



School of Psychology

Ysgol Seicoleg

Language-Eloquent Cortex Mapping in Paediatric Epilepsy Surgery

Thesis submitted in partial fulfilment of the requirement for

the degree of:

Doctorate of Clinical Psychology (DClinPsy)

South Wales Doctoral Programme in Clinical Psychology

Cardiff University

Alexander P. Marsh

Supervised by:

Dr Chris Hobson & Professor Ingram Wright

20th August 2021

Acknowledgments

I would like to thank Dr Chris Hobson, for his support and guidance throughout this project. Chris graciously accepted me as an additional supervisee, who was attempting to undertake an ambitious project in limited time. For his support and encouragement throughout this time, I am very grateful.

I would like to thank Professor Ingram Wright, for not only his invaluable guidance and supervision within this project but for the last seven years of mentorship. I would not be the psychologist I am today without the Ingram's generosity in sharing his wealth of knowledge and experience, fostering my passion for paediatric neuropsychology and research.

I would like to acknowledge, Jonathon Delve, who spent tireless hours pulling scan data for me from mysterious servers, and Holly Elbert, for her support and patience in my inept scripting skills. I owe you more than simply baked goods.

I would like to acknowledge the tremendous support of the 2018 South Wales DClinPsy cohort. Notably, Gemma, Jess, Chris and Rachael. You have been my calm space to let it all out and the grounding rock of normalisation throughout this journey.

Lastly, I would like to thank my family. My parents, who have always encouraged me and have vicariously become neuroscientists through my long monologues on FaceTime. My grandparents, for their constant pride, providing me with motivation when it had run out. Jake, Zoe and Lily, for providing joy and welcome distraction. My dog, Mylo, for your reminders to get out the house – though your motivations were likely, at least in part, for the sniffs. Finally, Darren, I don't have words of my own to express my feelings of love and adoration, so in the immortal words of Patsy Stone, "Cheers, thanks a lot!".

Contents

Acknowledgments	1
Preface	5
Paper 1: The Predictive Validity of Language Mapping for Post-operative Outcome in Paediatric Epilepsy Surgery: A Systematic Review	9
Abstract	10
Key Points	11
Key Words.....	11
Introduction.....	12
Aims	19
Methods.....	21
Search strategies	21
Inclusion and exclusion criteria	22
Data extraction.....	23
Assessing Methodological Quality.....	23
Data Presentation and Analysis.....	24
Results.....	24
Search Results.....	24
Full-text reliability concordance	25
Study Characteristics.....	27
Language Mapping Tasks	29
Quality Assessment.....	29
Attrition and Failure Rates.....	36
Predictive Validity Findings.....	37
Comparative Utility Synthesis.....	41
Discussion	43
Extent of Evidence.....	43
Quality of Evidence	45
Evidence for Predictive Validity	46
Strengths and Limitations of this Review	49
Recommendations for Future Research	50
Conclusion	52
Disclosures.....	52
References.....	53

Paper 2: Convergent Validity of Resting-State and Task-Based Functional Magnetic Resonance Imaging in Lateralising Language for Paediatric Epilepsy Surgery	64
Abstract	65
Key Points	66
Key Words.....	66
Introduction.....	67
Resting-State fMRI as an alternative	68
Resting-state fMRI and language.....	70
Current study.....	73
Methods Section	74
Subjects.....	74
Neuropsychological Testing	74
fMRI Paradigm.....	75
MRI Acquisition	76
Task-Based fMRI Analysis	76
Resting-State fMRI Analysis	77
Lateralisation Index (LI).....	78
Further Statistical Analysis.....	79
Results Section	82
Language Mapping	82
Index Concordance.....	87
Categorical Concordance.....	89
Validity with Post-operative outcome.....	90
Discussion	92
Feasibility of Resting-State Language Network Extraction in Routine Practice.....	92
Concordance of Language Lateralisation and Estimated Surgical Risk	92
Additional Limitations.....	98
Future Directions.....	99
Conclusions.....	101
Disclosures.....	102
References	103
Appendix.....	118
Submission Guidelines for Epilepsia	118
Research Team	129
QUADAS-2 Tool	130
Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence	138

Prospero Protocol.....	139
Empirical Project Approvals.....	141
Supplementary Material for Empirical Paper.....	142
Quality Control	142
BRCH Quality Control Guide	142
Quality Control Metrics for Task Based Scan	145
Quality Control Metrics for Resting state.....	149
Example of Case Excluded for Motion.....	149
Plans for Dissemination	152
Journal Publication.....	152
NHS Dissemination.....	152

Preface

Paediatric epilepsy surgery is an evidence-based treatment for focal onset epilepsy. An integral part of evaluation for surgical intervention is assessment of risk to cognitive skills, notably language. Previous work demonstrates language deficit can lead to academic underachievement and long-term social, professional, and neuropsychological problems. Post-operative language deficits can be estimated using clinical variables, such as age of seizure onset, presence of brain abnormalities, and age at time of surgery. However, these variables alone are insufficient to reliably predict outcome on an individual basis. Consequently, language mapping methods have been developed in order to improve estimation of risk for post-operative language deficits.

There is established evidence that a variety of language mapping methods are feasible in the paediatric population. Early reviews for individual techniques have shown promise in their predictive validity for post-operative outcome but have predominantly relied on mechanistic or case study evidence. More recently, group-level studies have examined predictive validity in the paediatric context, but the overall applicability of this evidence is remains unclear. Several works have investigated these techniques in a combined sample of adults and children. However, the evidence of these investigations is significantly restricted due to various methodological considerations, which are discussed.

A systematic review of the evidence in paediatric epilepsy populations was undertaken to:

- (1) To identify the current extent of group level studies that have explored the predictive validity of language mapping;
- (2) To assess the quality of these studies' evidence, utilising standardised metrics;

- (3) To synthesise the evidence of group level data for paediatric language mapping in predicting post-operative language decline in children considered for epilepsy resective surgery;
- (4) To use this synthesis to inform clinical practice guidance;
- (5) To provide guidance for any further research required.

A limited number (n=6) of studies met criteria for inclusion. Group-level paediatric evidence was restricted to diffusion weighted imaging (DWI), the Wada test and electrocortical stimulation (ESM) mapping. No paediatric group-level studies were found for any other mapping technique. The quality of the evidence was of unclear quality, predominantly due to the absence of reporting of key methodological aspects of studies and their samples. Promising preliminary evidence for DWI and ESM was reviewed. DWI demonstrated strong predictive validity and good success rates across three studies. Whereas ESM, had moderate quality evidence from only one study, demonstrating strong associations with post-operative outcome, but modest success rates. Due to limited number of records, further studies with improved methodology for both techniques are warranted before translation into routine practice. There was variable evidence for the Wada test in predicting outcome and confidence was compromised by methodological limitations. Recommendations for further research included: development and utilisation of open-access databases within paediatric epilepsy surgery in an effort to increase group-level evidence; improved reporting within studies of key demographic and methodological information; improved selection of reference standards and methods for determining post-operative decline; and further investigation of individual versus multimodal assessment of language mapping across methodologies, including functional magnetic resonance (fMRI) neuroimaging.

This systematic review is notably the first to explore efficacy in paediatric populations across modalities and benefits from robust quality review, utilising two clinically applicable tools. However, it is limited by lack of second review of all articles in inclusion. It also did not consider seizure freedom, alongside cognitive outcomes, nor the precise methodological variations within techniques. Despite this, it presents important clinical implications with respect to the limited evidence base to support routine clinical decision making in language lateralisation for epilepsy surgery.

fMRI is commonly employed as a technique for language mapping in paediatric epilepsy surgery. Task-based fMRI language mapping has demonstrated good concordance with 'gold standard' language mapping techniques, such as the Wada test and ESM. In adults, it has also shown good predictive validity for language outcome following epilepsy surgery. However, various limitations hinder its use within paediatric epilepsy populations. These are discussed. One technique that overcomes a number of these limitations is resting-state fMRI. Resting-state fMRI has shown good concordance to task-based fMRI in adult studies and preliminary evidence of concordance in adolescent epilepsy populations. However, there is limited evidence in samples that represent those presenting to epilepsy surgery programmes, notably for children under 12 years of age. Further studies are recommended due to limited sample sizes. The following study was performed, in order to:

- (1) Investigate whether generation of reliable resting-state language networks is feasible in routine practice in a diverse group of paediatric epilepsy patients;
- (2) explore concordance of language lateralisation and estimated surgical risk between resting-state fMRI language networks and conventional task-based fMRI methods;

(3) report surgical outcome in relation to language mapping in a small case series, exploring preliminary evidence for predictive validity.

Patients (n=26; n=10 under 12 years old), underwent language lateralisation with both resting-state and task-based fMRI. Surgical outcome was described in six cases. There was poor concordance found between resting-state and task-based fMRI in terms of degree of lateralisation and surgical risk, and for categorical (left, right, bilateral) language representation. However, categorical estimation of surgical risk demonstrated good concordance. There was promising evidence that resting-state fMRI may be useful in surgical decision making, but insufficient evidence was found for resting-state fMRI's use as a proxy to task-based fMRI for the lateralisation of language function in paediatric epilepsy surgery candidates.

This study built on previous findings examining validity in adolescent populations, extending the sample to pre-adolescent children. It also benefitted from a small predictive validity case series. A key strength was the utilisation of routine clinical data for validation, bolstering the translational feasibility and ecological validity of the findings. However, there are some limitations, including: use of concatenated, not dedicated resting-state scans; lack of behavioural scan data; and limited number of right-sided language cases, that warrant further research to validate the findings. Nonetheless, it importantly presents the clinician with evidence for caution in considering resting-state a proxy of task-based fMRI assessment.

Paper 1: The Predictive Validity of Language Mapping for Post-operative Outcome in Paediatric Epilepsy Surgery: A Systematic Review

Alexander P. Marsh

(MarshAP@Cardiff.ac.uk)

School of Psychology, Cardiff University, Cardiff, United Kingdom

This work is submitted in partial fulfilment of the requirements for the degree of
Doctor of Clinical Psychology.

(2021)

Prepared for submission to *Epilepsia* with expanded word count as permitted by
DClinPsy LSRP Guidance

Paper abstract word count: 294

Paper body word count: 7790

Total word count: 8084

Abstract

Objectives: Estimating risk of language decline is warranted in paediatric epilepsy surgery but there is limited knowledge of the validity of language mapping techniques in this context. This paper's main aims were to (1) evaluate the state of the evidence investigating the predictive validity of any language mapping technique used in epilepsy surgery and (2) provide a synthesis of available evidence to inform future research and practice.

Methods: A systematic search of MEDLINE, APA PsychNET, and Web of Science (up to October 2020) was conducted. Records of cohort studies that reported mapping techniques and surgical outcomes in paediatric epilepsy were identified. Quality appraisal was undertaken with QUADAS-2 and OCEBM evidence levels. Relationships between mapping techniques and post-operative language outcome were synthesised across studies.

Results: Six studies met criteria for inclusion. These studies examined diffusion weighted imaging, electrocortical stimulation and the Wada test. The quality of five records was assessed as 'unclear' due to the lack of reporting of post-operative outcome methods and one was classified as 'low' quality. Diffusion weighted image mapping was found to have a strong relationship with language outcome across three studies. Electrocortical stimulation also demonstrated robust predictive outcome in one record but with high failure rates (48%) in another. Wada had variable success depending on reference standard thresholding.

Significance: There is great need for further investigation of language mapping techniques in predicting post-operative outcome in paediatric epilepsy surgery. Preliminary evidence demonstrates promising value of diffusion weighted imaging and electrocortical stimulation techniques within paediatric practice; however, further work needs to be undertaken to

support adoption of techniques into routine practice. Future research should specifically improve reporting of outcome methodologies. Other common mapping techniques should be investigated due to the successful translation of DWI and ESM into paediatric populations from adult practice.

Key Points

- There is a significant lack of evidence investigating the validity of language-eloquent cortex mapping in paediatric epilepsy surgery.
- The quality of the evidence is unclear due to insufficient reporting of methodologies for assessing post-operative language outcome.
- Several techniques have potential, with preliminary evidence for diffusion weighted imaging and electrocortical stimulation.
- Due to particular considerations around progressive functional specialisation and plasticity in the context of intractable epilepsy, specific validation is needed in paediatric samples before these methods are adopted into routine practice guidance for risk estimation

Key Words

Language Lateralisation; Systematic Review; Eloquent Cortex Mapping; Childhood Epilepsy

Introduction

Epilepsy surgery is considered a highly effective treatment for children with focal, structural and refractory epilepsy¹⁻⁴. Compared to alternative treatment options (e.g., antiepileptic drugs, the ketogenic diet, or vagal nerve stimulation), which confer a <50% chance of seizure reduction, epilepsy surgery offers superior seizure freedom rates of >70% for up to two years post-operatively¹⁻³. Given the relative success and advances in this approach, the number of operative procedures has increased in recent years¹⁻⁴. With this increase, there is growing consideration that seizure freedom alone is an inadequate measure of outcome. Broader factors, notably cognitive outcome⁵⁻⁷, are being included in surgical success considerations^{8,9}. Meta-analytic data examining neuropsychological surgical outcomes demonstrated that surgery has the potential to lead to both decline and improvement of cognitive function^{6,10}. Although control of seizures is the primary goal of surgery, better seizure control alone does not necessarily guarantee better outcome¹¹. A recent meta-analysis¹¹ of adult outcomes for childhood epilepsy demonstrated that cognitive dysfunction, above seizure control, was the most reliable predictor of poor outcome in education, employment, relationships, finance, independent living, and social functioning. Whereas better seizure control alone did not necessarily guarantee better outcome¹¹.

Of the range of cognitive outcomes of concern in paediatric epilepsy, language is of prominent importance as language dysfunction can lead to academic underachievement and long-term social, professional, and psychological problems¹²⁻¹⁴. Several paediatric studies have demonstrated altered language function following resective surgery¹⁵⁻¹⁸. Resections within the dominant left hemisphere can lead to language deficits both in the immediate post-operative weeks and in the longer term¹⁹⁻²². Preservation of language function is therefore of

significant clinical importance. However, estimating operative risk to language is an ongoing clinical challenge.

Risk assessment of surgical resection to language is complex due to variability in cerebral language representation within the population. Longstanding evidence demonstrates that language is typically left-hemisphere dominant with known variation (i.e., bilateral or right-sided dominance) in both right- (5%) and left-handed (22%) individuals^{23–30}. However, in those with epilepsy localised to typical language-related areas, there is even greater variability; with atypical hemispheric and atypical regional language representation occurring in up to 70% of these individuals^{23,26,28,30–44}. Furthermore, damage to even anatomically distal but functionally related structures (e.g., mesial temporal and subcortical structures) can have negative impacts on language function and disrupt typical organisation within the brain, leading to neuroplastic change^{36,37,42,45–51}. Background clinical variables, such as pathology and epilepsy semiology, alongside neuropsychological profile, may provide some informed estimation of language localisation at a group level^{52,53}. However, these variables have not demonstrated sufficiently robust associations to prove useful in the individual case^{53–55}.

To respond to the challenge of mapping language, several techniques and methodologies have been investigated^{56,57}. In the last five years there has been a notable (circa 55%) increase in the number of publications relating to paediatric language mapping in epilepsy surgery⁵⁸. The characteristics of the most prominent techniques investigated are shown in Table 1 and more detailed review of their methodology is available elsewhere^{57,59,60}. Several reviews exploring functional language mapping techniques have examined the validity of various mapping approaches in determining language function against other metrics^{23,28,57,61–68}, such as in vivo behavioural validation (i.e. speech arrest associated with cortex knockout through

stimulation or anaesthesia) or alternative mapping methodologies (e.g. comparing fMRI against the Wada test⁴²). Whilst these investigations have clinical merit, there are circularity issues that limit their use as evidence for accurate identification of language and estimating post-operative decline. Recent work demonstrates that even stereo-electroencephalography language electrocortical stimulation mapping (considered to be the 'gold standard' mapping technique), may under detect areas of language function^{69,70}. Utilising imperfect reference standards when evaluating the diagnostic utility of a mapping methodology introduces verification bias that may fail to effectively determine the mapping method's accuracy⁷¹, thus introducing circularity. There is also reasonable evidence for concern that in these comparative studies, the index standard may outperform the reference standard. For example, several studies⁷²⁻⁷⁵ have compared the novel methodology of resting-state fMRI against task-based fMRI to validate its use. However, there are several considerations that mean resting-state may be a superior tool; for example, resting-state may be a more robust metric⁷⁶ due its better Signal-to-Noise Ratio (SNR)^{72,76,77}. Cochrane guidance advises against reviews of diagnostic test accuracy in such circumstances, where imperfect reference standards are employed⁷¹. Arguably the most robust reference standard for language mapping would be the postoperative language deficits assessed by neuropsychological evaluation, as these represent the true functional impact of neurosurgery⁶⁹. Whilst a recent systematic review⁶ has explored cognitive outcomes (including language) from paediatric surgery, it did not explore the utility of mapping techniques in predicting outcome.

Table 1. Summary of different techniques investigated to map language function (adapted from Sagar⁵⁷)

	Technique	Spatial Resolution	Temporal Resolution (ms)	Principle
Invasive	Direct Electrical Stimulation	5 mm	Instant	Electrical stimulation of cortex
	Wada	Hemisphere	N/A	Cortical disruption via anaesthetisation of functional hemisphere
	Electrocorticography Positron Emission Tomography	<10mm ~4mm	<1ms 45000	Electrical recording Metabolic (FDG) or Cerebral Perfusion (H ₂ ¹⁵ O)
Non-Invasive	Functional Magnetic Resonance Imaging	1mm	~3000	Blood-oxygen level dependent signal response
	Magnetoencephalography	2-3mm	1	Magnetic signal
	Diffusion Tensor Imaging/Diffusion Weighted Imaging	2.0-2.5mm	N/A	H ₂ O Diffusion
	Transcranial Magnetic Stimulation	5-10mm	>70	Magnetic Stimulation
	Electroencephalography	7-10mm	2	Scalp-surface electrical recording
	Functional Near Infra-red spectroscopy	1cm	100	Absorption spectrum

These pre-operative assessments can be important for clinical management in different ways. As previously outlined, there are certain situations in which clinical variables and presentation are clear indications of suitability for surgical intervention. However, suitability is just one of a number of considerations with regard to the potential costs and benefit of surgical treatment. Surgery should always consider the best available evidence around potential risk from multiple contributory sources. For example, in the instance of a young patient who is in refractory status epilepticus, surgery is likely

to take place irrespective of any overlap between eloquent language cortex and target surgical zone due to the risk of morbidity. In this instance, language mapping would not be considered essentially for surgical decision making but is additionally useful in terms of counselling operative risk and planning post-operative care. If eloquent cortex was considered to overlap with the surgical resection zone, it would be important to use this information to inform consent procedures, and to incorporate it into post-surgical care planning, such as the provision of speech and language therapy. In other instances, surgical decisions are less clear. For example, in a sixteen-year-old child, who had onset of a focal epilepsy at age eight, whose EZ is localised to the inferior frontal gyri regions, and there is modest language impairment on behavioural assessment and infrequent but modestly severe seizures. In this situation, language location may play a vital role in surgical decision making and in counselling the patient regarding potential risks. It will remain unclear from the demographic and clinical variables where language may be lateralised as there is significant potential for reorganisation given this history. If language has not benefitted from atypical reorganisation, resection of eloquent cortex may result in significant impairment with potential little recovery due to the age of the patient. Alternatively, should reorganisation have already occurred to the contralateral hemisphere, the patient would potentially benefit from a confident surgical decision with few if any risks to postoperative language ability.

There is notable evidence for the predictive validity of language mapping techniques in adult epilepsy surgery^{19–21,28,73,78–81}. However, there are significant limitations in generalising this to paediatric populations⁸². Even in healthy children, language mapping poses several specific challenges: from developmentally appropriate paradigm design and extensive patient preparation to specific paediatric considerations for data processing^{48,83}. These considerations compromise the generalisation of adult findings to paediatric studies.

One key factor impacting generalisation is the relevant structural and functional differences between children and adults, which may affect the acquisition, analysis, and interpretation of

paediatric language mapping data. For example, the maturational stage of myelin development and progression of synaptic pruning are associated with differences in metabolism and blood flow^{84,85}. These physiological differences can affect the detection, magnitude, and extent of BOLD response⁸⁴ and the functional sensitivity of electrical cortical stimulation⁸⁵. Anatomical differences between child and adult brains can distort the magnitude of signal and the functional localisation, notably when data are warped onto adult anatomical atlases for analysis⁸⁴, as commonly performed in fMRI, magnetoencephalography (MEG) and Positron Emission Topography (PET).

The differences in structure and physiology also relate to functional specialisation. Younger children demonstrate relatively reduced frontal activity compared to adults, whilst showing increased activity in the medial parietal cortex, posterior cingulate, and occipital cortex, in comparison with adults on a variety of language tasks⁸⁶⁻⁸⁸. Additionally, typically developing children under 10 years are more likely to have bilateral language lateralisation compared to adolescents and adults^{43,89-91}; indicating increasing specialisation of language to the left throughout brain development and complicating the generalisation of the adult evidence base to this younger sample.

These factors are amplified in the context of epilepsy, where both the seizures and anti-epileptic medication can have an impact. Atypical language representation is even more common in children with epilepsy^{22,44,56,92} and previous work has demonstrated reliability declines for certain mapping techniques in those with atypical representation²⁸. Additionally, children with developmental epilepsy demonstrate differing patterns of neuroplastic function to those with adult onset epilepsy^{22,44,92}; with interhemispheric reorganisation of language

being more common in childhood onset epilepsy, whilst intrahemispheric reorganisation and compensation is more common in adult onset epilepsy^{22,44,92}.

The capacity for children with epilepsy to engage and find the negative effects to be acceptable in mapping tasks has been shown to be compromised when compared to adults and individuals without epilepsy^{56,93}. Furthermore, the assumptions of the underlying principles of mapping techniques may be compromised. For example, previous work has demonstrated relationships between fMRI signal strength and language skills⁹⁴. Language skills are frequently compromised in many paediatric patients¹², potentially leading to poorer signal in epilepsy populations. Other neurodevelopmental sequelae of epilepsy may also impact, such as neurovascular decoupling that may further perturb localisation techniques^{95,96} by impacting the signal upon which they rely, i.e., those that rely on hemodynamic responses. Anti-epileptic medication is another example. Antiepileptic drugs have been shown to elevate thresholds for stimulation mapping⁹⁷, further interfering with the sensitivity of these techniques, as outlined above.

Combined, these factors explain previous findings that functional mapping of language in paediatric epilepsy often produces more unpredictable results^{85,98}, validating the need for a specific examination of the paediatric literature. Furthermore, children might be candidates for surgical approaches not viable in adults^{48,99,100}; and the potential benefits may outweigh the costs in paediatric surgeries, due to improved outcomes^{6,7}, which may not be justifiable in adult populations^{48,99,100}. This makes efforts to establish validated language mapping in children more pressing but also indicates the different context in which surgical decision making takes place, requiring dedicated exploration of paediatric research.

There is evidence for the validity of mapping techniques in paediatric epilepsy populations; however, much is in the form of case studies^{101–107}. Whilst these are useful to inform future research and practice considerations, they are not suitably robust to inform clinical guidance and so are of limited use in routine practice¹⁰⁸. Consequently, several studies have begun to examine predictive validity using recommended¹⁰⁸ cohort designs^{85,101,109–111}, in order to make practice recommendations. However, to date, this evidence has not been systematically explored or evaluated for its quality or utility in informing clinical guidance.

Aims

This primary aims of this paper are:

- (1) To identify the current extent of group level studies that have explored the predictive validity of language mapping in children undergoing epilepsy surgery;
- (2) To assess the quality of these studies' evidence, utilising standardised metrics^{112,113};
- (3) To synthesise the evidence of group level data for language mapping in predicting post-operative language decline in children considered for epilepsy resective surgery;
- (4) To use this synthesis to inform clinical practice guidance;
- (5) To provide guidance on the requirements for further research.

A secondary aim of this systematic review is to describe the acceptability of language mapping techniques. No previous systematic review of paediatric language mapping has, to the author's knowledge, explored the acceptability across techniques. This is clinically useful data for routine practice when weighing up the success of a technique. Whilst a technique may hold strong predictive validity, if it is not acceptable to the child, it is of little use in routine

practice. This will be explored through examining drop out and failure rate reporting within records.

This review will not consider task design or comparative efficacy of different tasks as this has been examined elsewhere⁶⁶. Similarly, the efficacy of the technique will be examined but it is beyond the scope of this review to explore specifics of technique, such as comparative utility of pharmacological regimes for the Wada test. Many works have considered both adults and children together and grouped them in their analyses of predictive validity of language mapping^{73,79,81}. This grouping compromises the results for the methodological reasons outlined above, therefore these works will not be included.

Methods

This review was conducted and reported in accordance with Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) statement¹¹⁴.

Search strategies

The specific search strategies were created in collaboration with an expert Health Sciences Librarian with skills in systematic review searching. Two main blocks of search terms (Table 2) were finalised and combined with the 'AND' function to search the following publication databases: MEDLINE, APA PsychNET, and Web of Science. Each were searched from each database's inception to 29/10/2020.

The specific search strategies were created in collaboration with an expert Health Sciences Librarian with skills in systematic review searching. Several search strategies were tested to ensure best sensitivity to target records. Originally, search terms made explicit reference to known lateralisation methods outlined in Table 1 within Block 1 (see Table 2). This was found to limit records. When these terms were introduced as a separate block, the records became too general to process effectively and were in the order of tens of thousands. Two main blocks of search terms (Table 2) were finalised and combined with the 'AND' function to search the following publication databases: MEDLINE, APA PsychNET, and Web of Science. Each were searched from each database's inception to 29/10/2020. The results of the finalised search were cross-referenced with prior searches that returned relevant records to ensure relevant records were not lost, as a quality control process.

Table 2. Blocks of search terms used to query publication databases in the review strategy

Block	Search Terms
Block 1: Mapping	(language laterali*/OR language mapping/OR language/OR (stereotaxic/AND atlas)/OR (brain AND map*)/OR (brain and mapping))
Block 2: Population	(epilepsy/OR epilep*/OR surgery*/OR surgical/OR surg*/OR)

In addition, where relevant, reference lists were reviewed, and additional studies included into screening. Search results were loaded into the web-based Covidence¹¹⁵ programme to remove duplications, and to complete screening, blinded full-text review, and data-extraction.

Inclusion and exclusion criteria

One reviewer (AM) independently screened all titles and abstracts and identified potentially relevant material to be reviewed. Full text review was undertaken by AM. Second reviewer (DQ) was blinded and reviewed 38% of full-text records. AM-DQ rating concordance was blindly computed through Covidence software.

A study was included if it conformed to all the following criteria:

1. The study reported original data published in a peer-reviewed scientific journal;
2. The study included an investigation of mapping of language functions;
3. The study included paediatric (<18 years of age) epilepsy patients who had undergone epilepsy surgery;
4. The study was a cohort design;
5. The study reported post-operative language outcomes in relation to mapping procedures.

