pinpointed important extratemporal regions with early electrical spread of seizures that were highly predictive of surgical failures, such as the anterior lateral temporal cortex, insulo-fronto-opercular areas, and perisylvian areas [9]. Another study demonstrated that ipsilateral temporo-polar hypometabolism is predictive of favorable surgical outcome in mesial temporal lobe epilepsy [4]. All these areas overlap with broader regions we identify in our study.

The most important distinguishing factors based on the list of quantitative features generated from feature selection involve different asymmetry-based features in regions known to play important roles in seizure spread of the mesial temporal lobe (MTL) epileptic network in both structural and functional investigations [4,28]. Asymmetry-based features in these regions and their contralateral equivalents, as well as deep subcortical volumes (such as the thalamus), were identified (Table S3).

Manno et al., 1994 [3] were the first to compute quantitative imaging features from [^{18}F]FDG-PET images to predict outcome at 1 year in anterior temporal lobectomies; it was determined that asymmetry of temporal hypometabolism (computed using asymmetry index) was a sensitive and specific predictor (78% sensitivity/73% specificity) of surgical outcome. Wong et al., 2010 [28] showed that hypometabolism was present in the ipsilateral frontal lobe, insula, and some ipsilateral occipital regions in patients with MTL epilepsy and hippocampal sclerosis (HS). Our feature space included asymmetry indices computed in these 3 regions, as well as additional frontoparietal regions. There were a few highly ranked extratemporal regional predictors in our models, but temporal asymmetry was still a better predictor than extratemporal predictors (e.g., superior temporal gyrus, transverse temporal gyrus). While past studies may have been ambivalent about using thalamic hypometabolism to predict seizure recurrence [28–31], our best-performing models suggest thalamic asymmetry to be a feature of consideration. Specifically, quantitative thalamic PET asymmetry regardless of hemisphere of surgery meant higher likelihood of seizure freedom, suggesting it may be a useful indicator of lateralization within a broader network.

Our study has some limitations. First, this is an exploratory study that aims to estimate the value of adding quantitative PET features using a cohort of patients with heterogeneous clinical features. Along these lines, the codification of clinical imaging reports may not fully capture the findings and impression as intended or communicated during routine clinical evaluation. Second, since our model was only trained on temporal lobe epilepsy patients, it cannot be generalized to extratemporal lobe epilepsy. Third, we did not consider the extent of resection and its effect on outcome, as has been attempted in a prior study [5]. Future studies incorporating quantitative image features from MRI such as extent of resection, adjacent gliosis, and overlap of MRI resection with quantitative PET abnormalities may provide additional benefit as has been shown in previous studies [8]. Co-registration of MRI and PET data or simultaneous PET/MRI acquisition with newer instruments is also worthy of future study.

The authors further recognize that our study cohort is imbalanced with a high proportion of Engel IB and worse outcomes (64/96). Machine learning techniques tend to underperform on the minority class, in this case representing Engel IA seizure freedom. Though beyond

the scope of this study, one way to overcome this imbalance is to oversample the minority class, a technique known as Synthetic Minority Oversampling Technique (SMOTE). This technique was recently applied to a machine learning study in a temporal lobe epilepsy cohort [32] and a point of consideration for future epilepsy studies. The authors also acknowledge that our demonstrated improvement in out-of-bag accuracy does not necessarily imply statistical significance. It is a well-known issue that it is possible to obtain results that appear better than chance with small datasets that are randomly generated [33]. One way to test this is to shuffle classifiers (e.g., Engel class outcomes) with each iteration and make a correction to the real dataset [34]. Though not performed in our study, it is an important consideration for future studies working with small neuroimaging samples.

Lastly, our models risk overfitting because of the large set of features computed, but the feature reduction step performed within each bootstrap mitigates that risk. Since the primary goal of this study is to measure the value of including quantitative imaging features over outcome models using clinical variables, this method of feature selection allowed us to build the most generalizable model within our sample size. These features should be investigated further in a larger prospective cohort, ideally including a large number of extratemporal lobe patients. Our randomization process created an equal distribution of suboptimal and good outcome patients in each bootstrap stratified cohort, ensuring the generalizability of this study. Furthermore, in distinction to prior studies focused on the temporal lobe epilepsy population, our investigation included a balance of lesional (55) and nonlesional (41) patients. A minor limitation of this study is that PET studies were performed across a variety of instruments, including PET only and PET/CT, however, this is unlikely to have influenced our algorithm which focused on asymmetry. Future studies investigating voxel-wise whole brain and ROI-based classifiers will likely be useful in translating PET machine learning imaging algorithms into clinical radiology decision support.

