Non-Invasive Brain Mapping Localizes Essential Language Function in Surgical Glioma Patients

Matthew Muir, BS[1], Kyle Noll, PhD[2], Sarah Prinsloo, PhD[1], Hayley Michener, MS[1], Priscella Asman, PhD[1], Jeffrey I. Traylor, MD[4], Vinodh A. Kumar, MD[5] , Chibawenye Ene, MD, PhD[1], Ho-Ling Liu, PhD[3], Sujit S. Prabhu, MD[1*]

[1]Department of Neurosurgery, The University of Texas M.D. Anderson Cancer Center, Houston, Texas (USA)

[2]Department of Neuro-Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas (USA)

[3]Department of Imaging Physics, The University of Texas M.D. Anderson Cancer Center, Houston, Texas (USA)

[4]Department of Neurological Surgery, The University of Texas Southwestern Medical Center, Dallas, Texas (USA)

[5]Department of Neuroradiology, The University of Texas MD Anderson Cancer Center, Houston, Texas (USA)

Keywords: glioma; eloquent; functional imaging; tractography; aphasia

*CORRESPONDENCE

Sujit S. Prabhu, MD

1515 Holcombe Boulevard

Room FC7.2000, Unit 442Houston, TX 77030-4009

Abstract

The essential neuroanatomical structure of language processing remains unclear. Here we show in a cohort of 79 consecutive glioma patients undergoing surgery in the dominant hemisphere that postoperative language deficits result from subcortical damage to individualized ROIs within normative white matter tracts. We integrate these findings into a data-driven, non-invasive brain mapping approach capable of using cortical transcranial magnetic stimulation (TMS) mappings and diffusion imaging to localize language-essential regions and predict the long-term functional outcome of personalized surgical strategies with an accuracy of 94%. We use this technique to create a group-level probabilistic atlas of functional white matter in glioma patients, revealing evidence for a novel, disease-specific pattern of subcortical connectivity deep to the temporo-parietal-occipital (TPO) junction that anatomically correlates with regional vulnerability to aphasic damage. This study provides novel insight into the functional anatomy of language processing while defining a reproducible approach for non-invasively mapping essential language function.

Introduction

Glioma patients undergoing surgery in the dominant hemisphere sometimes sustain permanent deficits in language function that significantly decrease both quality of life and survival.[1-3] Standard-of-care preoperative imaging fails to consistently localize meaningful function. Structural anatomy fails to reliably predict cortical function in a patient-specific manner due to significant neuroplasticity and inter-individual variability in glioma patients.[4-6] Functional MRI can interrogate individualized function but suffers from false negatives due to neurovascular decoupling and false positives from functionally correlated but non-essential regions.[7-9] At the subcortical level, tract infiltration, tumor mass effect, previous resections, unclear anatomo-functional correlations, and a lack of standardization in patient-specific tractography algorithms all contribute to distort the localization of clinically relevant subcortical anatomy. These limitations combine to obfuscate the potential neurological impact of planned resections.

Direct electrical stimulation (DES) has emerged as a robust technique to map patient-specific functional anatomy by correlating focal inhibitory stimulation with language task performance in real time.[5,10-12] While the invasive nature of the method precludes its use for patient selection and surgical planning, surgeons use it intraoperatively to optimize the "onco-functional" balance: maximizing the extent of resection while minimizing postoperative deficits. However, confounders (e.g., inconsistent speech arrest interpretations, variation in stimulation intensity, angle of the probe) may impact the consistency and reliability of the technique. Factors such as perioperative morbidity, intraoperative pain, seizure risk, and temporal constraints prohibit comprehensive mapping on every patient. Finally, it remains unclear whether intraoperatively identified sites of transient language disruption would lead to permanent impairment if resected.

Despite the widespread use of DES, subjective judgements often determine important surgical decisions to preserve language ability.[13] Understanding the neuroanatomical mechanisms that underpin long-term aphasic surgical deficits could provide data-driven surgical indications for patients with dominant hemisphere lesions. Preoperative localization of essential language function could transform patient selection, surgical planning, and intraoperative decisions. Transcranial magnetic stimulation (TMS) has recently emerged as a promising preoperative language mapping modality.[14,15] It relies on a similar mechanism to DES but can apply targeted inhibitory stimulation non-invasively in an outpatient setting.[16] We aim to use advanced preoperative brain mapping, resection volumetrics, and formal language evaluations to localize the anatomical origins of aphasic surgical deficits and predict language outcomes with non-invasive imaging.

Results

Table 1 shows patient demographics and patient characteristics. Around 90% of patients had lesions defined as eloquent or near-eloquent by according to Sawaya et al.[17] Around a third of patients (27%) presented with aphasia (WAB < 93.8) at baseline, about half (47%) experienced new or worsened aphasic deficits in the immediate postoperative period, and 10 patients (13%) sustained permanent language decline compared to their preoperative status. Of the 10 patients with long-term aphasia, three (30%) primarily had receptive deficits while seven (70%) primarily had expressive deficits. With regards to the primary domain of neurocognitive impairment, three patients exhibited deficits in verbal fluency, one patient specifically had a deficit in semantic fluency, two had deficits in lexical retrieval (anomia), and two had deficits in auditory comprehension. Baseline demographics and clinical characteristics were not associated with permanent postoperative aphasic decline. Supplementary Figure 1 shows the group-level distribution of tumor volumes in MNI space.

Predicting Aphasia with Cortical Data

Resection of TMS points did not correlate with long-term aphasic deficits. However, deficits significantly associated with resection of TMS point subgroups of subcortical connectivity at FA thresholds of 50% and above, with greatest odds ratio found at an FAT of 75% (OR=8.74 , p=<.01). Figures 2A and 2B demonstrate how different profiles of subcortical connectivity impact the predictive value of TMS points for functionality. Figure 2C shows the group-level cortical distribution of both TMS points and the most predictive subgroup of TMS points when stratified by subcortical connectivity. This analysis shows that subcortical connectivity stratifies TMS points into a group-level cortical distribution that closely aligns with canonical patterns derived from invasive stimulation: hotspots in the frontal operculum, ventral precentral gyrus, and posterior superior temporal gyrus.[18]

Characterization of TMS Tracts

We characterized the white matter tracts that were produced by TMS-based fiber tracking (at the most predictive FA threshold, 75%), showing that they recapitulate normative patterns of

subcortical connectivity at a group level. Figure 3A shows that these tracts, reconstructed in a completely data-driven manner, produce hotspots in MNI space at a group-level that anatomically co-localize with population-averaged white matter tracts traditionally associated with language. Figure 3B (left) demonstrates that the arcuate fasciculus is the most common pathway reconstructed by TMS tracts. Figure 3C illustrates the effect of progressive thresholding on the group-level TMS tracts and the HCP-defined probabilistic arcuate fasciculus, showing that the highest intensity voxels of both atlases localize to the posterior temporo-parietal segment of the traditional arcuate pathway. Figure 3D shows that the group-level TMS fiber tracts in MNI space reproduce the same voxel intensity pattern of the SLF as probabilistic group-level mapping of the SLF in healthy subjects from the HCP. Both have significantly lower intensity voxels anterior to the precentral gyrus.

Predicting Aphasia with Subcortical Data

Figure 4 shows the predictive value of TMS tracts. 4A shows an ROC curve for white matter tracts in the same format as the cortical ROC curve from Figure 1, demonstrating a much more robust Area Under the Curve (AUC=.90). 4B shows the results of multivariate analysis when using the best TMS point subgroup along with the best TMS tract model as predictors. Adding cortical data to the subcortical model demonstrated negligible improvement in predictive value, while adding subcortical data to the cortical model decreased its odds ratio from 10 to near 0 ($p<.001$). 4C shows a confusion matrix depicting the predictive value of TMS tracts reconstructed at the most predictive FA threshold. This model demonstrated a sensitivity of 90%, specificity of 87%, PPV of 50%, and an NPV of 98%. This analysis indicates that aphasic deficits anatomically localize at the subcortical level to TMS-derived fiber tracts, though the cortical sites themselves do not seem to contribute to functional decline.

Normalized Analyses

Given the significant number of false positives (TMS tract resection with no deficit), we normalized the resected portion of TMS tracts in each patient to MNI space in search of a generalizable method to classify true versus false positive TMS tract resections. Initial composite visualization of true positives (n=9) and false positives (n=9) in MNI space revealed that true positives cluster within the white matter of the anterior temporal lobe and posterior temporo-parietal-occipital (TPO) junction. In contrast, false positives distributed throughout the left hemisphere and did not show a discernable pattern. Figure 5A shows group-level visualizations of true and false positive tract-cavity segmentations in MNI space. Supplementary Figure 2 shows the spatial association between the smoothed true positive tract-cavity segmentations and normative tracts, specifically within areas of significant white matter convergence within the anterior temporal lobe and TPO junction.

Figure 5B shows that true positive segmentations preferentially co-localized with the normative volumes of HCP-defined, language-associated white matter tracts. True positive segmentations demonstrated statistically significant differences compared to false positives across all three measures of association with normative tracts: volume of overlap ($p<.05$), percentage of overlap ($p<.05$), and number of HCP tracts overlapped ($p<.05$). Figure 5C depicts

how each specific normative tract contributed to the classification between true and false positive tract-cavity segmentations. Additionally, every patient had at least one atlas-based tract segmentation resected (not shown), demonstrating the inadequacy of normative data and the importance of individualized mapping. Together, this analysis shows that non-compensable function localizes to individualized substructures within normative language-associated white matter tracts.

We hypothesized that measures of normative tract association (overlap volume, percentage, and number of tracts) could classify TMS tract resections and improve the predictive model. We computed the classification performance of every possible permutation of cutoffs for classifying TMS tract resections with these measures. We computed the average accuracy of each cutoff permutation, depicted in heatmaps in Figure 5D. We computed and highlighted the row or column of each heatmap that showed the highest average accuracy, indicating the best cutoff for each respective measure. These cutoffs, derived from group-level computations in MNI space, were then later applied to make predictions on an individual level. This process resulted in the selection of the following three parameters as the best predictors of persisting worsening of aphasia: over 1 cm^3 volume of infiltration, more than 3 tracts overlapped, and more than 50% of the tract-cavity segmentation overlapping.

Figure 6 further demonstrates the importance of individualized mapping as well as the significant inter-individual anatomical variability of essential language function. Given that co-localized segments of TMS tracts and atlas-based tract segmentations provide robust long-term deficit predictions when combined with measures of surgical damage, we used these regions in each patient as a marker for the anatomical substrate of non-compensable language function. In all patients (n=8) who exhibited a positive prediction in the final model, we created 8 different ROIs from these co-localized segments. We analyzed the group-level association of these ROIs with permanent deficits (if resected) with and without co-localized TMS tracts. Across the whole cohort, 92% of patients who underwent ROI resections that include TMS tracts sustained deficits while 34% of patients who underwent ROI resections irrespective of individualized data sustained deficits ($p<.05$). Figure 6B shows the respective anatomical locations and neurocognitive deficits associated with each ROI.