Case reports, case series, and 'grey' literature were excluded as these are rated as 'very low' quality of evidence^{108,113} noted to be of limited importance in clinical decision making¹⁰⁸. Additionally, papers who did not conduct a distinct statistical analysis on paediatric samples were excluded.

Data extraction

A standardised data extraction form was utilised within Covidence. For each study, the following information was extracted: country of origin, study design, study funding sources, possible conflicts of interest for study authors, population description, inclusion criteria, exclusion criteria, total number of participants, age of participants, sex, ethnicity, seizure types, age of epilepsy onset, seizure laterality, baseline IQ, baseline language function, dropout rate, mapping technique modality, language mapping task, language assessment, methods for assessing outcome, timing of post-operative assessment and predictive validity findings.

Assessing Methodological Quality

The Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence are a widely used tool that allows clinicians to assess the level of evidence to best inform practice¹¹³. The Oxford Centre for Evidence Based Medicine¹¹³ levels of evidence were used to assess the level of evidence of included studies. The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool¹¹² was used to evaluate the quality of evidence, as recommended by Cochrane guidance¹¹⁶. Two reviewers independently assessed quality, blind to the other's view. Disagreements were resolved by consensus.

Data Presentation and Analysis

A two-staged approach to analysis was planned. Stage one involved the tabulation of study characteristics. In the second stage, sensitivity and specificity data were due to be extracted into a 2x2 table for meta-analysis. However, due to lack of records, no reporting of true/false positive and true/false negative statistics, and heterogeneity in other statistical metrics (e.g., correlation coefficients, regression methods, odd ratios etc.), meta-analysis could not be conducted. Instead, a qualitative synthesis approach was taken. RevMan¹¹⁷ was used to compute a Study Flow Diagram, Risk of Bias and Applicability Concerns graph and summary table.

Results

Search Results

Searches and references review resulted in 1346 records after duplicates were removed. Many studies excluded in title and abstract were not concerned with the mapping of language function in the context of surgery and were in relation to language development or language functioning in injury or disease. Three hundred and seventy-four records were included for full text review, with six studies finally included for quality assessment and data extraction. Of the full text reviews, age of population was the most common reason for exclusion with 36% of studies looking at adults only; followed by: grouping together of paediatric and adult data in analysis, which accounted for 29% of full text exclusions; lack of post-operative data (12%); utilisation of case study or case series designs (9%); epilepsy surgery was not the described population (e.g. Glioma without epilepsy) (6%); did not contain empirical data (5%); classed as grey literature (2%); and had insufficient information to establish eligibility, for

example did not report age of the participants (1%). The selection process is summarised in Figure 1.

Full-text reliability concordance

Inter-rater reliability was examined in Covidence using Cohen's kappa and demonstrated significant concordance of full-text review ($k=1$, $p=0.002$, $n=145$).

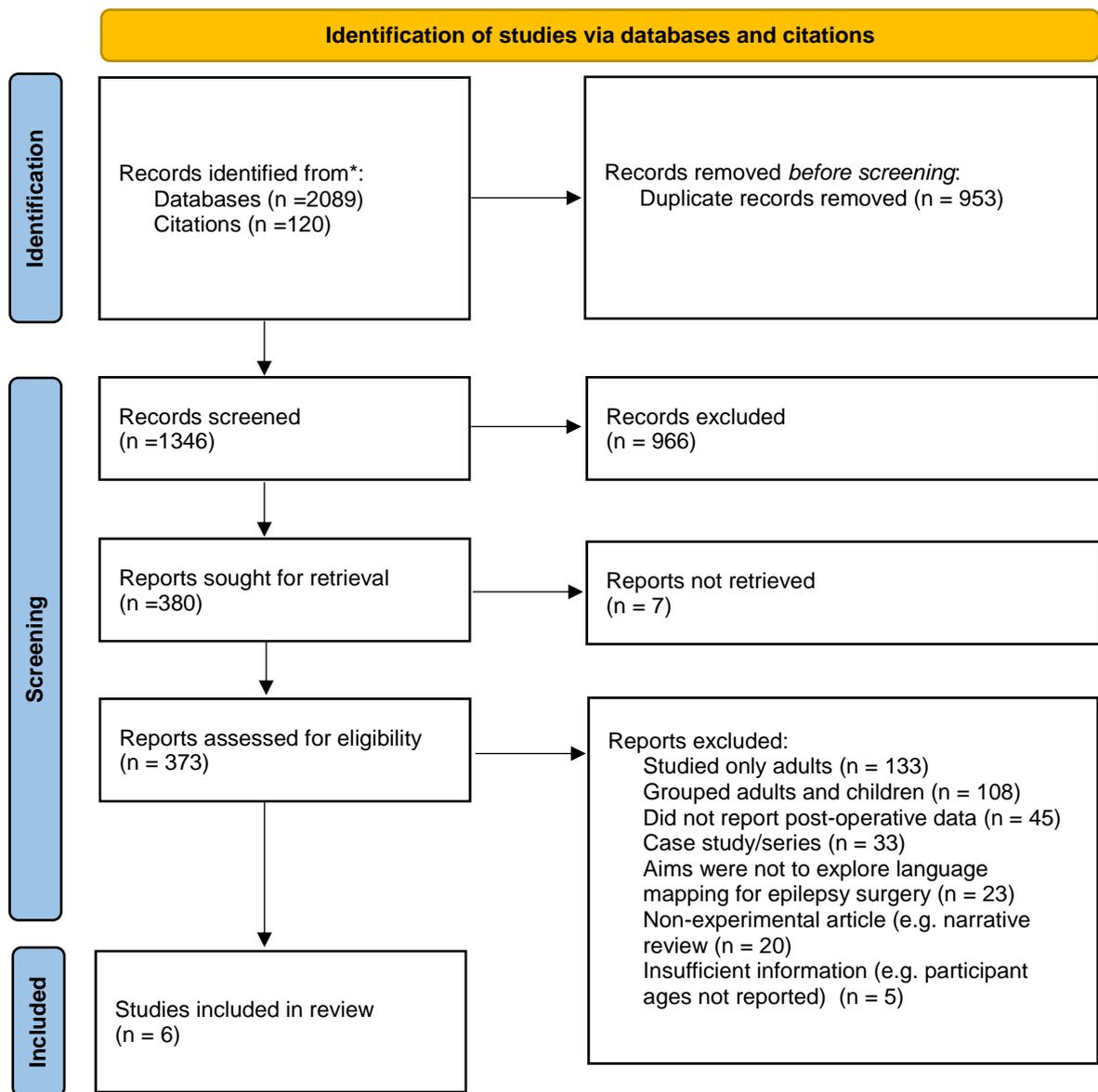


Figure 1. Study Flow Diagram summarising study selection process consistent with PRISMA guidance

Study Characteristics

Study characteristics and findings are tabulated in full (see Table 3). Of the studies meeting inclusion criteria, techniques included Diffusion Weighted Imaging (DWI), the Wada Test, and Electrocortical Stimulation Mapping (ESM). These mapping techniques are evaluated in turn.

Diffusion Weighted Imaging (DWI)

Within the six studies included, three studies related to diffusion weighted imaging^{110,118,119}. One utilised predetermined clinical mapping characteristics from electrocortical stimulation¹¹⁹. These records were identified as overlapping publications, coming from the same research group, under the same funding grant and recruited from the same hospital in the United States. The authors were contacted to establish if the cohorts were independent samples, but no response was received. As the demographics and analysis of the samples differed across studies, they were treated liberally as three independent records. Across the studies, n=308 patients were recruited, with n=98 (53 female) being included in final analysis. Ethnicity was not reported in any study. Populations all referenced children who met criteria for epilepsy surgery investigation. The mean age ranged from 7.1 to 11.1 years, with a range of 2-18 years of age. Seizure types, age of onset, baseline IQ function were not reported. One study¹¹⁸ had an even distribution of left-right seizure laterality, one¹¹⁰ had a 16:21 respective split and one¹¹⁹ did not report characteristics of seizure lateralisation. In one study¹¹⁹, baseline language function was assessed by the Clinical Evaluation of Language Function-Preschool (CELF-P) for children aged 2–5 years or the Clinical Evaluation of Language Function, Fourth Edition (CELF-4) for older children and was reported to be 84(±24)/80(±24) for expressive and receptive language, respectively.

Wada Mapping

Two studies included reported on Wada mapping^{85,109}. Across the studies, n=122 patients were recruited, with n=54 (30 female) being included in final analysis. Ethnicity was not reported. Studies were conducted in the Netherlands¹⁰⁹ and the United States⁸⁵. One study¹⁰⁹ only included complex partial seizures, the other did not report seizure type. The average age of seizure onset ranged from 5-11.1 years. The overall sample included n=41 with left lateralised seizures and n=14 right lateralised seizures. One study¹⁰⁹ reported baseline IQ average of 86, (SD=15; range 66-115) and average baseline language function, measured by verbal comprehension index, as 83 (SD=11; range 65-101).

Electrocortical Stimulation Mapping (ESM)

ESM was explored in two studies^{85,111}. Across the studies, n=369 patients were recruited, with n=138 (71 female) being included in final analysis. Ethnicity was not reported. Studies were conducted in the United States. Only one study¹¹¹ reported seizure types and included the following: Focal aware seizures (n=17); Focal seizures with impaired awareness (n=72); Focal to bilateral seizures (n=10); generalized tonic clonic (n=1); myoclonic (n=2); epileptic spasm (n=1). Across studies, 86 had left lateralised seizures, n=49 right. Notably, one study⁸⁵ only involved one right lateralised seizure focus. One study¹¹¹ reported broadly average baseline full scale IQ and verbal functioning (as measured by verbal comprehension index) as 79.7 and 86.6, respectively.

Language Mapping Tasks

Wada Test

In one study¹⁰⁹, children were asked to name five everyday objects such as spoon and toothbrush. The patient was requested to describe a picture featuring a “Cookie theft” and to execute four requests from a token test. In the other study⁸⁵, eight objects and four-line drawings were presented to the patient for naming and recall.

Electrocortical Stimulation Mapping

In one study, the children had to perform a picture naming task¹¹¹. In the other⁸⁵, at each stimulus site, the subject was asked to recite a well-known phrase, poem, or prayer (e.g., the Pledge of Allegiance) to evaluate for speech arrest. Adaptations were made for younger children who had difficulty with this task. These younger children were asked to count from 1 to 10 or requests were made to elicit spontaneous speech. The patient was also asked to name pictures or objects during stimulation. To assess comprehension, the patient was asked to complete simple phrases (e.g., “He became a clown and joined the circus”) or, in younger children, to follow one-step commands during stimulation (e.g., “point to the triangle”).

Quality Assessment

Oxford Centre for Evidence-Based Medicine Levels

For DWI, all three studies were ranked as good quality (Level 2). For Wada, one study¹⁰⁹ was ranked as good quality (Level 2), whilst the other⁸⁵ was ranked as poor quality (Level 4), due

to the lack of consistent use of standardised assessment as a reference standard. For ESM, similarly, one study¹¹¹ was ranked as good quality (Level 2), whilst the other⁸⁵ was ranked as poor quality (Level 4), due to the lack of consistent use of standardised assessment as a reference standard.

Quality Assessment with QUADAS-2

Based on predetermined criteria (see QUADAS-2 assessment template in appendix), studies were overall judged to be of ‘unclear’ methodologic quality, predominantly due to lack of clarity in reporting in reference standards (Figure 2).

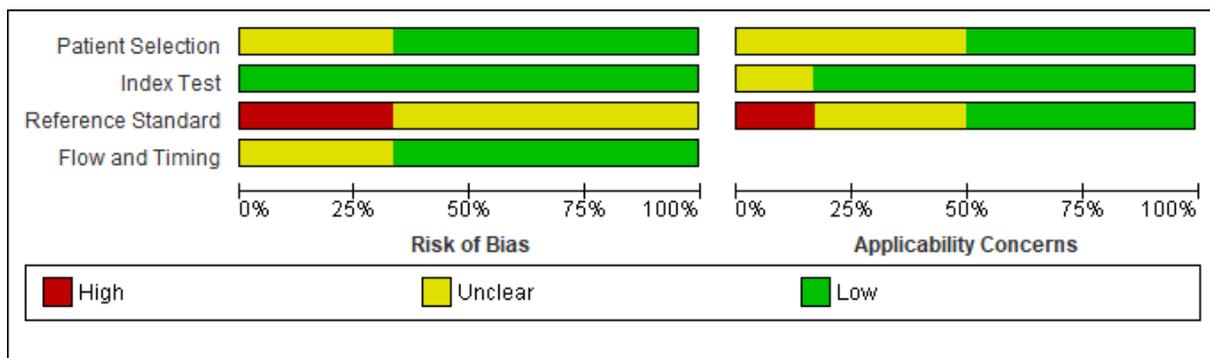


Figure 2. Risk of bias and applicability concerns graph: review of authors' judgements about each domain presented as percentages across included studies

Risk of systematic bias for patient selection was deemed to be ‘low’ in 67% of records and unclear in 33%. Concerns in applicability for patient selection were deemed to be ‘low’ in 17% of records and unclear in 83%. Notably, no studies reported ethnicity; whilst this was not deemed sufficient to change the overall rating in patient selection domain, it is noted to downgrade the quality of the records. Systematic bias in mapping technique was deemed to

be low in 100% of papers, fulfilling all the QUADAS criteria. Concerns applicability of mapping techniques in clinical practice were found to be low in all papers. Systematic bias in use of post-operative outcome assessment was found to be high in 33% of papers and unclear in 67%. Concerns of applicability of post-operative outcome assessment was found to be high in 17%, unclear in 33% and low in 50% of records. Systemic bias in the flow and timing was found to be low in 33% of papers and unclear in 67%.

Individual record quality ratings are presented in Figure 3.

Table 3. Included Study Characteristics

Study ID And Design	N (Included in final analysis)	Age of participant (years)	Sex	Seizure Types	Age of epilepsy onset (Years)	Seizure Laterality [#]	Baseline IQ	Baseline Language Function	Attrition	Mapping Technique	Language Assessment	Methods for assessing reliable change	Timing of post-op assessment	OCEBM Level of Evidence
deKoning 2009 <i>Cohort study</i>	28 (24)	Mean=11 Range 5-15	15 F 9 M	Complex partial	Average=11.1	11 L 13 R	Average = 86 SD=15 Range 66-115	VCI Average = 83 SD=11 Range 65-101	0.14	Wada	Language Tests for Dutch Children and Dutch and Bilingual Children Reynella Developmental Scales Schlichting Test of Language Production Peabody Picture Vocabulary Test	Change of scores at later session relative to scores before surgery divided by time period since test before surgery. Normalization applied because period after surgery varied from 1.8 - 2.5 years	Average = 2.1 years Range =1.8-2.5 years	2
Lee 2019 (1) <i>Cohort study</i>	96 (40)	Mean = 7.1 SD = 5.2 Range 2-17	22 F 18 M	NR*	NR*	20 L 20 R	NR*	NR*	0.58	DWI-MAP-ADFD	Not reported	Determined clinically by paediatric neurologist, physical, occupational, and speech therapists	2 weeks	2
Lee 2019 (2) <i>Cohort study</i>	84 (37)	Mean=11.1 SD=4.9 Range 2-18	20 F 20 M	NR*	NR*	16 L 21 R	NR*	NR*	0.07	DWI	Not reported	Determined clinically by paediatric neurologist and physical, occupational, and speech therapist, after surgery.	2-3 weeks	2
Lee 2020 (3) <i>Cohort study</i>	128 (21)	Mean =9.0 SD=4.9	11 F 10 M	NR*	NR*	NR*	NR*	CELF scores 84(±24)/80(±24) in preoperative expressive/receptive	0.84	NDLNA of Electrical stimulation mapping-driven diffusion MRI Tractography	CELF-P CELF-4	Change of postoperative language = (preoperative CELF score - postoperative CELF score)/preoperative CELF score. Threshold determined by clinician.	2-3 weeks	2
Sakpichaisakul 2020 <i>Cohort study</i>	275 (104)	Mean=12.2 Range 3-16	56 F 48 M	Focal aware Focal with impaired awareness	Average=6.4 Range=0.3-16	56 L 48 R	Average FSIQ=79.7	Average VIQ=86.6	0.16	ESM	VIQ of an age-appropriate version of the WPPSI WISC or WAIS	Analysis of covariance (ANCOVA) and multivariable linear regression models,	1 year	2

				Focal to bilateral GCTC Myoclonic Epileptic spasm							Spelling domain from WJ, PPVT; Verbal Learning (VLI/VLD) subtests of WRAML; Boston Naming Test	optimized using stepwise backward elimination.		
Schevon 2007	94 (30; n=21	Mean=9.9 Range 4.7-	15 F 15M	NR*	Average = 5	30 L 1 R	NR*	NR*	Wada = 71%	ESM and Wada	Neurologic examination and partial postoperative neuropsychological testing (not described)	NR*	NR*	4
Cohort study	Wada, 30 ESM)	14.9			SD=3.7 Range = 0.3-12				EMS =59%					

Abbreviations: ADFD = average direct-flip distance; Clinical Evaluation of Language Function = CELF; Clinical Evaluation of Language Function-Preschool = CELF-P; DWI = Diffusion-weighted imaging; ESM = Electrocortical Stimulation Mapping; Full Scale Intelligence Quotient = FSIQ; MAP = maximum a posteriori probability; Generalised Tonic Clonic = GCTC; NDINA = Novel Deep Learning Network Analysis; Verbal Intelligence Quotient = VIQ; Verbal Comprehension Index = VCI; Wechsler Preschool and Primary Scale of intelligence, 3rd/4th edition = WPPSI; Wechsler Intelligence Scale for children, 4th/5th edition = WISC; Wechsler Adult Intelligence Scale, 4th edition = WAIS; The Woodcock-Johnson Tests of Achievement, 3rd edition = WJ; Peabody Picture Vocabulary Test, 4th edition = PPVT; Wide Range Assessment of Memory and Learning, 2nd edition = WRAML; *NR = Not reported; % M = Male, F = Female; # L = Left, R = Right

	<u>Risk of Bias</u>				<u>Applicability Concerns</u>		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
de Koning,	?	+	?	+	+	+	+
Lee (1)	+	+	?	+	?	+	?
Lee (2)	+	+	?	+	?	+	?
Lee (3)	+	+	-	?	?	+	+
Sakpichaisakul,	+	+	?	+	+	+	+
Schevon,	?	+	-	?	+	?	-

 High	 Unclear	 Low
--	---	---

Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study. Lee (1)=Ref¹¹⁸; Lee (2)=Ref¹¹⁰; Lee (3)=Ref¹¹⁹

Diffusion Weighted Imaging

Overall, DWI studies were rated as unclear quality, predominantly due to lack of clarity of reporting of reference standards. There was partial reporting in all three studies about patient characteristics. Notably, seizure type, age of onset of epilepsy and baseline cognitive function were absent. Two studies did not describe the method of pre- and post-operative language assessment and their methodology for evaluating change^{110,118}. The description of post-operative assessment was limited to “clinically determined by a paediatric neurologist, and physical, occupational, and speech therapists”^{110,118}. Only one study¹¹⁹ described in detail the assessments used for pre- and post-operative assessment of language. The assessment of language function was conducted using the CELF, which is currently recommended as the

'gold-standard' in the UK¹²⁰. The methodology for assessing post-operative outcome in this study did not take account of many known variables that impact comparison of assessment scores across time in the individual case, such as practice effects, test-retest reliability, and regression to the mean¹²¹; therefore, could be vulnerable to several biases and was graded at high risk. Only one¹¹⁰ study mentioned blinding of index and reference standard. Lastly, one study¹¹⁹ provided insufficient detail about the timing of assessments, surgery, and outcome data to make quality decisions about flow bias.

Wada Test

Overall, the quality status for Wada is unclear. Studies^{85,109} did not provide reasons for exclusions of participants, therefore systematic bias cannot be assessed. Although both quality reviewers agreed this was not sufficient to regrade as high bias for applicability, one study¹⁰⁹ had an elevated number of representative cases of developmental pathology cases compared to the established demographic in the literature^{2,122,123}. One study⁸⁵ had insufficient reporting of post-operative methodology (i.e., did not report: neuropsychological tests used; methods for assessing difference between pre- and post-operative language; the thresholds used for determining language decline; or post-operative language score data). Post-operative assessment was inconsistent with only eight patients undergoing partial neuropsychology⁸⁵. Additionally, post-operative assessment was not applicable to routine clinical practice as it was based predominantly on neurologic examination⁸⁵ not neuropsychological assessment (as is routine practice^{120,124}); therefore, it was rated as high risk of bias. Timing of post-operative outcome was not reported in one study⁸⁵, therefore flow quality assessment was incomplete.

Electrocortical Stimulation Mapping

Of the two papers included, quality assessment was deemed to be unclear in one record¹¹¹ and poor in the other⁸⁵. One study⁸⁵ did not provide reasons for exclusions of participants, preventing full assessment of systematic bias. One study¹¹¹ reported the neuropsychologist who conducted the post-operative assessment was not blinded; however, due to the clearly defined statistical analysis, thresholding and use of standardised assessment in the methodology, both reviewers agreed this was unlikely to introduce significant bias. One study⁸⁵ had insufficient reporting of post-operative methodology (i.e., no reporting of neuropsychological tests used, methods for assessing difference between pre- and post-operative language, for the thresholds used for determining language decline, or of post-operative neuropsychology data). Additionally, post-operative assessment was inconsistent with only eight patients undergoing 'partial' neuropsychology⁸⁵. Lastly, post-operative assessment was not applicable to routine clinical practice as it was based predominantly on neurologic examination⁸⁵ not neuropsychological assessment, as is routine practice^{120,124}; therefore, was rated as high risk of bias.

Attrition and Failure Rates

For the DWI studies, attrition rates ranged from 16-84%. In the overall sample (n=308), 68% were excluded due to incomplete scan data. Of reported data (n=44), 16% were excluded due to poor quality scans.

For Wada, attrition rates ranged from 14-77%. Of the total sample, 63% patients were excluded for missing or unavailable patient data (i.e., Wada was not conducted,

neuropsychological assessment was not complete, patients did not attend follow up appointments etc). Of available data (n=21), 17% were not successfully lateralized with the Wada test. For reported records (n=21), there was a statistically significant difference (χ^2 test, $p < 0.05$) in the success rate of language lateralization by Wada testing in children split between younger and older than 10.2 years of age⁸⁵. More specifically, in children >10.2 years of age (n=7), 43% were not successfully lateralized versus 81% of children <10.2 years of age (n=16)⁸⁵.

For ESM, attrition rates ranged from 16-59%. Of total sample (n=369), 64% patients were excluded for missing or unavailable patient data (i.e., ESM was not conducted, neuropsychological assessment was not complete etc). Of available data (n=30), ESM failed to map language in 48% of patients.

Predictive Validity Findings

Diffusion Weighted Imaging

All studies reported DWI predictors were significantly associated with postoperative language deficit, with the correlation strength ranging from moderate to strong (n=98, r range= 0.62-0.91, $p < 0.05$)^{110,118,119}. In available records (n=61)^{118,119}, sensitivity and specificity data for DWI in accurate prediction of post-operative language decline both ranged from 69 to 100%, dependent on thresholds used for conservation of tracts. Of note, one study¹¹⁸ reported DWI prediction alone outperformed DWI prediction with integrated clinical variables (such as age and pathology), with average sensitivity and specificity of post-operative language prediction moving from 87%, when independent of clinical variables, down to 78%, when clinical

variables were incorporated into regression. There was consensus in the findings that non-resection of the defined functional language tracts utilising DWI accurately predicted the absence of postoperative language deficits, as defined within the studies^{110,118,119}.

Wada Test

Findings related to Wada testing were very limited in the available records^{85,109}, as outcome prediction was included only in these papers' secondary aims. Consequently, studies did not report sensitivity or specificity data; however, these were able to be extracted from one study's¹⁰⁹ data table¹ and computed (Table 4). Computed data are also graphically illustrated within a receiver operating characteristic (ROC) curve (Figure 4). When poor post-operative outcome is considered to include language delay, arrest or deterioration two years after surgery, the Wada demonstrates good sensitivity and moderate specificity (labelled as low threshold in Table 4). However, where poor post-operative outcome is considered only to be language deterioration (labelled as high threshold in Table 4), the Wada's ability to predict outcome was highly variable across all language domains, with generally poor ability (\bar{x} 0.27). There was a significant ($p < 0.05$) association between the presence of post-operative expressive language delay (over a two-year period) and temporal lobectomy in the hemisphere to which the Wada lateralised language¹⁰⁹. This association was not found to be true for receptive development of language. For those children in which the Wada lateralised language to the surgical hemisphere, language development was inordinately slow or arrested¹⁰⁹. Schevon et al.'s⁸⁵ findings are somewhat inconsistent, in that they found no post-operative language decline when Wada (alongside

¹de Koning et al.¹⁰⁹: Table 4 in article. Titled: Change in LAdif (language age minus chronological age) from before to 2 years after surgery.

clinical characteristics and ESM) contributed to surgical decision making. However, no detail was provided in the way it contributed to decision making and resection.

