Predicting outcome following resective brain surgery remains a high-priority goal in neurology and neurosurgery. Recently, a need for classifying surgical outcome on the basis of functional status, such as cognition, rather than just seizure control, has been articulated. A number of neuroimaging techniques give us insight into the neuroanatomical correlates of cognition. These techniques are increasingly being used as part of presurgical planning algorithms, providing an opportunity to not only gain insights into the functional neuroanatomy of “at-risk” cognition and how it relates to the underlying pathology, but also use such insights to predict cognitive outcome following surgery, such as anterior temporal lobectomy (ATL) for epilepsy.

The predictive power and the combined clinical utility of 3 major MRI techniques (functional MRI [fMRI], resting state fMRI [rsfMRI], and diffusion tensor imaging [DTI]) are unknown and untested. These 3 modalities can be conceptually combined to obtain a more complete view of the neural modules and connectivity networks implementing cognition and driving adaptive or maladaptive neuroplasticity responses in the brain. Functional MRI can be used to define an active network implementing a key function that might be “at risk” with temporal lobe surgery (e.g., verbal fluency), although the validity of this approach needs to be established in prospective studies. Preoperative fMRI has been shown to be sensitive to language and verbal memory laterality and is a known predictor of cog-
nitive outcome in the setting of epilepsy. For instance, a large fMRI study of verbal and visual memory in patients undergoing left or right anterior temporal lobectomy suggested that presurgical fMRI activation of the ipsilateral anterior temporal lobe was correlated with greater postsurgical memory decline. Resting state fMRI can be used to capture correlated neural signaling, identifying the network of gray matter regions that are biased toward neural communication, although no research has yet evaluated its surgical applicability. DTI can be used to evaluate the integrity of and specify the anatomical connectivity between the functionally relevant gray matter regions. One published study has reported that a change in fractional anisotropy correlates with good language outcome following temporal lobectomy, but this study lacked a normal control group. In short, rsfMRI can identify the cortical regions communicating within a network; DTI can verify that these regions are anatomically connected; and fMRI can define the cognitive function(s) the network may be communicating about. In practical terms, all three can be gathered during one MRI scanning session with a minimal increase in scanning time and no increase in patient discomfort.

Despite the growing use of task-driven fMRI, DTI, and even resting state imaging for clinical purposes, our knowledge of the comparative sensitivity of each technique to cognitive and behavioral change following surgery is limited. Nor do we understand how to use these 3 imaging techniques in an integrated fashion within our presurgical algorithms as a means of enhancing surgical outcomes. More specifically, to date no empirical study has examined all 3 modalities for their single or joint power in predicting cognitive or seizure outcome following surgery. It is important to note that the most reliable current predictors of postsurgical cognitive outcome remain baseline neurocognitive testing, the intracarotid amobarbital examination, the surgical hemisphere (dominant, non-dominant), and epilepsy duration, and no study has compared the predictive power of the emerging imaging modalities with some of the more established predictors.

In this report, we investigate the predictive power of these imaging modalities by calculating the degree of deviation from normal for each modality, focusing on a specific network involved in a key at-risk function, namely verbal fluency. The relative change in this deviation score from pre- to postsurgery is used to capture alterations and potential reorganization in the verbal fluency network following ATL. Although the role of the dominant mesial temporal lobe in expressive language is poorly understood, we chose verbal fluency as word-finding and expressive language difficulties are among the most common complaints made by patients after resective surgery.

Based on the unique information about brain networks offered by each imaging modality, we hypothesize that each will contribute important information about postsurgical cognitive outcome, making a combined, multimodal statistical model the most robust in terms of effectively predicting postsurgical cognitive outcome. We use fMRI to create a brain map of verbal fluency followed by characterizations of resting state functional connectivity and white matter anatomical connectivity utilizing the fMRI verbal fluency network as the seed. More specifically, we hypothesize that utilizing all 3 modalities will be superior at predicting verbal fluency outcome compared to any single imaging modality, as well as to established demographic and neuropsychological predictors. Finally, because we are focusing on verbal (semantic) fluency, we expect change in the fMRI networks to be the most predictive single modality, as this modality relies on a semantic fluency task as stimulation.