Final Predictive Model

To perform the analysis for the final iteration of the predictive model, we warped the normative volumes of the 7 language-associated tracts from MNI space to patient-space. The resection volume along with these two segmentations (and derived cutoffs) were used to generate outcome predictions in each patient. Figure 7A shows the confusion matrix for this final model. This resulted in 80% sensitivity, 97% specificity, 80% PPV, and 97% NPV. The odds ratio significantly increased with each iteration, culminating in an odds ratio of 134 ($p<.001$) (Figure 7B). Figure 7D illustrates the workflow for prospective application of this data to guide surgical planning and execution, specifically for a patient with a glioma in the TPO junction.

Figure 7D demonstrates precise prediction of the extent of resection by showing that TMS tracts compose the deep functional boundary of the resection cavity. The complete

preservation of TMS tracts (along with TMS-HCP co-localized segmentations) correlated with intact postoperative language function, demonstrating accurate outcome prediction for this patient. This flow chart also shows that the TMS-HCP co-localized tract segmentation predicted the real-time linguistic performance (error vs task completion) of this patient in response to direct electrophysiological stimulation of subcortical brain. As the resection progressed, the depth of the cavity did not elicit errors until the TMS-HCP co-localized segmentation was reached. Here, subcortical stimulation at 5 milliamps elicited strikingly reproducible semantic errors associated with an object naming paradigm. The positive stimulation point co-localized with the TMS-HCP segmentation at a distance of 0 millimeters. The negative stimulation points (n=6) were all more than 5 mm away from the TMS-HCP segmentation.

In contrast, Figure 7E shows that state-of-the-art automated fiber tracking (TractSeg) did not accurately predict intraoperative mapping or extent of resection/postoperative outcome. The left image shows how the positive stimulation point localizes distantly from the AF segmentation (in contrast to the TMS-HCP co-localized segmentation). The middle image shows that although the positive stimulation point did co-localize with the MdLF, two negative points also co-localized with this tract. The right image indicates that a significant portion of the MdLF was resected, indicating an inaccurate outcome prediction based on TractSeg for this patient.

Probabilistic Atlas of Functional White Matter

We then constructed a group-level visualization from these individual markers of essential function to produce an anatomo-functional probabilistic atlas of the white matter that underpins language function in glioma patients, shown in Figure 8A. Voxel distributions of each individual tract demonstrate that these four pathways have hotspots around the white matter of the TPO junction. Corresponding probabilistic volumes of normative data for the same tracts, shown in Supplementary Figure 3, do not show the same pattern. Figure 8B shows the average voxel intensity of these tract segmentations within the TPO ROI generated in Figure 5 versus outside of this ROI. This plot shows that every tract segmentation that penetrates the TPO (AF, ILF, IFOF, MdLF) has significantly higher voxel intensities within this region. This demonstrates that the novel pattern of subcortical connectivity in glioma patients anatomically correlates with vulnerability to aphasic damage.

Discussion

These results constitute the first, to our knowledge, robust prediction of surgical outcomes using preoperative imaging and resection volumetrics. Multiple previous studies have predicted language outcomes with preoperative imaging while others have documented associations between resections and postoperative language decline.[19-25] However, in our view, this study has unique clinical value. Here we report novel data that can directly inform operative decisions by predicting the long-term functional impact of individual tumor-removal strategies with strong accuracy (>90%). The method could also target and optimize common non-invasive treatments for glioma patients such as radiation therapy for tumor control or brain stimulation for language rehabilitation.

Additionally, this study suggests an anatomo-functional explanation for the origins of aphasic surgical deficits. Figure 4 reveals that non-compensable function lies at the subcortical

level while Figure 5 shows that aphasic deficits localize specifically to subregions of atlas-based segmentations of major association white matter tracts. Figure 6 shows that these substructures require individualized mapping for accurate localization, explaining the confusing anatomical heterogeneity of postoperative aphasic decline reported by neurosurgeons. Despite this patient-specificity, a group-level pattern emerged among the "true positive" tract resections. The resected portion of tracts in patients who sustained deficits clustered within the white matter of the TPO junction and anterior temporal lobe. This finding remarkably recapitulates results reported in both the stroke and resection lesion mapping literature that likewise identify these two distinct anterior/posterior subcortical regions as strongly correlated with long-term aphasia.[21,26-28] These findings have significant implications for anatomical stratification and prognostication for surgical glioma patients irrespective of advanced brain mapping techniques.

We found that subcortical connectivity plays a significant role in defining both the function and functional anatomy of the cortical language network. While TMS points alone showed no functional relevance, certain subgroups defined by subcortical connectivity demonstrated consistent functionality. Additionally, group-level visualization in MNI space revealed that TMS points distribute across the left hemisphere with rough clusters in the perisylvian region. In contrast, clear clusters in the frontal operculum, ventral precentral gyrus, and posterior superior temporal gyrus emerged when selecting only for the points with robust connecting tracts. Notably, this non-invasively captured pattern closely recapitulates the cortical distribution of DES-positive language sites reported in the largest study of intraoperative language mapping to date.[18]

Figure 4 not only shows that subcortical tract injury independently predicts deficits but also demonstrates that adding cortical resection data does not improve the predictive value of the subcortical model. Despite the disproportionate clinical interest in cortical language mapping compared to subcortical stimulation, our results fit well with previous studies. Duffau et al have repeatedly written about the importance of preserving subcortical connectivity to maintain capacity for long term function. Studies have shown that while cortical topology has significant potential for neuroplastic reorganization and functional compensation, subcortical circuitry seems to lack this dynamic flexibility.[29-32] Perhaps cortical mapping demonstrates efficacy because it acts as an indirect proxy to guide the resection away from the non-compensable, surgically relevant subcortical circuitry that connects these cortical sites to the global language network.

Prospective application of the TMS-HCP reconstructions precisely predicted intraoperative subcortical language mapping in a patient with recurrent glioblastoma. Stimulating the surgical cavity during the resection did not produce errors until the preoperatively identified region was reached. Here, at the intersection of TMS tracts and the normative HCP arcuate fasciculus, strikingly reproducible semantic errors were elicited. The precise anatomo-functional correlation in a patient with severely distorted anatomy illustrates the ability of this imaging approach to accommodate for the significant inter-individual variability found in a cohort of surgical glioma patients. TractSeg, a sophisticated deep learning approach to data-driven fiber tracking, demonstrated clear deficiencies in localizing intraoperatively identified subcortical sites of functionality, underscoring the necessity of integrating clinically relevant outcome measures into the brain mapping development process.

Many approaches to tractography seem to use arbitrary outcome measures by prioritizing structural visualization. For example, some employ algorithms to uncover tracts previously hidden by lesion infiltration.[33] However, this assumes that any uncovered fibers are functionally or clinically relevant. While tumor invasion does decrease local fractional anisotropy, this infiltration usually renders the tract non-functional as well. Figure 4 provides evidence for a strong correlation between language function and relative (patient-specific) fractional anisotropy with an ROC curve (AUC=.90), adding to our previous work that showed similar direct correlations between the FA threshold and motor function.[34-36] This correlation also likely reflects the direct connection between increased neural activity, increased myelination profiles, and increased fractional anisotropy demonstrated in important previous work.[37] We discuss this in more detail later.

The neurosurgical literature has extensively described the phenomenon of "brain shift"-intraoperative mechanical deformation that can distort image-guided navigation.[38-40] Some assert that this confounds the position of preoperative localizations relative to the postoperative resection volume, obfuscating which structures were removed. However, the conclusions of our study rely specifically on the consistent perioperative localization of white matter tracts, which reside in subcortical regions much less susceptible to deformation than the cortical surface.[41-43] Additionally, multiple recent studies have shown that elastic, non-linear co-registration algorithms efficaciously correct for intraoperative deformation.[43-45] These data, combined with the robust predictive models reported in this study, indicate that computational techniques can sufficiently address any perioperative anatomical distortions in glioma patients that may confound long-term outcome predictions based on preoperative imaging and the resection volume.

This study also has significant neuroscientific implications. While functional reorganization in glioma patients is well established at the cortical level, corollary white matter activity remains unclear. The prevailing consensus asserts that subcortical white matter lacks the capacity for "biological plasticity", or macrostructural rearrangement on a time scale from days to months.[29] Data from this study supports this notion. Figure 5 shows that resected tract segments from patients who sustained aphasic deficits uniformly co-localize in MNI space with normative structures of major association white matter tracts, adding to accumulating evidence that white matter does not engage in biological plasticity by forming entirely novel pathways to sustain language function.

However, it remains unclear how dynamic superficial cortical hubs interface with static deep subcortical circuitry to orchestrate language function in a lesioned brain. Figure 3 shows that data-driven fiber tracking based on peritumoral cortical TMS language sites reconstructs local subcortical fascicles from all 7 of the language-associated white matter tracts at a group-level. Figure 4 demonstrates that these non-invasively acquired cortical sites localize functional peritumoral fibers in a comprehensive manner by showing that tract preservation accurately predicts a safe outcome in over 98% of cases. These findings apply to regions across the left hemisphere, indicating that functional fiber segments posing surgical risk have significant local cortical connectivity, irrespective of anatomical location, that render them localizable purely from peritumoral cortical data.

Figure 2 also illustrates the anatomical heterogeneity of white matter cortical connectivity by showing a broad distribution of subcortically connected TMS points. While traditional conceptualizations regarding the cortical connections of association white matter tracts focus on the group-level origin and termination of major pathways, this study provides evidence that functional tracts have individualized cortical connections distributed throughout their structure. Figures 8 also shows that functional white matter clusters homogenously into the main bodies of normative tracts at a group-level yet displays individualized patterns closer to the cortex. These results provide evidence that confirms the hypothesis of Duffau et al who speculate that white matter structure becomes more individualized towards the cortical surface and beyond the core "skeleton" of main pathways.[29] These widespread structural interfaces may signify the missing link between a dynamic cortical language network and static subcortical pathways in glioma patients.

While white matter tracts do not seem to undergo macrostructural changes, subcortical fascicles may contribute to adaptive plasticity in a more subtle manner. Previous work has shown that increased neural activity drives myelination.[46] DWI data can detect these microstructural changes through fractional anisotropic measurements after a few weeks of altered activity.[37] We demonstrated that patient-specific measures of cortical connectivity (TMS mapping data) and microstructural integrity (FA threshold selection approach) can differentiate functional and non-functional white matter through data-driven fiber tracking. We can reasonably speculate that regional asymmetry in one or both measures are exclusive to glioma patients and driven by increased activity from glioma-induced plasticity. If so, this finding sheds light on the nature of subcortical plasticity in glioma patients by suggesting that white matter adaptations that correlate with cortical reorganization manifest through changes in the microstructural properties or cortical connectivity of pre-existing pathways.