In conclusion, when combined with typical clinical variables (EEG, MRI lesion status, and qualitative PET reads), quantitative FDG-PET imaging features derived from asymmetry using established and readily available software can predict long-term seizure recurrence after epilepsy surgery in MRI lesional and nonlesional patients with temporal onset. Imaging features derived from a machine learning approach show promise in elucidating the underlying pathophysiologic networks of epilepsy and lay groundwork for incorporating computer-assisted modeling into clinical practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Grant Support/Acknowledgements: Virtual cortical resection R01: 1 R01 NS099348–01, Thornton Foundation, K23 NS092973. Additional grant funding provided by NIH 1-T32-NS-091006–01 (Training Program in Neuroengineering and Medicine). JGD was supported by K23 DA038726. AAT received grant funding from NIH 5T32EB004311–09 (Penn Radiology Research Track Residency). RTS was partially supported by R01NS085211. The authors would like to acknowledge Preya Shah and Dr. Dorian Pustina for useful discussion and feedback on the manuscript.

Disclosure of conflicts of interest

Dr. Kini reports grants from University of Pennsylvania during the conduct of the study; Dr. Thaker reports grants from University of Pennsylvania during the conduct of the study; Dr. Hadar reports grants from University of Pennsylvania during the conduct of the study; Dr. Shinohara reports grants from NIH, NMSS, and the Race to Erase MS Foundation during the conduct of the study; personal fees from Genentech/Roche, outside the submitted work; Dr. Brown has nothing to disclose; Dr. Dubroff reports grants from the NIH; Dr. Davis reports grants from University of Pennsylvania and NIH during the conduct of the study.

Appendix A

Name	Location	Role	Contribution
Lohith G Kini, MD, PhD	University of Pennsylvania, Philadelphia, PA	Author	conception and design of the study; acquisition and analysis of data; drafting a significant portion of the manuscript or figures
Ashesh A Thaker, MD	University of Colorado Anschutz Medical Campus, Aurora, CO	Author	conception and design of the study; acquisition and analysis of data; drafting a significant portion of the manuscript or figures
Peter N Hadar, MD	University of Pennsylvania, Philadelphia, PA	Author	conception and design of the study; acquisition and analysis of data; drafting a significant portion of the manuscript or figures
Russell T	Shinohara, PhD		University of Pennsylvania, Philadelphia, PA
Author	acquisition and analysis of data.		
Mesha-Gay Brown, MD	University of Colorado Anschutz Medical Campus, Aurora, CO	Author	acquisition and analysis of data.
Jacob G Dubroff, MD, PhD	University of Pennsylvania, Philadelphia, PA	Author	conception and design of the study; acquisition and analysis of data; drafting a significant portion of the manuscript or figures
Kathryn A Davis, MD	University of Pennsylvania, Philadelphia, PA	Author	conception and design of the study; acquisition and analysis of data; drafting a significant portion of the manuscript or figures

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yebeh.2020.107714.

Abbreviations:

PET	Positron emission tomography
FDG	fluorodeoxyglucose
CV	cross validation

References

[1]. Kwan P, Schachter SC, Brodie MJ. Drug-resistant epilepsy. N Engl J Med 2011;365:919–26. 10.1056/NEJMra1004418. [PubMed: 21899452]

[2]. Bernasconi N, Bernasconi A. Imaging the epileptic brain—time for new standards. Nat Rev Neurol 2014;10:133–4. 10.1038/nrneurol.2013.280. [PubMed: 24419681]

[3]. Manno EM, Sperling MR, Ding X, Jaggi J, Alavi A, O'Connor MJ, et al. Predictors of outcome after anterior temporal lobectomy: Positron emission tomography. Neurology 1994;44:2321. 10.1212/WNL.44.12.2321.

[4]. Dupont S, Semah F, Clémenceau S, Adam C, Baulac M, Samson Y, et al. Accurate prediction of postoperative outcome in mesial temporal lobe epilepsy. Arch Neurol 2000;57:789–94. 10.1001/archneur.57.9.1331. [PubMed: 10867774]

[5]. Higo T, Sugano H, Nakajima M, Karagiozov K, Iimura Y, Suzuki M, et al. The predictive value of FDG-PET with 3D-SSP for surgical outcomes in patients with temporal lobe epilepsy. Seizure 2016;41:127–33. 10.1016/j.seizure.2016.07.019. [PubMed: 27552380]

[6]. Chassoux F, Semah F, Bouilleret V, Landre E, Devaux B, Turak B, et al. Metabolic changes and electro-clinical patterns in mesio-temporal lobe epilepsy: A correlative study. Brain 2004;127:164–74. 10.1093/brain/awh014. [PubMed: 14534161]

[7]. Waxman AD, Herholz K, Lewis DH, Herscovitch P, Minoshima S, Ichise M, et al. Society of Nuclear Medicine Procedure Guideline for FDG PET Brain Imaging n. d

[8]. Cahill V, Sinclair B, Malpas CB, McIntosh AM, Chen Z, Vivash LE, et al. Metabolic patterns and seizure outcomes following anterior temporal lobectomy. Ann Neurol 2019;85:241–50. 10.1002/ana.25405. [PubMed: 30609109]

[9]. Chassoux F, Artiges E, Semah F, Laurent A, Landré E, Turak B, et al. ^{18}F-FDG-PET patterns of surgical success and failure in mesial temporal lobe epilepsy. Neurology 2017;88:1045–53. 10.1212/WNL.0000000000003714. [PubMed: 28188304]