**Methods**

**Participants**

A total of 15 patients with left temporal lobe epilepsy were recruited from the Jefferson Comprehensive Epilepsy Center. A total of 20 age-matched healthy controls were recruited from the Thomas Jefferson University community. All patients underwent a standardized en bloc resection of the anterior temporal lobe, including amygdalohippocampectomy, as treatment for medically intractable temporal lobe epilepsy (TLE). Patients underwent neuroimaging before (mean 45 days) and after (mean 219 days) surgery. For details of the Jefferson Comprehensive Epilepsy Center algorithm for surgical decision making see Sperling et al. All patients met the following inclusion criteria: unilateral temporal lobe seizure onset (as localized through surface video/electroencephalography recordings); MRI and/or PET confirmation of epileptogenic temporal lobe pathology; and a Full Scale IQ of at least 75. All controls were free of any psychiatric or neurological disorders. The study was approved by the Institutional Review Board for Research with Human Subjects at Thomas Jefferson University, and all participants provided written informed consent. Table 1 outlines the demographic and clinical characteristics of the participants.

**Scanning Parameters**

All patients underwent MRI with a 3-T X-series Philips Achieva clinical MRI scanner using an 8-channel head coil during a 4-year period (2007–2011). There were no changes to the scanner software (reconstructor version 8), hardware, or sequence during that time. Participant data were collected in a pseudo-randomized, interleaved fashion (i.e., not all preoperative scans were completed before all postoperative scans), so as to not introduce any bias based on temporally dependent scanner calibration.

Each participant’s data were acquired utilizing the following scan order: survey, field reference inhomogeneity scan, T1 (for details on T1 data acquisition, see Osipowicz et al.), and experimental scans (fMRI, rsfMRI, and DTI). The experimental scans were collected in positions identical to and prescribed by the MPRage (anterior commissure–posterior commissure line). All data were collected and analyzed using standard and established methods. For details and the specific parameters related to collection and analysis of each image modality, see Osipowicz et al. and Tracy et al. and Supplemental Data.

**Neurocognitive Measure of Cognitive Outcome and Cognitive Task Network Definition Through fMRI**

Patients were classified according to cognitive outcome,
TABLE 1. Summary of patient demographic and clinical characteristics