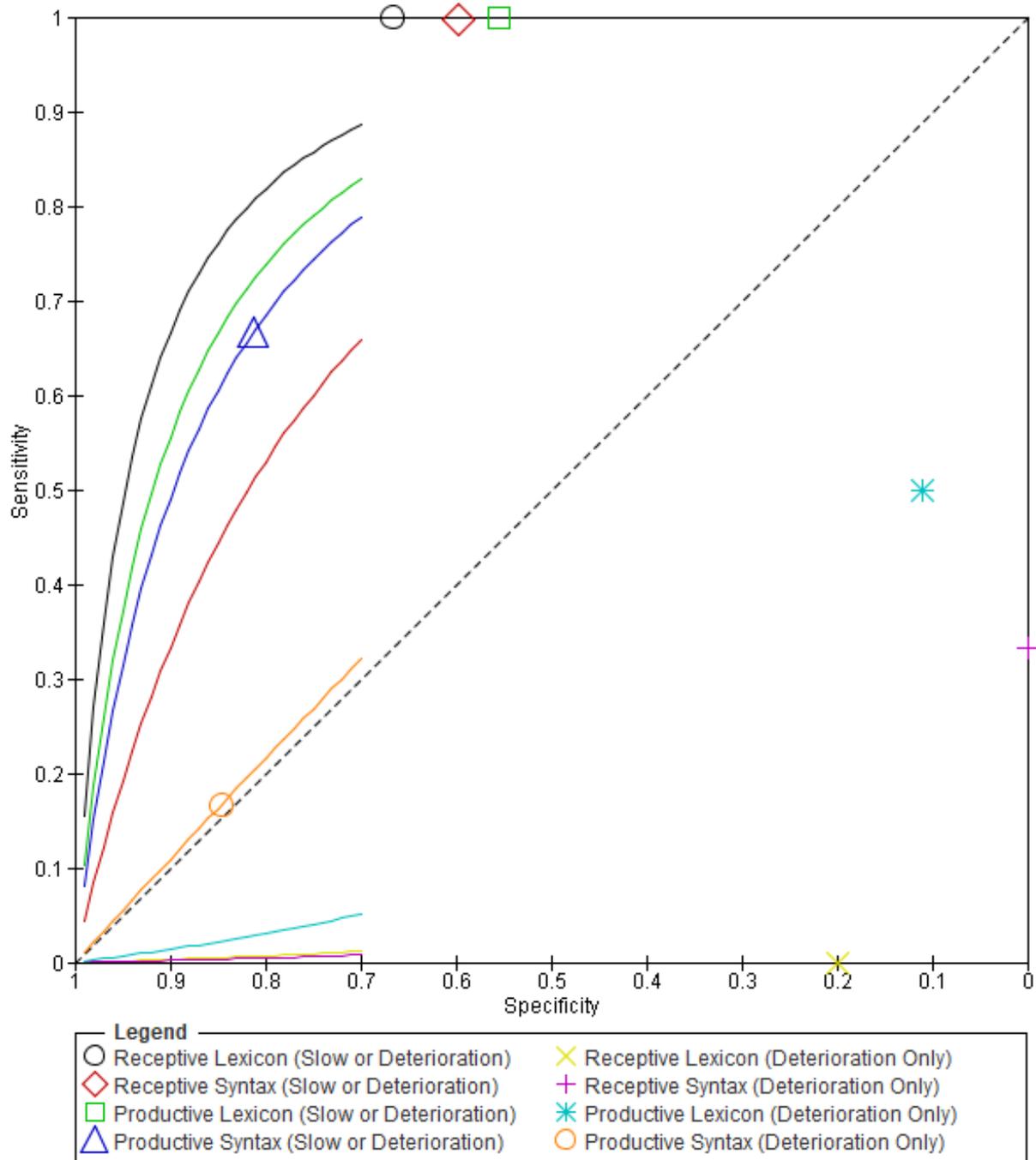


Figure 4. ROC curve analysis of Wada sensitivity and specificity ratings extracted from de Koning et al.¹⁰⁹.

Table 4. Sensitivity and Specificity data extracted from de Koning et al.¹⁰⁹

	Language Domain	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Low Threshold	Receptive Lexicon	4	5	0	10	1.00 [0.40, 1.00]	0.67 [0.38, 0.88]
	Receptive Syntax	1	6	0	9	1.00 [0.03, 1.00]	0.60 [0.32, 0.84]
	Productive Lexicon	4	8	0	10	1.00 [0.40, 1.00]	0.56 [0.31, 0.78]
	Productive Syntax	2	3	1	13	0.67 [0.09, 0.99]	0.81 [0.54, 0.96]
High Threshold	Receptive Lexicon	0	12	4	3	0.00 [0.00, 0.60]	0.20 [0.04, 0.48]
	Receptive Syntax	1	13	2	0	0.33 [0.01, 0.91]	0.00 [0.00, 0.25]
	Productive Lexicon	2	16	2	2	0.50 [0.07, 0.93]	0.11 [0.01, 0.35]
	Productive Syntax	1	2	5	11	0.17 [0.00, 0.64]	0.85 [0.55, 0.98]

TP = True Positive, FP = False Positive, FN = False Negative, TN = True Negative. Low threshold was set as patients who were reported to either decline or demonstrate slowed development post-operatively; high threshold was set as patients who were only reported to decline.

Electrocortical Stimulation Mapping

One study¹¹¹ reported ESM increased the odds of an improved post-surgical language outcome at 1-year by 1.85 times. However, there was a noted difference of age and ESM's ability to determine outcome. In children younger than five, no added effect of language ESM (over no ESM) was seen on post-surgical outcome scores; whereas a significant effect ($p < 0.05$) was seen across language (naming ability), IQ and memory outcome scores for children over five years of age. In children over five years, there was a significantly increased rate ($p = 0.001$) of post-operative decline in general cognitive function for those who did not undergo ESM language mapping compared with those who did; 68% versus 30% rates of decline,

respectively. Notably this effect was more pronounced in right hemisphere resections, where decline in overall cognitive scores was seen in 76% patients who did not have ESM compared to 25% that did. This finding may be an artefact of cognitive skill. Children with more severe epilepsy and poorer cognitive abilities maybe less likely to find acceptable ESM procedures and may have worse outcomes as a consequence of these background factors, not ESM procedures.

Of note, those children who had ESM showed an improvement in general cognition (+1.2 standard deviation); whereas those without ESM showed a decline (-1.5 standard deviation). Schevon et al.'s⁸⁵ findings are consistent, in that they found no post-operative language decline when ESM (alongside clinical characteristics, and in a subset of patients, Wada testing) contributed to surgical decision making. However, no detail was provided in the way it contributed to decision

Comparative Utility Synthesis

Overall, within the studies included, the strongest evidence exists for diffusion weighted imaging in predicting post-operative language deficits, with synthesis of findings across three studies demonstrating a significant predictive relationship between defined language-related tracts and post-operative deficits. Additionally, successful DWI-mapping rates were as high as 93%. However, the quality of DWI evidence was rated to be overall 'unclear', with one study¹¹⁹ having a high rating of risk of bias with respect to its reference standard. The quality of the ESM and Wada records was variable with one study⁸⁵ demonstrating high risk of bias for the reference standard and unclear bias about the use of patient populations. Similar to DWI, ESM-mapped language areas demonstrated robust relationships with post-operative

outcome¹¹¹; however, ESM success was much more modest, with a 52% success rate⁸⁵.
Extracted and analysed data from Wada studies demonstrated variable outcomes depending on threshold selected and strong conclusions were again compromised due to the unclear¹⁰⁹ and poor⁸⁵ quality of the studies.

Discussion

This systematic review aimed to identify the current extent of group level studies that have explored the predictive validity of language mapping and assess their quality. It also aimed to provide an evidence synthesis of individual and comparative validity for mapping techniques in predicting post-operative outcome; and aimed to suggest guidance for further research and clinical practice, where applicable. As a secondary aim, acceptability of language mapping techniques in paediatric populations was investigated.

Extent of Evidence

Despite an increasing number of relevant indexes on MEDLINE⁵⁸, this review found the extent of robust evidence examining the predictive validity of language mapping in childhood epilepsy surgery is limited. No studies for Positron Emission Tomography, Functional Magnetic Resonance Imaging, Magnetoencephalography, Transcranial Magnetic Stimulation, Electroencephalography or Functional Near Infra-red spectroscopy mapping techniques met inclusion for this review. This was most commonly due to grouping of adults and children in analysis or case series design. This finding demonstrates there is currently limited evidence at level 3 or above, as evaluated by the Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence, for these techniques to predict language outcome in paediatric epilepsy surgery populations. Of those techniques that met criteria for inclusion (DWI, Wada and ESM), there were a very limited number of records, failing minimum recommended criteria for a meta-analysis¹¹⁶.

There were a number of case studies^{101–106,125}^{101–106,125} and group-level reports (where adults and children were grouped in the analysis^{36,52,69,126–143}¹⁰⁸). There are several advantages of utilising case studies. Case studies allow for in-depth, multi-factorial investigation of complex cases and clinical questions in a ‘real-life’ setting (Crowe et al, 2011). Case studies offer an opportunity to provide in-depth insights into aspects of a case and this detailed understanding can be used to generate hypotheses that may be relevant to broader clinical practice. Therefore, case studies have significant implications for clinical theory development and testing (Crowe et al., 2011). Case studies can offer unique insights into care delivery and any issues relevant to complex or idiosyncratic cases within the application of practice. This is particularly useful in heterogenous clinical groups, such as those within epilepsy. However, there are limitations to the generalisability of case studies to inform routine practice. Routine guidance is applied to groups who will have more varied presentations and contexts than that of an individual (or even a series of individuals) presented in a case study. Therefore, there may be factors that prevent the demonstrated efficacy to generalise to this wider group. It is considered scientifically and clinically inappropriate to take what has been shown to work for one person or a small series of individuals and apply it to a diverse group without further validation of the diagnostic or prognostic methodology works for the wider clinical group. Following generation of a clinical theory and practice method from case study data, the next step is to systematically explore if these principles are effective at a group level, for example through phased trials. This allows for their adoption into routine guidance. Therefore, this review solely focussed on this stage of evidence for routine guidance development. Once guidance is established, case studies can be used to validate their utility in-depth at the individual case level and offer ecologically valid assessment and future research directions.

The lack of findings for records examining cognitive outcome data is consistent with previous systematic reviews⁶, which found after screening 8189 titles and 127 full text papers, only five studies had explored post-operative cognitive outcomes at a group level. This highlights a significant gap in our understanding of the cognitive outcomes of paediatric epilepsy surgery more broadly. The lack of available data may reflect the recency of attention in this topic, following the reclassification of epilepsy to include the cognitive effects of seizures¹⁴⁴, as well as the shift in broader operative outcome consideration to include cognition⁵⁻⁷, following the recent rise in paediatric procedures¹⁻⁴.

Quality of Evidence

The OCEBM level-based ratings found indicated that the evidence is of overall good quality; however, on more detailed analysis by QUADAS-2 assessment, quality assurance was unclear. There was a lack of reporting of key demographic variables within some studies. Notably, seizure type, age of onset of epilepsy, baseline cognitive and language function were not consistently reported. This information is not only relevant to ensure generalisability of findings to clinical practice, but also is key in considering likelihood of atypical language representation and language localisation within the population studied^{36,37,42,45-53}. As mentioned in the introduction, epilepsy surgery candidacy is a diverse population, with a range of ages, seizure presentations, and onset ages^{2,122,123}. Research on a homogenous subsection of this group may not generalise well to the heterogenous population. Furthermore, age of onset, baseline function and seizure characteristics all have bearing on language outcome^{52,53}; therefore, are essential contextual information and any biases in these factors would limit the conclusions of the study. Of the two included records that did

report comprehensive clinical variables, predictive associations were found between these factors and language localisation^{111,119}. It is important for research to ensure mapping techniques add information above and beyond background clinical variables^{53-55,119}. One notable weakness of the existing research is the lack of reporting of ethnicity within research. The reporting and use of references standards were the greatest quality concern within the review. There was a noted lack of description of methodology for assessing pre- and post-operative language function, for determining reliable change across the two time points and determining the threshold for post-operative decline. This absence of reporting not only prevents quality appraisal¹¹² and meta-analysis¹¹⁶ but also study replication, adding to the evidence crisis within clinical psychology and neurosciences^{145,146}.

Evidence for Predictive Validity

This review demonstrates limited but overall moderate quality evidence for the use of DWI in predicting language outcome from paediatric epilepsy surgery. All records found a significant relationship with post-operative outcome. The findings of this review are consistent with previous systematic review¹⁰⁷, which focused predominantly on case studies and mechanistic evidence. Furthermore, based on attrition data, DWI appears to be clinically feasible and successful. DWI demonstrated only a 16% failure rate and regularly took place in routine clinical practice, with around 1 in 3 patients enrolled in our samples' resective surgery programmes undergoing DWI. Despite DWI's promise, there are limitations that need to be addressed before adopting into routine practice. It is unclear whether there is overlap between the record's samples, limiting the strength of conclusions drawn from the meta-synthesis. Additionally, the partial reporting of patient characteristics (e.g., seizure type, age

of onset of epilepsy and baseline cognitive function) limits generalisability to routine practice. Lastly, the lack of clear reference standard assessment and the timing of a post-operative assessment at two-weeks limits confidence in the validity of the data. Further work needs to be conducted with additional patient samples, longer term follow-up and operationalised outcome assessment before DWI can be adopted into routine practice. Although noted to be beyond the scope of this review, it is important to note the tractography methods were heterogenous across studies and warrant further exploration.

This study found evidence for Wada's predictive validity to be of unclear quality and applicability. The findings demonstrate variability in predictive validity outcome depending on the thresholds that are applied. When determining poor language outcome in the broad sense (i.e., any slowed or arrested development post-operatively), Wada was found to have good sensitivity (~ 0.92) and reasonable specificity (~ 0.66). This is generally consistent with reports in adults^{28,78}. However, the significantly small sample size and large confidence intervals on accuracy estimations warrant significant caution in generalising these findings to routine practice. Additionally, when a more conservative threshold of predicting post-operative decline (without considering slow language development an operative outcome) is applied, the Wada's predictive ability falls significantly (~ 0.27). Furthermore, there was a significant absence of reporting in terms of outcome variables that may have confounded post-operative language function, such as: post-operative seizure control, ongoing use of anti-epileptic medications, rehabilitation etc. All these factors limit long-term post-operative outcome conclusions. Lastly, consistent with clinical practice observations, the Wada is not well tolerated accepted by children, with failure rates of 43% in over 10s and 81% in children under tens. Therefore, there is limited evidence for the use of Wada in routine practice to

predict language outcome in paediatric surgery. The Wada procedure comes with significant clinical risk and in adult practice has largely been replaced by functional MRI^{28,147}.

With respect to ESM, this review indicated variable quality of evidence. One record⁸⁵ was rated poor quality (Level 4) with high risk of bias and applicability concerns. Therefore, the outcomes of this study should not be considered for routine practice. The second study¹¹¹ was rated to be good quality (Level 2) evidence and had low risk of bias and applicability concern. This study reported odds of an improved post-surgical language outcome at 1-year by 1.85 times when ESM was utilised in surgical decision making and demonstrated a strong relationship with language (and IQ & memory) in over five-year-olds. However, this effect did not extend to under five-year-olds. It is important to note, the outcome reference standards did differ by age, with those under five having their language assessment scores incorporated into a wider intellectual component. This decision was based on a principal component analysis designed to mitigate the impact of autocorrelation seen in cognitive assessment scores. In those over five, language naming was found to be an independent cognitive component and so was explored as a separate outcome. Through this approach of attempting to mitigate the inter-dependencies of cognitive assessment, sensitivity to detect language change may have been lost with additional variance introduced by other cognitive domains. Additionally, there was a notably smaller sample (n=15) of under-fives, compared to over-fives (n=89). Therefore, limited conclusions can be drawn about language outcome in the under 5s sample. The findings of robust predictive value of ESM for language are inconsistent with adult findings, that demonstrate ESM does not prevent post-operative naming deficit¹⁴⁸⁻¹⁵⁰. Evidence has shown that the risk of naming deficit increases with age of onset¹⁴⁹, and so younger children may have underlying neuroplastic mechanisms that afford the greater post-operative protection than adults, which may explain the difference in outcome. The authors

offer their superior sample size as an alternate explanation between their findings and the adult literature¹¹¹. One other explanation may be related to bias in the sample selection. Children with more severe epilepsy and poorer cognitive abilities maybe less likely to tolerate find ESM procedures acceptable and may have worse outcomes because of these background factors, not lack of ESM procedures. Prospective evidence with blind group allocation is required to overcome this. Overall, there is moderate quality evidence of ESM. Although promising results are seen, the failure rate of ESM was noted to be 48% in our sample. Further prospective studies are required before it can be recommended for routine practice in predicting post-operative outcome. It would be beneficial for these studies to consider the acceptability of the procedure and blinded group allocation.

It is important to note, the comparison of DWI and ESM is useful in addressing the specific comparative utility of DWI in those discrete cases warranting invasive cortical mapping (ESM). However, it is possibility that DWI has more ubiquitous utility in a broader population and therefore its utility should be judged solely with reference to ESM.

Strengths and Limitations of this Review

A notable strength of this review is that it is the first systematic review to explore all modalities of language mapping in paediatric populations and their predictive validity. It has identified key gaps in the evidence base both for cognitive outcomes in paediatric surgery and the validation of current language mapping techniques in this setting. This has allowed for important recommendations to be made for future research. A second advantage this review holds is the use of two robust and widely used quality tools. The adherence to published standards for protocol development and reporting also afford credibility^{114,116}.

The review holds several methodological limitations. Notably, the lack of comprehensive second review on all records in screening due to resource constraints. However, quality checks were adopted to limit potential bias and second review on 38% of full text records demonstrated 100% concordance. Furthermore, data was not extracted or cross-referenced by a second reviewer, which increases the chance of bias. These factors are somewhat mitigated by the use of a standardised data extraction template.

This review did not consider seizure freedom or baseline seizure severity alongside language preservation. A key aspect of the surgical decision making is balancing the risk of eloquent cortex (i.e., language) against sufficiently disrupting the epileptogenic network through resection. In certain cases, the decision to risk language disruption is warranted in the context of severe and debilitating seizures. Therefore, consideration of seizure burden and post-operative seizure freedom should be reported to provide context to any potential post-operative language risk taking.

Lastly, though it was noted to be beyond scope of this review, the precise methodological considerations of mapping techniques vary considerably and may impact post-operative prediction. Consideration of the optimal mapping methodologies (e.g., current strength, language stimulation task etc) should be undertaken before adopting any mapping techniques into practice.

Recommendations for Future Research

A significant focus of future research should be improving the evidence base with respect to cognitive outcome data of paediatric epilepsy surgery. The success of other clinical databases, such as ABIDE¹⁵¹, have demonstrated the benefit of national professional networks in

populating open-source databases for research. Existing patient databases, such as Orion, held by the National Children's Epilepsy Surgery Network, that contain pre- and post-operative neuropsychological and comprehensive clinical demographic data should be considered in future research applications. This would enable retrospective cohort studies to take place and offer techniques (often used clinically on a case-by-case basis) validation at the group level, enabling recommendations to be made for routine practice.

Research should seek to include key demographic information in reporting, including ethnicity. There is a noted bias generally in clinical research towards white males; research fails to represent other ethnic backgrounds despite an overrepresentation of minority groups within health services, when compared to the general population¹⁵². It should ensure the inclusion of key clinical characteristics and ensure that mapping methodologies are adding value above and beyond what is readily available.

Further studies should involve a clear rationale and reporting of reference standards. Formal standardised assessment by an appropriately qualified neuropsychologist would be considered 'gold standard'^{153,154}. Comprehensive assessment of more than one language domain is also recommended¹⁰⁹ and would be adequately met by clinical tools, such as the CELF. Analysis at the individual case level would be the most clinically relevant as this has the highest application to routine practice. There are limitations to group analysis of predictive validity as these studies may benefit from the additional discriminant power afforded by larger numbers and improved noise reduction in variance^{53-55,121}. These may not be replicable in a clinical setting, where post-operative predictions are being made in the individual case⁵³⁻⁵⁵. Outcome thresholds should be clearly justified, and cognitive outcome assessment should account for intrasubject variability (e.g., developmental trajectories), variability of instrument

(i.e. test-retest, practice effects), and statistical considerations (e.g. regression to the mean)^{121,155}.

Lastly, research should attempt to simultaneously investigate multimodal language mapping in prospective cohorts to best ascertain which, standalone or combined, methodologies offer superior predictive validity in the mapping of language and minimising post-operative decline in paediatric epilepsy surgery.

Conclusion

This review has found a significant need for further investigation of language mapping tools in predicting post-operative outcome. Preliminary evidence demonstrates promising value of DWI and ESM within paediatric practice, demonstrating different language mapping methods are ~~tolerable~~ acceptable and may be successful in predicting outcome in paediatric populations. However, further work needs to be undertaken before these techniques are adapted into routine practice as the quality of the evidence is 'unclear' due to lack of reporting of reference standard methodologies. In addition, other promising mapping techniques (such as fMRI) should be further investigated due to the successful translation of the aforementioned modalities into paediatric populations from the adult literature.

Disclosures

Alexander Marsh has no conflicts of interest to disclose. Alexander Marsh confirms that he has read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

1. Braun, K. P. J. Influence of epilepsy surgery on developmental outcomes in children. *European Journal of Paediatric Neurology* **24**, 40–42 (2020).
2. Lamberink, H. J., Boshuisen, K., Rijen, P. C. van, Gosselaar, P. H. & Braun, K. P. J. Changing profiles of pediatric epilepsy surgery candidates over time: A nationwide single-center experience from 1990 to 2011. *Epilepsia* **56**, 717–725 (2015).
3. Baud, M. O. *et al.* European trends in epilepsy surgery. *Neurology* **91**, e96–e106 (2018).
4. Knight, E. M. P., Schiltz, N. K. & Bakaki, P. M. Increasing utilization of pediatric epilepsy surgery in the United States between 1997 and 2009. *Epilepsia* **56**, 375–381 (2015).
5. Hoppe, C. *et al.* Parents' view of the cognitive outcome one year after pediatric epilepsy surgery. *Epilepsy & Behavior* **101**, 106552 (2019).
6. Flint, A. E. *et al.* Neuropsychological outcomes following paediatric temporal lobe surgery for epilepsies: Evidence from a systematic review. *Seizure* **52**, 89–116 (2017).
7. Helmstaedter, C. *et al.* Cognitive outcome of pediatric epilepsy surgery across ages and different types of surgeries: A monocentric 1-year follow-up study in 306 patients of school age. *Seizure* **77**, 86–92 (2020).
8. Wrench, J. M., Matsumoto, R., Inoue, Y. & Wilson, S. J. Current challenges in the practice of epilepsy surgery. *Epilepsy Behav* **22**, 23–31 (2011).
9. Osipowicz, K., Sperling, M. R., Sharan, A. D. & Tracy, J. I. Functional MRI, resting state fMRI, and DTI for predicting verbal fluency outcome following resective surgery for temporal lobe epilepsy. *Journal of Neurosurgery* **124**, 929–937 (2016).
10. Sherman, E. M. S. *et al.* Neuropsychological outcomes after epilepsy surgery: Systematic review and pooled estimates. *Epilepsia* **52**, 857–869 (2011).
11. Puka, K., Tavares, T. P. & Speechley, K. N. Social outcomes for adults with a history of childhood-onset epilepsy: A systematic review and meta-analysis. *Epilepsy & Behavior* **92**, 297–305 (2019).
12. Baumer, F. M., Cardon, A. L. & Porter, B. E. Language Dysfunction in Pediatric Epilepsy. *The Journal of Pediatrics* **194**, 13–21 (2018).
13. Overvliet, G. M. *et al.* Nocturnal epileptiform EEG discharges, nocturnal epileptic seizures, and language impairments in children: Review of the literature. *Epilepsy Behav* **19**, 550–558 (2010).
14. Overvliet, G. M. *et al.* Early onset of cortical thinning in children with rolandic epilepsy. *Neuroimage Clin* **2**, 434–439 (2013).

15. Lendt, M., Helmstaedter, C. & Elger, C. E. Pre- and postoperative neuropsychological profiles in children and adolescents with temporal lobe epilepsy. *Epilepsia* **40**, 1543–1550 (1999).
16. Jambaqué, I. *et al.* Memory functions following surgery for temporal lobe epilepsy in children. *Neuropsychologia* **45**, 2850–2862 (2007).
17. Bulteau, C. *et al.* Language recovery after left hemispherotomy for Rasmussen encephalitis. *Epilepsy & Behavior* 51–57 (2015).
18. Vega, C. *et al.* Lexical retrieval pre- and posttemporal lobe epilepsy surgery in a pediatric sample. *Epilepsy & Behavior* **1**, 61–65 (2015).
19. Trimmel, K. *et al.* Naming fMRI predicts the effect of temporal lobe resection on language decline. *Annals of Clinical & Translational Neurology* **6**, 2186–2196 (2019).
20. Audrain, S., Barnett, A. J. & McAndrews, M. P. Language network measures at rest indicate individual differences in naming decline after anterior temporal lobe resection. *Human Brain Mapping* **39**, 4404–4419 (2018).
21. You, X. *et al.* fMRI prediction of naming change after adult temporal lobe epilepsy surgery: Activation matters. *Epilepsia* **60**, 527–538 (2019).
22. Schwarz, M., Pauli, E. & Stefan, H. Model based prognosis of postoperative object naming in left temporal lobe epilepsy. *Seizure* **14**, 562–568 (2005).
23. Dym, R. J., Burns, J., Freeman, K. & Lipton, M. L. Is functional MR imaging assessment of hemispheric language dominance as good as the Wada test?: a meta-analysis. [Review]. *Radiology* **261**, 446–455 (2011).
24. Pujol, J., Deus, J., Losilla, J. M. & Capdevila, A. Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology* **52**, 1038–1043 (1999).
25. Rasmussen, T. & Milner, B. The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Sciences* **1**, 355–369 (1977).
26. Springer, J. A. *et al.* Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Journal of Neurology* **122**, 2033–2046 (1999).
27. Szaflarski, J. P. *et al.* Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology* **59**, 238–244 (2002).
28. Szaflarski, J. P. *et al.* Practice guideline summary: Use of fMRI in the presurgical evaluation of patients with epilepsy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* **88**, 395–402 (2017).
29. Woods, R. P., Dodrill, C. B. & Ojemann, G. A. Brain injury, handedness, and speech lateralization in a series of amobarbital studies. *Ann Neurol* **23**, 510–518 (1988).

30. You, X. *et al.* A decisional space for fMRI pattern separation using the principal component analysis-A comparative study of language networks in pediatric epilepsy. *Human Brain Mapping* **34**, 2330–2342 (2013).
31. Binder, J. R. *et al.* Determination of language dominance using functional MRI: A comparison with the Wada test. *Neurology* **46**, 978–984 (1996).
32. Wellmer, J. *et al.* Unilateral intracarotid amobarbital procedure for language lateralization. *Epilepsia* **46**, 1764–1772 (2005).
33. Fernández, G. *et al.* Language mapping in less than 15 minutes: real-time functional MRI during routine clinical investigation. *Neuroimage* **14**, 585–594 (2001).
34. Frost, J. A. *et al.* Language processing is strongly left lateralized in both sexes. Evidence from functional MRI. *Brain* **122 (Pt 2)**, 199–208 (1999).
35. Conradi, N. *et al.* Hemispheric language lateralization in presurgical patients with temporal lobe epilepsy: Improving the retest reliability of functional transcranial Doppler sonography. *Epilepsy & Behavior* 48–52 (2019).
36. Gaillard, W. D. *et al.* fMRI language task panel improves determination of language dominance. *Neurology* **63**, 1403–1408 (2004).
37. Gaillard, W. D. *et al.* Atypical language in lesional and nonlesional complex partial epilepsy. *Neurology* **69**, 1761–1771 (2007).
38. Gaillard, W. D. *et al.* Language dominance in partial epilepsy patients identified with an fMRI reading task. *Neurology* **59**, 256–265 (2002).
39. Ramsey, N. F., Sommer, I. E. C., Rutten, G. J. & Kahn, R. S. Combined Analysis of Language Tasks in fMRI Improves Assessment of Hemispheric Dominance for Language Functions in Individual Subjects. *NeuroImage* **13**, 719–733 (2001).
40. Spreer, J. & Quiske, A. Unsuspected Atypical Hemispheric Dominance for Language as Determined by fMRI: Reply. *Epilepsia* **44**, 734–735 (2003).
41. Spreer, J. *et al.* Determination of hemisphere dominance for language: comparison of frontal and temporal fMRI activation with intracarotid amytal testing. *Neuroradiology* **44**, 467–474 (2002).
42. Woermann, F. G. *et al.* Language lateralization by Wada test and fMRI in 100 patients with epilepsy. *Neurology* **61**, 699–701 (2003).
43. Yuan, W. *et al.* fMRI Shows Atypical Language Lateralization in Pediatric Epilepsy Patients. *Epilepsia* **47**, 593–600 (2006).
44. Hadac, J., Brozova, K., Tintera, J. & Krsek, P. Language lateralization in children with pre and postnatal epileptogenic lesions of the left hemisphere: An fMRI study. *Epileptic Disorders* S19–S27 (2007).