[10]. Yang P-F, Pei J-S, Zhang H-J, Lin Q, Mei Z, Zhong Z-H, et al. Long-term epilepsy surgery outcomes in patients with PET-positive, MRI-negative temporal lobe epilepsy. Epilepsy Behav 2014;41:91–7. 10.1016/j.yebeh.2014.09.054. [PubMed: 25461196]

[11]. Téllez-Zenteno JF, Dhar R, Wiebe S. Long-term seizure outcomes following epilepsy surgery: a systematic review and meta-analysis. Brain 2005;128:1188–98. 10.1093/brain/awh449. [PubMed: 15758038]

[12]. de Tisi J, Bell GS, Peacock JL, McEvoy AW, Harkness WF, Sander JW, et al. The long-term outcome of adult epilepsy surgery, patterns of seizure remission, and relapse: a cohort study. Lancet 2011;378:1388–95. 10.1016/S0140-6736(11)60890-8. [PubMed: 22000136]

[13]. Jutila L, Immonen A, Mervaala E, Partanen J, Partanen K, Puranen M, et al. Long term outcome of temporal lobe epilepsy surgery: analyses of 140 consecutive patients. J Neurol Neurosurg Psychiatry 2002;73:486–94. 10.1136/JNNP.73.5.486. [PubMed: 12397139]

[14]. Breiman L. Random forests. Mach Learn 2001;45:5–32. 10.1023/A:1010933404324.

[15]. Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam C-P, et al. Blood-based protein biomarkers for diagnosis of Alzheimer disease. Arch Neurol 2012;69:1318. 10.1001/archneurol.2012.1282. [PubMed: 22801742]

[16]. Theodore WH, Sato S, Kufta CV, Gaillard WD, Kelley K. FDG-positron emission tomography and invasive EEG: seizure focus detection and surgical outcome. Epilepsia 1997;38:81–6. 10.1111/j.1528-1157.1997.tb01081.x. [PubMed: 9024188]

[17]. Pustina D, Avants B, Sperling M, Gorniak R, He X, Doucet G, et al. Predicting the laterality of temporal lobe epilepsy from PET, MRI, and DTI: A multimodal study. NeuroImage Clin 2015;9:20–31. 10.1016/j.nicl.2015.07.010. [PubMed: 26288753]

[18]. Delbeke D, Lawrence SK, Abou-Khalil BW, Blumenkopf B, Kessler RM. Postsurgical outcome of patients with uncontrolled complex partial seizures and temporal lobe hypometabolism on 18FDG-positron emission tomography. Invest Radiol 1996;31:261–6. [PubMed: 8724123]

[19]. Vinton AB, Carne R, Hicks RJ, Desmond PM, Kilpatrick C, Kaye AH, et al. The extent of resection of FDG-PET hypometabolism relates to outcome of temporal lobectomy. Brain 2007;130:548–60. 10.1093/brain/awl232. [PubMed: 16959818]

[20]. Rathore C, Dickson JC, Teotónio R, Ell P, Duncan JS. The utility of 18F-fluorodeoxyglucose PET (FDG PET) in epilepsy surgery. Epilepsy Res 2014;108:1306–14. 10.1016/j.eplepsyres.2014.06.012. [PubMed: 25043753]

[21]. Chugani HT, Shields WD, Shewmon DA, Olson DM, Phelps ME, Peacock WJ. Infantile spasms: I. PET identifies focal cortical dysgenesis in cryptogenic cases for surgical treatment. Ann Neurol 1990;27:406–13. 10.1002/ana.410270408. [PubMed: 2353794]

[22]. Takahashi M, Soma T, Kawai K, Koyama K, Ohtomo K, Momose T. Voxel-based comparison of preoperative FDG-PET between mesial temporal lobe epilepsy patients with and without

postoperative seizure-free outcomes. Ann Nucl Med 2012;26:698–706. 10.1007/s12149-012-0629-9. [PubMed: 22810894]

[23]. Sakamoto S, Takami T, Tsuyuguchi N, Morino M, Ohata K, Inoue Y, et al. Prediction of seizure outcome following epilepsy surgery: Asymmetry of thalamic glucose metabolism and cerebral neural activity in temporal lobe epilepsy. Seizure 2009;18:1–6. 10.1016/j.seizure.2008.05.004. [PubMed: 18550392]

[24]. LoPinto-Khoury C, Sperling MR, Skidmore C, Nei M, Evans J, Sharan A, et al. Surgical outcome in PET-positive, MRI-negative patients with temporal lobe epilepsy. Epilepsia 2012;53:342–8. 10.1111/j.1528-1167.2011.03359.x. [PubMed: 22192050]

[25]. Joo EY, Hong SB, Han HJ, Tae WS, Kim JH, Han SJ, et al. Postoperative alteration of cerebral glucose metabolism in mesial temporal lobe epilepsy. Brain 2005;128:1802–10. 10.1093/brain/awh534. [PubMed: 15872014]

[26]. Carne RP, O'Brien TJ, Kilpatrick CJ, MacGregor LR, Hicks RJ, Murphy MA, et al. MRI-negative PET-positive temporal lobe epilepsy: a distinct surgically remediable syndrome. Brain 2004;127:2276–85. 10.1093/brain/awh257. [PubMed: 15282217]