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Sex</th>
<th>Age (yrs)*</th>
<th>Ep Dur (yrs)</th>
<th>BL</th>
<th>Postop</th>
<th>Cognitive Outcome†</th>
<th>Path</th>
<th>Sz Freq/ Mo</th>
<th>Sz Type</th>
<th>EC‡</th>
<th>AEDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>26</td>
<td>21</td>
<td>24</td>
<td>32</td>
<td>Good Atrophy</td>
<td>1</td>
<td>CPS, SPS, CPS/G</td>
<td>A</td>
<td>Levetiracetam, lamotrigine</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>28</td>
<td>12</td>
<td>19</td>
<td>12</td>
<td>Poor Enlarge (horn)</td>
<td>2</td>
<td>CPS, CPS/G</td>
<td>A</td>
<td>Levetiracetam, oxcarbazepine</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>41</td>
<td>36</td>
<td>11</td>
<td>3</td>
<td>Poor MTS</td>
<td>16</td>
<td>CPS, SPS</td>
<td>C</td>
<td>Levetiracetam, valproic acid, phenobarbital</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>30</td>
<td>11</td>
<td>21</td>
<td>20</td>
<td>Good MTS</td>
<td>2</td>
<td>CPS, GTC</td>
<td>A</td>
<td>Levetiracetam</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>50</td>
<td>31</td>
<td>21</td>
<td>13</td>
<td>Poor Atrophy</td>
<td>7</td>
<td>CPS, CPS/G</td>
<td>C</td>
<td>Phenytoin</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>47</td>
<td>42</td>
<td>21</td>
<td>11</td>
<td>Poor Hippo atrophy</td>
<td>15</td>
<td>CPS</td>
<td>A</td>
<td>Carbamazepine, irbesartan</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>42</td>
<td>4</td>
<td>18</td>
<td>8</td>
<td>Poor MTS</td>
<td>1</td>
<td>CPS</td>
<td>A</td>
<td>Lamotrigine, memantine</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>32</td>
<td>8</td>
<td>17</td>
<td>18</td>
<td>Good MTS</td>
<td>1</td>
<td>CPS, CPS/G</td>
<td>C</td>
<td>Lamotrigine, carbamazepine</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>53</td>
<td>40</td>
<td>13</td>
<td>18</td>
<td>Good MTS</td>
<td>5</td>
<td>CPS, GTC</td>
<td>A</td>
<td>Lamotrigine, carbamazepine</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>65</td>
<td>45</td>
<td>16</td>
<td>20</td>
<td>Good Hippo atrophy</td>
<td>7</td>
<td>CPS</td>
<td>A</td>
<td>Topiramate, olanzapine</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>39</td>
<td>26</td>
<td>16</td>
<td>21</td>
<td>Good Hippo atrophy</td>
<td>1</td>
<td>CPS</td>
<td>A</td>
<td>Valproic acid, phenobarbital, carbamazepine</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>60</td>
<td>56</td>
<td>18</td>
<td>11</td>
<td>Poor MTS</td>
<td>2</td>
<td>CPS, SPS</td>
<td>A</td>
<td>Levetiracetam</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>52</td>
<td>32</td>
<td>23</td>
<td>17</td>
<td>Poor Enlarge (horn)</td>
<td>1</td>
<td>CPS, GTC</td>
<td>A</td>
<td>Lamotrigine, carbamazepine</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>41</td>
<td>37</td>
<td>26</td>
<td>25</td>
<td>Good MTS</td>
<td>2</td>
<td>CPS, GTC</td>
<td>A</td>
<td>Venlafaxine, lamotrigine, levetiracetam</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>26</td>
<td>5</td>
<td>14</td>
<td>17</td>
<td>Good Atrophy</td>
<td>20</td>
<td>CPS</td>
<td>A</td>
<td>Diazepam, levetiracetam, topiramate</td>
<td></td>
</tr>
</tbody>
</table>

Mean (SD) 42 (12) 27 (16) 18 (4) 16 (7) 5 (6)

* Age at baseline scan.
† Based on RCI.
‡ Based on Engel Class I subscore; enlarge = enlargement; Ep Dur = epilepsy duration; GTC = generalized tonic-clonic seizures; Hippo = hippocampal; MTS = mesial temporal sclerosis; Path = temporal lobe pathology; SPS = simple partial seizures; Sz = seizure; VFL = verbal (semantic) fluency.

Table: Multimodal prediction of verbal outcome

utilizing a verbal fluency measure collected as part of the neuropsychological assessment of language conducted on a presurgical (within 3 months of surgery) and postsurgical (1 year postsurgery) basis. The Controlled Oral Word Association Test subtest involving Semantic [animal] Fluency was administered under standard conditions and instructions.43 During the task, the participant must generate members of the category “four-legged animals” for 1 minute, with the total number generated used as the variable of interest. The task requires semantic processing and lexical access of common nouns. A Reliable Change Index (RCI)17 was computed for this measure, revealing that a change in either direction of 4 raw score points on the semantic fluency measure reflected reliable change. In terms of neuropsychological outcome following brain surgery, evidence of no change or improvement in performance is generally considered to constitute a good outcome, as both patterns represent a preservation of the “at-risk” skill or function. Therefore, based on these RCI data, we defined 2 groups of cognitive outcome, capturing either a good (improved or unchanged performance on the Semantic Fluency measure) or a poor outcome A total of 8 of 15 patients obtained a good outcome, and 7 of 15 obtained a poor outcome.