Perhaps, to sustain language processing and support cortical reorganization, the brain adapts the route of information flow at the subcortical level to the specific constraints imposed by the tumor. We speculate that the fundamental architecture of connectivity does not rearrange in response to an invasive lesion but instead facilitates atypical patterns of information flow within certain segments that display local signatures of functionality in the form of individualized cortical connections and/or robust myelination profiles. These biomarkers of recent activity may indicate which particular pathways the language network depends on in each patient. An extensive literature on language acquisition supports this notion by routinely identifying correlated microstructural white matter changes yet likewise detecting no evidence for macrostructural rearrangement.[47-49] This idea provides a possible mechanistic explanation for the efficacy of TMS guided fiber tracking (standardized to the patient-specific fractional anisotropic profile) in localizing functional white matter.

Previous work can draw limited conclusions regarding the structure of functional subcortical networks. While many established neuroimaging techniques can interrogate global cortical function, insights into subcortical network underpinnings derive from either functionally ambiguous DTI tractography or anatomically confined direct stimulation limited by the surgical window.[50] This limits the ability of researchers to map the global structural pathway of functional white matter tracts, obfuscating the 3D spatial organization of functional subcortical fascicles.[51] We address these technical neuroscientific constraints by applying a double layer of functional integration to white matter tract reconstruction from diffusion imaging.

First, cortical functional sites from focal brain stimulation were used as regions of interest (ROIs) for tractography. Next, we functionally optimized the resulting reconstructions by finding the postprocessing pipeline that maximizes the association between network damage and postoperative aphasic decline across a large cohort. This pipeline includes the individualized FA threshold selection approach along with data-driven streamline filtering using normative HCP data. This process culminated in the development of a novel, data-driven brain mapping approach capable of non-invasively mapping the functionally essential components of subcortical networks. In addition to being a robust clinical tool for informing surgical decisions, the reproducibility, accessibility, and scalability of this technique make it a promising approach for researchers to investigate cognitive functional structure in the context of gliomas. The method significantly improves upon existing white matter reconstruction techniques by providing global representations of functional white matter without the hazards of invasive procedures.

We apply this technique to our cohort to create a novel anatomo-functional probabilistic atlas, providing unprecedented insight into the group-level 3D organization of functional white matter tracts in glioma patients. This atlas suggests that glioma patients have a distinct group-level pattern of meaningful subcortical connectivity by showing that segments of multiple tracts have significantly higher voxel intensities within the white matter of the TPO junction, a phenomenon not observed in population-averaged fiber tracking data from healthy subjects (Supplementary Figure 3). Data from both Figure 5 and Figure 8 indicate, with independent measures, that the white matter of the TPO junction serves as a hub for critical language-processing circuitry in glioma patients. Although disproportionate regional cortical seeding could represent a significant confounder that potentially accounts for this pattern, tumors (and as a result, cortical seeding patterns- Figure 1) in this cohort most frequently localized to the anterior temporal lobe as shown in Supplementary Figure 1.

While many studies have generated computational models to predict the outcome of epilepsy surgery and inform clinical decisions, this work represents a first-of-its-kind for glioma surgery.[52] We show that preoperative TMS mapping, combined with diffusion imaging and normative data, can accurately predict the functional outcome of volumetric surgical strategies. The predictive model can precisely measure one of the most important and elusive[53,54] characteristics for dominant hemisphere gliomas: proximity to regions of non-compensable language function. The novel imaging method could significantly improve patient prognostication and stratification in clinical trials, contributing to the scaled development of new personalized treatment strategies for glioma patients. Irrespective of TMS, this study addresses significant weaknesses in neurosurgical tractography by using functional outcomes to develop and optimize data-driven postprocessing protocols.

We chose a broad outcome classification scheme (“aphasia” vs “no aphasia”) to capture surgical relevance instead of more clinically ambiguous impairments of higher-order cognitive processes. While we intentionally designed our approach to selectively capture gross language deficits that universally impact quality-of-life and survival[3,55], predicting specific higher-order cognitive impairments would have significant clinical utility for more personalized surgical care of glioma patients. However, cognitive testing categorizes impairments into domains that may

not precisely align with the brain's underlying functional architecture. For example, errors classified as "semantic fluency", while specific in terms of the phenotype, may reflect deficits in both phonological and semantic processing. This concern, as well as a historical precedent in neurosurgical studies to use clinical outcome measures that selectively capture basic language function[3,10,17], informed our approach.

This study has many limitations that have been discussed throughout the manuscript. Because limited accessibility of TMS in neurosurgical settings stalls widespread translation of the current models outlined here, future work should focus on producing the same quality of predictive data with task-independent modalities. In addition to limiting accessibility, the dependence of these models on TMS mapping data represents one of the most significant barriers to standardized predictions capable of generalizing to other cohorts and institutions. To facilitate rapid clinical application, we chose to use a widely available clinical neuronavigation software for fiber tracking (Brainlab Elements). Future iterations of this work will integrate more flexible fiber tracking software to optimize and standardize methodologies. Future work should also utilize cohorts large enough to be divided into training and validation subsets to integrate measures of generalizability into the development process. Additionally, the retrospective nature of this study makes it vulnerable to possible selection bias, although the comprehensive standard-of-care employed at our institution lends itself well to rigorous retrospective analysis.

Conclusion

We found that essential brain regions of language processing localize at the subcortical level to patient-specific ROIs within normative association white matter pathways. TMS-derived structural networks can map these individualized segments and accurately predict the impact of personalized surgical strategies on long-term language function. This study addresses key technical deficits in neurosurgery and neuroscience by defining a reproducible, non-invasive mapping approach that accurately identifies brain regions essential to language function. We use this technique to create a novel probabilistic atlas of functional white matter in glioma patients that provides evidence for a disease-specific pattern of subcortical connectivity.

Methods

This retrospective study analyzed the prospectively collected data of 79 consecutive adult presurgical glioma patients. The study included all patients with gliomas in the dominant hemisphere who underwent TMS language mapping as part of our institution's standard preoperative protocol for eloquent tumors. Exclusion criteria were limited to contraindications to TMS mapping, such as the presence of a pacemaker or cochlear implant. As part of the standard-of-care for eloquent tumors, patients underwent an awake craniotomy with intraoperative language

mapping as well as formal pre and postoperative language evaluations by a trained neuropsychologist or speech pathologist. The patients in this study underwent TMS mapping and surgery between July 1st 2017 and August 1st 2023.

Aphasia Definition

Patients were administered formal language evaluations by a trained neuropsychologist or speech pathologist preoperatively, immediately postoperatively, and 1-3 months postoperatively. A retrospective binary classification approach (new or worsened aphasia vs preserved language function) was designed to create an outcome measure that captures gross language deficits induced by the surgery. Patients with preserved language function upon immediate postoperative testing (determined by the trained evaluator in real-time compared to the preoperative evaluation) were classified as "preserved language function" without further review of records. Patients with immediate postoperative decline (47%) were followed to the 1-3 month formal evaluation. Patients with postoperative decline at the 1-3 month evaluation (13%) were followed clinically through the 6 month mark to qualitatively verify persistence of the deficit compared to the 1-3 month evaluation. The specific domain of impairment was derived from the 1-3 month formal evaluation compared to preoperative baseline.

The formal language evaluations consisted of either the Western Aphasia Battery (WAB) or the following battery of neuropsychological testing: animal fluency (semantic fluency), letter fluency (phonemic fluency), and the Boston Naming Test. For patients with an immediate postoperative deficit who also had a 1-3 month WAB, the preoperative aphasia quotient (AQ) was compared to the 1-3 month postoperative AQ to define the primary outcome. Continuous AQ scores were classified in terms of aphasic severity based on published cut-points: 0-24, 26-50, 51-75, and 76-93.8, and 93.8 and above.[56] A deficit was recorded if the postoperative AQ was decreased to the next category. For patients with an immediate postoperative deficit who did not have a 1-3 month WAB, the neuropsychological battery was used to classify the outcome. Given the more sensitive nature of the specific neuropsychological batteries compared to the WAB, an aphasic deficit was defined as a decline of at least 1.5 standard deviations on the 1-3 month evaluation (compared to preop) in at least 2/3 tests.

Transcranial Magnetic Stimulation Protocol

All patients underwent mapping using an object naming paradigm. To define the stimulation targets for each tumor, an evenly spaced grid composed of targets 1 cm apart was overlayed onto the lesion and perilesional region. A subset of these targets were labeled language positive (i.e., corresponding with speech disruption) as determined by the patient-specific mapping. Figure 1A-C illustrates the TMS mapping process. The protocol specifics are detailed below. Correlations with intraoperative direct cortical stimulation are detailed in previous work.[14]

Navigated TMS language mapping was performed (NBS System 3.2; Nexstim, Helsinki, Finland). TMS stimulation onset was simultaneous with presentation of line drawing objects (picture to trigger interval was 0.0 s). Time-locked rapid rate TMS trains were given for 5 Hz/5 pulses for 140 stimulations. The display time was 1,000 ms, and interpicture interval was 3,000 ms. We performed TMS language mapping according to a conventionally accepted protocol for

TMS language mapping.[57,58] The most likely location of the hand knob was identified anatomically. This area was then stimulated in a random pattern while systematically varying the rotation, tilt, and yaw of the magnetic field. The location of maximal motor evoked potential was identified. Resting motor threshold (RMT) was identified using this position.[59] Once determined, stimulator output typically began at 110% RMT for the repetitive TMS language mapping procedure. We used a common anterior/posterior coil orientation where the coil is oriented horizontally between the nasion and external auditory meatus. Baseline object naming without stimulation was performed, which consisted of presentation of line drawings of common objects and required the patient to verbally respond with the name of the pictured object. Pictures that elicited misnaming or hesitation were discarded. Baselining was repeated until a reliable stimulus set was obtained. Target dots 1 cm apart from each other were created on the surface of the three-dimensional (3D) brain reconstruction, based on the preoperative MRI, in a grid-like pattern. One centimeter was selected because a 1 cm radius of tissue is typically attempted to be preserved around intraoperatively identified language-positive sites. The grid was created in the tumoral and peritumoral areas. The pictures acquired from baselining were presented time locked to the TMS pulses while the coil was moved over the cerebral hemisphere. During this process, the patient was audiovisually recorded. A site was considered a positive "hit" for language eloquence if disruption (e.g., no response, performance error, and semantic error) was observed on any of the trials. A hesitation was counted when the patient's response came after the TMS stimulation had ended for that picture. Positive sites were marked on the 3D brain surface as white dots. These sites were exported in digital imaging and communications in medicine (DICOM) files to be uploaded in the Neuronavigation system (Elements; BrainLab, Munich, Germany).