45. You, X. *et al.* Sub-patterns of language network reorganization in pediatric localization related epilepsy: a multisite study. *Hum Brain Mapp* **32**, 784–799 (2011).
46. Binder, D. K. & Schramm, J. Transylvian functional hemispherectomy. *Child's Nervous System* **22**, 960–966 (2006).
47. Rosenberger, L. R. *et al.* Interhemispheric and intrahemispheric language reorganization in complex partial epilepsy. *Neurology* **72**, 1830–1860 (2009).
48. Wilke, M. *et al.* Clinical functional MRI of the language domain in children with epilepsy. *Human Brain Mapping* **32**, 1882–1893 (2011).
49. Staudt, M. *et al.* Early left periventricular brain lesions induce right hemispheric organization of speech. *Neurology* **57**, 122–125 (2001).
50. Staudt, M. *et al.* Right-hemispheric organization of language following early left-sided brain lesions: functional MRI topography. *Neuroimage* **16**, 954–967 (2002).
51. Hamberger, M. J. Cortical language mapping in epilepsy: A critical review. *Neuropsychology Review* **17**, 477–489 (2007).
52. Berl, M. M. *et al.* Characterization of atypical language activation patterns in focal epilepsy. *Annals of Neurology* **75**, 33–42 (2014).
53. Gelinias, J. N., Fitzpatrick, K. P. V., Kim, H. C. & Bjornson, B. H. Cerebellar language mapping and cerebral language dominance in pediatric epilepsy surgery patients. *NeuroImage: Clinical* **6**, 296–306 (2014).
54. Anderson, D. P. *et al.* fMRI lateralization of expressive language in children with cerebral lesions. *Epilepsia* **47**, 998–1008 (2006).
55. Baxendale, S. & Thompson, P. Beyond localization: the role of traditional neuropsychological tests in an age of imaging. [Review]. *Epilepsia* **51**, 2225–2230 (2010).
56. de Ribaupierre, S., Wang, A. & Hayman-Abello, S. Language Mapping in Temporal Lobe Epilepsy in Children: Special Considerations. *Epilepsy Research and Treatment* **2012**, e837036 (2012).
57. Sagar, S., Rick, J., Chandra, A., Yagnik, G. & Aghi, M. K. Functional brain mapping: overview of techniques and their application to neurosurgery. *Neurosurg Rev* **42**, 639–647 (2019).
58. National Library of Medicine. PubMed Search. *PubMed* [https://pubmed.ncbi.nlm.nih.gov/?term=\(language+mapping\)+and+\(epilepsy+surgery\)&sort=pubdate&filter=age.allchild](https://pubmed.ncbi.nlm.nih.gov/?term=(language+mapping)+and+(epilepsy+surgery)&sort=pubdate&filter=age.allchild) (2021).
59. Raichle, M. E. A brief history of human brain mapping. *Trends in Neurosciences* **32**, 118–126 (2009).

60. Pelletier, I., Sauerwein, H. C., Lepore, F., Saint-Amour, D. & Lassonde, M. Non-invasive alternatives to the Wada test in the presurgical evaluation of language and memory functions in epilepsy patients. [Review]. *Epileptic Disorders* **9**, 111–126 (2007).
61. Collinge, S. *et al.* Pre-surgical mapping of eloquent cortex for paediatric epilepsy surgery candidates: Evidence from a review of advanced functional neuroimaging. [Review]. *Seizure* **1**, 136–146 (2017).
62. Schwartz, E. S. *et al.* Magnetoencephalography for pediatric epilepsy: how we do it. [Review] [50 refs]. *Journal of Neuroradiology* **29**, 832–837 (2008).
63. Otsubo, H. & Snead, O. C. Magnetoencephalography and magnetic source imaging in children. *J Child Neurol* **16**, 227–235 (2001).
64. Lefaucheur, J.-P. & Picht, T. The value of preoperative functional cortical mapping using navigated TMS. [Review]. *Neurophysiologie Clinique* **46**, 125–133 (2016).
65. Trebuchon, A. & Chauvel, P. Electrical Stimulation for Seizure Induction and Functional Mapping in Stereoelectroencephalography. [Review]. *Journal of Clinical Neurophysiology* **33**, 511–521 (2016).
66. Bradshaw, A. R., Thompson, P. A., Wilson, A. C., Bishop, D. V. M. & Woodhead, Z. V. J. Measuring language lateralisation with different language tasks: a systematic review. *PeerJ* **5**, e3929 (2017).
67. Gallentine, W. B. & Mikati, M. A. Intraoperative electrocorticography and cortical stimulation in children. [Review]. *Journal of Clinical Neurophysiology* **26**, 95–108 (2009).
68. Arya, R., Horn, P. S. & Crone, N. E. ECoG high-gamma modulation versus electrical stimulation for presurgical language mapping. *Epilepsy & Behavior* **1**, 26–33 (2018).
69. Arya, R. *et al.* Electrical stimulation mapping of language with stereo-EEG. *Epilepsy & Behavior* **1**, (2019).
70. Roux, F.-E., Durand, J.-B., Djidjeli, I., Moyse, E. & Giussani, C. Variability of intraoperative electrostimulation parameters in conscious individuals: language cortex. *Journal of Neurosurgery* **126**, 1641–1652 (2017).
71. Reitsma, J. B. *et al.* Chapter 9: Assessing methodological quality. in *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (The Cochrane Collaboration, 2009).
72. Fox, M. D. & Greicius, M. Clinical applications of resting state functional connectivity. *Frontiers in systems neuroscience* **4**, 19 (2010).
73. Rolinski, R. *et al.* Language lateralization from task-based and resting state functional MRI in patients with epilepsy. *Human Brain Mapping* **41**, 3133–3146 (2020).
74. Lang, S., Duncan, N. & Northoff, G. Resting-State Functional Magnetic Resonance Imaging: Review of Neurosurgical Applications. *Neurosurgery* **74**, 453–465 (2014).

75. Lee, M. H. *et al.* Clinical Resting-state fMRI in the Preoperative Setting: Are We Ready for Prime Time? *Topics in Magnetic Resonance Imaging* **25**, 11–18 (2016).
76. Agarwal, S. *et al.* Demonstration of Brain Tumor-Induced Neurovascular Uncoupling in Resting-State fMRI at Ultrahigh Field. *Brain Connect* **6**, 267–272 (2016).
77. Branco, P. *et al.* Resting-State Functional Magnetic Resonance Imaging for Language Preoperative Planning. *Frontiers in Human Neuroscience* **10**, (2016).
78. Sabsevitz, D. S. *et al.* Use of preoperative functional neuroimaging to predict language deficits from epilepsy surgery. *Neurology* **60**, 1788–1792 (2003).
79. Rolinski, R. *et al.* Functional MRI and direct cortical stimulation: Prediction of postoperative language decline. *Epilepsia* **60**, 560–570 (2019).
80. Karami, M. *et al.* Presurgical Language Mapping in Patients With Intractable Epilepsy: A Review Study. *Basic and Clinical Neuroscience* **12**, 163–176 (2021).
81. Barba, C. *et al.* Patterns and predictors of language representation and the influence of epilepsy surgery on language reorganization in children and young adults with focal lesional epilepsy. *PloS one* **15**, e0238389 (2020).
82. Smith, M. L. Neuropsychology in epilepsy: Children are not small adults. *Epilepsia* **51**, 68–69 (2010).
83. O’Shaughnessy, E. S., Berl, M. M., Moore, E. N. & Gaillard, W. D. Pediatric functional magnetic resonance imaging (fMRI): Issues and applications. *Journal of Child Neurology* **23**, 791–801 (2008).
84. Gaillard, W. D. *et al.* Cortical localization of reading in normal children: an fMRI language study. *Neurology* **57**, 47–54 (2001).
85. Schevon, C. A. *et al.* Pediatric language mapping: sensitivity of neurostimulation and Wada testing in epilepsy surgery. *Epilepsia* **48**, 539–545 (2007).
86. Schlaggar, B. L. *et al.* Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science* **296**, 1476–1479 (2002).
87. Brown, T. T. *et al.* Developmental changes in human cerebral functional organization for word generation. *Cereb Cortex* **15**, 275–290 (2005).
88. Wood, A. G. *et al.* Language cortex activation in normal children. *Neurology* **63**, 1035–1044 (2004).
89. Holland, S. K. *et al.* Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage* **14**, 837–843 (2001).
90. Schapiro, M. B. *et al.* BOLD-fMRI signal increases with age in selected brain regions in children. *Neuroreport* **15**, 2575–2578 (2004).
91. Szaflarski, J. P., Holland, S. K., Schmithorst, V. J. & Byars, A. W. fMRI study of language lateralization in children and adults. *Hum Brain Mapp* **27**, 202–212 (2005).

92. Kadis, D. S. *et al.* Intrahemispheric reorganization of language in children with medically intractable epilepsy of the left hemisphere. *Journal of the International Neuropsychological Society* **13**, 505–516 (2007).
93. Yerys, B. E. *et al.* The fMRI success rate of children and adolescents: typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. *Hum Brain Mapp* **30**, 3426–3435 (2009).
94. Booth, J. R. *et al.* Relation between brain activation and lexical performance. *Human brain mapping* **19**, 155–169 (2003).
95. Ulmer, J. L. *et al.* Lesion-induced Pseudo-dominance at Functional Magnetic Resonance Imaging: Implications for Preoperative Assessments. *Neurosurgery* **55**, 569–581 (2004).
96. Cho, N. S., Peck, K. K., Zhang, Z. & Holodny, A. I. Paradoxical Activation in the Cerebellum During Language fMRI in Patients with Brain Tumors: Possible Explanations Based on Neurovascular Uncoupling and Functional Reorganization. *Cerebellum* **17**, 286–293 (2018).
97. Chen, R., Samii, A., Canos, M., Wassermann, E. M. & Hallett, M. Effects of phenytoin on cortical excitability in humans. *Neurology* **49**, 881–883 (1997).
98. Berger, M. S., Kincaid, J., Ojemann, G. A. & Lettich, E. Brain mapping techniques to maximize resection, safety, and seizure control in children with brain tumors. *Neurosurgery* **25**, 786–792 (1989).
99. Bast, T., Ramantani, G., Seitz, A. & Rating, D. Focal cortical dysplasia: prevalence, clinical presentation and epilepsy in children and adults. *Acta Neurol Scand* **113**, 72–81 (2006).
100. Jonas, R., Nguyen, S. & Hu, B. Cerebral hemispherectomy: hospital course, seizure, developmental, language and motor outcomes. *Neurology* **62**, 1712–1721 (2004).
101. Jeong, J.-W., Asano, E., Juhasz, C. & Chugani, H. T. Localization of specific language pathways using diffusion-weighted imaging tractography for presurgical planning of children with intractable epilepsy. *Epilepsia* **56**, 49–57 (2015).
102. Liegeois, F. *et al.* Language reorganization in children with early-onset lesions of the left hemisphere: An fMRI study. *Journal of Neurology* **127**, 1229–1236 (2004).
103. Biega, T. J., Levy, L. M., Bannout, F., Gaillard, W. D. & Potolicchio, S. J. Bilateral and independent Broca areas confirmed by Wada test and functional magnetic resonance imaging. *Journal of Computer Assisted Tomography* **33**, 560–561 (2009).
104. Gil Robles, S. *et al.* Discrepancies between preoperative stereoencephalography language stimulation mapping and intraoperative awake mapping during resection of focal cortical dysplasia in eloquent areas. *Stereotactic & Functional Neurosurgery* **86**, 382–390 (2008).
105. Ogawa, H. *et al.* Clinical Impact and Implication of Real-Time Oscillation Analysis for Language Mapping. *World Neurosurgery* **1**, 123–131 (2017).

106. Alonso, F., Sweet, J. & Miller, J. Speech mapping using depth electrodes: The 'electric Wada'. *Clinical Neurology & Neurosurgery* **1**, 88–90 (2016).
107. Szmuda, M. *et al.* Diffusion tensor tractography imaging in pediatric epilepsy – A systematic review. *Neurologia i Neurochirurgia Polska* **50**, 1–6 (2016).
108. Guyatt, G. *et al.* GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *Journal of clinical epidemiology* **64**, 383–394 (2011).
109. de Koning, T. *et al.* Language development before and after temporal surgery in children with intractable epilepsy. *Epilepsia* **50**, 2408–2419 (2009).
110. Lee, M.-H. *et al.* Novel diffusion tractography methodology using Kalman filter prediction to improve preoperative benefit-risk analysis in pediatric epilepsy surgery. *Journal of Neurosurgery* **1**, 1–13 (2019).
111. Sakpichaisakul, K. *et al.* Neuropsychological outcomes after pediatric epilepsy surgery: Role of electrical stimulation language mapping. *Seizure* **1**, 183–191 (2020).
112. Whiting, P., Rutjes, A. W. S., Reitsma, J. B., Bossuyt, P. M. M. & Kleijnen, J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* **3**, 25 (2003).
113. OCEBM Levels of Evidence Working Group. 'The Oxford 2011 Levels of Evidence.' Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653> (2011).
114. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Group, P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine* **6**, e1000097 (2009).
115. Veritas Health Innovation. *Covidence systematic review software*. (2018).
116. Macaskill, P., Gatsonis, C., Deeks, J., Harbord, R. & Takwoingi, Y. Cochrane handbook for systematic reviews of diagnostic test accuracy. *Version 0.9. 0. London: The Cochrane Collaboration* **83**, (2010).
117. Cochrane, T. C. Review Manager (RevMan) 5.3. *Copenhagen: The Nordic Cochrane Centre* (2008).
118. Lee, M.-H. *et al.* Prediction of postoperative deficits using an improved diffusion-weighted imaging maximum a posteriori probability analysis in pediatric epilepsy surgery. *J Neurosurg Pediatr* 1–12 (2019) doi:10.3171/2018.11.PEDS18601.
119. Lee, M.-H. *et al.* Novel Deep Learning Network Analysis of Electrical Stimulation Mapping-Driven Diffusion MRI Tractography to Improve Preoperative Evaluation of Pediatric Epilepsy. *IEEE Trans. Biomed. Eng.* **67**, 3151–3162 (2020).
120. Paediatric Epilepsy Neuropsychology Working Group. Revised Schedule of Neuropsychological Assessment for use in routine pre-operative assessment. in *Neuropsychological Management of Children undergoing Surgical Investigation* 1–6 (2016).

121. Crawford, J. R. & Garthwaite, P. H. Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology* **21**, 611–620 (2007).
122. Cross, J. H. & Vezyroglou, K. Selecting the right paediatric patient for epilepsy surgery. *Paediatrics and Child Health* **27**, 287–291 (2017).
123. Cross, J. H. *et al.* Proposed Criteria for Referral and Evaluation of Children for Epilepsy Surgery: Recommendations of the Subcommittee for Pediatric Epilepsy Surgery. *Epilepsia* **47**, 952–959 (2006).
124. Baker, G. A. Psychological and neuropsychological assessment before and after surgery for epilepsy: implications for the management of learning-disabled people. *Epilepsia* **1**, 41–43 (2001).
125. Van Poppel, M. *et al.* Passive language mapping with magnetoencephalography in pediatric patients with epilepsy. *Journal of Neurosurgery* **10**, 96–102 (2012).
126. Ojemann, J. G., Ojemann, G. A. & Lettich, E. Cortical stimulation mapping of language cortex by using a verb generation task: effects of learning and comparison to mapping based on object naming. *Journal of Neurosurgery* **97**, 33–38 (2002).
127. Marusic, P. *et al.* Focal cortical dysplasias in eloquent cortex: functional characteristics and correlation with MRI and histopathologic changes. *Epilepsia* **43**, 27–32 (2002).
128. Kurthen, M. The Intracarotid Amobarbital Test: Uses - procedures - results. [German]. *Der Nervenarzt* **63**, 713–724 (1992).
129. Kurthen, M. *et al.* Interhemispheric dissociation of expressive and receptive language functions in patients with complex-partial seizures: An amobarbital study. *Brain and Language* **43**, 694–712 (1992).
130. Kurthen, M. *et al.* Quantitative and qualitative evaluation of patterns of cerebral language dominance. An amobarbital study. *Brain & Language* **46**, 536–564 (1994).
131. Knake, S. *et al.* Language lateralization in patients with temporal lobe epilepsy: a comparison of functional transcranial Doppler sonography and the Wada test. *Neuroimage* **19**, 1228–1232 (2003).
132. Kirsch, H. E. *et al.* Limitations of Wada memory asymmetry as a predictor of outcomes after temporal lobectomy. *Neurology* **65**, 676–680 (2005).
133. Jones, S. E., Mahmoud, S. Y. & Phillips, M. D. A practical clinical method to quantify language lateralization in fMRI using whole-brain analysis. *Neuroimage* **54**, 2937–2949 (2011).
134. Jabbour, R. A., Hempel, A., Gates, J. R., Zhang, W. & Risse, G. L. Right hemisphere language mapping in patients with bilateral language. *Epilepsy & Behavior* **6**, 587–592 (2005).
135. Ivanova, A. *et al.* Intrinsic functional organization of putative language networks in the brain following left cerebral hemispherectomy. *Brain Structure & Function* **222**, 3795–3805 (2017).

136. Gartus, A., Foki, T., Geissler, A. & Beisteiner, R. Improvement of clinical language localization with an overt semantic and syntactic language functional MR imaging paradigm. *Journal of Neuroradiology* **30**, 1977–1985 (2009).
137. Fontoura, D. R. da, Branco, D. de M., Anes, M., Costa, J. C. da & Portuguese, M. W. Language brain dominance in patients with refractory temporal lobe epilepsy: a comparative study between functional magnetic resonance imaging and dichotic listening test. *Arquivos de Neuro-Psiquiatria* **66**, 34–39 (2008).
138. Findlay, A. M. *et al.* Dynamics of hemispheric dominance for language assessed by magnetoencephalographic imaging. *Annals of Neurology* **71**, 668–686 (2012).
139. DeSalvo, M. N. *et al.* Resting-State Functional MR Imaging for Determining Language Laterality in Intractable Epilepsy. *Radiology* **281**, 264–269 (2016).
140. Brazdil, M., Zakopcan, J., Kuba, R., Fanfrdlova, Z. & Rektor, I. Atypical hemispheric language dominance in left temporal lobe epilepsy as a result of the reorganization of language functions. *Epilepsy & Behavior* **4**, 414–419 (2003).
141. Bowyer, S. M. *et al.* Magnetoencephalographic localization of the basal temporal language area. *Epilepsy & Behavior* **6**, 229–234 (2005).
142. Benjamin, C. F. *et al.* Presurgical language fMRI: Mapping of six critical regions. *Human Brain Mapping* **38**, 4239–4255 (2017).
143. Alarcon, G. *et al.* The Single Word Auditory Comprehension (SWAC) test: A simple method to identify receptive language areas with electrical stimulation. *Epilepsy & Behavior* **1**, 266–272 (2019).
144. Wilson, S. J. & Baxendale, S. The new approach to classification: Rethinking cognition and behavior in epilepsy. *Epilepsy & Behavior* **41**, 307–310 (2014).
145. Tackett, J. L., Brandes, C. M., King, K. M. & Markon, K. E. Psychology’s Replication Crisis and Clinical Psychological Science. *Annual Review of Clinical Psychology* **15**, 579–604 (2019).
146. Huber, D. E., Potter, K. W. & Huszar, L. D. Less “Story” and more “Reliability” in cognitive neuroscience. *Cortex* **113**, 347–349 (2019).
147. Qadri, S., Dave, H., Das, R. & Alick-Lindstrom, S. Beyond the Wada: An updated approach to pre-surgical language and memory testing: An updated review of available evaluation techniques and recommended workflow to limit Wada test use to essential clinical cases. *Epilepsy Research* **174**, 106673 (2021).
148. Ives-Deliperi, V. L. & Butler, J. T. Naming outcomes of anterior temporal lobectomy in epilepsy patients: A systematic review of the literature. *Epilepsy & Behavior* **24**, 194–198 (2012).
149. Davies, K. G., Risse, G. L. & Gates, J. R. Naming ability after tailored left temporal resection with extraoperative language mapping: Increased risk of decline with later epilepsy onset age. *Epilepsy & Behavior* **7**, 273–278 (2005).

150. Hermann, B., Davies, K., Foley, K. & Bell, B. Visual confrontation naming outcome after standard left anterior temporal lobectomy with sparing versus resection of the superior temporal gyrus: a randomized prospective clinical trial. *Epilepsia* **40**, 1070–1076 (1999).
151. ABIDE. http://fcon_1000.projects.nitrc.org/indi/abide/.
152. Caplan, A. & Friesen, P. Health disparities and clinical trial recruitment: Is there a duty to tweet? *PLOS Biology* **15**, e2002040 (2017).
153. Gaillard, W. D. *et al.* Establishing criteria for pediatric epilepsy surgery center levels of care: Report from the ILAE Pediatric Epilepsy Surgery Task Force. *Epilepsia* (2020) doi:10.1111/epi.16698.
154. Wilson, S. J. *et al.* Indications and expectations for neuropsychological assessment in routine epilepsy care: Report of the ILAE Neuropsychology Task Force, Diagnostic Methods Commission, 2013-2017. *Epilepsia* **56**, 674–681 (2015).
155. Wright, I. & Marsh, A. A Robust Statistical Approach to Classification of Developmental Change in Cognitive Test Scores. Presented at Forty Fifth Annual Meeting International Neuropsychological Society. *Journal of the International Neuropsychological Society* **23(S1)**, 298 (2017).

Paper 2: Convergent Validity of Resting-State and Task-Based Functional Magnetic Resonance Imaging in Lateralising Language for Paediatric Epilepsy Surgery

Alexander P. Marsh

(MarshAP@Cardiff.ac.uk)

School of Psychology, Cardiff University, Cardiff, United Kingdom

This work is submitted in partial fulfilment of the requirements for the degree of Doctor of Clinical Psychology.

(2021)

Prepared for submission to *Epilepsia* with expanded word count as permitted by DClInPsy LSRP Guidance

Paper abstract word count: 297

Paper body word count: 6014

Total word count: 6311

Abstract

Objectives: Task-based fMRI is routine for language mapping in paediatric epilepsy surgery but is not well received by children. Resting-state fMRI is a potential alternative technique, having no task demands. In adults, it has good concordance with task-based fMRI and validity in estimating surgical risk to language. This study aimed to investigate the concordance of resting-state state and task-based fMRI in lateralising language function in paediatric epilepsy surgery. It also aimed to investigate predictive validity in a case series.

Methods: A retrospective cohort study of thirty-five patients undergoing epilepsy surgery work-up and language mapping was undertaken. Task-based fMRI and resting-state fMRI language activation maps were used to compute lateralisation indexes (LIs). Language lateralisation was classed as: left, bilateral or right based on LI. A continuous surgical risk index (SII) based on LIs was computed. SII were classified into 'at risk' or 'no risk', based on conservative and liberal thresholds. Regression analysis of individual cases determined presence of language decline, post-operatively.

Results: Twenty-six patients (12 female) were included (mean age=12.8, range=7-18 years). No statistically significant concordance ($p>0.05$) was found between resting-state and task-based fMRI in the variables of: LI, SII and categorical language lateralisation. Categorical surgical risk assessment (i.e. 'at risk', 'no risk') was significantly concordant between resting-state and task-based fMRI at both thresholds. Task-based fMRI accurately predicted post-operative language outcome in 50% of cases ($n=6$), regardless of threshold. Resting-state predicted outcome in 66% and 83% of cases, at liberal and conservative thresholds, respectively.

Significance: There is insufficient evidence for use of resting-state fMRI as a proxy to task-based fMRI for the lateralisation of language function in paediatric epilepsy surgery candidates. However, it may be of use as a standalone instrument based on the predictive data in our small sample. Further investigation of resting-state fMRI is warranted in paediatric epilepsy surgery.

Key Points

- Task-based fMRI language mapping is not well accepted in paediatric epilepsy patients; resting-state fMRI may be an appropriate substitute.
- In a broad sample of paediatric epilepsy presentations, resting-state and task-based fMRI were found not to agree on language laterality.
- However, resting-state and task-based fMRI were shown to agree on estimated surgical risk to language outcome.
- Resting-state fMRI demonstrated better predictive validity of language outcome than task-based fMRI in a small case series
- Further validation work is needed before resting-state fMRI is adopted into routine practice guidance for risk estimation.

Key Words

Eloquent Cortex Mapping; Childhood Epilepsy; connectivity fMRI

Introduction

Epilepsy is a condition affecting central nervous system networks, where abnormalities in the structural and functional organization of the brain's systems leads to persistent neuronal hyper-excitability and subsequent seizures¹. The International League Against Epilepsy recently proposed a new classification for epilepsies. This included three main types of epilepsy: (1) focal onset, (2) generalised onset, (3) unknown onset². Focal onset epilepsy is defined as having a predisposition to seizures that originate "within networks limited to one hemisphere. They may be discretely localised or more widely distributed"². One of the most effective treatments for focal epilepsy is resective surgery of the epileptogenic zone (EZ) tissue³⁻⁵, however this comes with a potential cost of post-operative cognitive decline if the resected tissue is still subserving cognitive function³⁻⁶. One key goal of pre-operative assessment is to determine the suitability of a resective approach. This involves balancing the potential benefit of removing epileptogenic tissue with the risk of removing potentially eloquent cortex causing post-operative deficit. The risk of post-operative language decline is of particular importance as language impairment can lead to academic underachievement and long-term social, professional, and neuropsychological problems in persons with epilepsy⁷⁻⁹.

Current methods to predict risk of post-operative language decline use a combination of demographic data, structural imaging, cognitive testing, and task-based fMRI to estimate the contribution of the EZ to current function¹⁰⁻¹². In many cases multi-disciplinary clinical teams can make informed predictions of cognitive outcome¹³⁻¹⁷. However, many children with epilepsy cannot effectively engage in formal cognitive evaluation or task-based fMRI¹⁸⁻²⁰.