A verb-generation task was used as task-driven stimulation during fMRI. This task requires semantic processing of a visually presented common noun and lexical access of verbs and action words associated with that noun (see supplemental material online for details on the verb-generation task). Prior work has demonstrated that this task reliably establishes expressive language lateralization and shows activation in Broca’s area.11,22,31

Definition of the Normal Control Templates for Quantifying Network Deviations in Left ATL Patients

The above verb-generation task was administered to our sample of normal healthy controls. The group activation map based on statistical comparison of the verb-generation condition with the visual control condition revealed activation in the middle frontal gyrus (BA47); bilateral medial frontal gyri extending superiorly to include the superior frontal gyri (BA6); left middle temporal gyrus (BA22) extending inferiorly to include the fusiform gyrus (BA37); and the right inferior frontal gyrus (BA44) (see Fig. 1). These areas of significant activation defined the normative functional network that served as a seed for subsequent rsfMRI and DTI analyses.

Analyses (see Supplemental Data for methods) of our normal control rsfMRI data using the above fMRI-defined verb-generation network as seed revealed a large degree of functional connectivity within the left hemisphere. The significantly correlated regions include the left inferior frontal gyrus, extending posteriorly to include pre-
K. Osipowicz et al.

motor and primary motor areas and extending inferiorly to include the anterior temporal lobe. From the anterior temporal lobe this cluster extended posteriorly to include the posterior middle temporal gyrus, extending superiorly to include the superior temporal and supramarginal gyri. Also, bilateral medial frontal gyri functional connectivity was observed, quite similar to the bilateral medial frontal activations observed under fMRI. Lastly, in the right hemisphere, we saw a connectivity cluster that included the inferior frontal lobe and anterior and middle temporal lobes (see Fig. 1).

Using the same fMRI seeds for DTI tractography (see supplemental data online for methods), we saw white matter anatomical connectivity that involves the inferior longitudinal fasciculus, uncinate fasciculus, inferior fronto-occipital fasciculus, arcuate fasciculus, forceps major, corpus callosum (posterior splenium), and superior longitudinal fasciculus, all bilaterally. It should be noted that these tracks were more robust in the left hemisphere, with the left dominance most prominent in the forceps major and the superior longitudinal fasciculus (see Fig. 1).

Deriving a Single Measure of Change From Pre- and Postsurgical Deviation Scores

Methodological problems in measuring how to detect real versus random variation over time are important considerations for this type of project. To avoid problems we operationalize change by comparing each patient’s imaging data with a modality-specific index of their deviation from normal at each point in time (presurgery, to be referred to as baseline, and postsurgery). This procedure provides a systematic means of quantifying the typicality of both pre- and postsurgical neurocognitive network organization, a procedure that alleviates concern about “double dipping” of analysis results. Our methodology also minimizes reliance on random, session-specific fluctuations as part of the reference or comparison point for determining change.

For each imaging modality, we used the average of the maps for the normal controls (fMRI activation, rsfMRI connectivity, DTI tracts) as a template of the normal network and computed the goodness of fit between this template and each patient’s individual map. We replicated the template-matching procedure used by Gross et al. In short, this is a nonlinear procedure that takes the number of the patient’s voxels falling within the template relative to the total number of voxels in the template and subtracts the number of the patient’s voxels falling outside of the template relative to the total number of brain voxels, yielding a goodness of fit for each individual map (Fig. 2). It should be noted that the type of score present in each voxel varies

![Diagram](image)