Fiber Tracking

Structural MRIs were performed using a 3T MRI scanner (GE Healthcare, Waukesha, Wisconsin) with an eight-channel head coil. DWI data was acquired using a diffusion-weighted spin-echo echo-planar imaging sequence (repetition time/echo time = 10 000/62 ms, matrix size = 128 × 128, field of view = 22 × 22 cm, slice thickness = 2.5 mm with no intersection gap, number of diffusion-weighting directions = 32, *b* value = 1000 s/mm^2). In total, 44 slices were acquired, covering the medulla to the top of the brain. High-resolution 3D spoiled gradient-echo T1-weighted sequences were acquired for anatomic reference. Diffusion-weighted spin-echo echo-planar imaging sequences were uploaded to Brainlab Elements (Munich, Germany), and tractography was generated using a deterministic FACT and TEND method provided by the software.[60]

Positive TMS points were defined as one ROI for fiber tracking in each patient. The minimum fiber lengths of 30 mm and a maximum angulation of 50 degrees were used. Fiber tracts generated using TMS points as the seed are referred to as "TMS tracts" in this study. Given the significant inter-individual variability in fractional anisotropy (FA), a relative approach for FA thresholding was used to standardize the selection to the individual profile of the patient.[61] For each patient, the FA threshold (FAT) was progressively increased by .01 after starting at 0. 100% was defined on an individual basis as the absolute value at which tracts are no longer generated. We then analyzed four different percentages of FATs across the whole cohort: 25%, 50%, 75%,

85%. **Figure 1B** illustrates TMS-based fiber tracking. Atlas-based segmentations from the Human Connectome Project (HCP842)[62] were used to obtain population-averaged volumes in MNI space. Both the binary masks of population-averaged tracts were obtained as well as the probabilistic templates that reflect the inter-individual variability of the cohort. TractSeg[63] was used to produce data-driven fiber tract segmentations for the case example in Figure 7.

Resection Analyses

The resection cavity was segmented in a data-driven manner.[64] Overlap between preoperative reconstructions (TMS points, fiber tracts) and the resection cavity was quantitatively analyzed. For Figures 2 and 4, any overlap between the TMS points/tracts and the resection cavity was included as a positive predictor in binary logistic regression models that predict pre to postoperative long-term aphasic decline. For Figures 6 and 7, positive or negative predictions were based on the cutoff parameters derived in Figure 5.

Image Processing

Preprocessing was performed using Brainlab Elements (Brainlab AG; Munich, Germany), Matlab (The Mathworks Inc., Massachusetts, USA), SPM (https://www.fil.ion.ucl.ac.uk/spm/), and Python. Brainlab Elements' "Image Fusion" and "Distortion Correction" modules were used to non-linearly co-register the preoperative T1 MRI used for TMS mapping and intraoperative navigation with the postoperative T1 MRI (1 day postop). This elastic co-registration approach has been validated in multiple previous studies for correcting imaging distortions caused by surgically-induced brain shift.[43,65-68] Gaussian smoothing was performed with a full-width at half maximum of 5, 7.5, and 10, specified in figure legends.

Cortical data (TMS points and associated connectivity-based subgroups of TMS points) and subcortical data (fiber tracts-"TMS tracts") were segmented in Brainlab and exported. SPM was used to spatially normalize patient-specific data to MNI space. Matlab was used for voxel-wise summations to produce composite probabilistic atlases of TMS tracts and cortical points after normalization to MNI space. For Figure 8, different portions of the composite TMS tract probabilistic atlas were segmented according to their overlap with normative tract volumes. The ROI from Figure 5 was overlayed onto these probabilistic volumes and a voxel-wise analysis computed the average voxel intensity of tract segmentations within this ROI versus outside of this ROI.

In patients who had TMS tracts resected, Brainlab Elements was used to automatically segment the overlap between the tracts and resection cavity to produce tract-cavity segmentations. These segmentations were then normalized to MNI space to identify the optimal approach for classifying "true positive" (resection of tract with corresponding deficit) and "false positive" (resection of tract without corresponding deficit) tract-cavity segmentations. A grid-search with k-fold cross validation was used to define the optimal set of cutoffs for classification across three different measures of spatial association between tract-cavity segmentations and normative tract volumes.

The following cut-offs were analyzed: Volume of overlap (cm^3): [.5, 1, 1.5, 2, 2.5] Number of normative tracts overlapped: [2, 3, and 4] and Percentage of overlap (% of tract-cavity segmentation overlapped by normative tract volumes): [40, 50, 60, 70, and 80]. The advanced normalization technique (ANT) software package[69] was used to warp normative data to patient

space to apply these cutoffs to individualized predictions for the final iteration of the predictive model. The normalization transformations to MNI space were first computed. The inverse transformation was then used to warp HCP tracts from MNI space to patient space. The cutoffs for the three measures derived in MNI space were used to generate predictions within the final iteration of the predictive model reported in Figure 7.

In Figure 6, the regions of co-localized TMS tracts and normative HCP tracts ("TMS-HCP") that corresponded with a postoperative deficit (n=8) were analyzed to understand their group-level features. Each region was segmented in MNI space to produce 8 ROIs. Each patient-specific ROI was analyzed for its degree of overlap with the normalized resection cavities from the whole cohort. If the spatial association between the patient-specific resection cavity and ROI exceeded the cutoffs defined in Figure 5, this was counted as a resection of that ROI for that patient. The same analysis was then performed along with the patient-specific TMS tract volumes. A resection was recorded if the cavity-ROI overlap included the TMS tract volume for that patient. The percentage of patients who had positive ROI resections that included TMS tracts versus patients who had positive ROI resections irrespective of TMS tracts were analyzed according to the percent who sustained long term deficits in each group.

Statistical Analysis

Statistical analyses were performed using Python (Version 3.12.1). Scipy.stats, statsmodels, and pandas were used for basic computations such as t-tests and regression. Independent, one-tailed t-tests were utilized. Wald tests were used to evaluate the statistical significance of regression models. Likelihood ratio test (LRT) was used to evaluate the statistical significance of multivariate nested models compared to corresponding univariate models. Seaborn and matplotlib were used for more complex visualizations such as scatter plots, box plots, and heatmaps. Scikit-learn was used to generate ROC curves, confusion matrices, and k-fold cross validations. Confusion matrix elements were used to analyze the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the predictive models based on mapping and the resection volume. A true positive was defined as one or more resected sites (TMS points and/or white matter tracts (WMTs)) with a permanent deficit. A false positive was defined as one or more resected sites that did not have a corresponding permanent deficit. A true negative resulted from no sites being resected and no permanent deficit. A false negative was defined as having no sites resected but incurring a deficit.

Author Contributions

MM: Study concept and design; acquisition, analysis, and interpretation of data; drafting and revising of the manuscript. KN: analysis and interpretation of data; revising manuscript. SP: data acquisition, administration, and revising manuscript. HM: acquisition of data. PA: revising manuscript JT: revising manuscript. VK: revising manuscript. CE: revising manuscript. HL: revising manuscript. SS: Study concept and design, acquisition of data, revising manuscript.

Data Availability

The institutional-specific raw data from this study is available upon reasonable request. The publicly available normative tract volumes can be accessed at: https://brain.labsolver.org/hcp_trk_atlas.html

References

1. Ius T, Isola M, Budai R, et al. Low-grade glioma surgery in eloquent areas: volumetric analysis of extent of resection and its impact on overall survival. A single-institution experience in 190 patients: clinical article. *J Neurosurg.* 2012; 117(6):1039-1052.

2. Marko NF, Weil RJ, Schroeder JL, Lang FF, Suki D, Sawaya RE. Extent of resection of glioblastoma revisited: personalized survival modeling facilitates more accurate survival prediction and supports a maximum-safe-resection approach to surgery. *J Clin Oncol.* 2014; 32(8):774-782.

3. Rahman M, Abbatematteo J, De Leo EK, et al. The effects of new or worsened postoperative neurological deficits on survival of patients with glioblastoma. *J Neurosurg.* 2017; 127(1):123-131.

4. Herholz K, Thiel A, Wienhard K, et al. Individual functional anatomy of verb generation. *Neuroimage.* 1996; 3(3 Pt 1):185-194.

5. Ojemann G, Ojemann J, Lettich E, Berger M. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J Neurosurg.* 1989; 71(3):316-326.

6. Ojemann GA, Whitaker HA. Language localization and variability. *Brain Lang.* 1978; 6(2):239-260.

7. Binder JR. Functional MRI is a valid noninvasive alternative to Wada testing. *Epilepsy Behav.* 2011; 20(2):214-222.

8. Chang EF, Raygor KP, Berger MS. Contemporary model of language organization: an overview for neurosurgeons. *J Neurosurg.* 2015; 122(2):250-261.

9. Pillai JJ, Mikulis DJ. Cerebrovascular reactivity mapping: an evolving standard for clinical functional imaging. *AJNR Am J Neuroradiol.* 2015; 36(1):7-13.

10. Sanai N, Mirzadeh Z, Berger MS. Functional outcome after language mapping for glioma resection. *N Engl J Med.* 2008; 358(1):18-27.

11. Duffau H, Capelle L, Sichez J, et al. Intra-operative direct electrical stimulations of the central nervous system: the Salpêtrière experience with 60 patients. *Acta Neurochir (Wien).* 1999; 141(11):1157-1167.

12. Haglund MM, Berger MS, Shamseldin M, Lettich E, Ojemann GA. Cortical localization of temporal lobe language sites in patients with gliomas. *Neurosurgery.* 1994; 34(4):567-576; discussion 576.

13. Lakhani DA, Sabsevitz DS, Chaichana KL, Quiñones-Hinojosa A, Middlebrooks EH. Current State of Functional MRI in the Presurgical Planning of Brain Tumors. *Radiol Imaging Cancer.* 2023; 5(6):e230078.

14. Muir M, Patel R, Traylor J, et al. Validation of Non-invasive Language Mapping Modalities for Eloquent Tumor Resection: A Pilot Study. *Front Neurosci.* 2022; 16:833073.

15. Ille S, Sollmann N, Hauck T, et al. Combined noninvasive language mapping by navigated transcranial magnetic stimulation and functional MRI and its comparison with direct cortical stimulation. *J Neurosurg.* 2015; 123(1):212-225.