Even in typically developing children, task-based fMRI poses several specific challenges: from developmentally appropriate paradigm design to extensive patient preparation^{21–24}. The paradigms used need to be stratified by age and appropriate for the stage of language development for the child²⁴. For example, tasks requiring written stimuli will not be appropriate for a child who has not yet developed reading literacy. Similarly, tasks providing little cognitive challenge may result in poorer signal^{25–27}. Therefore, tasks appropriate for younger children or children with cognitive delay, would not be as effective in more cognitively developed individuals. Additionally, paradigms need to take careful account of task-related signal confounds, such as the act of button pressing during a decision task²⁸. Such confounds can introduce non-language related signal, which can be somewhat mitigated by paired behaviours in the rest phase and general linear modelling in analysis but leave the approach vulnerable to false-positive results^{28,29}. These problems are further exacerbated in clinical populations. Lower rates of success in task-based fMRI have been shown in neurodevelopmental and epilepsy populations, when compared to healthy controls²³. This effect is even more pronounced in primary age children^{23,24}. Reasons for low acceptability include: increased refusal to enter scanner, cognitive impairment interfering with task engagement and increased motion^{23,28}. The lack of acceptability of task-based fMRI significantly limits post-operative estimation of risk. Therefore, developing alternative methodologies for assessment of post-operative risk is imperative for these populations.

Resting-State fMRI as an alternative

Recent advances in resting-state connectivity imaging may provide an alternative for children who cannot undergo current routine assessment with task-based fMRI. Resting-state fMRI

has demonstrated reliable low frequency (<0.1 Hz) blood-oxygen level dependent (BOLD) signal changes in several spatially distinct brain networks, including language^{18,29–31}. Resting-state fMRI overcomes many of the difficulties experienced conducting task-based fMRI. The most notable advantage is that resting state fMRI has the absence of the requirement for the participant to undertake a task. This eliminates the difficulties with age-limiting factors of behavioural assessment and poor engagement with task due to cognition, anxiety, or sensory issues. The absence of task requirements also overcomes the challenge of producing valid and developmentally appropriate paradigms, which do not have inherent confound effects of task on observed BOLD activity²⁹.

There is growing evidence that resting-state fMRI does not need to be conducted whilst the participant is awake³². Several recent studies of resting-state fMRI network correlation patterns through various sleep states^{32,33} indicate that the networks are relatively stable. However, decreased metabolic rates in the frontal and parietal cortices have been observed in deep sleep³⁴ and general anaesthesia³⁵. Whilst further evidence needs to validate this approach in functional mapping, resting-state fMRI would overcome one of the biggest barriers to obtaining accurate scans by reducing motion and behavioural confounds.

Another advantage of resting state fMRI is that the BOLD signal oscillations can provide up to three times higher signal-to-noise ratio than task-related signal increases^{18,29}. Multiple works suggest BOLD signal has potential to be reduced in young and pathological populations^{26,36–38}. The final clinical advantage of resting-state fMRI is its improved time-efficiency over task-based fMRI. The imaging protocols can be collected more rapidly (~5minutes) than many task-based protocols. One resting-state scan serves multiple mapping purposes³⁹, thus placing less demands on limited scanner time¹⁸.

There are several approaches to resting-state fMRI, which are described in detail elsewhere⁴⁰ and summarised in Table 1.

Table 1. Functional Integration Methods for Identifying Neural Networks (Based on Ref⁴⁰)

Analysis Method	Description
Functional Connectivity Density (FCD)	Most basic measure of functional connectivity and does not require a priori information ⁴⁰ . Measures the association of an individual voxel's time-series and every other voxel in the brain. Through this it identifies functional hubs, however, does not determine which hubs are connected to each other ⁴⁰ .
Seed-Based (ROI) Functional Connectivity	Examines the association between the time-course of a seed region and other regions of the brain. Is also known as region-of-interest-based analysis. Significant correlations between regions infers functional association ⁴⁰ . Is driven by a priori information to define seed region.
Independent Component Analysis (ICA)	Separates BOLD signal through multivariate decomposition into several statistical components, which represent functionally associated networks ⁴⁰⁻⁴³ . Several networks are commonly generated using this approach: default mode network; auditory network; salience network; executive control network; medial visual network; lateral visual network; sensorimotor network; frontoparietal attentional network; limbic network; precuneus network ⁴⁰

Resting-state fMRI and language

There is evidence of reliable language network lateralisation in adult healthy controls^{30,44-47}.

Resting-state language networks and their lateralised functional asymmetry have been reliably reproduced across institutions in 970 healthy controls⁴⁶ and demonstrated temporal reliability in both the short (30 minutes) and long-term (5-11 months)⁴⁵. When comparing

fluctuations in Broca's and Wernicke's area on resting-state fMRI to the laterality index of task-based fMRI, a strong positive correlation is demonstrated, indicating resting-state fMRI is comparably effective in predicting language laterality in adult healthy controls^{44,47}.

There is good evidence for the use of resting-state fMRI with adults with intractable epilepsy. The measurement of hemispheric language dominance with resting-state functional MR imaging is highly concordant with other standard lateralisation procedures, such as the intracarotid amobarbital procedure, demonstrating up to 96% accuracy, sensitivity, and specificity⁴⁸. Additionally, in adult epilepsy patients, resting-state fMRI derived language network maps were found to be comparable to that derived from direct cortical stimulation⁴⁹. Lastly, resting-state language laterality has demonstrated better predictive validity of post-operative naming decline than task-based fMRI⁵⁰.

The above data supports the utility of resting-state fMRI in language lateralisation for both healthy control and clinical samples; however, validation needs to take place as there are important physiologic and anatomic differences in children, varying with age, which may affect the acquisition, analysis, and interpretation of paediatric fMRI data⁵¹. For example, synaptic pruning and brain myelination with associated changes in glucose consumption and blood flow may affect the detection, magnitude, and extent of BOLD response⁵¹. Morphological differences between paediatric and adult brains may influence the magnitude of signal and the location of task activation, notably when data are warped onto adult atlases for analysis⁵¹. These morphological differences also bear relation to functional specialisation. Younger children demonstrate relatively less frontal activity, whilst showing increased activity in the medial parietal cortex, posterior cingulate, and occipital cortex, in comparison with adults on a variety of language tasks³⁶⁻³⁸. Additionally, typically developing children under 10 years are more likely to have atypical language lateralisation compared to adolescents and

adults^{52–55}; indicating increasing specialisation of language to the left throughout brain development and complicating the generalisation of the adult evidence base to this younger sample. When the complexity of pathology is introduced alongside already challenging translation issues; with developmental epilepsy and related lesions demonstrating differing patterns of neuroplastic function to adult onset epilepsy^{56,57}, there is good cause for paediatric sample validation. The challenge to translation is exemplified in the finding that functional mapping of language in paediatrics often produces more unpredictable results^{6,58}. Therefore, thorough investigation into the comparable efficacy of resting-state fMRI to routine practice measures (e.g., task-based fMRI, electrographic cortical stimulation etc.) is warranted in paediatric samples, prior to its use in routine clinical practice.

A recent study of children with epilepsy demonstrated differences in seed-based resting-state fMRI networks between paediatric epilepsy patients with left- and right-hemispheric language lateralization in a proportion of their sample, demonstrating the discriminant potential of resting-state fMRI in paediatric language lateralisation⁵⁹. A second study⁶⁰ found an overall concordance rate of 0.93 (95% CI 0.76–0.99) between task-based fMRI and resting-state fMRI in a sample of 29 adolescents with epilepsy. Although these studies show promising early evidence, they were carried out in limited samples with relatively homogeneous pathology and included mainly adolescents. These samples do not represent those who typically present for epilepsy surgery work-up, with the ages ranging from 0-18 years (median age 10.4 years) and a range of epilepsy presentations^{4,12,61}. As the majority of the current methodological challenges apply mainly to pre-adolescent children, it is vital studies are replicated including younger age (e.g., <12 years) samples, as discussed above. Additionally, as patients with epilepsy are a heterogenous group, it is imperative that studies

are conducted that demonstrate reliability of these methodologies in a diverse range of pathology, most notably a range of left-right lateralised EZ and left-right language dominance.

Current study

This study sought to build on existing evidence by utilising a paediatric retrospective, anonymised, and systematically collected clinical dataset. The study had three main aims. The first aim was to evidence generation of reliable resting-state language networks is feasible in routine practice in a diverse group of paediatric epilepsy patients, including those under 12 years old. The second aim was to report on concordance of language lateralisation and estimated surgical risk between resting-state fMRI language networks and conventional task-based fMRI methods. This was examined both on a continuous spectrum and in terms of categorical concordance. The final aim was to report surgical outcome in relation to language mapping in a small case series, exploring preliminary evidence for predictive validity.

A moderate relationship between task-based and resting-state fMRI lateralisation was hypothesised. Although a statistically significant relationship was expected, a weaker association than previously described was predicted due to the inclusion of more children under 10 years old, and the many factors outlined above that may introduce variance in signal strength and language laterality in these younger cases across modalities.

Methods Section

Subjects

Forty-nine patients with focal epilepsy underwent fMRI and neuropsychological assessment as part of their presurgical investigations at Bristol Royal Hospital for Children (BRCH) between 2015-2020. Ten patients completed surgery and had subsequent 1-year follow-up neuropsychological assessment. Twenty-three patients were excluded; 14 did not complete scanning and nine failed quality assessment⁶² of scan data due to significant motion. The reasons provided in the records for incomplete scanning were anxiety and behavioural disengagement. All patients were English speakers. Demographic and clinical data of included participants are summarised in Table 2. This project was reviewed and approved by The NHS Research and Development Division at University Hospitals Bristol and Weston Trust (Registration reference: PNEUS/CA/2020-21/,01; see Appendix).

Neuropsychological Testing

Neuropsychological testing was performed in accordance with the National Children's Epilepsy Surgery Protocol. Intellectual functioning was assessed using the Wechsler Intelligence Scales⁶³⁻⁶⁵. These include a measure of language within the Verbal Comprehension Index (VCI)⁶³⁻⁶⁵, which is reported and used in analysis as an estimate of behavioural language function, as in previous literature⁶⁶. Difference was computed by subtracting post-operative VCI from baseline VCI. The reliability of the changes in VCI was determined by using Crawford and Garthwaite's⁶⁷ regression analysis of the individual case.

Clinically significant decline was determined by both a statistical difference ($p < 0.05$) and that the difference was of low estimated prevalence ($< 5\%$).

fMRI Paradigm

A covert verb generation task, described in previous studies^{53,68,69}, was used to assess language laterality (Figure 1). In this task, participants were presented both visually and verbally with a series of single nouns and asked to silently generate an associated verb (e.g., chair -> sit). Presentation® (Neurobehavioural Systems Inc, www.neurobs.com) was used to present the stimulus and was synchronised with the MR pulse.

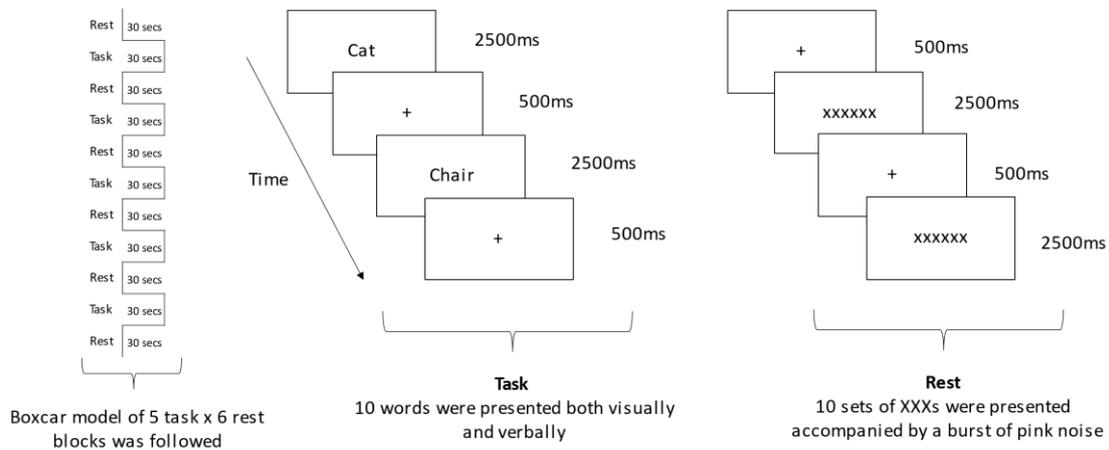


Figure 1. Covert-Verb Paradigm

MRI Acquisition

MRI data was obtained using a 3T Siemens Skyra MRI Scanner with standard 20 channel head and neck coil. The functional scans consisted of 110 volumes collected using gradient-echo-planar imaging (EPI) with a TR=2800 ms, TE=30ms and 90-degree flip angle. Forty contiguous axial slices were collected, with 3mm thickness, 192 mm field of view and a voxel size of 3x3x3mm. T1 structural scans were obtained (192 volumes) with a 256 x 256 matrix, voxel size 1x1x1 mm.

Task-Based fMRI Analysis

Image processing was conducted using SPM12⁷⁰ through MATLAB R2021a⁷¹. Functional images were realigned^{70,72} and slice-timing correction⁷³ was applied. Images were co-registered to the structural T1⁷⁰. Functional and anatomical data were normalised into standard Montreal Neurological Institute (MNI) space and segmented into grey matter, white matter, and CSF tissue classes^{70,74}; functional and anatomical data were resampled to the default 180x216x180mm bounding box, with 2mm isotropic voxels for functional data and 1mm for anatomical data. Functional data was smoothed using spatial convolution with a Gaussian kernel of 8mm FWHM⁷⁵. Quality control inspection of co-registration, segmentation and movement parameters was undertaken, and poor-quality scans excluded from the analysis (see supplementary material for quality control process). Language task-related activity was modelled by convolving a vector of block onsets for each condition with a canonical hemodynamic response function (HRF) to create regressors of interest. Single patient level analyses were performed using a general linear model^{70,76}, contrasting task and control ('rest') conditions.

Resting-State fMRI Analysis

Sixty volumes of rest were extracted from secondary task-based scans (e.g., fingertip-tapping motor fMRI², foot-tapping motor fMRI¹ etc.). The motor tasks were based on tapping either the contralateral foot or hand of the surgical hemisphere. No explicit rest instructions were provided, and no stimulus was presented during this time. Pre-processing of images was carried out using SPM12⁷⁰ as above. Data were analysed using the CONN toolbox⁶² (www.nitrc.org/projects/conn) within SPM12⁷⁰ and MATLAB R2021a following predetermined default settings⁶². Additional pre-processing was performed on the derived connectivity data, again following default pre-processing pipeline defined by CONN⁶². As recommended⁶², an outlier identification procedure using a framewise displacement cut-off of >0.9mm or global BOLD signal changes >5 s.d. was undertaken to produce a first-level 'scrubbing' covariate. CONN was used to apply an anatomical component-based noise correction procedure (aCompCor) and included noise components from cerebral white matter and cerebrospinal areas^{77,78}, estimated patient-motion parameters⁷⁹, identified outlier scans or scrubbing⁸⁰ and BOLD signal trends within each session⁶². As standard, temporal band-pass filtering was applied and set to exclude <0.008Hz and >0.09Hz⁸¹. Quality Control (QC) plots⁶² were examined at each stage and poor-quality scans excluded from the analysis (See supplementary material for example). To extract language relevant functional-connectivity networks, CONN was used to undertake a weighted general linear model (GLM) applying seed-to-voxel based analysis using pre-determined regions of interest (ROI; bilateral brain regions including inferior frontal gyrus and superior temporal gyrus seeded at peak co-

² The rest in these conditions involved no stimuli presentation, as in the covert verb fMRI paradigm

ordinates, as defined by CONN⁶²). Seed regions were determined using CONN's pre-set language network processing pipeline, as in Whitfield-Gabrieli and Nieto-Castanon⁶². These regions are consistent with prior literature investigating language lateralisation⁵⁹⁻⁶⁰. Bivariate correlation was applied within the GLM with haemodynamic response function convolved weighting. ROIs were determined using an automated anatomical labelling atlas^{62,70,82}. Left and Right ROI-generated networks were combined into a final single parametric map, as in Rolinski⁸³.

Lateralisation Index (LI)

The LI- toolbox^{84,85} was used to calculate lateralisation indexes. A boot-strapping approach⁸⁴ was utilised using default processing pipeline. The inferior frontal gyrus pars opercularis, pars triangularis, and superior temporal gyrus were inclusively masked (Figure 2), and the midline (± 5 mm) was excluded. As is conventional in the literature⁸⁶, positive LIs reflect a greater degree of left-sided (i.e., typical) language representation, negative LIs reflect a greater degree of right-sided (i.e., atypical) language representation.

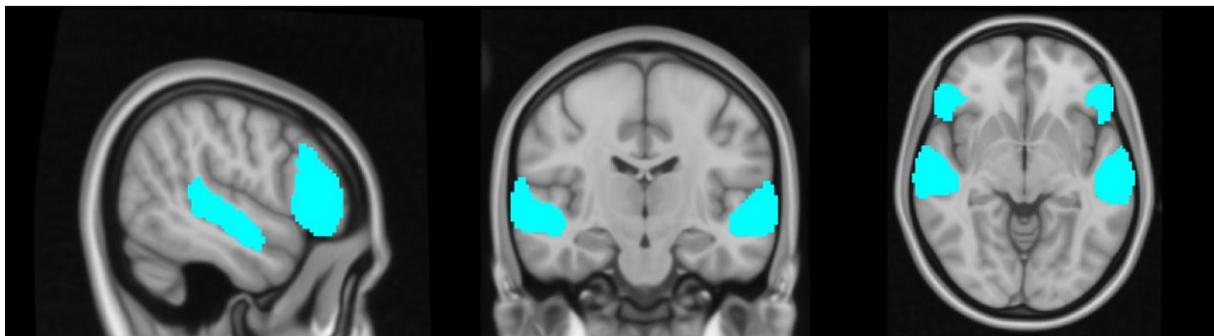


Figure 2. Lateralisation inclusive masks (in blue) overlaid onto MNI T1 Standard Brain

Further Statistical Analysis

The following statistical analysis was conducted in IBM SPSS Statistics Version 26.

Index Concordance

Lateralisation Indices

Bland-Altman diagrams were produced to visually depict any potential bias and the degree of agreement between task-based and resting-state lateralisation indices, as recommended by Kwiecien and colleagues⁸⁷.

Due to the non-normative distribution of the data and small sample sizes, non-parametric Spearman's Rho was applied to examine if there is a statistically significant relationship between task-based and resting-state lateralisation indices.

Surgical Ipsilateral Indices

Surgical Ipsilaterally Indices (SII) were calculated based on surgical target hemisphere. These were computed by reversing the sign of the LI, if pathology was on the right (e.g., when surgical target is right, LI = -0.6, then SII = +0.6; when surgical site is left, LI = -0.6, then SII = -0.6). With this metric, the greater the positivity of the SII, the greater the potential surgical risk to language; the greater the negativity of the SII, the lower the risk is to language.

Due to the non-normative distribution of the data and small sample sizes, non-parametric Spearman's Rho was applied to examine if there is a statistically significant relationship between task-based and resting-state SII.

Categorical concordance

Lateralisation concordance

To allow for calculation of categorical concordance, Lis >0.2 were coded as indicating left hemisphere dominance for language function, Lis <-0.2 were coded as right and all values between were coded as bilateral, as is conventional in the literature⁸⁶. Cohen's kappa was applied to examine agreement between task-based and resting-state lateralisation indices, as recommended by Kwiecien⁸⁷.

Surgical Ipsilateral Indices concordance

The SIIs were coded based on two thresholds. The low threshold was coded <0 = no risk, >0 = risk; and the high threshold was coded <-0.2 = no risk, >-0.2 = risk. -0.2 was chosen as this is the established cut off for non-bilateral representation. Cohen's kappa was applied to examine agreement between task-based and resting-state SIIs, as recommended by Kwiecien⁸⁷.

Validity

Post-operative outcome

Descriptive statistics of post-operative outcome are described and tabulated.

Results Section

Demographic variables and baseline characteristic data are presented in Table 2.

Table 2. Summary of baseline clinical and demographic characteristics

Sex		Handedness	
Female	12 (46%)	Left	5 (20%)
Male	14 (54%)	Right	20 (80%)
Age (Years)		ILAE Seizure Onset Type	
Mean	12.8	Generalised	1 (4%)
Median	12.6	Focal	17 (65%)
Range	7-18	Mixed Focal & Generalised	8 (31%)
Baseline Language (VCI)		Surgical Target Hemisphere	
Mean	87	Left	15 (58%)
Median	85	Right	11 (42%)
Range	60-136		

Summary of baseline clinical and demographic characteristics are summarised in full in Table 3.

Language Mapping

Reliable resting-state and task-based language networks were extracted in 26 patients (Figure 3) and indices by patient are presented in Table 4.

Table 3. Baseline clinical and demographic characteristics of individual patients

Participant	Age (years)	Gender	Handedness	Onset Age	Sz Type	Presumed Epileptogenic Focus	Surgical Target Hemisphere	Baseline Language (VCI)
1	8.5	F	R	5	Focal motor and focal with impaired awareness	Left Temporal	Left	84
2	10.1	F	L	11	focal onset with paroxysmal symptoms	Right Frontal	Right	98
3	18.2	F	R	5	nocturnal focal with impaired awareness	Right Frontal, Left Hippocampal, Amygdala	Right	136
4	8.9	M	R	0.8	focal, historical generalised	Right Temporal	Right	95
6	8.3	M	L	15	focal	Left Temporal	Left	62
7	17.9	M	R	11	focal and nocturnal generalised	Right Temporal	Right	95
8	16.2	F	L	12	focal leading to bilateral tonic clonic, focal vacant events	Bilateral Temporal	Left	81
9	16.4	F	R	4	focal onset with secondary generalise	Right Frontal	Right	76
10	14.2	F	R	3	focal	Multiple Right Lesions	Right	89
11	13.5	F	R	13	focal with impaired awareness, nausea, automatisms, and tonic R hand	Right Temporal, Right Occipital	Right	95
12	14.6	M	R	14	focal with altered awareness	Left Temporal	Left	121
13	17.3	M	R	2	focal onset with impaired awareness	Left Temporal	Left	105
14	8.8	M	R	9	generalised tonic clonic and absence events	Left Temporal, Bilateral Frontal	Left	99
15	10.5	M	R	7	focal	Left Frontal	Left	81
16	12.2	M	R	3	focal	Multiple Left Foci (TS)	Left	92
17	10.2	M	R	12	focal	Right frontal and parietal	Right	75
20	13.1	F	R	11	partial onset seizures with secondary generalisation (tonic clonic), absence	Left posterior	Left	70
21	12	F	R	1.5	generalised (absences) plus focal onset plus complex partial seizure aura	Right temporal	Right	87
22	15.1	M	R	2	focal onset tonic with secondary generalisation	Left temporal	Left	84

24	12.9	M	L (right sided weakness)	2	focal seizures with staring	Left temporal	Left	63
29	7	F	R	0	frontal seizures with secondary generalisation	Left hemisphere	Left	81
30	8.1	F	R	5	focal	Right Frontal	Right	88
31	11.7	M	L	6	focal onset with retained awareness, occasional bilateral tonic clonic	Left Frontal	Left	60
33	12.3	F	L	0.5	focal onset motor seizures sometimes secondary bilateral convulsion	Right Frontal	Right	98
34	17.1	M	R	7	complex partial (focal)	Left Frontal	Left	72
35	17.6	M	R	5	Focal motor and focal with impaired awareness	Left Temporal	Left	83

Sz= Seizure; GS = Generalised Seizure, FS = Focal Seizure; VCI = Verbal Comprehension Index

(Healthy Control Normative mean = 100, Standard Deviation = 15); TS = Tuberos Sclerosis

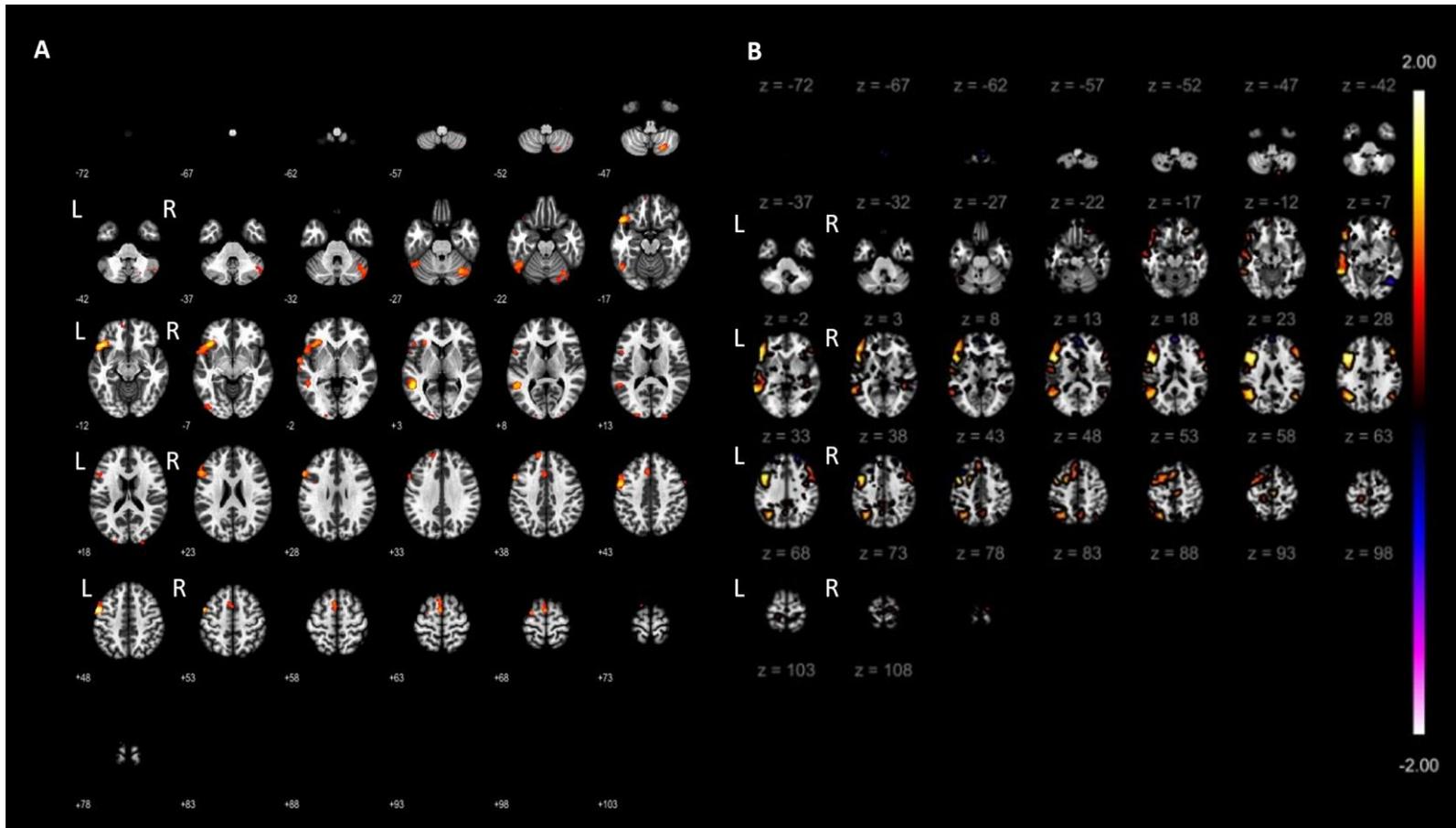


Figure 3. Illustration of visualised statistical parametric T-score (hot) maps overlaid onto International Consortium for Brain Mapping High Resolution T1 Adult Template.