**FIG. 2.** Calculation of the deviation measure: A represents the number of voxels in the relevant template; D represents the number of voxels in the brain absent the template (therefore, total brain volume equals A + D); B represents the number of the patient’s voxels that fall within the template (i.e., overlap between patient and template); C represents the number of the patient’s voxels that do not fall within the template. Figure is available in color online only.
for the different types of maps. For instance, fMRI utilizes a t-score, rsfMRI a Z-score, and DTI an inclusion variable. In interpreting the goodness-of-fit index, a score of 1.0 indicates a perfect fit with the normative template (all voxels within the template and none outside), and a score of zero indicates a poor fit (no voxels within the template; all outside the template) (See Fig. 3). For the logistic regression, once the goodness of fit is computed, this is transformed to a normalized deviation and expressed as a normal Z-score.

The change in this deviation score between the post-and presurgical sessions was computed (i.e., postsurgical goodness of fit minus presurgical goodness of fit; referred to as the change score). For each modality, the closer the change score is to zero, the less change is observed across the 2 scanning sessions.

Statistical Analyses
To predict cognitive outcome, we built logistical regression statistical models comparing the change measures for each modality for their ability to reliably distinguish between good and poor cognitive outcome. We also assessed the predictive power of the baseline (presurgical) deviation of each of the imaging modalities. As a second level of analyses, we reexamined the ability of any significant predictors yielded by the above analyses for their ability to outperform rival hypotheses in terms of predictive power. These rival hypotheses included baseline measures of neuropsychological function (baseline semantic fluency), demographic variables (age, baseline Verbal IQ), and clinical measures representing important characteristics of epilepsy. All results reported obtained an alpha (p value) of ≤ 0.05.

Results
Table 1 displays the clinical and demographic data for all participants. There was no observable difference between the groups on any of the variables listed, as assessed by Student t-test or chi-square test, as appropriate. Figure 3 displays the goodness-of-fit index for all 3 imaging modalities for the good and poor outcome groups at baseline (presurgery) and postsurgery.

Comparing the Imaging Modalities for Their Ability to Predict Cognitive Outcome
In the logistic regression (unique sum of squares) on cognitive outcome containing the change scores of all 3 modalities, this 3-variable model was significant (see Table 2 and Fig. 4). The model explained 52% of the variance in outcome and correctly classified 87% of the patients. While no single effect in the model was significant, the standardized beta values suggested that DTI may have contributed the most in terms of predictive power. A classification plot (see Fig. 4) shows the good spread between good and poor outcome provided by the 3-modality predictive model.

To investigate the effect each of the change measures had on the model, we computed the incremental R² for each change measure. This showed that DTI accounted for 32%; resting state for 15%; and fMRI for 4% of the observed variance. While no single-variable model significantly predicted outcome, the 2-variable model involving rsfMRI and DTI was significant, accounting for 49% of the variance in outcome.

We examined the 3 baseline deviation scores in a separate logistic regression model (unique sum of squares) of cognitive outcome. Here, the 3-variable model produced a statistical trend (p < 0.075), with 37% of the variance in outcome explained and 73% of the patients correctly classified on outcome. No single effect in the model was significant, although the standardized beta values suggested that baseline DTI deviation may have contributed the most in terms of predictive power.

Next, we sought to determine whether the 3-variable model involving change scores was a better predictor of cognitive outcome after accounting for the effects of the baseline imaging data. The 3-variable logistic regression was rerun using each modality’s baseline (presurgical) deviation score as a covariate, with separate logistic regressions for each covariate. In each case, the 3-variable model remained significant, with the variance explained comparable to the original logistic regressions (51.9% for the original model; 52.3% for the model with the fMRI baseline; 52.1% with the rsfMRI baseline; 52.3% with the DTI baseline). Again, among the predictor variables, no standardized beta was significant. In no case did the standardized beta value for baseline deviation score exceed that of the change scores, indicating that after accounting for these baseline covariates, the original 3-variable model containing the change scores for all 3 imaging modalities remained significant.