16. Ruohonen J, Karhu J. Navigated transcranial magnetic stimulation. *Neurophysiol Clin.* 2010; 40(1):7-17.

645 **17.** Sawaya R, Hammoud M, Schoppa D, et al. Neurosurgical outcomes in a modern series of 400
646 craniotomies for treatment of parenchymal tumors. *Neurosurgery.* 1998; 42(5):1044-1055;
647 discussion 1055-1046.
648 **18.** Lu J, Zhao Z, Zhang J, et al. Functional maps of direct electrical stimulation-induced speech arrest
649 and anomia: a multicentre retrospective study. *Brain.* 2021; 144(8):2541-2553.
650 **19.** Binding LP, Dasgupta D, Taylor PN, et al. Contribution of White Matter Fiber Bundle Damage to
651 Language Change After Surgery for Temporal Lobe Epilepsy. *Neurology.* 2023; 100(15):e1621-
652 e1633.
653 **20.** Zigiotto L, Vavassori L, Annicchiarico L, et al. Segregated circuits for phonemic and semantic
654 fluency: A novel patient-tailored disconnection study. *Neuroimage Clin.* 2022; 36:103149.
655 **21.** Tuncer MS, Salvati LF, Grittner U, et al. Towards a tractography-based risk stratification model for
656 language area associated gliomas. *Neuroimage Clin.* 2021; 29:102541.
657 **22.** Ille S, Zhang H, Sogerer L, et al. Preoperative function-specific connectome analysis predicts
658 surgery-related aphasia after glioma resection. *Hum Brain Mapp.* 2022; 43(18):5408-5420.
659 **23.** Cui M, Guo Q, Chi Y, et al. Predictive model of language deficit after removing glioma involving
660 language areas under general anesthesia. *Front Oncol.* 2022; 12:1090170.
661 **24.** Caverzasi E, Hervey-Jumper SL, Jordan KM, et al. Identifying preoperative language tracts and
662 predicting postoperative functional recovery using HARDI q-ball fiber tractography in patients
663 with gliomas. *J Neurosurg.* 2016; 125(1):33-45.
664 **25.** Emblem KE, Due-Tonnessen P, Hald JK, et al. Machine learning in preoperative glioma MRI:
665 survival associations by perfusion-based support vector machine outperforms traditional MRI. *J*
666 *Magn Reson Imaging.* 2014; 40(1):47-54.
667 **26.** Corbetta M, Ramsey L, Callejas A, et al. Common behavioral clusters and subcortical anatomy in
668 stroke. *Neuron.* 2015; 85(5):927-941.
669 **27.** Griffis JC, Nenert R, Allendorfer JB, Szaflarski JP. Damage to white matter bottlenecks contributes
670 to language impairments after left hemispheric stroke. *Neuroimage Clin.* 2017; 14:552-565.
671 **28.** Mirman D, Zhang Y, Wang Z, Coslett HB, Schwartz MF. The ins and outs of meaning: Behavioral
672 and neuroanatomical dissociation of semantically-driven word retrieval and multimodal
673 semantic recognition in aphasia. *Neuropsychologia.* 2015; 76:208-219.
674 **29.** Duffau H. A two-level model of interindividual anatomo-functional variability of the brain and its
675 implications for neurosurgery. *Cortex.* 2017; 86:303-313.
676 **30.** Sarubbo S, De Benedictis A, Merler S, et al. Towards a functional atlas of human white matter.
677 *Hum Brain Mapp.* 2015; 36(8):3117-3136.
678 **31.** Sarubbo S, Tate M, De Benedictis A, et al. Mapping critical cortical hubs and white matter
679 pathways by direct electrical stimulation: an original functional atlas of the human brain.
680 *Neuroimage.* 2020; 205:116237.
681 **32.** Herbet G, Maheu M, Costi E, Lafargue G, Duffau H. Mapping neuroplastic potential in brain-
682 damaged patients. *Brain.* 2016; 139(Pt 3):829-844.
683 **33.** Henderson F, Jr., Parker D, Vijayakumari AA, et al. Enhanced Fiber Tractography Using Edema
684 Correction: Application and Evaluation in High-Grade Gliomas. *Neurosurgery.* 2021; 89(2):246-
685 256.
686 **34.** Muir M, Prinsloo S, Michener H, et al. TMS Seeded Diffusion Tensor Imaging Tractography
687 Predicts Permanent Neurological Deficits. *Cancers (Basel).* 2022; 14(2).
688 **35.** Muir M, Prinsloo S, Michener H, et al. Transcranial magnetic stimulation (TMS) seeded
689 tractography provides superior prediction of eloquence compared to anatomic seeded
690 tractography. *Neurooncol Adv.* 2022; 4(1):vdac126.

691 **36.** Muir M, Gadot R, Prinsloo S, et al. Comparative study of preoperative functional imaging
692 combined with tractography for prediction of iatrogenic motor deficits. *J Neurosurg.* 2023;
693 139(1):65-72.
694 **37.** Scholz J, Klein MC, Behrens TE, Johansen-Berg H. Training induces changes in white-matter
695 architecture. *Nat Neurosci.* 2009; 12(11):1370-1371.
696 **38.** Roberts DW, Hartov A, Kennedy FE, Miga MI, Paulsen KD. Intraoperative brain shift and
697 deformation: a quantitative analysis of cortical displacement in 28 cases. *Neurosurgery.* 1998;
698 43(4):749-758; discussion 758-760.
699 **39.** Nimsky C, Ganslandt O, Cerny S, Hastreiter P, Greiner G, Fahlbusch R. Quantification of,
700 visualization of, and compensation for brain shift using intraoperative magnetic resonance
701 imaging. *Neurosurgery.* 2000; 47(5):1070-1079; discussion 1079-1080.
702 **40.** Hartkens T, Hill DL, Castellano-Smith AD, et al. Measurement and analysis of brain deformation
703 during neurosurgery. *IEEE Trans Med Imaging.* 2003; 22(1):82-92.
704 **41.** Elias WJ, Fu KM, Frysinger RC. Cortical and subcortical brain shift during stereotactic procedures.
705 *J Neurosurg.* 2007; 107(5):983-988.
706 **42.** Petersen EA, Holl EM, Martinez-Torres I, et al. Minimizing brain shift in stereotactic functional
707 neurosurgery. *Neurosurgery.* 2010; 67(3 Suppl Operative):ons213-221; discussion ons221.
708 **43.** Ille S, Schwendner M, Zhang W, Schroeder A, Meyer B, Krieg SM. Tractography for Subcortical
709 Resection of Gliomas Is Highly Accurate for Motor and Language Function: ioMRI-Based Elastic
710 Fusion Disproves the Severity of Brain Shift. *Cancers (Basel).* 2021; 13(8).
711 **44.** Zhang W, Ille S, Schwendner M, Wiestler B, Meyer B, Krieg SM. Tracking motor and language
712 eloquent white matter pathways with intraoperative fiber tracking versus preoperative
713 tractography adjusted by intraoperative MRI-based elastic fusion. *J Neurosurg.* 2022:1-10.
714 **45.** Ille S, Schroeder A, Wagner A, et al. Intraoperative MRI-based elastic fusion for anatomically
715 accurate tractography of the corticospinal tract: correlation with intraoperative neuromonitoring
716 and clinical status. *Neurosurg Focus.* 2021; 50(1):E9.
717 **46.** Mount CW, Monje M. Wrapped to Adapt: Experience-Dependent Myelination. *Neuron.* 2017;
718 95(4):743-756.
719 **47.** Wei X, Gunter TC, Adamson H, et al. White matter plasticity during second language learning
720 within and across hemispheres. *Proc Natl Acad Sci U S A.* 2024; 121(2):e2306286121.
721 **48.** Lee B, Park JY, Jung WH, et al. White matter neuroplastic changes in long-term trained players of
722 the game of "Baduk" (GO): a voxel-based diffusion-tensor imaging study. *Neuroimage.* 2010;
723 52(1):9-19.
724 **49.** Schlegel AA, Rudelson JJ, Tse PU. White matter structure changes as adults learn a second
725 language. *J Cogn Neurosci.* 2012; 24(8):1664-1670.
726 **50.** Duffau H. Stimulation mapping of white matter tracts to study brain functional connectivity. *Nat*
727 *Rev Neurol.* 2015; 11(5):255-265.
728 **51.** Duffau H, Thiebaut de Schotten M, Mandonnet E. White matter functional connectivity as an
729 additional landmark for dominant temporal lobectomy. *J Neurol Neurosurg Psychiatry.* 2008;
730 79(5):492-495.
731 **52.** Matarrese MAG, Loppini A, Fabbri L, et al. Spike propagation mapping reveals effective
732 connectivity and predicts surgical outcome in epilepsy. *Brain.* 2023; 146(9):3898-3912.
733 **53.** Ius T, Angelini E, Thiebaut de Schotten M, Mandonnet E, Duffau H. Evidence for potentials and
734 limitations of brain plasticity using an atlas of functional resectability of WHO grade II gliomas:
735 towards a "minimal common brain". *Neuroimage.* 2011; 56(3):992-1000.
736 **54.** Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. Probabilistic map of critical functional
737 regions of the human cerebral cortex: Broca's area revisited. *Brain.* 2014; 137(Pt 10):2773-2782.

55. Chaichana KL, Halthore AN, Parker SL, et al. Factors involved in maintaining prolonged functional independence following supratentorial glioblastoma resection. Clinical article. *J Neurosurg.* 2011; 114(3):604-612.

56. Kertesz A. Western Aphasia Battery--Revised. 2007.

57. Lioumis P, Zhdanov A, Mäkelä N, et al. A novel approach for documenting naming errors induced by navigated transcranial magnetic stimulation. *J Neurosci Methods.* 2012; 204(2):349-354.

58. Krieg SM, Lioumis P, Mäkelä JP, et al. Protocol for motor and language mapping by navigated TMS in patients and healthy volunteers; workshop report. *Acta Neurochir (Wien).* 2017; 159(7):1187-1195.

59. Krieg SM, Shiban E, Buchmann N, et al. Utility of presurgical navigated transcranial magnetic brain stimulation for the resection of tumors in eloquent motor areas. *J Neurosurg.* 2012; 116(5):994-1001.

60. Mori S, Crain BJ, Chacko VP, van Zijl PC. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol.* 1999; 45(2):265-269.

61. Frey D, Strack V, Wiener E, Jussen D, Vajkoczy P, Picht T. A new approach for corticospinal tract reconstruction based on navigated transcranial stimulation and standardized fractional anisotropy values. *Neuroimage.* 2012; 62(3):1600-1609.

62. Yeh FC, Panesar S, Fernandes D, et al. Population-averaged atlas of the macroscale human structural connectome and its network topology. *Neuroimage.* 2018; 178:57-68.

63. Wasserthal J, Neher P, Maier-Hein KH. TractSeg - Fast and accurate white matter tract segmentation. *Neuroimage.* 2018; 183:239-253.

64. Henker C, Kriesen T, Glass Ä, Schneider B, Piek J. Volumetric quantification of glioblastoma: experiences with different measurement techniques and impact on survival. *J Neurooncol.* 2017; 135(2):391-402.