Comparison between (A) Task-based maps; and (B) Resting-state maps. Left-Right orientation displayed in column 1.

Table 4. Lateralisation and Surgical Ipsilateral Indices by patient

Participant	TB LI (SD)	TB Category	RS LI (SD)	RS Category	TB SII	TB Risk Category (LT)	TB Risk Category (HT)	RS SII	RS Risk Category (LT)	RS Risk Category (HT)
1	-0.55 (0.26)	Right	0.73 (0.15)	Left	-0.55	No Risk	No Risk	0.73	Risk	Risk
2	-0.04 (0.17)	Bilateral	-0.37 (0.23)	Right	0.04	Risk	Risk	0.37	Risk	Risk
3	0.66 (0.14)	Left	-0.37 (0.22)	Right	-0.66	No Risk	No Risk	0.37	Risk	Risk
4	0.45 (0.13)	Left	0.68 (0.18)	Left	-0.45	No Risk	No Risk	-0.68	No Risk	No Risk
6	0.30 (0.23)	Left	0.27 (0.089)	Left	0.3	Risk	Risk	0.27	Risk	Risk
7	-0.29 (0.15)	Right	-0.12 (0.14)	Bilateral	0.29	Risk	Risk	0.12	Risk	Risk
8	0.77 (0.14)	Left	0.36 (0.2)	Left	0.77	Risk	Risk	0.36	Risk	Risk
9	0.12 (0.15)	Bilateral	0.00 (0.14)	Bilateral	-0.12	No Risk	Risk	-0.00	No Risk	Risk
10	0.80 (0.1)	Left	0.35 (0.16)	Left	-0.8	No Risk	No Risk	-0.35	No Risk	No Risk
11	0.74 (0.16)	Left	0.48 (0.24)	Left	-0.74	No Risk	No Risk	-0.48	No Risk	No Risk
12	0.59 (0.14)	Left	0.36 (0.12)	Left	0.59	Risk	Risk	0.36	Risk	Risk
13	0.79 (0.12)	Left	0.26 (0.21)	Left	0.79	Risk	Risk	0.26	Risk	Risk
14	0.76 (0.13)	Left	0.36 (0.18)	Left	0.76	Risk	Risk	0.36	Risk	Risk
15	0.92 (0.051)	Left	0.24 (0.046)	Left	0.92	Risk	Risk	0.24	Risk	Risk
16	0.19 (0.22)	Bilateral	0.22 (0.21)	Left	0.19	Risk	Risk	0.22	Risk	Risk
17	0.72 (0.12)	Left	0.02 (0.1)	Bilateral	-0.72	No Risk	No Risk	-0.02	No Risk	Risk
20	0.76 (0.14)	Left	0.41 (0.16)	Left	0.76	Risk	Risk	0.41	Risk	Risk
21	0.73 (0.15)	Left	0.35 (0.23)	Left	-0.73	No Risk	No Risk	-0.35	No Risk	No Risk
22	0.66 (0.19)	Left	-0.22 (0.12)	Right	0.66	Risk	Risk	-0.22	No Risk	No Risk
24	-0.07 (0.23)	Bilateral	0.61 (0.2)	Left	-0.07	No Risk	Risk	0.61	Risk	Risk
29	-0.78 (0.17)	Right	0.54 (0.22)	Left	-0.78	No Risk	No Risk	0.54	Risk	Risk
30	0.41 (0.32)	Left	0.46 (0.17)	Left	-0.41	No Risk	No Risk	-0.46	No Risk	No Risk
31	0.66 (0.21)	Left	-0.45 (0.15)	Right	0.66	Risk	Risk	-0.45	No Risk	No Risk
33	0.17 (0.39)	Bilateral	0.30 (0.18)	Left	-0.17	No Risk	Risk	-0.30	No Risk	No Risk
34	0.59 (0.21)	Left	-0.04 (0.011)	Bilateral	0.59	Risk	Risk	-0.04	No Risk	Risk
35	-0.032 (0.33)	Bilateral	-0.02	Bilateral	-0.03	No Risk	Risk	-0.02	No Risk	Risk

Index Concordance

The Bland–Altman plot is presented in Figure 4. There was a notable cluster above 0 towards the right, indicating a potential error proportional to size of measure. This was in the direction of left-lateralised cases. A simple linear regression was computed to estimate proportional bias and indicated no effect ($F(1,24)=2.76$, $p=.11$, $R^2=.103$, $R^2_{\text{adjusted}}=.07$). As is illustrated in Figure 4., 92% of the differences lie between the limits of agreement from mean (d) – 1.96 standard deviation (SD) and d + 1.96 SD.

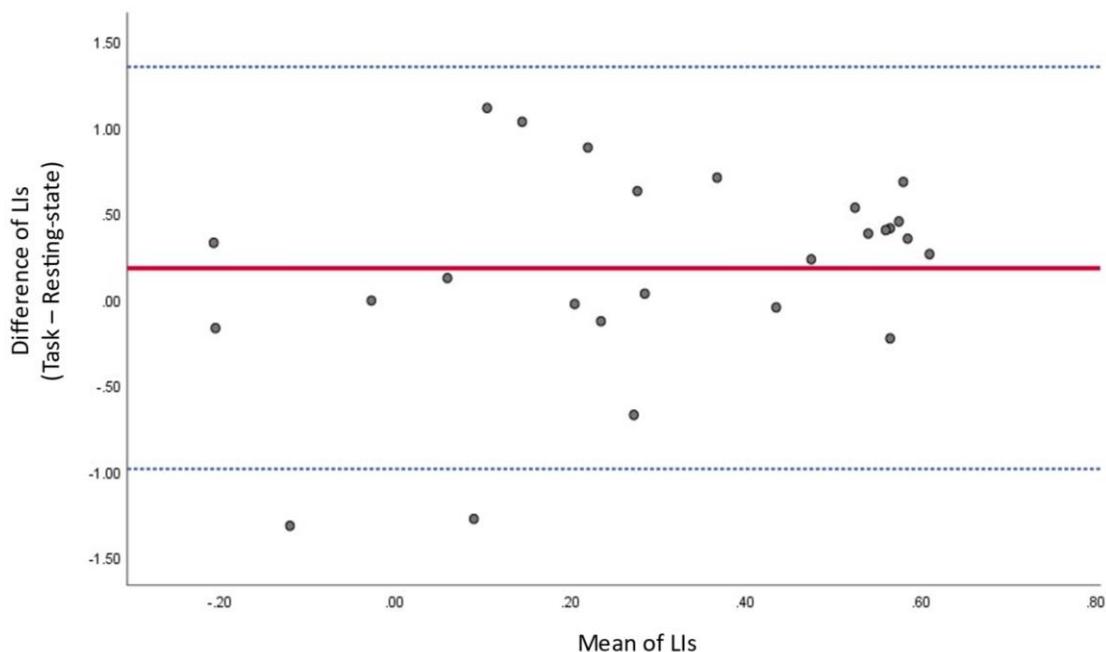


Figure 4. Bland–Altman plot was computed to illustrate the differences between the task-based and resting-state lateralisation indices.

LI = Lateralisation Index. Red continuous lines represent mean LI difference; blue dotted lines represent 95% lower and upper confidence interval limits.

There was no significant relationship between task-based and resting-state fMRI lateralisation indices, $r_s=-0.03$, 95% CI [-0.42,0.37], $p=0.88$, $n=26$. (See Figure 5.) There was no significant relationship between task-based and resting-state fMRI SII, $r_s=-0.23$, 95% CI [-0.18,0.58], $p=0.25$, $n=26$. (See Figure 6.)

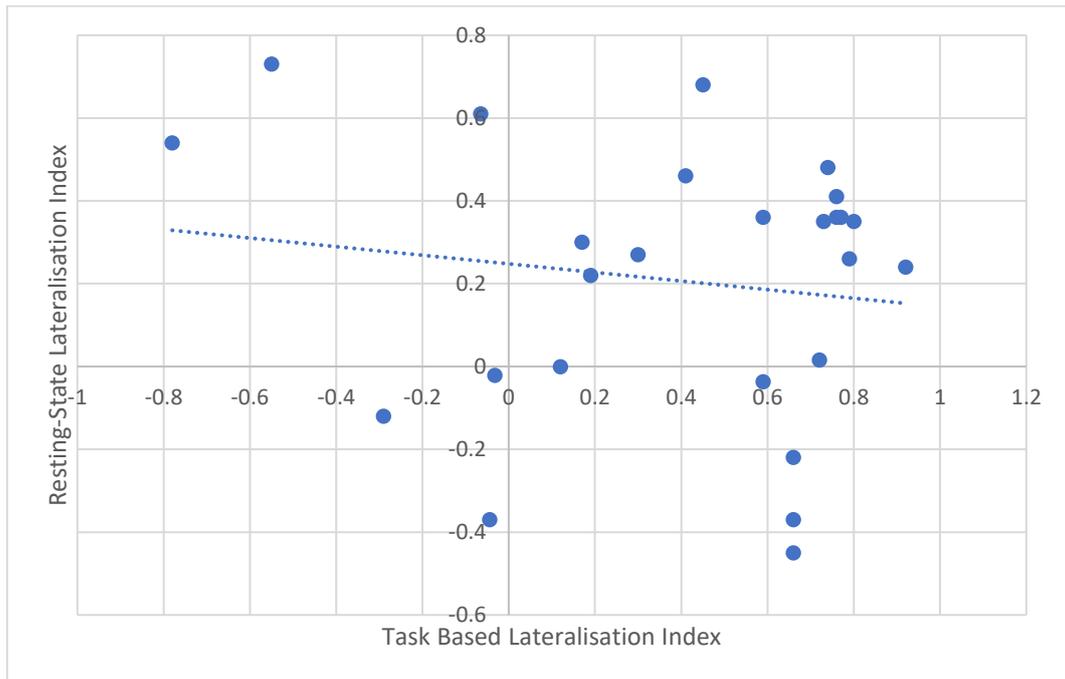


Figure 5. Scatter plot to illustrate the relationship between the task-based and resting-state lateralisation indices

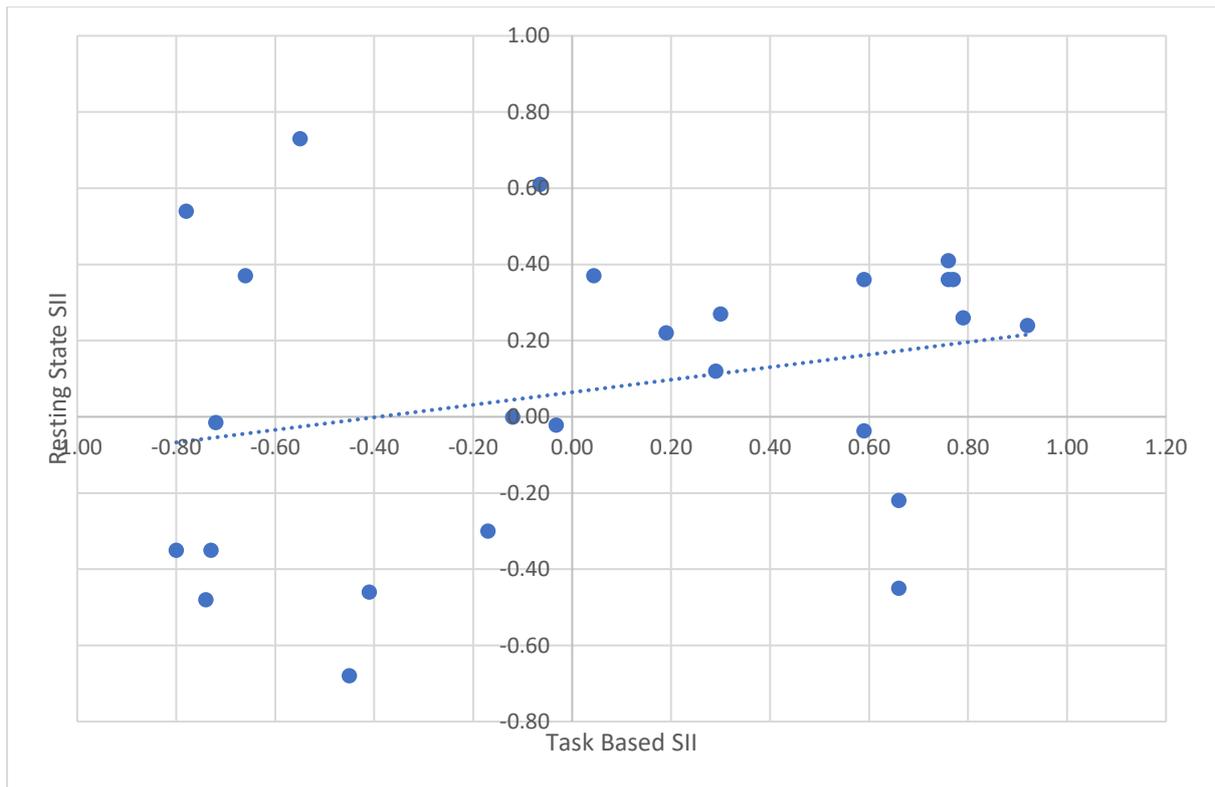


Figure 6. Scatter plot to illustrate the relationship between the task-based and resting-state Surgical Ipsilateral Indices (SII).

Categorical Concordance

Language Lateralisation

Task-based LI defined 17 patients (65.4%) as left, six (23.1%) as bilateral, and three as right (11.5%). Resting-state LI defined 17 as left (65.4%), 5 (19.2%) as bilateral, and four (15.4%) as right.

There was slight agreement (based on published criteria⁸⁸) between task-based and resting-state LIs, $\kappa=.096$, but agreement did not reach statistical significance, $p=.516$.

Surgical Risk

Task-based SII defined 13 (50%) as no risk of language decline post-surgery, 13 at risk (50%) at low threshold; nine (35%) at no risk and 17 (65%) at risk at a high threshold. Resting-state SII defined 12 (46%) at no risk, 14 at risk (54%) at low threshold; eight (31%) at no risk and 18 (69%) at risk at a high threshold.

Based on published criteria⁸⁸, there was moderately significant agreement between task-based and resting-state SIIs at the lower threshold ($\kappa=.46$, $p=.018$), and fair (approaching moderate) significant agreement between task-based and resting-state SIIs at the higher threshold ($\kappa=.39$, $p=.046$).

Validity with Post-operative outcome

Outcome data are summarised in Table 5. Task-based fMRI accurately predicted outcome in 50% of cases at low and high threshold categories. Resting-state fMRI accurately predicted outcome 66% and 83% of the time at the low (SII >0.0 = risk) and high (SII >-0.2 = risk) threshold categories, respectively.

Table 5. Post-Operative Outcome by patient with language lateralisation and surgical ipsilateral indices/categories (n=6)

Participant	Surgical Hemisphere	Pre-VCI	Post-VCI	VCI Change	Reliable Change	TB LI Cat	RS LI Cat	TB SII	RS SII	TB SII Cat LT	RS SII Cat LT	TB SII Cat HT	RS SII Cat HT
21	Right	87	87	0	No	Left	Left	-0.73	-0.35	No Risk	No Risk	No Risk	No Risk
29	Left	81	78	-3	No	Right	Left	-0.78	+0.54	No Risk	Risk	No Risk	Risk
30	Right	88	84	-4	No	Left	Left	-0.41	-0.46	No Risk	No Risk	No Risk	No Risk
31	Left	60	57	-3	No	Left	Right	0.66	-0.45	Risk	No Risk	Risk	No Risk
34	Left	72	76	4	No	Left	Bilateral	0.59	-0.04	Risk	No Risk	Risk	Risk
35	Left	83	76	-7	Yes	Bilateral	Bilateral	-0.03	-0.02	No Risk	No Risk	Risk	Risk

VCI= Verbal Comprehension Index, TB = task-based fMRI; RS = Resting-state fMRI; LI =

Language Lateralisation Index; SII = Surgical Ipsilateral Index; Cat = Category (i.e., as

categorised); LT = Low threshold (<0=no risk, >0=risk); HT = High threshold, (<-0.2=no risk, >-

0.2=risk)

Discussion

The study had three key aims: (1) to evidence generation of reliable resting-state language networks is feasible in routine practice in a diverse group of paediatric epilepsy patients; (2) examine concordance of language lateralisation and estimated surgical risk between resting-state fMRI language networks and conventional task-based fMRI methods; (3) present a small case series to explore preliminary evidence for predictive validity of resting-state fMRI in estimating language outcome.

Feasibility of Resting-State Language Network Extraction in Routine Practice

Resting-state language networks were reliably extracted from 26 of 35 patients utilising secondary task-based fMRI data sets. This finding supports previous works^{59,60,60}, demonstrating resting-state language network extraction is feasible in routine practice in a diverse group of paediatric epilepsy patients, satisfying the first aim of the paper.

Concordance of Language Lateralisation and Estimated Surgical Risk

With respect to the second aim, variable results were found for agreement between resting-state and task-based fMRI outcomes. This aim was examined both on a continuous (LI & SII) and a categorical (left/bilateral/right & risk/no risk) dimension. With relation to the continuous dimension, there was no significant association between task-based and resting-state language lateralisation indices. Additionally, there was no association between resting-state and task-based fMRI when LI was converted to a surgical ipsilateral index (SII; which quantifies the estimated extent of eloquent language cortex lateralised to the proposed surgical hemisphere).

With respect to the categorical dimension, both task-based and resting-state lateralised the majority (65%) of the patients as left-hemisphere dominant for language, bilateral was the next common categorisation in both methodologies (23%, 19% respectively), followed by right lateralisation (12%, 15% respectively). There was no statistically significant agreement between task-based and resting-state language lateralisation categorisation. However, as hypothesised, there was moderately significant agreement between task-based and resting-state SII at the lower threshold, and fair (approaching moderate) statistically significant agreement between task-based and resting-state SII at the higher threshold. Notably, the application of a risk threshold at high and low values did not change the statistical significance of this agreement but did moderate the degree. The data indicated that the resting-state SII had more conservative risk estimation than the task-based fMRI SII.

The finding of surgical risk concordance is important as it bears impact on surgical decision making and the ultimate utility of resting-state in the surgical assessment process. It is consistent with adult works, which demonstrate both task-based and resting-state fMRI have predictive validity in surgical risk assessment of language function^{50,89} The difference in findings of concordance between SII and language LI across modalities is surprising. This difference may be due to the imposed thresholding and dichotomising of outcome variables, minimising the variance displayed in the LI categorisation and the continuous scales of SII and LI. Consequently, further interrogation of larger datasets is warranted. The more conservative tolerance of risk in the data with resting-state may also represent a finding that resting-state is a more sensitive predictor of post-surgical language outcome than task-based fMRI in adult epilepsy populations⁵⁰, as was demonstrated in the post-operative data.

The statistically significant lack of concordance between resting-state and task-based fMRI LIs was unexpected. These findings of poor concordance are in contrast with findings of robust task-based and resting-state LI correlation in adults with epilepsy⁸³. There are several considerations that need to be accounted for when interpreting this discrepancy between the findings of this study and previous reports. One consideration is the difference in LI calculation. Our study's method for computing an LI may possess greater reliability (see^{84,85} for examples of robustness of bootstrap analysis approach), than the simple LI calculation used in previous adult work⁸³. In addition, the lack of association may also be due to the physiological and anatomical differences between adults and children. Maturation changes lead to differential levels of glucose consumption and blood flow across adults and children⁵¹. These developmental differences may affect the detection, magnitude, and extent of the resting-state fMRI language network in paediatric samples when compared to adults⁵¹. Challenges associated with the use of valid and developmentally appropriate paradigms for task-based fMRI, which may also have inherent confound effects of task on observed BOLD activity²⁹; and the challenges of engagement with fMRI tasks^{23,28}, may result in poorer language lateralisation in children than in adults with conventional task-based fMRI. These challenges do not apply to resting state and are less pertinent to adult task-based fMRI, and therefore concordance may be poorer in paediatric populations in comparison to adults. Also, previous studies have demonstrated differential signal intensity across resting-state and task-based fMRI^{18,29}, which may account for additional variance and poorer association in LI in the context of more atypical language representation in paediatric than in adult populations. Although other evidence demonstrates that the BOLD signal is reliable across ages⁵¹, this finding was in healthy children, and it may not generalise to children with neuropathology. Additional neuropathic developmental impacts, such as neurovascular decoupling, may have

confounded BOLD signal^{90,91}, which may differentially affect task-based and resting-state signal^{92,93}, leading to increased variance in signal across modalities. Previous studies have demonstrated differential signal intensity across resting-state and task-based fMRI^{18,29}, which may account for additional variance and poorer association in LI. The finding of poor translation to paediatric samples from adult studies would not be unique to fMRI and has been demonstrated even in 'gold-standard' techniques, such as extra-operative neurostimulation⁶ and Wada¹³.

With respect to categorical concordance, the finding of poor agreement between resting-state and task-based fMRI is inconsistent with previous results in paediatric samples⁴⁸, which report significant concordance (0.93). There are several methodological inconsistencies which may also account for lack of categorical concordance of language. The categorisation in our study was based on LI computation, whilst Desai *et al.*⁶⁰ utilised expert visual inspection to categorise laterality. Previous work has demonstrated reliably good concordance between expert visual inspection and lateralisation index categorisation^{94,95}, therefore expert analysis was not included in this study. However, the difference may originate in our methodology of categorising hemispheric LI. The use of categorical simplistic left-right-bilateral hemispheric categorisation will inevitably lead to loss of information, as will restricted inclusive masks of only Broca's and Wernicke's area, excluding other aspects of the language network. These methodologies are in contrast with the awareness that language laterality exists on a continuum^{56,96} and with recent work that has demonstrated different regional and intra-hemispheric language representation can exist^{66,97-99}. This information may have been accounted for by visual inspection, potentially accounting for discrepancy in findings. Another potential reason for lack of categorical concordance is the metric of concordance differed across studies. Desai *et al.*⁶⁰ utilised Fisher's exact test. This statistic assumes that the

comparator (task-based fMRI) is correct 100% of the time; however, the data does not support this assumption¹³. The significance of this is a proportion of the time task-based fMRI may be correct (or not) by chance, potentially artificially elevating concordance. Cohen's kappa, however, accounts for this by computing concordance rates above chance and thus may produce lower rates of concordance⁸⁷. Another explanation for the lower rates of concordance may be the younger sample included in this study. Desai *et al.*⁶⁰ only included adolescents. Previous studies have demonstrated a stratification of language dominance across age, notable with those younger than ten having a higher prevalence of atypical language representation, whilst those over ten demonstrate patterns more consistent with adults⁵²⁻⁵⁵. Whilst this study was insufficiently powered to explore age as a covariate in the analysis, or to examine differences between pre-adolescent and adolescent groups, it may be a reasonable explanation for inconsistencies with previous reports⁶⁰.

The lack of concordance does not necessarily mean resting-state is an inferior technique in paediatric samples. In the case of aberrated BOLD signal, it is supposed resting-state may be a more robust metric⁹³ due its better SNR^{18,29,93}. Functionally, resting-state is also more able to detect subtle differences in language network configuration, that might not be available through standard task-based fMRI investigations, as suggested in previous works^{47,100,101}. These language network configuration differences are more likely to be seen in paediatric samples³⁶⁻³⁸. Additionally, although the covert-verb generation paradigm has previously been demonstrated to activate both expressive (i.e., frontal) and receptive (i.e., temporal) language systems (as it did in our sample, Figure 3), it is biased towards the frontal lobes and may not adequately assess the contribution of posterior language systems^{47,53,68,69,102}. Collectively, these may lead to significant variance in LI, leading to poor associations and agreement between the measures.

In the small case series of post-operative patients, task-based fMRI accurately predicted outcome 50% of the time at the low and high threshold categories. Resting-state fMRI accurately predicted outcome 66% of the time at the low threshold category, and 83% of the time at the high threshold category.

There are several reasons why resting-state fMRI has demonstrated better predictive ability in this small case series, over conventional task-based fMRI, and why it may be a superior tool in clinical practice. Resting-state fMRI BOLD signal oscillations can provide up to three times higher signal-to-noise ratio than task-related signal increases^{18,29}, suggesting a potential better sensitivity to language-based activity. The absence of task requirements also overcomes the challenge of finding valid and developmentally appropriate paradigms, which do not have inherent confound effects of task on observed BOLD activity²⁹; and the challenges of engagement with fMRI tasks^{23,28}. Resting-state is also more able to detect subtle differences across the entire language network, which might not be available through standard task-based fMRI investigations, as suggested in previous works^{47,100,101}. Resting-state fMRI holds improved time-efficiency over task-based fMRI, with scans collected more rapidly (~5minutes) than many task-based protocols and one resting-state scan serves multiple mapping purposes³⁹, thus placing less demands on limited scanner time¹⁸.

Although this finding demonstrates promise as a tool in surgical decision making and gives merit to the further study of resting-state methodology, the conclusions here are significantly limited due to the small number and only a single case demonstrating post-operative decline allowing for the quantification of true positive results. This will likely be an ongoing issue as if eloquent cortex is identified it is often the case a more conservative resection is undertaken, if any, reducing the prevalence of post-operative decline.

Additional Limitations

There are some limitations in the task-based methodology. Although coaching and pre/post-scan assessment of engagement took place, measures of task participation were not taken due to the covert nature of the paradigm, and therefore it is uncertain whether patients were able to perform the task correctly during scanning. Previous data has suggested in-scanning performance is worse than practice¹⁰³. However, children generally perform well at this task, with robust activation being demonstrated^{53,68}. In a recent study, the task was performed successfully by young children with epilepsy, with at least 95% of the trials correctly completed⁶⁶. Notwithstanding, best practice does advocate for overt speech, as task engagement has been found to be a greater confound than motion in paediatric epilepsy samples¹⁰³. This study was limited in paradigm selection due to the retrospective nature of the data. Additionally, only one task was utilised to extract a language network. Whilst previous work has demonstrated verb generation is one of the most reliable language tasks to predict hemispheric dominance^{104,105}, conferring the validity of this study's approach, several authors have demonstrated the need for multiple tasks to demonstrate reliable networks in the individual case¹⁰⁶⁻¹⁰⁹.

The use of extracted 'rest' from task-based scans is a methodological limitation. Although there is growing use and evidence to support this approach demonstrating extracted rest blocks are suited to resting-state connectivity analysis¹¹⁰⁻¹¹², several quantitative differences have been found within these networks when compared to continuous resting-state scans¹¹⁰. Extracted rest scans have been shown to have poorer network reliability, notably in the individual case¹¹¹, as is the approach adopted here. Additionally, residuals from task have still

been demonstrated following rest extraction¹¹¹, which may confound signal. For example, in this paper, the involvement of motor residuals may confound resting-state signal and misrepresent language network localisation.