[27]. Bell ML, Rao S, So EL, Trenerry M, Kazemi N, Matt Stead S, et al. Epilepsy surgery outcomes in temporal lobe epilepsy with a normal MRI. Epilepsia 2009;50:2053–60. 10.1111/j.1528-1167.2009.02079.x. [PubMed: 19389144]

[28]. Wong CH, Bleasel A, Wen L, Eberl S, Byth K, Fulham M, et al. The topography and significance of extratemporal hypometabolism in refractory mesial temporal lobe epilepsy examined by FDG-PET. Epilepsia 2010;51:1365–73. 10.1111/j.1528-1167.2010.02552.x. [PubMed: 20384730]

[29]. Newberg AB, Alavi A, Berlin J, Mozley PD, O'Connor M, Sperling M. Ipsilateral and contralateral thalamic hypometabolism as a predictor of outcome after temporal lobectomy for seizures. J Nucl Med 2000;41:1964–8. [PubMed: 11138679]

[30]. Choi JY, Kim SJ, Hong SB, Seo DW, Hong SC, Kim B-T, et al. Extratemporal hypometabolism on FDG PET in temporal lobe epilepsy as a predictor of seizure outcome after temporal lobectomy. Eur J Nucl Med Mol Imaging 2003;30:581–7. 10.1007/s00259-002-1079-8. [PubMed: 12557048]

[31]. Hashiguchi K, Morioka T, Yoshida F, Kawamura T, Miyagi Y, Kuwabara Y, et al. Thalamic hypometabolism on ^{18}FDG-positron emission tomography in medial temporal lobe epilepsy. Neurol Res 2007;29:215–22. 10.1179/174313206X153851. [PubMed: 17439707]

[32]. Gleichgerrcht E, Keller SS, Drane DL, Munsell BC, Davis KA, Kaestner E, et al. Temporal lobe epilepsy surgical outcomes can be inferred based on structural connectome hubs: A machine learning study. Ann Neurol 2020;88:970–83. 10.1002/ana.25888. [PubMed: 32827235]

[33]. Combrisson E, Jerbi K. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. J Neurosci Methods 2015;250:126–36. 10.1016/j.jneumeth.2015.01.010. [PubMed: 25596422]

[34]. Del Gaizo J, Fridriksson J, Yourganov G, Hillis AE, Hickok G, Misic B, et al. Mapping language networks using the structural and dynamic brain connectomes. ENeuro 2017;4:1–14. 10.1523/ENEURO.0204-17.2017.