65. Retif P, Djibo Sidikou A, Mathis C, et al. Evaluation of the ability of the Brainlab Elements Cranial Distortion Correction algorithm to correct clinically relevant MRI distortions for cranial SRT. *Strahlenther Onkol.* 2022; 198(10):907-918.

66. Calvo-Ortega JF, Mateos J, Alberich Á, et al. Evaluation of a novel software application for magnetic resonance distortion correction in cranial stereotactic radiosurgery. *Med Dosim.* 2019; 44(2):136-143.

67. Negwer C, Hiepe P, Meyer B, Krieg SM. Elastic Fusion Enables Fusion of Intraoperative Magnetic Resonance Imaging Data with Preoperative Neuronavigation Data. *World Neurosurg.* 2020; 142:e223-e228.

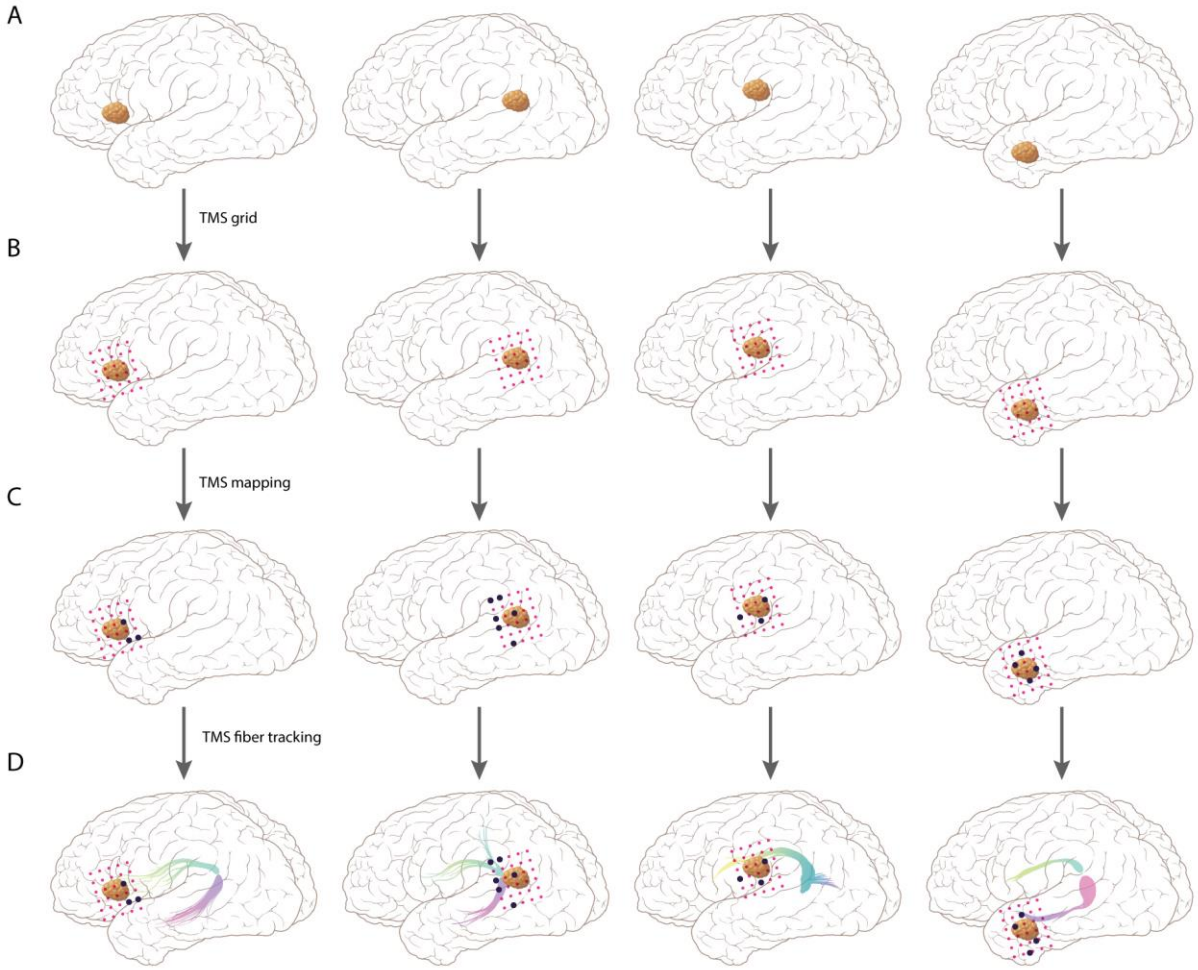
68. Gerhardt J, Sollmann N, Hiepe P, et al. Retrospective distortion correction of diffusion tensor imaging data by semi-elastic image fusion - Evaluation by means of anatomical landmarks. *Clin Neurol Neurosurg.* 2019; 183:105387.

69. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage.* 2011; 54(3):2033-2044.

Table 1: Patient demographics and baseline characteristics

Demographics	Number	Percentage (%)
Gender		
Male	47	60%
Female	32	40%
Age		
<60	53	67%
>60	26	33%
Ethnicity		
White	64	81%
Black	3	3.4%
Hispanic	5	6.3%
Asian	3	3.4%
Other	4	5.1%
Handedness		
Right	72	91%
Left	4	5.1%
Mixed	3	3.8%
Tumor Type		
Low grade glioma	33	42%
High grade glioma	46	58%
Location		
Frontal	31	39%
Temporal	31	39%
Parietal	10	13%
Insular	7	8.9%
Eloquence		
Eloquent	36	46%
Near Eloquent	34	43%
Non Eloquent	9	11%
Tumor Size		
<10 cm^3	21	34%
>10 cm^3 & <40 cm^3	31	39%
>40 cm^3	26	33%
Extent of Resection		
GTR	51	65%
STR	23	29%
<50%	5	6.3%
Preoperative Aphasia		
Yes	21	27%
No	58	73%
Transient Postop Aphasia		
Yes	37	47%
No	42	53%
Permanent Postop Aphasia		
Yes	10	13%
No	69	87%
Type of Permanent Aphasia		
Expressive	7	70%
Receptive	3	30%

Figure 1: TMS fiber tracking in glioma patients. The four columns represent tumors in different locations. (B) shows placement of the grid used to guide TMS stimulation. (C) shows examples of TMS mapping results. (D) shows examples of fiber tracking based on TMS points.



© 2024 The Board of Regents of the University of Texas System

Figure 2: Subcortical connectivity determines the functionality of TMS data and defines a canonical distribution of cortical language sites at a group level. (A) left shows the odds ratios for different subgroups of TMS points defined by their subcortical connectivity profile. (B) shows the predictive value for TMS points stratified by the FA thresholds used to define their subcortical connectivity. (C) shows the cortical distribution of TMS points (left) and TMS points with the most predictive FA threshold of subcortical connectivity (right). (D) shows the result of Gaussian smoothing (FWHM=10) and thresholding of cortical distributions of TMS points alone (left) and TMS points with robust subcortical connectivity (right). (n=79. Wald test: *P < 0.05, **P < 0.01)

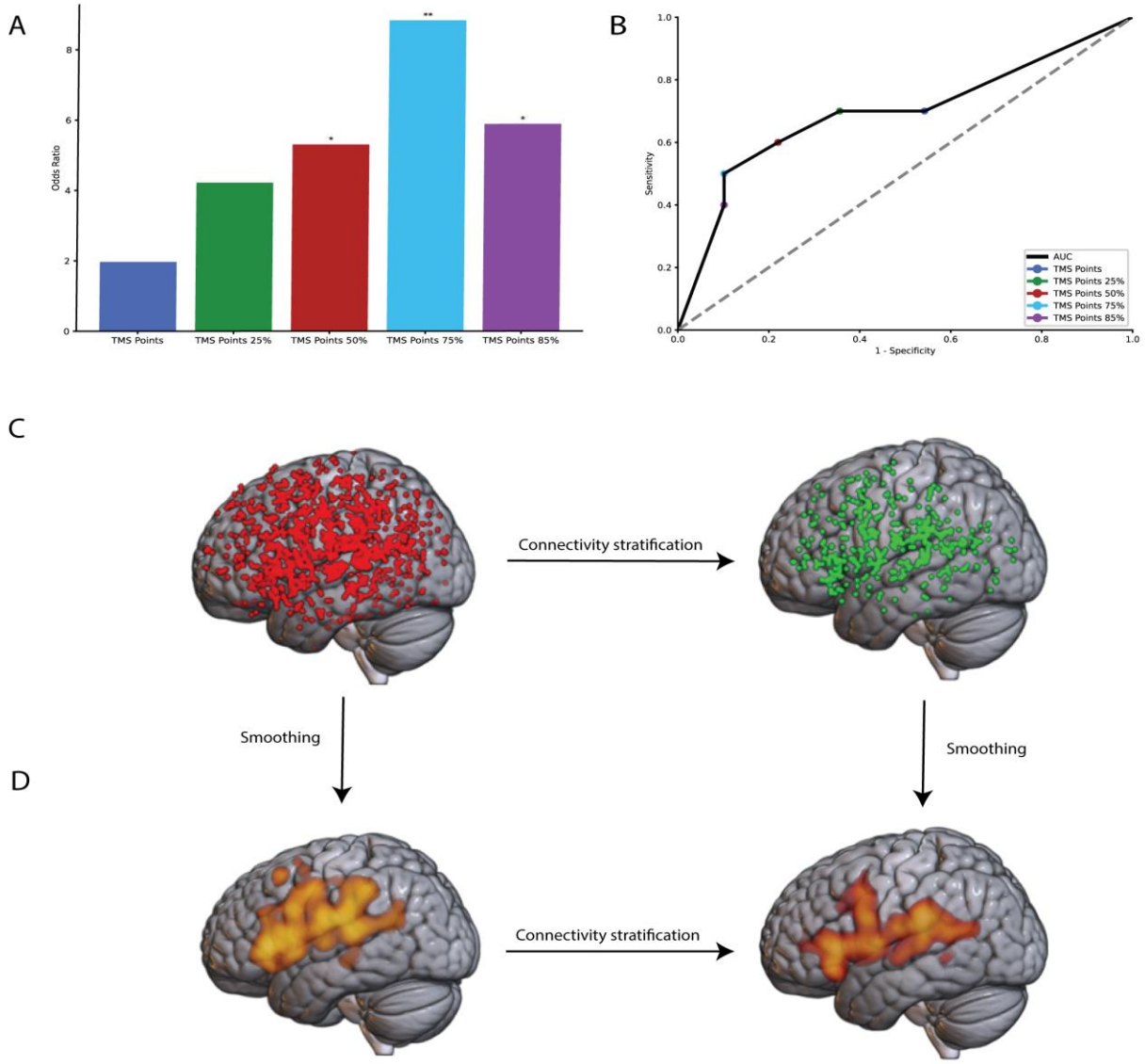


Figure 3: Data-driven fiber tracking based on cortical language mapping localizes patient-specific white matter tracts that recapitulate normative patterns of subcortical connectivity at a group level. (A) shows the probabilistic distribution of TMS tracts normalized to MNI space and added together with a voxel-wise summation. Top row shows the overlay of language associated white matter tracts in MNI space from the HCP project, depicting how TMS tract hotspots co-localize with normative tract segmentations. (B) shows the average voxel intensity of the probabilistic TMS tracts within the volumes of normative HCP tracts. (right) The left plot shows the average TMS tract voxel value within HCP tract volumes compared to all other intra-cerebral voxels. (C) shows progressive thresholding for probabilistic TMS tracts from this study compared to progressive thresholding for the probabilistic arcuate fasciculus from the HCP project. Both converge along the posterior parieto-temporal segment of the arcuate fasciculus. (D) Shows the comparison between the probabilistic SLF from the HCP project and the SLF portion of the probabilistic TMS tracts. The visualization shows that hotspots localize to the posterior segment of the tract in both cases. The scatter plot shows the significant correlation between voxel intensities of the probabilistic HCP SLF and the voxel intensities of the TMS SLF. The box plot shows the average voxel intensities for the anterior (anterior to the precentral gryus) and posterior (precentral gyrus and posterior) segments of the SLF in both the HCP and TMS probabilistic reconstructions. (n=79. One-tailed t-tests, linear regression: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. Error bars represent SD)