Another limitation is the use of adult template brains. Anatomical differences between child and adult brains can distort the magnitude of signal and the functional localisation, notably when data are warped onto adult anatomical atlases for analysis⁸⁴

A further limitation of this study is the restricted number of right-sided (<~4) cases. Although the number of left lateralised cases by both resting-state and task-based LI didn't reach statistical significance for proportional bias in the logistic regression, it limits clinical applicability of the findings. The finding of low atypical representation incidence in our sample is inconsistent with previous prevalence rates of atypical hemispheric and atypical regional language representation occurring in up to 70% of individuals with epilepsy^{13,52,56,94,96,106,113–124} and is likely a consequence of the small sample size. Thus, there is limited generalisability of the conclusions of this work to those who are right hemisphere dominant for language, which represent a large proportion of epilepsy patients.

Future Directions

Inevitably, replication and studies with larger sample sizes and dedicated continuous resting-state scans will be needed to explore the reliability of agreement between the measures, due to potential systematic bias and proportionally an unacceptable number of outliers¹²⁵ (8% falling outside the confidence limits of the Bland-Altman plot) within our small dataset. In

larger samples, these factors may be normalised. Further studies should be completed in order to contribute to future meta-analytic review and inform practice guidelines.

Future research should focus on recruiting significant representation in younger (<12 years) age groups and those with atypical language patterns. Work to explore differences between pre-adolescent, adolescent and adult groups will aid our understanding of the limitations of resting-state and underlying mechanisms between potential differences in these groups.

The use of computational methods for determining language laterality holds benefit over visual inspection due to the time needed to develop the expertise required for expert visual inspection. Future work should continue to adopt this approach but should consider the use of mapping language individually using multiple seed points that tap the whole language network and examine intrahemispheric connectivity, as in recent studies^{83,126,127}.

Lastly, the systematic collection and reporting of post-operative outcome data is essential in examining the predictive validity of these methodologies and should be a key focus of future work. Further study should explore both the ability to predict decline but also the predictive validity of improved function following surgery, so as best to inform patient, family, and professional decision-making.

Post-operative assessment should evaluate all aspects of cognitive and behavioural function assessed prior to surgery¹²⁸, which should include post-operative fMRI. This will allow assessment of language network integrity post-operatively, in comparison to pre-operative status and would help the clinician consider (in the context of wider clinical factors), the likelihood of ongoing language deficit and treatment planning. Areas of cognitive and behavioural domains to be assessed, with associated tools, are outlined in the Children's Epilepsy Surgery Service Protocol in the Appendix. Post-operative assessment should include

parental or caregiver report of behaviour, and cognitive and academic ability¹²⁸. Teacher/educator evaluations are also important to ensure ongoing academic attainment and evaluate skills across settings¹²⁸. Epilepsy-related quality of life measures are an essential part of the assessment¹²⁸, and dedicated tools are available, which consider the various impacts on quality of health, activity, and participation in society (see Appendix).

Alongside the pre-operative factors (e.g., length of epilepsy, age of onset, etc.), there are several additional factors that need to be considered in interpreting post-operative outcome results, including: the nature, timing and extent of the surgery, seizure control, rehabilitation, educational support and family functioning¹²⁸. The timing of the postoperative assessment will have a bearing on the clinical interpretation and significance of the results. Due to many factors (e.g., Wallerian degeneration, resolution of post-surgical neuroinflammation, brain-blood barrier rupture and repair etc), the longer-term the follow-up assessment, the more accurate the results are at reflecting the long-term post-operative outcome¹²⁸. In some instances, it can take at least 5 years after the surgery for quantifiable changes in health-related quality of life and cognitive changes to emerge¹²⁸.

Conclusions

Though there is promising evidence that resting-state fMRI may be useful in surgical decision making, at the current time there is insufficient evidence for use of resting-state fMRI as a proxy to task-based fMRI for the lateralisation of language function in paediatric epilepsy surgery candidates.

The limitations of the current study and previous evidence restrict current recommendation for its use in routine clinical practice. However, this study does provide preliminary evidence that the predictive validity of resting-state for post-operative language outcome warrants further investigation.

Disclosures

Alexander Marsh has no conflicts of interest to disclose. Alexander Marsh confirms that he has read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

1. Richardson, M. P. Large scale brain models of epilepsy: dynamics meets connectomics. *J. Neurol. Neurosurg. Psychiatry* **83**, 1238–1248 (2012).
2. Fisher, R. S. *et al.* Instruction manual for the ILAE 2017 operational classification of seizure types. *Epilepsia* **58**, 531–542 (2017).
3. Braun, K. P. J. Influence of epilepsy surgery on developmental outcomes in children. *Eur. J. Paediatr. Neurol.* **24**, 40–42 (2020).
4. Lamberink, H. J., Boshuisen, K., Rijen, P. C. van, Gosselaar, P. H. & Braun, K. P. J. Changing profiles of pediatric epilepsy surgery candidates over time: A nationwide single-center experience from 1990 to 2011. *Epilepsia* **56**, 717–725 (2015).
5. Baud, M. O. *et al.* European trends in epilepsy surgery. *Neurology* **91**, e96–e106 (2018).
6. Schevon, C. A. *et al.* Pediatric language mapping: sensitivity of neurostimulation and Wada testing in epilepsy surgery. *Epilepsia* **48**, 539–545 (2007).
7. Baumer, F. M., Cardon, A. L. & Porter, B. E. Language Dysfunction in Pediatric Epilepsy. *J. Pediatr.* **194**, 13–21 (2018).
8. Overvliet, G. M. *et al.* Nocturnal epileptiform EEG discharges, nocturnal epileptic seizures, and language impairments in children: Review of the literature. *Epilepsy Behav.* **19**, 550–558 (2010).
9. Overvliet, G. M. *et al.* Early onset of cortical thinning in children with rolandic epilepsy. *NeuroImage Clin.* **2**, 434–439 (2013).

10. Beers, C. A. & Federico, P. Functional MRI applications in epilepsy surgery. [Review]. *J. Neurol. Sci.* **39**, 271–285 (2012).
11. Collinge, S. *et al.* Pre-surgical mapping of eloquent cortex for paediatric epilepsy surgery candidates: Evidence from a review of advanced functional neuroimaging. [Review]. *Seizure* **1**, 136–146 (2017).
12. Cross, J. H. & Vezyroglou, K. Selecting the right paediatric patient for epilepsy surgery. *Paediatr. Child Health* **27**, 287–291 (2017).
13. Szaflarski, J. P. *et al.* Practice guideline summary: Use of fMRI in the presurgical evaluation of patients with epilepsy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* **88**, 395–402 (2017).
14. Alpherts, W. C. J. *et al.* Standard versus tailored left temporal lobe resections: differences in cognitive outcome? *Neuropsychologia* **46**, 455–460 (2008).
15. Arya, R. *et al.* Neuropsychological outcomes after resection of cortical sites with visual naming associated electrocorticographic high-gamma modulation. *Epilepsy Res.* **151**, 17–23 (2019).
16. Babajani-Feremi, A. *et al.* Predicting postoperative language outcome using presurgical fMRI, MEG, TMS, and high gamma ECoG. *Clin. Neurophysiol.* **129**, 560–571 (2018).
17. Benjamin, C. F. A. *et al.* Presurgical language fMRI: Clinical practices and patient outcomes in epilepsy surgical planning. *Hum. Brain Mapp.* **39**, 2777–2785 (2018).
18. Branco, P. *et al.* Resting-State Functional Magnetic Resonance Imaging for Language Preoperative Planning. *Front. Hum. Neurosci.* **10**, (2016).

19. Barnett, A., Audrain, S. & McAndrews, M. P. Applications of Resting-State Functional MR Imaging to Epilepsy. [Review]. *Neuroimaging Clin. N. Am.* **27**, 697–708 (2017).
20. de Ribaupierre, S., Wang, A. & Hayman-Abello, S. Language Mapping in Temporal Lobe Epilepsy in Children: Special Considerations. *Epilepsy Res. Treat.* **2012**, e837036 (2012).
21. Wilke, M. *et al.* Clinical functional MRI of the language domain in children with epilepsy. *Hum. Brain Mapp.* **32**, 1882–1893 (2011).
22. O’Shaughnessy, E. S., Berl, M. M., Moore, E. N. & Gaillard, W. D. Pediatric functional magnetic resonance imaging (fMRI): Issues and applications. *J. Child Neurol.* **23**, 791–801 (2008).
23. Yerys, B. E. *et al.* The fMRI success rate of children and adolescents: typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. *Hum. Brain Mapp.* **30**, 3426–3435 (2009).
24. Byars, A. W. *et al.* Practical Aspects of Conducting Large-Scale Functional Magnetic Resonance Imaging Studies in Children. *J. Child Neurol.* **17**, 885–889 (2002).
25. Meyer, M., Friederici, A. D. & von Cramon, D. Y. Neurocognition of auditory sentence comprehension: event related fMRI reveals sensitivity to syntactic violations and task demands. *Cogn. Brain Res.* **9**, 19–33 (2000).
26. Booth, J. R. *et al.* Relation between brain activation and lexical performance. *Hum. Brain Mapp.* **19**, 155–169 (2003).
27. Herting, M. M., Gautam, P., Chen, Z., Mezher, A. & Vetter, N. C. Test-retest reliability of longitudinal task-based fMRI: Implications for developmental studies. *Dev. Cogn. Neurosci.* **33**, 17–26 (2018).

28. Kesavadas, C. & Thomas, B. Clinical applications of functional MRI in epilepsy. *Indian J. Radiol. Imaging* **18**, 210–217 (2008).
29. Fox, M. D. & Greicius, M. Clinical applications of resting state functional connectivity. *Front. Syst. Neurosci.* **4**, 19 (2010).
30. Damoiseaux, J. S. *et al.* Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci.* **103**, 13848–13853 (2006).
31. Fox, M. D. & Raichle, M. E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* **8**, 700–711 (2007).
32. Horowitz, S. G. *et al.* Low frequency BOLD fluctuations during resting wakefulness and light sleep: A simultaneous EEG-fMRI study. *Hum. Brain Mapp.* **29**, 671–682 (2008).
33. Fukunaga, M. *et al.* Large-amplitude, spatially correlated fluctuations in BOLD fMRI signals during extended rest and early sleep stages. *Magn. Reson. Imaging* **24**, 979–992 (2006).
34. Steriade, M. Impact of network activities on neuronal properties in corticothalamic systems. *J. Neurophysiol.* **86**, 1–39 (2001).
35. John, E. R. *et al.* Invariant reversible QEEG effects of anesthetics. *Conscious. Cogn.* **10**, 165–183 (2001).
36. Schlaggar, B. L. *et al.* Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science* **296**, 1476–1479 (2002).
37. Brown, T. T. *et al.* Developmental changes in human cerebral functional organization for word generation. *Cereb. Cortex N. Y. N 1991* **15**, 275–290 (2005).

38. Wood, A. G. *et al.* Language cortex activation in normal children. *Neurology* **63**, 1035–1044 (2004).
39. Lang, S., Duncan, N. & Northoff, G. Resting-State Functional Magnetic Resonance Imaging: Review of Neurosurgical Applications. *Neurosurgery* **74**, 453–465 (2014).
40. Lv, H. *et al.* Resting-State Functional MRI: Everything That Nonexperts Have Always Wanted to Know. *Am. J. Neuroradiol.* ajnr;ajnr.A5527v1 (2018) doi:10.3174/ajnr.A5527.
41. Kiviniemi, V., Kantola, J.-H., Jauhiainen, J., Hyvärinen, A. & Tervonen, O. Independent component analysis of nondeterministic fMRI signal sources. *NeuroImage* **19**, 253–260 (2003).
42. van de Ven, V. G., Formisano, E., Prvulovic, D., Roeder, C. H. & Linden, D. E. J. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. *Hum. Brain Mapp.* **22**, 165–178 (2004).
43. Beckmann, C. F., DeLuca, M., Devlin, J. T. & Smith, S. M. Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 1001–1013 (2005).
44. Smitha, K. A., Arun, K. M., Rajesh, P. G., Thomas, B. & Kesavadas, C. Resting-state seed-based analysis: an alternative to task-based language fMRI and its laterality index. *Am. J. Neuroradiol.* **38**, 1187–1192 (2017).
45. Zhu, L. *et al.* Temporal Reliability and Lateralization of the Resting-State Language Network. *PLoS ONE* **9**, e85880 (2014).
46. Tomasi, D. & Volkow, N. D. Network connectivity predicts language processing in healthy adults. *Hum. Brain Mapp.* **41**, 3696–3708 (2020).

47. Doucet, G. E. *et al.* Resting-state functional connectivity predicts the strength of hemispheric lateralization for language processing in temporal lobe epilepsy and normals. *Hum. Brain Mapp.* **36**, 288–303 (2015).
48. DeSalvo, M. N. *et al.* Resting-State Functional MR Imaging for Determining Language Laterality in Intractable Epilepsy. *Radiology* **281**, 264–269 (2016).
49. Mitchell, T. J. *et al.* A Novel Data-Driven Approach to Preoperative Mapping of Functional Cortex Using Resting-State Functional Magnetic Resonance Imaging. *Neurosurgery* **73**, 969–983 (2013).
50. Audrain, S., Barnett, A. J. & McAndrews, M. P. Language network measures at rest indicate individual differences in naming decline after anterior temporal lobe resection. *Hum. Brain Mapp.* **39**, 4404–4419 (2018).
51. Gaillard, W. D. *et al.* Cortical localization of reading in normal children: an fMRI language study. *Neurology* **57**, 47–54 (2001).
52. Yuan, W. *et al.* fMRI Shows Atypical Language Lateralization in Pediatric Epilepsy Patients. *Epilepsia* **47**, 593–600 (2006).
53. Holland, S. K. *et al.* Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage* **14**, 837–843 (2001).
54. Schapiro, M. B. *et al.* BOLD-fMRI signal increases with age in selected brain regions in children. *Neuroreport* **15**, 2575–2578 (2004).
55. Szaflarski, J. P., Holland, S. K., Schmithorst, V. J. & Byars, A. W. fMRI study of language lateralization in children and adults. *Hum. Brain Mapp.* **27**, 202–212 (2005).

56. Hadac, J., Brozova, K., Tintera, J. & Krsek, P. Language lateralization in children with pre and postnatal epileptogenic lesions of the left hemisphere: An fMRI study. *Epileptic. Disord.* S19–S27 (2007).
57. Kadis, D. S. *et al.* Intrahemispheric reorganization of language in children with medically intractable epilepsy of the left hemisphere. *J. Int. Neuropsychol. Soc.* **13**, 505–516 (2007).
58. Berger, M. S., Kincaid, J., Ojemann, G. A. & Lettich, E. Brain mapping techniques to maximize resection, safety, and seizure control in children with brain tumors. *Neurosurgery* **25**, 786–792 (1989).
59. Nath, A. *et al.* Determination of Differences in Seed-Based Resting State Functional Magnetic Resonance Imaging Language Networks in Pediatric Patients with Left- and Right-Lateralized Language: A Pilot Study. *J. Epilepsy Res.* **9**, 93–102 (2019).
60. Desai, V. R. *et al.* Language lateralization with resting-state and task-based functional MRI in pediatric epilepsy. *J. Neurosurg. Pediatr.* **23**, 171–177 (2019).
61. Cross, J. H. *et al.* Proposed Criteria for Referral and Evaluation of Children for Epilepsy Surgery: Recommendations of the Subcommission for Pediatric Epilepsy Surgery. *Epilepsia* **47**, 952–959 (2006).
62. Whitfield-Gabrieli, S. & Nieto-Castanon, A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* **2**, 125–141 (2012).
63. Wechsler, D. *Wechsler Preschool & Primary Scale of Intelligence - Fourth UK Edition (WPPSI-IV UK)*. (Pearson Clinical UK, 2013).

64. Wechsler, D. *Wechsler Intelligence Scale for Children*® - Fourth UK Edition (WISC-IV UK). (Pearson Clinical UK, 2004).
65. Wechsler, D. *Wechsler Intelligence Scale for Children*® - Fifth UK Edition (WISC-V UK). (Pearson Clinical UK, 2016).
66. Foley, E. *et al.* MEG Assessment of Expressive Language in Children Evaluated for Epilepsy Surgery. *Brain Topogr.* **32**, 492–503 (2020).
67. Crawford, J. R. & Garthwaite, P. H. Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology* **21**, 611–620 (2007).
68. Liegeois, F. *et al.* Language reorganization in children with early-onset lesions of the left hemisphere: An fMRI study. *J. Neurol.* **127**, 1229–1236 (2004).
69. Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* **331**, 585–589 (1988).
70. Ashburner, J. *et al.* SPM12 manual. *Wellcome Trust Cent. Neuroimaging Lond. UK* **2464**, (2014).
71. Matlab. *Matlab Version R2021a*. (The Mathworks Inc., 2021).
72. Andersson, J. L., Hutton, C., Ashburner, J., Turner, R. & Friston, K. Modeling geometric deformations in EPI time series. *NeuroImage* **13**, 903–919 (2001).
73. Henson, R. N. A., Buechel, C., Josephs, O. & Friston, K. J. The slice-timing problem in event-related fMRI. *NeuroImage* **9**, 125- (1999).

74. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).
75. Hagler, D. J., Saygin, A. P. & Sereno, M. I. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage* **33**, 1093–1103 (2006).
76. Friston, K. J. *et al.* Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2**, 189–210 (1994).
77. Behzadi, Y., Restom, K., Liu, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007).
78. Chai, X. J., Castañón, A. N., Ongür, D. & Whitfield-Gabrieli, S. Anticorrelations in resting state networks without global signal regression. *NeuroImage* **59**, 1420–1428 (2012).
79. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. & Turner, R. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* **35**, 346–355 (1996).
80. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341 (2014).
81. Hallquist, M. N., Hwang, K. & Luna, B. The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *NeuroImage* **82**, 208–225 (2013).
82. Rolls, E. T., Huang, C.-C., Lin, C.-P., Feng, J. & Joliot, M. Automated anatomical labelling atlas 3. *NeuroImage* **206**, 116189 (2020).
83. Rolinski, R. *et al.* Language lateralization from task-based and resting state functional MRI in patients with epilepsy. *Hum. Brain Mapp.* **41**, 3133–3146 (2020).

84. Wilke, M. & Lidzba, K. LI-tool: a new toolbox to assess lateralization in functional MR-data. *J. Neurosci. Methods* **163**, 128–136 (2007).
85. Wilke, M. & Schmithorst, V. J. A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *NeuroImage* **33**, 522–530 (2006).
86. Seghier, M. L. Laterality index in functional MRI: methodological issues. *Magn. Reson. Imaging* **26**, 594–601 (2008).
87. Kwiecien, R., Kopp-Schneider, A. & Blettner, M. Concordance Analysis. *Dtsch. Ärztebl. Int.* **108**, 515–521 (2011).
88. Landis, J. R. & Koch, G. G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **33**, 159 (1977).
89. You, X. *et al.* fMRI prediction of naming change after adult temporal lobe epilepsy surgery: Activation matters. *Epilepsia* **60**, 527–538 (2019).
90. Ulmer, J. L. *et al.* Lesion-induced Pseudo-dominance at Functional Magnetic Resonance Imaging: Implications for Preoperative Assessments. *Neurosurgery* **55**, 569–581 (2004).
91. Cho, N. S., Peck, K. K., Zhang, Z. & Holodny, A. I. Paradoxical Activation in the Cerebellum During Language fMRI in Patients with Brain Tumors: Possible Explanations Based on Neurovascular Uncoupling and Functional Reorganization. *The Cerebellum* **17**, 286–293 (2018).
92. Wang, J. *et al.* Altered coupling between resting-state glucose metabolism and functional activity in epilepsy. *Ann. Clin. Transl. Neurol.* **7**, 1831–1842 (2020).

93. Agarwal, S. *et al.* Demonstration of Brain Tumor-Induced Neurovascular Uncoupling in Resting-State fMRI at Ultrahigh Field. *Brain Connect.* **6**, 267–272 (2016).
94. You, X. *et al.* A decisional space for fMRI pattern separation using the principal component analysis-A comparative study of language networks in pediatric epilepsy. *Hum. Brain Mapp.* **34**, 2330–2342 (2013).
95. Anderson, D. P. *et al.* FMRI lateralization of expressive language in children with cerebral lesions. *Epilepsia* **47**, 998–1008 (2006).
96. Springer, J. A. *et al.* Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *J. Neurol.* **122**, 2033–2046 (1999).
97. Tailby, C., Abbott, D. F. & Jackson, G. D. The diminishing dominance of the dominant hemisphere: Language fMRI in focal epilepsy. *NeuroImage Clin.* **1**, 141–150 (2017).
98. Balter, S., Lin, G., Leyden, K. M., Paul, B. M. & McDonald, C. R. Neuroimaging correlates of language network impairment and reorganization in temporal lobe epilepsy. *Brain Lang.* **193**, 31–44 (2019).
99. Berl, M. M. *et al.* Characterization of atypical language activation patterns in focal epilepsy. *Ann. Neurol.* **75**, 33–42 (2014).
100. Doucet, G., Osipowicz, K., Sharan, A., Sperling, M. R. & Tracy, J. I. Hippocampal Functional Connectivity Patterns During Spatial Working Memory Differ in Right Versus Left Temporal Lobe Epilepsy. *Brain Connect.* **3**, 398–406 (2013).
101. Waites, A. B., Briellmann, R. S., Saling, M. M., Abbott, D. F. & Jackson, G. D. Functional Connectivity Networks Are Disrupted in Left Temporal Lobe Epilepsy. *Ann. Neurol.* **59**, 335–343 (2006).

102. Fiez, J. A., Raichle, M. E., Balota, D. A., Tallal, P. & Petersen, S. E. PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cereb. Cortex* **6**, 1–10 (1996).
103. Croft, L. J. *et al.* To speak, or not to speak? The feasibility of imaging overt speech in children with epilepsy. *Epilepsy Res.* **107**, 195–199 (2013).
104. Harrington, G. S., Buonocore, M. H. & Farias, S. T. Intrasubject reproducibility of functional MR imaging activation in language tasks. *Am. J. Neuroradiol.* **27**, 938–944 (2006).
105. Ruff, I. M. *et al.* Predictors of naming decline after dominant temporal lobectomy: age at onset of epilepsy and age of word acquisition. *Epilepsy Behav.* **10**, 272–277 (2007).
106. Gaillard, W. D. *et al.* Language dominance in partial epilepsy patients identified with an fMRI reading task. *Neurology* **59**, 256–265 (2002).
107. Wilke, M. *et al.* Why one task is not enough: Functional MRI for atypical language organization in two children. *Eur. J. Paediatr. Neurol.* **14**, 474–478 (2010).
108. Niskanen, E. *et al.* The effect of fMRI task combinations on determining the hemispheric dominance of language functions. *Neuroradiology* **54**, 393–405 (2012).
109. Alemi, R., Batouli, S. A. H., Behzad, E., Ebrahimipoor, M. & Oghabian, M. A. Not single brain areas but a network is involved in language: Applications in presurgical planning. *Clin. Neurol. Neurosurg.* **165**, 116–128 (2018).
110. Fair, D. A. *et al.* A METHOD FOR USING BLOCKED AND EVENT-RELATED FMRI DATA TO STUDY “RESTING STATE” FUNCTIONAL CONNECTIVITY. *NeuroImage* **35**, 396–405 (2007).

111. Ganger, S. *et al.* Comparison of continuously acquired resting state and extracted analogues from active tasks: Extracting Resting-State From Task Paradigms. *Hum. Brain Mapp.* **36**, 4053–4063 (2015).
112. Braun, U. *et al.* Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11678–11683 (2015).
113. Dym, R. J., Burns, J., Freeman, K. & Lipton, M. L. Is functional MR imaging assessment of hemispheric language dominance as good as the Wada test?: a meta-analysis. [Review]. *Radiology* **261**, 446–455 (2011).
114. Binder, J. R. *et al.* Determination of language dominance using functional MRI: A comparison with the Wada test. *Neurology* **46**, 978–984 (1996).
115. Wellmer, J. *et al.* Unilateral intracarotid amobarbital procedure for language lateralization. *Epilepsia* **46**, 1764–1772 (2005).
116. Fernández, G. *et al.* Language mapping in less than 15 minutes: real-time functional MRI during routine clinical investigation. *NeuroImage* **14**, 585–594 (2001).
117. Frost, J. A. *et al.* Language processing is strongly left lateralized in both sexes. Evidence from functional MRI. *Brain J. Neurol.* **122 (Pt 2)**, 199–208 (1999).
118. Conradi, N. *et al.* Hemispheric language lateralization in presurgical patients with temporal lobe epilepsy: Improving the retest reliability of functional transcranial Doppler sonography. *Epilepsy Behav.* 48–52 (2019).
119. Gaillard, W. D. *et al.* fMRI language task panel improves determination of language dominance. *Neurology* **63**, 1403–1408 (2004).

120. Gaillard, W. D. *et al.* Atypical language in lesional and nonlesional complex partial epilepsy. *Neurology* **69**, 1761–1771 (2007).
121. Ramsey, N. F., Sommer, I. E. C., Rutten, G. J. & Kahn, R. S. Combined Analysis of Language Tasks in fMRI Improves Assessment of Hemispheric Dominance for Language Functions in Individual Subjects. *NeuroImage* **13**, 719–733 (2001).
122. Spreer, J. & Quiske, A. Unsuspected Atypical Hemispheric Dominance for Language as Determined by fMRI: Reply. *Epilepsia* **44**, 734–735 (2003).
123. Spreer, J. *et al.* Determination of hemisphere dominance for language: comparison of frontal and temporal fMRI activation with intracarotid amytal testing. *Neuroradiology* **44**, 467–474 (2002).
124. Woermann, F. G. *et al.* Language lateralization by Wada test and fMRI in 100 patients with epilepsy. *Neurology* **61**, 699–701 (2003).
125. Riffenburgh, R. H. & Gillen, D. L. Techniques to Aid Analysis. in *Statistics in Medicine* 631–649 (Elsevier, 2020). doi:10.1016/B978-0-12-815328-4.00027-9.
126. Joliot, M., Tzourio-Mazoyer, N. & Mazoyer, B. Intra-hemispheric intrinsic connectivity asymmetry and its relationships with handedness and language Lateralization. *Neuropsychologia* **93**, 437–447 (2016).
127. Pravatà, E. *et al.* Functional Connectivity MR Imaging of the Language Network in Patients with Drug-Resistant Epilepsy. *Am. J. Neuroradiol.* **32**, 532–540 (2011).
128. Baxendale, S., Wilson, S.J., Baker, G.A., Barr, W., Helmstaedter, C., Hermann, B.P., Langfitt, J., Reuner, G., Rzezak, P., Samson, S. and Smith, M.L., 2019. Indications and expectations for neuropsychological assessment in epilepsy surgery in children and adults:

Report of the ILAE Neuropsychology Task Force Diagnostic Methods Commission: 2017–2021

Neuropsychological assessment in epilepsy surgery. *Epileptic Disorders*, 21(3), pp.221-234.