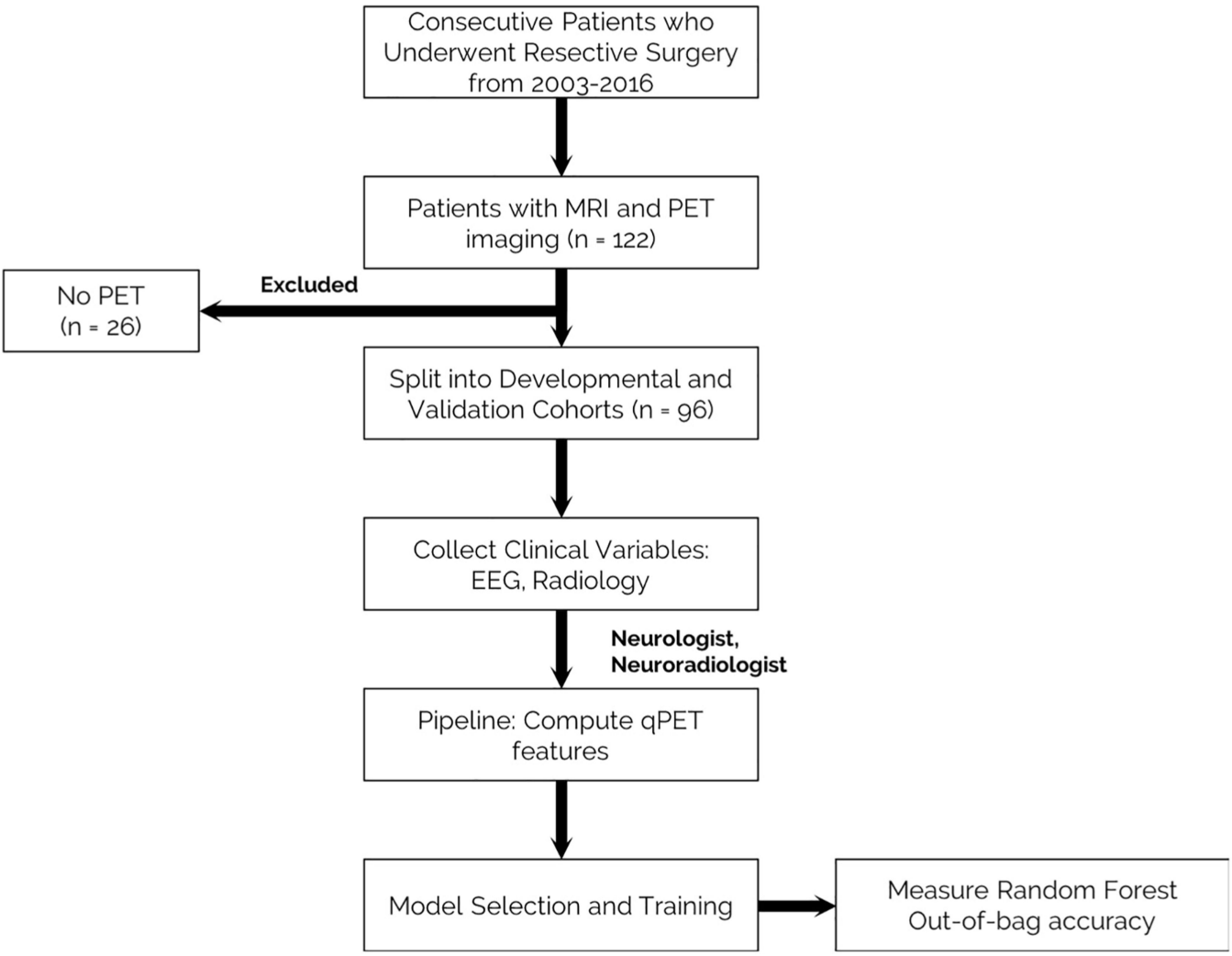

Fig. 1.

Workflow of data collection and analysis. All patients since 2003 who under surgical resective therapy for epilepsy were retrospectively screened according to inclusion and exclusion criteria. The included patients were then screened for availability of preoperative PET and MRI images. Clinical variables were determined by retrospective study of surgical case conference notes and medical records where available. Surgical outcome, scalp-EEG, and neuroradiology reads were confirmed by board-certified neurologists (M.G-B., K.D.) and radiologists (A.T., J.D.). All PET imaging was then processed using the computational pipeline outlined in Fig. 2, filtering out any imaging that was corrupt or unable to be processed. The final cohort of $n = 96$ patients (89 temporal) was then used as the cohort for models to subsample from and predict surgical outcome.