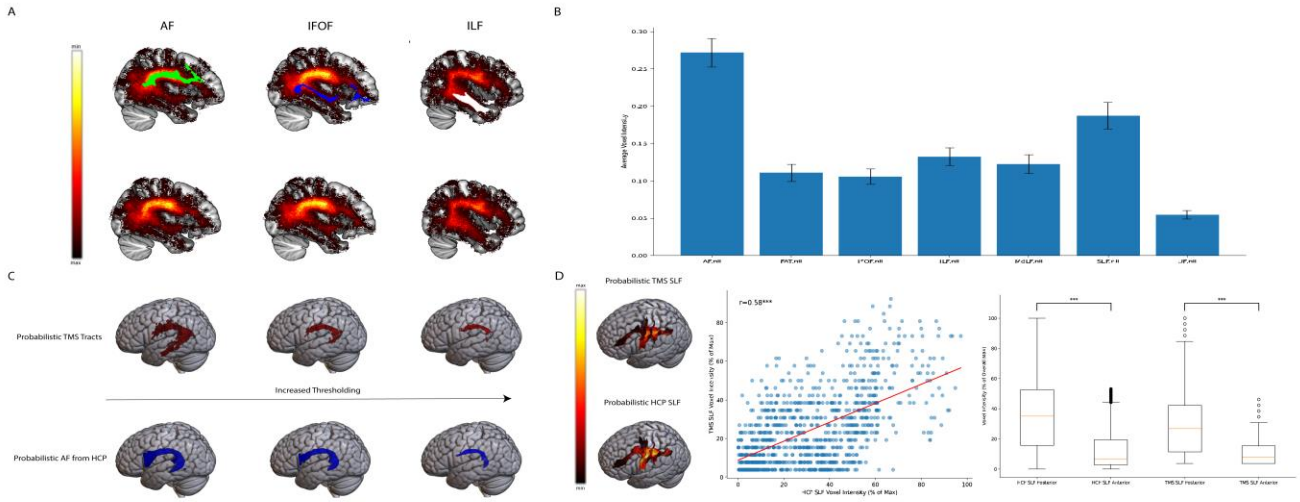


Figure 4: Non-compensable language function localizes at the subcortical level. (A) shows an ROC curve that models the predictive value of TMS seeded white matter tracts across a spectrum of FA thresholds. (B) depicts univariate versus multivariate analysis for prediction of postoperative deficits. Multivariate analysis including both subcortical tracts and subcortically connected TMS points decreased the odds ratio of the cortical data to less than 1 and increased the odds ratio of TMS tracts from 60 to 80. This plot shows that adding cortical data to the subcortical model does not improve the predictive value in a statistically significant manner while adding subcortical data to the cortical model does significantly decrease its predictive value. (C) shows the confusion matrix for the predictive value of TMS tracts at the best performing FA threshold (75%). (n=79. Wald test for logistic regression predictor significance, LRT for the statistical significance of the difference between univariate and nested multivariate models: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$)

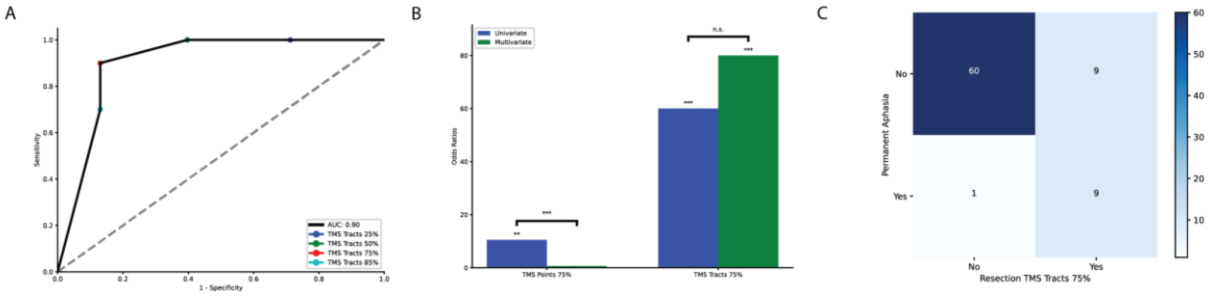


Figure 5: MNI space. The resected portion of TMS tracts in patients who sustained deficits (true positives) cluster in the white matter of the anterior temporal lobe and TPO junction and co-localize with normative HCP language-associated tracts in MNI space compared to tract-cavity segmentations in patients without deficits (false positives). (A) shows the tract-cavity segmentation and spatial normalization workflow (top left) along with a composite visualization of the true positive and false positive tract-cavity segmentations in MNI space (top right and bottom). (B) shows the preferential co-localization with HCP language associated tracts of true positive tract-cavity segmentations (n=9) compared to false positive tract cavity segmentations (n=9) across three measures: number of HCP tracts overlapped, volume of overlap, and percentage of tract-cavity segmentation overlap with HCP tracts. (C) shows how each HCP tract contributes to true positive versus false positive classification. (D) shows heatmaps that illustrate the grid-search process to find the optimal cutoffs across the three measures that most accurately classify true positives versus false positives. The black rectangle indicates the row or column that showed the highest accuracy for that heatmap. This process resulted in the selection of these three parameters: over 1 cm^3 volume of infiltration, more than 3 tracts overlapped, and more than 50% of the tract-cavity segmentation overlapped. (n=18. One tailed t-test: $*P < 0.05$)

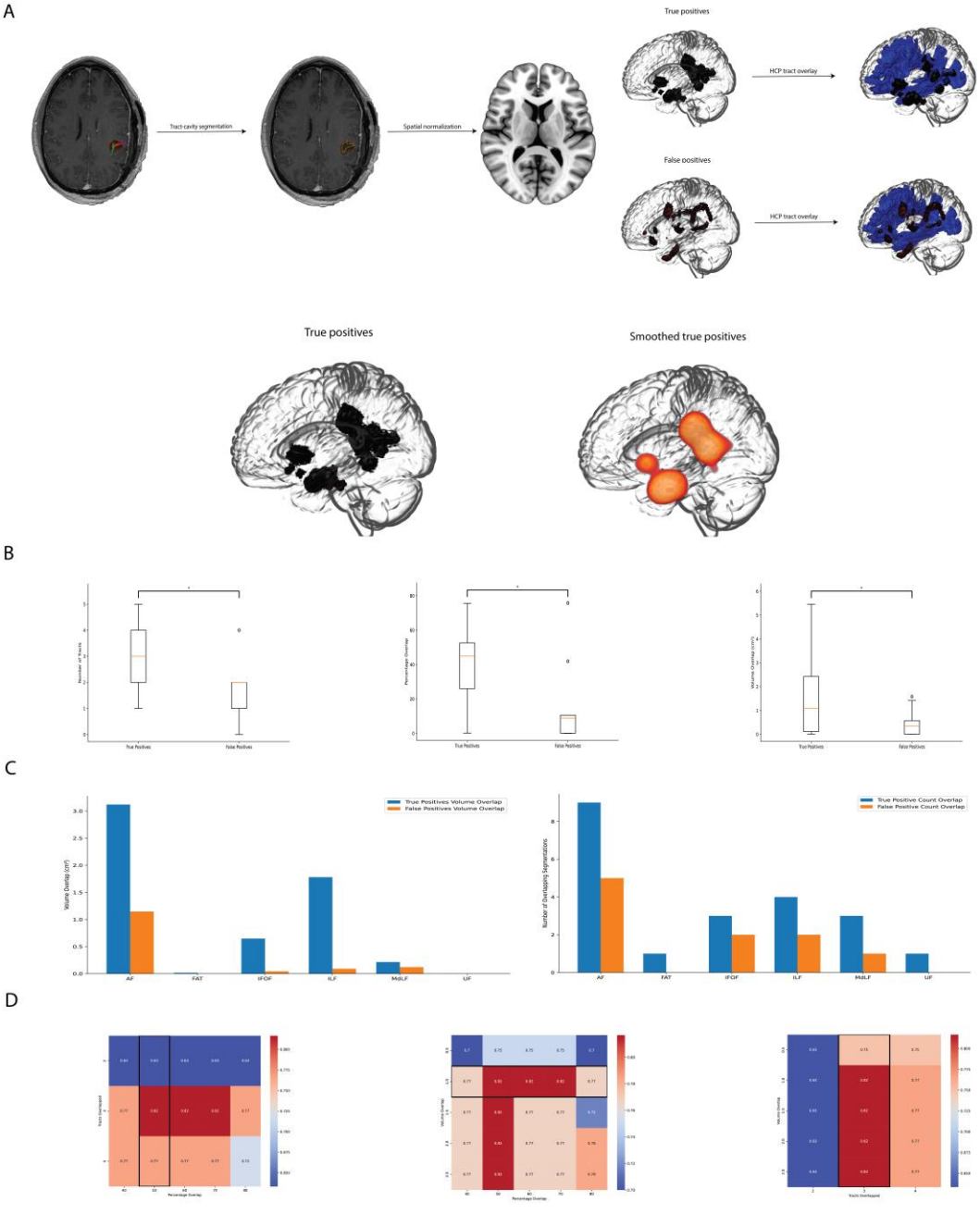


Figure 6: MNI space. Non-compensable language function localizes to individualized substructures of normative white matter tracts. (A) Resections of normative white matter substructures (ROIs 1-8) only produce deficits if they include TMS tracts. This demonstrates that these ROIs are not associated with deficits at a group-level without individualized data. (B) Patient-specific anatomical locations of the resected co-localized TMS tract-HCP segmentations (smoothed, FWHM=5) in patients with deficits. (n=8. One-tailed t-test: $*P < 0.05$)