### Table 2. Results of the main logistical regression models

<table>
<thead>
<tr>
<th>Model</th>
<th>Change Score</th>
<th>Baseline</th>
<th>Change &amp; Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model significance</td>
<td>0.012</td>
<td>0.075</td>
<td>0.002</td>
</tr>
<tr>
<td>Chi-square</td>
<td>10.96</td>
<td>6.89</td>
<td>20.7</td>
</tr>
<tr>
<td>Cox &amp; Snell R²</td>
<td>0.519</td>
<td>0.369</td>
<td>0.749</td>
</tr>
<tr>
<td>% correct classification</td>
<td>86.7</td>
<td>73.3</td>
<td>100</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>1.4</td>
<td>1.63</td>
<td>—†</td>
</tr>
<tr>
<td>Strongest predictor*</td>
<td>DTI (-1.74)</td>
<td>DTI (2.1)</td>
<td>—</td>
</tr>
</tbody>
</table>

* The strongest predictor is noted with its beta weight in parentheses. 
† Because there was no error in correct classification, the odds ratio and strongest predictor could not be estimated for the model containing both change and baseline deviations.
Lastly, to better understand the strength of the above 3-modality model in predicting cognitive outcome, we examined demographic and clinical measures to determine if they constitute rival or alternate explanatory factors. Similar to the logistic regression analysis with the baseline deviation scores, we reran the 3-variable logistic regression model, in separate analyses, with each of the above measures inserted as a single covariate. The specific covariates tested were epilepsy duration, age, baseline animal naming score, and baseline Verbal IQ. For each covariate, the 3-variable logistic regression model remained significant with the proportion of variance explained in cognitive outcome comparable to the initial model (52.7% for model with epilepsy duration; 52.9% with age; 59.3% with animal naming; 56.7% with Verbal IQ). In no case did the standardized beta value for covariate exceed that of the change scores, indicating that after accounting for these demographic and clinical measures, the original 3-variable model containing the change scores remained significant.

Discussion

The purpose of this study was to test whether the changes observed on 3 brain mapping modalities are predictive of postsurgical cognitive outcome, specifically verbal fluency, following dominant-side ATL for intractable TLE. The data show that a statistical model combining 3 change variables, one from each modality, reliably predicts verbal fluency outcome, outperforming any model containing change scores from any single or duplue pairing of the modalities. We found that a 2-variable model with rsfMRI and DTI (with fMRI excluded) also produced a statistically significant model of prediction. However, as was the case with the 3-variable model, no individual modality was predictive of outcome. In the 3-variable model, DTI accounts for the largest single amount of observed variance in verbal fluency outcome.

The direction of change scores indicated that for the patients with good outcomes, all 3 imaging modalities displayed a better match to the normative template postsurgery, suggesting that normalization had occurred (see goodness-of-fit data, Fig. 3). In quantitative terms, the degree of fit relative to the normative template changed most dramatically for the resting state functional connectivity measure. This pattern suggests that achieving a good outcome involves engaging normative regions to complete the task, regions that were previously less involved. Yet, to do so appears to require achieving more normative forms of connectivity for both gray and white matter. This indicates that when the anterior and mesial temporal lobe are absent, supporting prior or improved levels of performance requires a new, more typical patterning of white and gray matter connectivity.