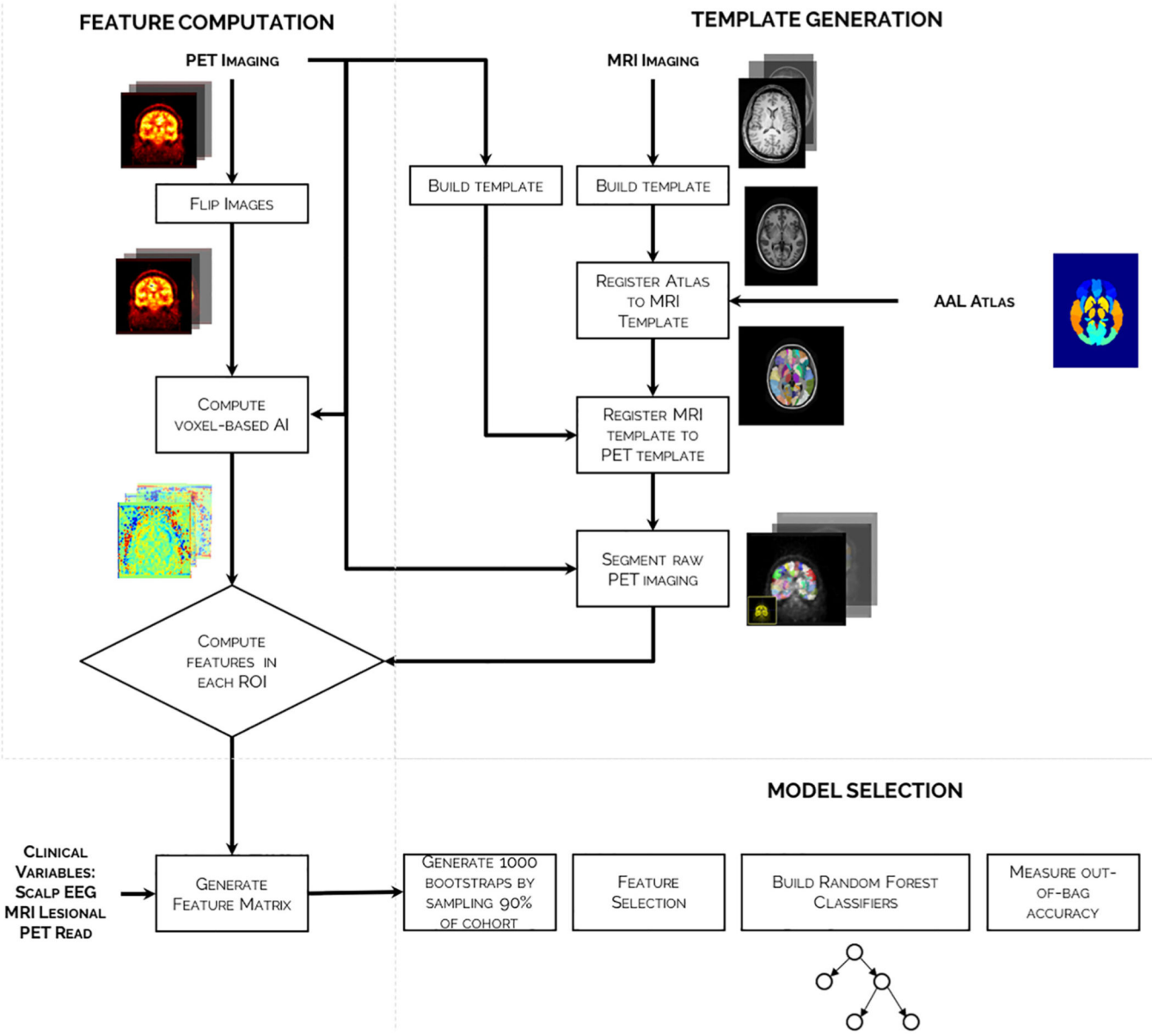

Fig. 2.

Computational pipeline. Computational pipeline that processes raw PET to generate quantitative imaging features across AAL region of interests (ROIs). First, a small subset of patient MRI and PET images are used to create a group-specific template in order to increase registration and segmentation accuracy when warping the AAL atlas to the patient's native PET domain. Second, all patient raw PET images are registered to its mirror image in order to generate a voxel-based asymmetry index measure. These processed results are then averaged and quantified across all gray matter AAL ROIs in order to generate the feature vector for each patient. All feature vectors across patients, along with one-hot encoded clinical variables, are merged to create the final feature matrix for model investigation.

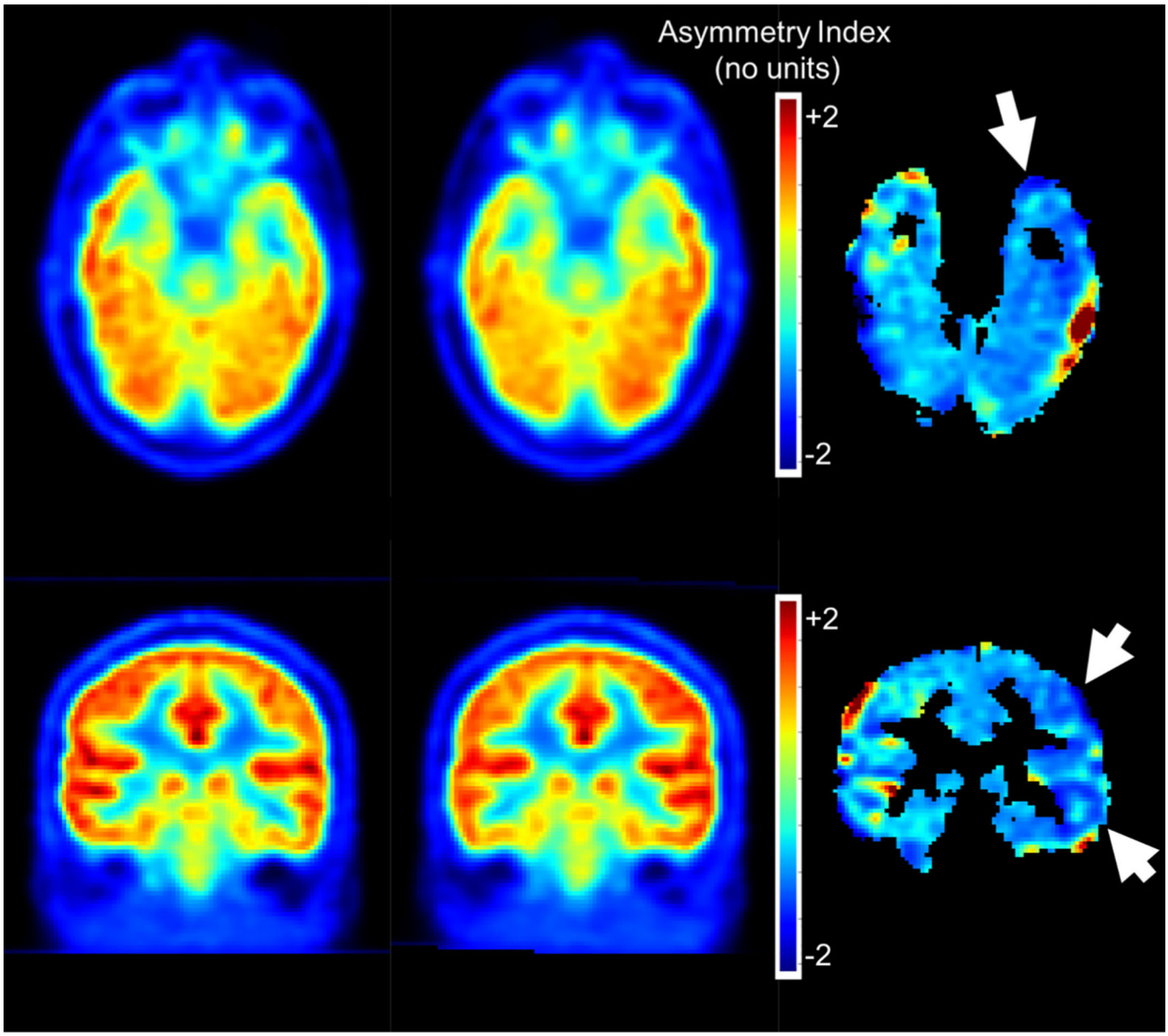

Fig. 3.