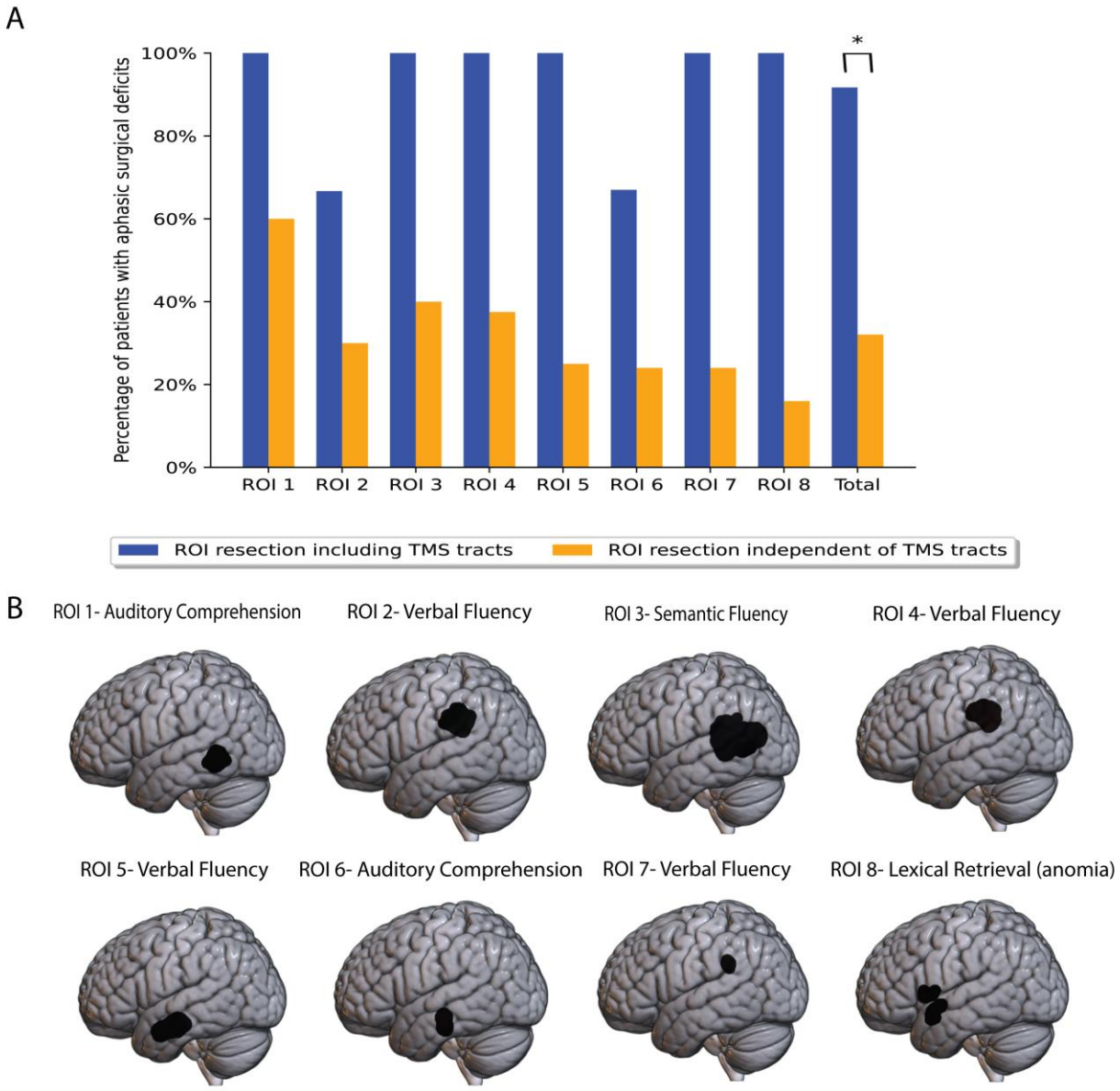


Figure 7: Overview of the development process, final performance of predictive model, and prospective application in a surgical glioma patient. (A) shows a confusion matrix for the best performing final model in predicting aphasic deficits. (n=79). (B) shows how the odds ratio significantly increases with each iteration, culminating in an odds ratio of 134 in the final comprehensive model. (C) shows the predictive value of each of the developing components of the model in terms of an ROC plot. (D) illustrates the prospective application of the data for surgical planning and execution with a flow chart. Patient presents with a dominant hemisphere lesion and undergoes TMS mapping and fiber tracking. HCP tracts are overlayed and a segmentation of the TMS-HCP tract overlap is made. These data accurately predicted subcortical stimulation responses (n=6), extent of resection (n=1), and surgical outcome (n=1). (E) shows the poor performance of TractSeg white matter segmentations in predicting the same metrics. Left shows the AF (red) with the negative (blue) and positive (black) stimulation point. Middle shows the MdLF (red) with the negative and positive stimulation points. Right shows the significant overlap between the resection cavity (blue) and MdLF. (Wald test: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$)

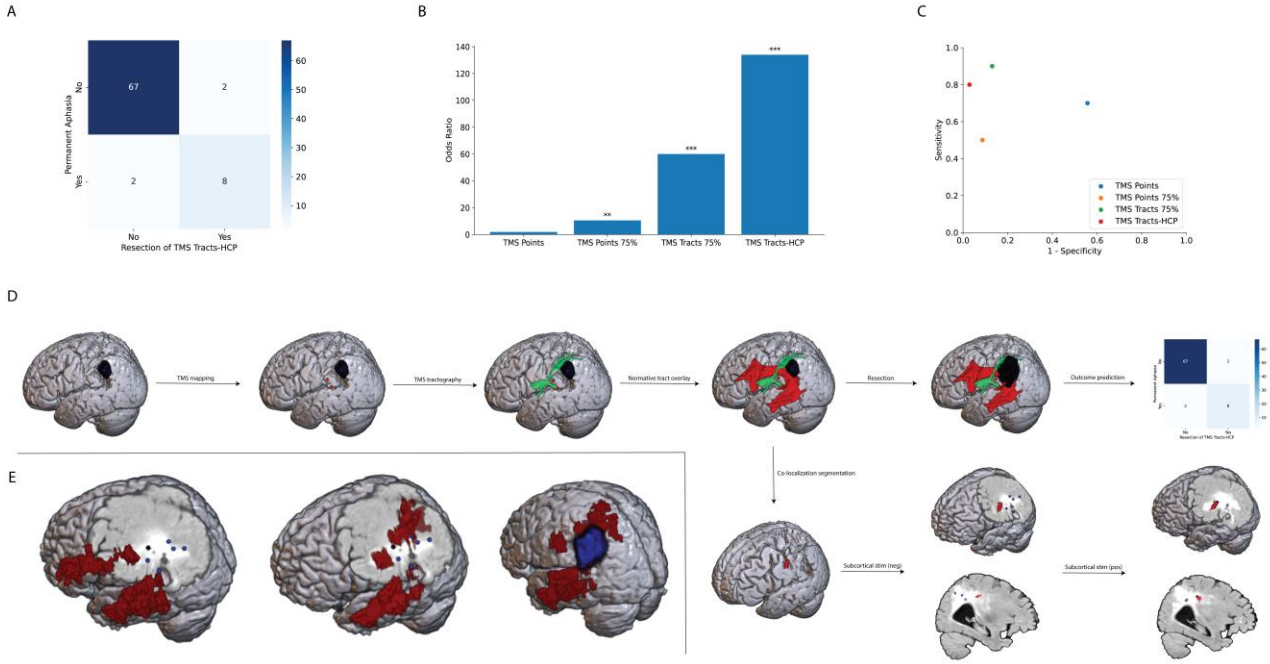
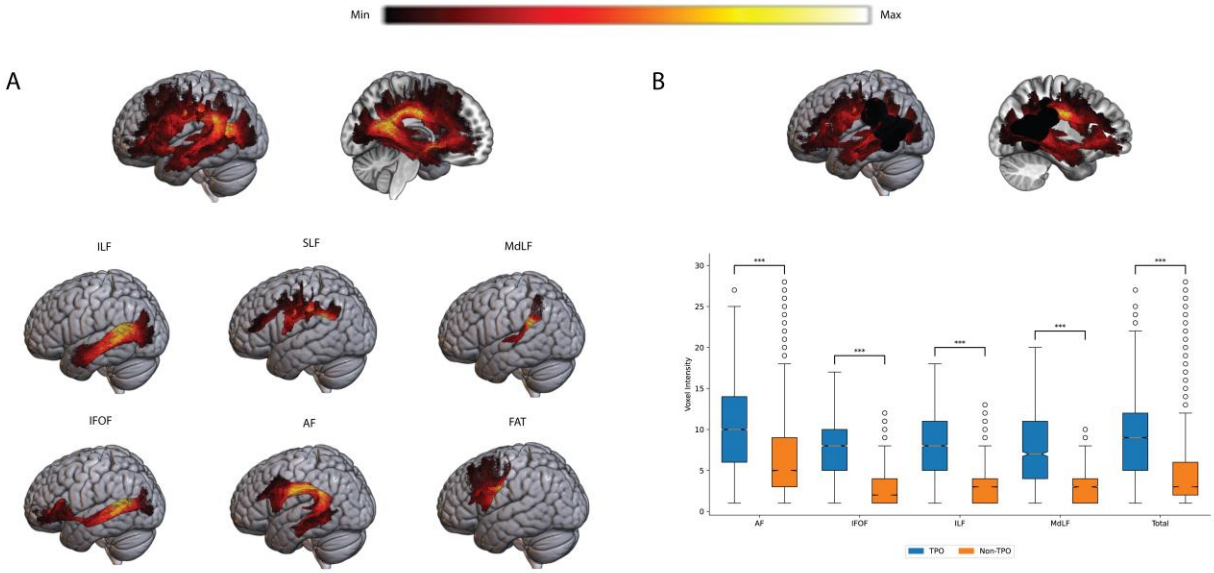



Figure 8: Novel pattern of subcortical connectivity anatomically correlates with regional vulnerability to aphasic damage: visualization and voxel-wise analysis of probabilistic atlas for language-essential white matter in glioma patients. (A) shows visualizations of probabilistic atlas of white matter in glioma patients that leads to long-term aphasic deficits when surgically damaged. This visualization shows hotspots in the white matter deep to the TPO in the tracts that penetrate this region. (B) shows spatial overlap analysis of the tracts from this atlas that penetrate this region with the ROI from Figure 5 that represents the cluster of "true positive" (resection corresponds with long-term aphasic decline) tract-cavity segmentations that localized to the TPO. (independent t-test: $***P < 0.001$)