In contrast, for the patients with poor outcomes the pattern of change across the 3 modalities was mixed. The task-fMRI measure showed a shift toward the template, a better fit postsurgery, but the resting state and DTI measures shifted away from the template. Our data suggest that good cognitive outcomes, even in the setting of more normative fMRI activation, require changes in functional and white matter connectivity before cognitive change can be effected. Gray and white matter connectivity, less normative in our poor-outcome patients, may not have been providing sufficient scaffolding to implement the neural communication necessary for the task. Said differently, the regional reorganization to implement better verbal fluency was present postsurgery, but the neuronal networks to support communication within the verbal fluency network were not in place. As to whether these resting state and white matter networks would, with time, reorganize, normalize, and come to implement a good verbal fluency outcome remains an open question. An intriguing possibility is that connectivity and task-related fMRI activations are on different trajectories postsurgery, reorganizing at different rates in good- and poor-outcome patients. Both the poor- and good-outcome groups were scanned at a mean interval of approximately 5 months postsurgery, but the poor-outcome group may have needed more time to recovery. For instance, the good-outcome group may recover and reorganize rapidly, implementing the task with the same areas seen in controls, areas that have renewed health as they have been unburdened by the impact of seizures. The poor-outcome group may be on a slower or less-effective trajectory, perhaps because changes in functional and white matter connectivity are hampered by the now-missing temporal lobe in ways that they are not in the good-outcome group. We should note that we sought to uncover meaningful group differences between the poor- and good-outcome groups in hopes of establishing the clinical features that might support their very differ-
multimodal prediction of verbal outcome

Multimodal prediction of verbal outcome

ent imaging patterns postsurgery, but we observed no reliable group differences. Another possibility is that the more aberrant connectivity characteristics of the poor-outcome group postsurgery suggest that these patients did not have the reserve capacity to reorganize network connectivity and normalize cognitive performance in the face of ATL.

Importantly, we found that the model containing the baseline (presurgical) measures from the 3 modalities was a weak and nonsignificant predictor of cognitive outcome. The combined 3-modality model of change remained significant even after accounting for these baseline measures. We also found that the key 3-modality model of change was a superior predictor of verbal fluency when competing against established predictors of cognitive outcome such as baseline neuropsychological measures of verbal fluency or IQ, as well as demographic and clinical measures. The 3-modality model of change remained significant regardless of whether these rival hypotheses were examined singly or together against the 3-variable model. This is important, as it demonstrates that these neuroimaging modalities have added value compared with extant models of predication.

Our data also make it clear that verbal (semantic) fluency is an at-risk function in patients with left temporal lobe epilepsy, as a significant proportion of our patients (roughly half of the sample) did experience a decline in this skill despite having a good surgical outcome in terms of seizure control (n.b., all of our patients were seizure free postsurgery). Our data show that what best distinguished the good- from the poor-outcome patients was the change profile, not the baseline status, of their neuroimaging. While change was key, and superior to the baseline findings, there was still a statistical trend suggesting that with larger samples baseline neuroimaging characteristics may also become a reliable predictor. Such data suggest that presurgical deviation from normal (a measure of reorganization in response to pathology) has the potential to become an important index of potential cognitive outcome.

We were surprised to find that fMRI deviation was not a significant predictor of cognitive outcome, as we had initially hypothesized, and that the model containing only rsfMRI and DTI was a significant predictor of cognitive outcome. We believe this may be caused by 2 factors. Both rsfMRI and DTI are measures of connectivity, and it would make sense that changes in connectivity might be more sensitive to reorganization than a measure more sensitive to the nodes in the network arising from the implementation of a specific task. Said differently, rsfMRI and DTI may capture network changes that are part of a more general reorganization of language, with the assumption that it may be more parsimonious for the brain to reorganize a network to preserve functions for a wide range of tasks, rather than link the reorganization to a single task. In this sense, our fMRI results may be too task specific. Perhaps utilizing a different language measure or multiple language tasks would increase the strength of the relationship between fMRI and cognitive outcome. We used a verb-generation task because it is a reliable language-lateralization task, and because we have observed deficits in semantic fluency in TLE patients. Perhaps an fMRI task that is designed to specifically activate the anterior temporal lobe, such as the task described by Binder et al. and Hamberger et al. would provide change scores that are predictive of verbal fluency. It is important to note that since our analyses focused on networks defined by fMRI seeds, even though fMRI was not the best predictor, it is still necessary to define the other networks.