Quantitative imaging features. Quantitative asymmetry features can be visualized to highlight areas of asymmetry. In this figure, we show sample asymmetry feature maps derived from the pipeline for a patient with focal hypometabolism in the left-temporal lobe and parietal regions. The original PET (axial and coronal) is shown on the left. The flipped version of the same PET is shown in the middle. And, the voxel-based asymmetry map is shown on the right panel with blue indicating negative asymmetry and red indicating positive asymmetry (larger than corresponding contralateral region). Only gray matter regions in the AAL parcellation scheme were averaged and used in our models. White arrows indicate areas of hypometabolism that were noted in clinical PET reads. In this case, there were multiple areas of hypometabolism, most notably in the anterior temporal,

posterior inferior temporal, and parietal regions. The asymmetry map is a normalized ratio and is unitless.

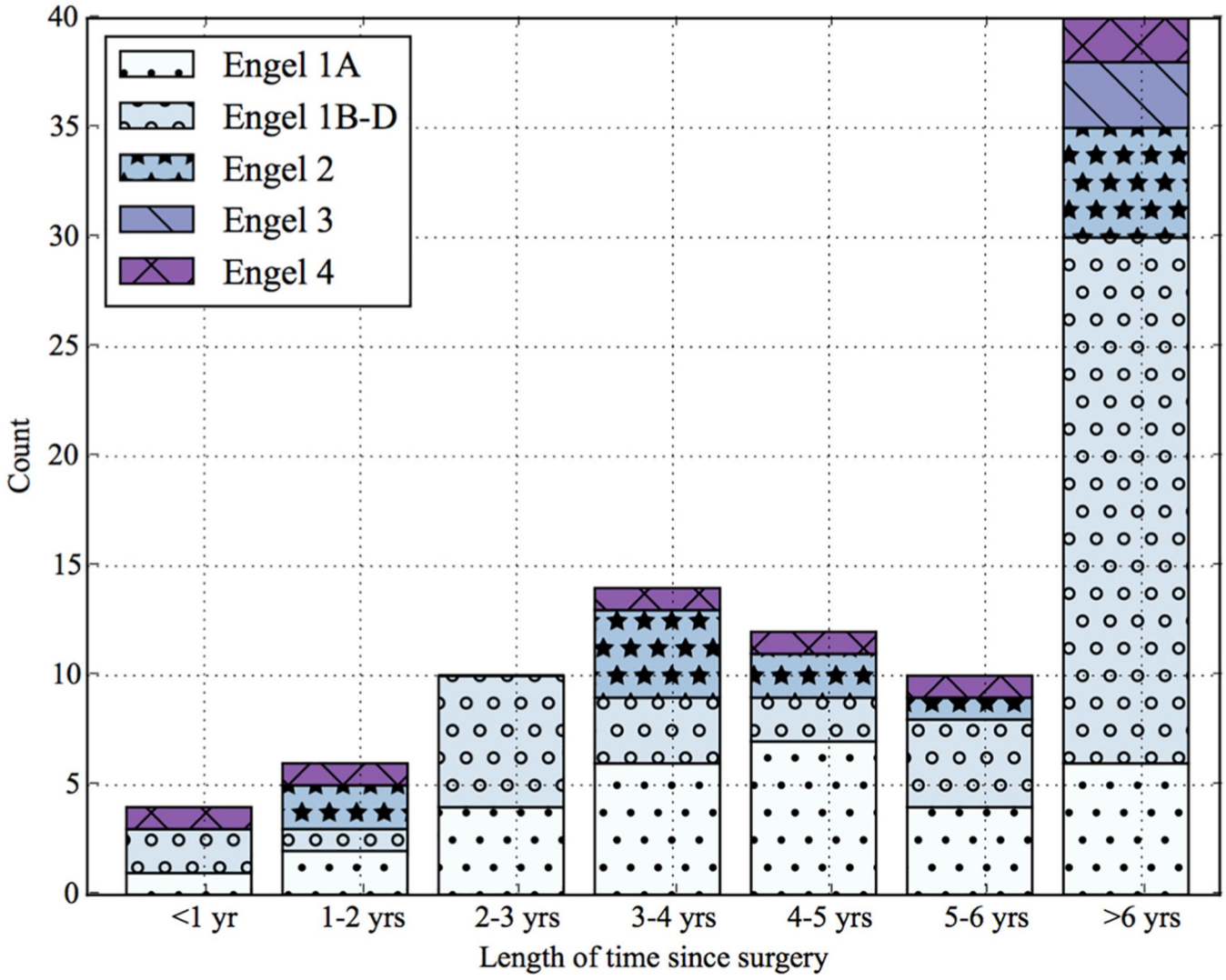

Fig. 4.

Distribution of surgical outcomes as a function of time since surgery. Histogram of surgical outcomes for all patients as a function of time of Engel outcome evaluation in months, with major timepoints in years following surgery and color-coded according to the outcome at the most recent office visit.

Table 1

Summary of Demographic Data. Patients in the cohort were grouped by the surgical outcome. First column shows patients who had complete seizure freedom (Engel IA). Second and third columns show patients who had Engel I seizure recurrence (Engel IB–D) and other Engel poor surgical outcome (Engel II–IV) respectively.

	Engel IA	Engel IB-ID	Engel II-IV	*p*-value[a]
Total number of subjects	32	33	31	
Age at surgery				
Mean/std dev	38+/−13	37+/−13	39+/−13	
Sex				0.45
Male	11	10	14	
Female	21	23	17	
Resected regions				0.29
LTL	16	15	16	
RTL	16	16	10	
LFL/RFL	0	1	3	
LPL/RPL	0	1	2	
MRI				0.38
Lesional	21	19	15	
Non-Lesional	11	14	16	
Pathology				0.97
HS/MTS	19	20	18	
Gliosis	7	6	4	
MCD	2	1	3	
Tumor/Vascular	1	2	1	
Dual Pathology	2	2	2	
Normal/Not Available	1	2	3	
PET Read				0.01
Focal	14	17	11	
Subtle	11	6	2	
Diffuse or Multifocal	3	3	10	
Normal	4	7	8	

[a]Pearson chi-square test.

FCD, focal cortical dysplasia; HS, hippocampal sclerosis; MTS, mesial temporal sclerosis; LTL, left-temporal lobe; RTL, right-temporal lobe; LFL, left frontal lobe; RFL, right frontal lobe; RPL, right parietal lobe; RFPL, right frontoparietal lobe.

Table 2

Comparison of models using qualitative PET to models using quantitative PET imaging features.

Model A Variables	Model A Out of Bag Accuracy (OOB_1)	Quantitative PET Model features (OOB_2)	Quantitative PET Model Out of Bag Accuracy (OOB_2)	$OOB_2 - OOB_1$ 95% CI
Prediction of seizure recurrence in ALL OUTCOMES (Engel IA vs. IB-ID, II-IV)				
Temporal Lobe patients (*n* = 89)				
Model A: EEG + MRI + Qualitative PET	0.61	**Model B:** EEG + MRI + Quantitative Asymmetry	0.71	(0.01, 0.18)
Model A: EEG + MRI + Qualitative PET	0.61	**Model C:** EEG + MRI + Qualitative PET + Quantitative Asymmetry	0.71	(0.01, 0.20)
Right-Temporal Lobe patients (*n* = 42)				
Model A: EEG + MRI + Qualitative PET	0.52	**Model B:** EEG + MRI + Quantitative Asymmetry	0.76	(0.09, 0.39)
Model A: EEG + MRI + Qualitative PET	0.52	**Model C:** EEG + MRI + Qualitative PET + Quantitative Asymmetry	0.72	(0.06, 0.36)
Left-Temporal Lobe patients (*n* = 47)				
Model A: EEG + MRI + Qualitative PET	0.60	**Model B:** EEG + MRI + Quantitative Asymmetry	0.65	(−0.05, 0.19)
Model A: EEG + MRI + Qualitative PET	0.60	**Model C:** EEG + MRI + Qualitative PET + Quantitative Asymmetry	0.64	(−0.11, 0.19)
Prediction of seizure recurrence in Engel I ONLY (Engel IA vs. IB-ID)				
Temporal Lobe patients (*n* = 89)				
Model A: EEG + MRI + Qualitative PET	0.40	**Model B:** EEG + MRI + Quantitative Asymmetry	0.72	(0.14, 0.52)
Model A: EEG + MRI + Qualitative PET	0.40	**Model C:** EEG + MRI + Qualitative PET + Quantitative Asymmetry	0.73	(0.14, 0.52)
Right-Temporal Lobe patients (*n* = 42)				
Model A: EEG + MRI + Qualitative PET	0.38	**Model B:** EEG + MRI + Quantitative Asymmetry	0.69	(0.08, 0.60)
Model A: EEG + MRI + Qualitative PET	0.38	**Model C:** EEG + MRI + Qualitative PET + Quantitative Asymmetry	0.76	(0.20, 0.64)
Left-Temporal Lobe patients (*n* = 47)				
Model A: EEG + MRI + Qualitative PET	0.41	**Model B:** EEG + MRI + Quantitative Asymmetry	0.70	(0.04, 0.54)
Model A: EEG + MRI + Qualitative PET	0.41	**Model C:** EEG + MRI + Qualitative PET + Quantitative Asymmetry	0.69	(0.00, 0.54)

Confidence intervals are computed on the difference of out-of-bag accuracy on 1000 bootstraps of stratified samples of the cohort. Of the 89 patients with temporal onset, 42 were right-temporal. The table shows the results of the model when looking at all temporal patients as well as only right- and left-temporal cases.