These results support our original hypothesis that a multimodal predictive model would be superior to any single modality in predicting verbal fluency outcome. This suggests that cognitive reorganization occurred in our ATL patients, with each modality capturing a unique aspect of that reorganization, compelling and recommending the use of all 3 modalities when trying to obtain the most reliable index of change. Importantly, our method quantified the deviation of the neural connectivity and the functional and anatomical neural networks of our patients from normal, examining the relative difference in those deviations as our index of change. In so doing, we avoided the use of our presurgical neuroimaging results as the template for comparing and determining change, minimizing concerns of double dipping of analysis results.

Our data indicate that ATL does reorganize networks. Prior research has demonstrated that TLE with an early age of onset or long duration can reorganize the neural representation of language. Ours is the first report that ATL, in the setting of a good seizure outcome, also results in a change in the neural representation of language evident both in terms of regional activation and the functional and white matter connectivity between these regions. Furthermore, our data also make clear that removal of the mesial temporal lobe alters language functioning and underlying language networks. Prior neuropsychological studies of language outcome following ATL have shown mixed results, with some showing a persistence of deficits, and others showing normalization. A prior neuroimaging study investigating the effect of ATL on language reorganization demonstrated the potential negative impact on language network organization. None of the above studies, however, presents evidence of reorganization from other imaging modalities. Thus, ours is the first study to demonstrate that fMRI language network reorganization co-occurs with functional connectivity changes on resting state fMRI and white matter changes on DTI. We must emphasize again, however, that our data make clear that what most reliably distinguishes good-versus poor-outcome patients is the change and reorganization observed on all 3 levels of network analysis.

There are a number of study limitations that need to be addressed. Our study suffered from a small sample size, limiting power, perhaps making it more impressive that we were still able to predict cognitive outcome. Although we focused on cognitive outcome, we had also hoped to report the ability of our 3 imaging modalities to predict seizure outcome, the ultimate test of the success of an ATL. However, we observed no variability in seizure outcome—all patients in our sample had a good outcome at the point of our assessment.

To our knowledge, this is the first study to examine the power of 3 neuroimaging techniques for predicting cognitive outcome. Our data suggest that by calculating the presurgical deviation (from controls) of results of 3 neu-
roimaging techniques we can predict cognitive outcome following ATL for medically intractable TLE. We propose that the best predictive models are those that contain rs-fMRI, DTI, and tMRI. Functional MRI is necessary to generate the seeds for resting state networks and DTI tracts, and then using all 3 modalities synergistically is the best way to predict postsurgical cognitive outcome. Future research should investigate whether other language tasks, or tasks from other cognitive domains such as memory, improve or alter the relative or combined predictive power of the 3 modalities.

Acknowledgment
This work was supported, in part, by a grant from the National Institute for Neurological Disorders and Stroke (grant no. R21 NS056071-01A1) to Dr. Joseph I. Tracy.

References

Disclosure
Dr. Sharan reports an ownership interest in ICVRX and ICP and a consultant relationship with Medtronic and SJM. Dr. Sperling reports a consultant relationship with Accorda Therapeutics and electroCore.

Author Contributions
Conception and design: all authors. Acquisition of data: Osipowicz, Sharan, Tracy. Analysis and interpretation of data: Osipowicz, Sperling, Tracy. Critically revising the article: Osipowicz, Tracy. Reviewed submitted version of manuscript: Osipowicz, Tracy. Approved the final version of the manuscript on behalf of all authors: Osipowicz. Statistical analysis: Osipowicz, Tracy. Administrative/technical/material support: all authors. Study supervision: Osipowicz, Tracy.

Supplemental Information
Online-Only Content
Supplemental material is available with the online version of the article.

Correspondence
Karol Osipowicz, Department of Psychology, Drexel University, 276 Stratton Hall, 3521 Chestnut St., Philadelphia, PA 19104. email: dr.o@drexel.edu.