Appendix

Submission Guidelines for *Epilepsia*

Full-length Original Research.

These articles should be limited in length to 4000 words, 50 references, and no more than 6 figures and tables (combined). Additional figures and tables will be permitted at the discretion of the Editors or can be submitted for “online only” Supporting Information (which will be linked to the online version of the published article). Authors should aim for presenting material clearly and completely, in the most concise and direct form possible; the Introduction section should be brief (typically less than 600 words), and the Discussion section should be restricted to issues directly relevant to the Results (typically less than 1200 words).

General Style Guidelines

Manuscripts are to be submitted (and will be published) in English. Writers not fluent in English should seek assistance to ensure proper grammar and syntax and to help generate a manuscript organization that facilitates reader understanding. Authors for whom English is a second language may choose to have their manuscript professionally edited before submission to improve the English. A list of independent suppliers of editing services can be found at <https://wileyeditingservices.com/en/>. All services are paid for and arranged by the author and use of one of these services does not guarantee acceptance or preference for publication. The Editors will not rewrite papers submitted in unacceptable English and will return such manuscripts for revision before sending them out for review. Use international nonproprietary (generic) names when referring to drugs; avoid proprietary (brand) names. All acronyms should be spelled out at first mention. Make sure to spell out all abbreviations at

first use in summary and again in the body of the manuscript. Also spell out any abbreviations in figures and tables in legends and footnotes, respectively. Spell out numbers below 10 and all numbers that are used to begin a sentence; use Arabic numerals for numbers 10 or larger and for units of measure. Confirm that the correct names of tests, agencies, organizations, and manufacturers are being provided. Confirm that data that are presented in the manuscript are consistent in all parts of the manuscript: numbers, percentages, and so on. Numbers should be checked to be sure they add up correctly. Confirm that all tables and figures are correctly cited in text and numbered in the order that they appear and that all references are correctly cited in text. Locations for manufacturers are not required. Manuscript text should be double spaced with at least a 1-inch margin on all sides using size 12 font. Word limits for each type of submission will generally be enforced unless there are good reasons not to do so. If manuscripts exceed these guidelines, authors should submit a cover letter explaining why the additional length is necessary. Authors are encouraged to use the most recent terminology of seizures and epilepsy.

Full-Length Original Research, Special Report, and Brief Communication

Title Page (The LSRP Title page template has been adopted in lieu of these recommendations, as per LSRP guidance)

Include the following information: Full title of the manuscript, which should be as concise and precise as possible; authors' names (first and last names, middle initial when commonly used by that author); institutional affiliation for each author named in English language and not in a national language (use superscripted numbers after each author's name, and a corresponding superscripted number before each institutional affiliation; names of institutions should be spelled out, but the abbreviation can be provided in parentheses);

contact information for the corresponding author (name, address, telephone number, fax number, e-mail address; ensure name matches that given in author list); Keywords for use by abstracting services (same as following summary); number of text pages; number of words; number of references; number of figures; number of tables; ORCID number for the first and senior authors, and any authors designated as corresponding.

Summary and Keywords

Provide a summary of no more than 300 words (200 words for Brief Communication). The summary for Full Length Original Research should consist of four sections, labelled: Objective; Methods; Results; Significance. This structured summary should concisely and specifically describe why and how the study was performed, the essential results, and what the authors conclude from the results. To promote brevity, authors may use phrases rather than complete sentences. The summary for Special Reports, Invited Commentaries, and Brief Communications is not structured, but should cover the same topics as the structured summary. The summary (structured or unstructured) should be followed by 3 to 6 Keywords (see above). A second short summary (less than 100 words) is required for Brief Communications that can be used in the print issue Table of Contents. Submit the second short summary as a Supporting Document.

Key Points

Box Include 3 to 5 key bullet points that summarize your article after the main body of text. Please ensure that each bullet point is no longer than 140 characters. (Brief Communications do not require a Key Point box.)

Introduction

State the objectives of the study clearly and concisely and provide a context for the study by referring judiciously to previous work in the area. Do not attempt to present a comprehensive view of the field. Provide a statement about the significance of this research for understanding and/or treating epilepsy.

Methods

Describe the research methods in sufficient detail that the work can be duplicated; alternatively, give references (if they are readily accessible) to previous comprehensive descriptions. Identify the statistical procedures that were used and the rationale for choosing a particular method, especially if it is not standard. Reports of experimental studies on humans must explicitly certify that the research received prior approval by the appropriate institutional review body and that informed written consent was obtained from each volunteer or patient. Studies involving animals must include an explicit statement that animal care and use conformed to institutional policies and guidelines. When animals are subjected to invasive procedures, details must be provided regarding the steps taken to eliminate/minimize pain and suffering, including the specific anaesthetics, analgesics, or other drugs used for that purpose (amounts, mode of delivery, frequency of administration). If extensive descriptions of methods are needed, provide basic information within the text and submit supplementary information for online Supporting Information.

Results

Results should be reported fully and concisely, in a logical order. Do not repeat methodologic details from the Methods section. Where possible, use figures and/ or tables to present the

data in a clear and concise format. Do not repeat data in the text that are given in a table but refer to the table. Provide textual explanations for all figures, with clear reference to the figure(s) under discussion. Descriptive information provided in figure legends need not be repeated in the text; use the text, however, to describe key features of the figures. When appropriate, give sample numbers, the range and standard deviation (or mean error) of measurements, and significance values for compared populations. q Discussion Provide an interpretation of the results and assess their significance in relation to previous work in the field. Do not repeat the results. Do not engage in general discussion beyond the scope of the experimental results. Conclusions should be supported by the data obtained in the reported study; avoid speculation not warranted by experimental results, and label speculation clearly. Discuss the significance of the data for understanding and/or treating epilepsy. q Statistical Methods The following guidelines assume familiarity with common statistical terminology and methods. We recommend that authors consult a biostatistician during the planning stages of their study, with further consultations during the analytical and interpretational stages.

1. Analysis guidelines:

- Use robust analytic methods when data are skewed.
- Use Kaplan-Meier methods, Cox proportional hazards, and mixed models analyses for longitudinal data.
- Account properly for statistical outliers.
- Use exact methods as much as possible in analyses of categorical data.
- Use appropriate correction procedures to account for multiple comparisons and conduct post hoc comparisons with statistically appropriate methods.

2. Presentation guidelines:

- Report means accompanied by standard deviations; standard errors should not be used.
- Present results with only as much precision as is appropriate.
- Present confidence intervals, whenever possible, including in figures.
- Describe quantity of missingness and methods used for handling such missingness.
- In general, present two-sided P values. P values larger than 0.01 should be reported to two decimal places, those between 0.01 and 0.001 to three decimal places, and those smaller than 0.001 should be reported as $P < 0.001$.
- In reporting clinical trials, include a flow diagram, a completed trial checklist, and trial registration information. The CONSORT flow diagram and checklist are recommended (<http://www.consortstatement.org/>).

Acknowledgments

Acknowledge sources of support (eg, grants from government agencies and private foundations), including funds obtained from private industry. Also acknowledge (consistent with requirements of courtesy and disclosure) participation of contributors to the study who are not included in the author list. q Disclosure of Conflicts of Interest In addition, each author should provide full disclosure of any conflicts of interest. One of the following sentences must be included at the end of the paper: either “Author A has received support from, and/or has served as a paid consultant for; Author B has received support from. The remaining authors have no conflicts of interest.” Or “None of the authors has any conflict of interest to disclose.”

Note: Disclosure is needed for financial income/payment from commercial sources, the interests of which are relevant to this research activity. Please identify sources from which

financial assistance/income was obtained during the period of the research activity and generation of the current report. Grants from government and/or private agencies should be identified in the Acknowledgments section.

Ethical Publication Statement

All papers must include the following statement to indicate that the authors have read the Journal's position on issues involved in ethical publication (see below) and affirm that their report is consistent with those guidelines: "We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines."

References

Authors are responsible for the accuracy of their references. References should follow a modified Vancouver style format. Refer to PubMed to ensure accurate and complete reference information. Citation of references in the text should be in superscript numbers (including those in figure legends and tables). When names are given with reference citations, check the reference list to make sure spelling is consistent. Cite the end references in numerical order. The first six authors should be listed and followed by et al. Use PubMed abbreviations for journals in the reference list at the end of the paper (as opposed to journal names being written out in full). Reference program patches are available on the Epilepsia ScholarOne (<https://mc.manuscriptcentral.com/epilepsia>); in the "Instructions and Forms" link. Number of references is limited to the following:

Full Length Original Research – 50

Brief Communication – 18

Review – 100

Special Report – 100

Figure legends

Number each legend sequentially to conform to the figure number (eg, Figure 1, Figure 2). The legend should provide a brief description of the figure, with explanation of all symbols and abbreviations. Written permission to use nonoriginal material must be obtained by the authors (from the original authors [where possible] and publishers). Credit for previously published material (author(s), date, journal/book title, and publisher) must be included in the legend. A figure legend should be listed at the end of the manuscript following the list of references. When references are made in the text to items within a figure (arrows, inserts, etc), make sure they are in the figure.

Tables

Tables should be formatted in the manner that the authors wish the table to appear in print. Present all tables together at the end of the main text document or as separate table files. Do not embed tables in the main text file or upload tables in image formats. Each table should be given a number and a descriptive title. Provide notes and explanations of abbreviations below the table and provide clear headings for each column and row. Do not duplicate data given in the text and/or in figures. Written permission to use nonoriginal material must be obtained by the authors (from the original authors [where possible] and publishers). Credit for previously published material (author(s), date, journal/book title, and publisher) must be included in the table notes.

Figures

All figures should be prepared with care and professionalism. Submissions that do not comply with the following formatting requirements will be returned for correction and resubmission. Figures should be submitted as TIF files in the size expected for final publication—approximately 3 inches (7-8 cm) for half columns and 6 to 7 inches (15-17 cm) for double columns. Do not embed figures within the main text document. Submit black and white figures with a minimum of 300 dpi (MRI scans) and for line drawings or figures that include embedded text (bar graphs with numbers) at least 600 dpi. Complex figures (including photographs, micrographs, and MRrelated images), either in colour, in halftones, or in black and white, should also be submitted in TIF format with a resolution of at least 600 dpi. We recommend saving the TIF files with LZW compression (an option when you “save as” in packages like Photoshop), which will make the files smaller and quicker to upload without reducing the resolution/quality. Save each TIF file with a name that includes the first author’s last name and the figure number as referenced in the text (eg, Smith-fig1.tif). Provide clear labels on the ordinate and abscissa. Figures with more than one part should be combined by the authors in the correct orientation and labelled with A, B, C, and so on. When relevant, include calibration information. Label figures using Calibri font and ensure that all labels are large enough to be clearly legible when the figure is reduced to fit onto a journal page. The maximum size of any figure is 7x9 inches (17x22.5 cm) and 40 megapixels; the total number of pixels for each figure (i.e., heightxwidth) must be less than 40 megapixels, otherwise the image will not convert to PDF format for review. There is no charge for colour figures. We strongly encourage authors to generate figures in colour (to enhance clarity of presentation and aesthetic appeal), using the colour palette below. Photographs or videos of patients should not reveal patient identity; masking eyes and/or other identifiers is compulsory unless

the eyes are essential to the meaning of the photograph or video. In addition, such photographs and videos must be accompanied by a letter stating that signed consent forms authorizing publication have been obtained for all identifiable patients, and that the consents will be maintained by the author for 7 years or until the patient reaches 21 years of age, whichever is longer. Do not send *Epilepsia* the consent forms; U.S. Federal privacy rules prohibits sending signed consent forms to *Epilepsia* or Wiley Publishing without written permission from the patient to do so. A sample signed consent form can be found on the *Epilepsia* ScholarOne site ([https:// mc.manuscriptcentral.com/epilepsia](https://mc.manuscriptcentral.com/epilepsia)); Click “Instructions and Forms” at the top right-hand corner of the homepage.

Supporting Information

Supporting information, to be published online only, can be submitted for review. Such material may include additional figures, large tables, videos, and so on that cannot be accommodated within the normal printed space allocation for an article but provide important complementary information for the reader. As determined by the reviewers and Editors, supporting information will be posted on the Wiley Online Library *Epilepsia* server and integrated directly into the full-text HTML article. Explicit reference to the supporting information in the main body of the text of the article is recommended, and the material must be captioned at the foot of the text, below the reference list. Citations should be in the following format: Figure S1, Table S1, Appendix S1, etc. Supporting information will be published as submitted and will not be corrected or checked for scientific content, typographical errors, or functionality. Although this material is hosted on Wiley Online Library, the responsibility for scientific accuracy and file functionality remains entirely with the authors. A disclaimer will be displayed to this effect with any supporting information

published. Supporting Information files should be accompanied by detailed information (if relevant) about what they are and how they were created (e.g., a native dataset from a specific piece of apparatus). Acceptable formats for Supporting Information include: General – Standard MS Office format (Word, Excel, PowerPoint, Project, Access, and so on); PDF Graphics – GIF; TIF (or TIFF); EPS; PNG; JPG (or JPEG); BMP; PS (postscript); embedded graphics (e.g. a GIF pasted into a Word file) are also acceptable.)

Research Team

Internal Supervisor

Dr Chris Hosbon, PhD. DClinPsy.

Senior Research Tutor, Cardiff University.

<https://www.researchgate.net/profile/Christopher-Hobson-3>

External Supervisor

Professor Ingram Wright, PhD. DClinPsy.

Consultant Paediatric Neuropsychologist, University Hospitals Bristol NHS Foundation Trust

Professor of Clinical Neuropsychology, University of Bristol

<https://www.researchgate.net/profile/Ingram-Wright>

Second Reviewer

Dr Darren Quelch, BSc (Hons). MB. ChB. PhD.

Junior Doctor, Aneurin Bevan University Health Board.

<https://www.researchgate.net/profile/Darren-Quelch-2>

QUADAS-2 Tool

DOMAIN 1: PATIENT SELECTION - Category A: Risk of Bias - Could the selection of patients have introduced bias?

Was a consecutive or random sample of patients enrolled?

Yes/No/Unclear

Was a case-control design avoided?

Yes/No/Unclear

Did the study avoid inappropriate exclusions?

Yes/No/Unclear

1. Low

2. High

3. Unclear

DOMAIN 1: PATIENT SELECTION - Category B: Concerns regarding applicability - Are there concerns that the included patients do not match the review question?

- Include children who have seizures that are uncontrolled by medical treatment (i.e., failure of two or three appropriate drugs) or are disabling (including medication side effects)
- Include exclusively the spectrum of patients suitable to surgical assessment, i.e.:
 - o Children with catastrophic early onset epilepsy with evidence of lateralisation of the seizure onset

- All children under 24 months old with evidence of focality of seizure onset, with or without an MRI evident lesion
 - Children of any age with evident focal epilepsy, or lateralised seizures associated with congenital hemiplegia, resistant to two appropriate anti-epileptic drugs (AEDs)
 - Children who have epilepsy associated with a lateralised abnormality seen on a brain scan
 - Children with epilepsy associated with Sturge Weber syndrome, benign tumours with developmental issues and/or ongoing seizures, or Rasmussen's syndrome
 - Children of any age with epilepsy associated with tuberous sclerosis resistant to two AEDs where seizures may arise from a single focus (probably from a single tuber)
 - Children who have 'drop attacks' as part of a more complex epilepsy
 - Children with epilepsy associated with hypothalamic hamartoma
- Surgical syndromes and aetiologies are more diverse in children than in adults 62, studies should attempt to include a range of pathologies that are suitable for surgical assessment (e.g. cortical dysplasia, Tuberous sclerosis complex, polymicrogyria, hypothalamic hamartoma, hemispheric syndromes, Sturge-Weber syndrome, Rasmussen syndrome, Landau–Kleffner syndrome and others i.e. dysembryoplastic neuroepithelial tumor, cerebrovascular insults etc.).
- Research studies would not have to cover all these conditions but should be generally have proportionate subgroups to those of the general surgical population

- Malformations of cortical development (MCD)=27%,
 - tumours=35%,
 - Hippocampal sclerosis=14%
 - Vascular=12%
 - Neurocutaneous=8%
 - Electroclinical syndromes=9%
 - Other=5%
 - Neuroplasticity of language is age-dependent, studies should include a range of ages that is representative of those with pathology that may be surgically resected.
- Median age of surgery 10.4 (range 0-18) – 2
 - Median age of seizure onset 2 (range 0-17) – 2
- EZ Targets for surgery are also variable and should be captured proportionate to the surgical population 2
 - Temporal = 71%
 - Frontal = 22%
 - Parietal = 5%
 - Occipital = 3%
 - A range of epilepsy severity should also be included, considering factors around
 - Number of seizures
 - Seizure burden/severity
 - The healthcare setting should be an established epilepsy surgery centre

1. Low

2. High

3. Unclear

DOMAIN 2: INDEX TEST - Category A: Risk of Bias - Could the conduct or interpretation of the index test have introduced bias?

Were the index test results interpreted without knowledge of the results of the reference standard? This item is similar to “blinding” in intervention studies.

Yes/No/Unclear

If a threshold was used, was it pre-specified? Selecting the test threshold to optimise sensitivity and/or specificity may lead to overoptimistic estimates of test performance, which is likely to be poorer in an independent sample of patients in whom the same threshold is used.

Yes/No/Unclear

Were the language mapping tests used the same as the post-operative tests?

Yes/No/Unclear

If more than one mapping technique was used, were they both assessed blinded?

Yes/No/Unclear

Were uninterpretable/ intermediate test results reported? Yes/No/Unclear

1. Low

2. High

3. Unclear

DOMAIN 2: INDEX TEST - Category A: Concerns regarding applicability - Are there concerns that the index test, its conduct, or interpretation differ from the review question?

Index tests are those that are being assessed (i.e. the mapping).

Did index tests methods vary from those specified in the review question?

Yes/No/Unclear

Was this language mapping test applied in the same way it would be in clinical practice?
e.g. was relevant clinical information (such as age of patient, epilepsy onset, handedness, etc) available to the person interpreting the results?

Yes/No/Unclear

Did the language mapping tests not form part of the pre- or post-operative language assessment?

1. Low

2. High

3. Unclear

DOMAIN 3: REFERENCE STANDARD - Category A: Risk of Bias - Could the reference standard, its conduct, or its interpretation have introduced bias?

Did the methods section of the paper describe the reference standards that were used?

Yes/No/Unclear

Were test results (mapping or post-op data) interpreted blind to the results of the other test, or blinding is dictated by the test order?

Yes/No/Unclear

Were uninterpretable/ intermediate test results reported?

Yes/No/Unclear

1. Low

2. High

3. Unclear

DOMAIN 3: REFERENCE STANDARD - Category B: Concerns regarding applicability - Are there concerns that the language function as evaluated by the pre/post-operative testing does not match the review question?

Is the post-operative assessment likely to classify the language function correctly?

Yes/No/Unclear

Has a standardised and reliable language measure been used?

Yes/No/Unclear

Has reliable change from pre- to post-op been examined?

Yes/No/Unclear

Are the post-operative assessments of language consistent with those in standard epilepsy practice?

Yes/No/Unclear

1. Low

2. High

3. Unclear

DOMAIN 4: FLOW AND TIMING - Category A: Risk of Bias - Could the patient flow have introduced bias?

Was there an appropriate interval between language mapping and post-operative assessment? (E.g. 12 months)

Yes/No/Unclear

Was the interval between language mapping and post-operative assessment consistent across participants?

Yes/No/Unclear

Did all patients receive a post-operative follow-up?

Yes/No/Unclear

Did all patients receive the same post-operative follow-up?

Yes/No/Unclear

Were all patients included in the analysis?

Yes/No/Unclear

Could the patient flow have introduced bias?

Yes/No/Unclear

1. Low

2. High

3. Unclear

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standards**	Case-control studies, or "poor or non-independent reference standard**"	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or <i>n</i> -of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, <i>n</i> -of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)*	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or <i>n</i> -of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non-randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

* Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.

** As always, a systematic review is generally better than an individual study.

How to cite the Levels of Evidence Table

OCEBM Levels of Evidence Working Group*. "The Oxford 2011 Levels of Evidence".

Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653>

* OCEBM Table of Evidence Working Group = Jeremy Howick, Iain Chalmers (James Lind Library), Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard and Mary Hodgkinson

Children's Epilepsy Surgery Service Protocol Exerts

ASSESSMENT BATTERY PROTOCOL				
	Under 2:6	2:6-5:11	6:0-15:11	Over 16:0
IQ	Bayley Scales	WPPSI-IV VCI FRI WMI PSI	WISC-V VCI (SI, VC) FRI (MR, FW) WMI (DS, PS) PSI (CD, SS) VSI (BD)	WAIS-IV VCI PRI WMI PSI
Memory		ChAMP (>5:0) Immediate subtests	ChAMP Immediate subtests	WMS IMI DMI AMI VMI vWMI
Language	Bayley Scales	CELF (>5:0) Sentence Comprehension Word Structure Recalling Sentences Formulated Sentences WPPSI-IV (<5) Receptive Language	CELF (6:0-8:11) Sentence Comprehension Word Structure Recalling Sentences Formulated Sentences <i>Or</i> CELF (9:0-12:11) Semantic Relationships Recalling Sentences Formulated Sentences Word Classes 2 <i>Or</i> CELF (13:0-16:11) Recalling Sentences Formulated Sentences Understanding Spoken Paragraphs Semantic Relationships	CELF (<16:11) Recalling Sentences Formulated Sentences Understanding Spoken Paragraphs Semantic Relationships
Attainment		WIAT Numerical Operations (>5:0) Word Reading (>4:00)	WIAT Numerical Operations Word Reading	WIAT Numerical Operations Word Reading
Attention		NEPSY (>5:0) Auditory Attention	CPT 3 <i>Or</i> NEPSY (5:0-7:11) Auditory Attention and Response Set	CPT 3
Executive Functioning		NEPSY (>3:0) Word Generation	NEPSY (<8:0) Word Generation <i>Or</i> DKEFS (>8:0) Trail Making Verbal Fluency	DKEFS Trail Making Verbal Fluency

Questionnaires	CBCL ABAS PRECAA PEDQoL Parent Report (>2:0)	CBCL ABAS PRECAA PEDQoL Parent Report PEDQoL Self Report (>5:0)	CBCL ABAS PRECAA PEDQoL Parent Report PEDQoL Self Report QoLIE-AD-48 (>11:0) <i>Or</i> ELDQOL (informant) **brief alongside the	CBCL ABAS PRECAA PEDQoL Parent Report PEDQoL Self Report QoLIE-AD-48 (<18:0) <i>Or</i> ELDQOL (informant)
Other things to consider	Vineland	WRAVMA BPVS (>3:0) TROG (>4:0) MFVPT (>4:0) NEPSY FTT (>5:0) NEPSY Oromotor Seq. (>3:0) NEPSY Visuomotor Prec. (>3:0) Vineland	WRAVMA BPVS TROG MFVPT NEPSY FTT NEPSY Oromotor Seq. (6:0-12:11) NEPSY Visuomotor Prec. (6:0-12:11) Vineland	WRAVMA BPVS TROG MFVPT NEPSY FTT (<16:11) Vineland

Empirical Project Approvals



Alexander Marsh
Department of Clinical Psychology
11th Floor, Tower Building
Cardiff University
70 Park Place
Cardiff
CF10 3AS

19th November 2020

Dear Alexander,

RE: Project Approval (PNEUS/CA/2020-21/01)

I am writing to confirm I have received confirmation from the Research Committee that your project "Convergent Validity of Seed-Based Resting State Functional and Structural Connectivity Magnetic Resonance Imaging Language Networks to Task-based Methods in the Bristol Paediatric Epilepsy Service" has been reviewed, approved and registered by the Research and Development Department, University Hospital Bristol and West NHS Foundation Trust on 19th November 2020. The approvals reference for your records is: PNEUS/CA/2020-21/01.

This project has been filed under "Special Priority" within the Neurosurgery Speciality. It is listed as a multi-speciality project with the following team members:

Name	Job Title	Speciality	Role within Project (data collection, Supervisor etc)
Alexander Marsh	Trainee Clinical Psychologist	Neuropsychology	Lead
Dr Ingram Wright	Consultant Neuropsychologist	Neuropsychology	Supervisor
Dr Marcus Likeman	Consultant Neuroradiologist	Neuroradiology	Support with imaging data and protocols
Dr Amy Spray	Assistant Psychologist	Neuropsychology	Support with analysis
Kathryn Isaacs	Radiographer	Radiology	Support with data aggregation

The project aims were registered as follows:

Aim 1:
(1) Examine the convergent validity (=the degree to which two measures of a construct are related) of functional resting state networks in lateralising language against our current task-based fMRI language paradigm
(2) To compare our level of convergent validity to latest research outcomes

Aims 2:
(1) Examine the convergent validity (=the degree to which two measures of a construct are related) of structural covariance networks in lateralising language against our current task-based fMRI language paradigm
(2) To compare our level of convergent validity to latest research outcomes

Aim 3:
(1) Examine the convergent validity (=the degree to which two measures of a construct are related) of combined functional resting state networks and structural covariance networks in lateralising language against our current task-based fMRI language paradigm
(2) To compare our level of convergent validity to latest research outcomes

The following data collection methods that have been approved: case note review; data from existing database(s).

The sample criteria approved were:

- Have a formal diagnosis of epilepsy
- Between the ages of 4-18 years old
- Have gone through the Children's Epilepsy Surgery Service Evaluation and have completed the following as part of that assessment:
 - Completed functional imaging
 - Neuropsychological Assessment
- Pass MRI quality control criteria

Your data extraction tool has been reviewed. Please ensure the use of a "Participant Number" is a code, not Hospital Number, for data protection.

The following timeline has been registered:

- You have requested access to records from: 01/01/2012 to 31/11/2020.
- You have confirmed that patient identifiable will not be shared outside the trust.
- The proposed start date for this project is: 15/11/2020
- The proposed end date for this project is: 31/01/2022

Please inform us of any anticipated changes to this timeline as your Research and Development Convener: Richard Hancock, will need to be informed and potentially review permissions.

Templates for presenting and reporting your findings are available on the intranet. You have also indicated that you have an intention to publish and present your results outside of the trust. The following information has been reviewed and approved:

The results will be of interest to other epilepsy centres across the UK for the purposes of benchmarking and service evaluation/improvement projects. We aim for them to be published in a peer reviewed journal (e.g. epilepsy) and presented at relevant conferences (e.g. National CESS Conference). This project is also being undertaken as part of the requirements of a Doctor of Clinical Psychology professional training programme and will be submitted to the programme for review.



University Hospitals Bristol NHS Foundation Trust
Tel 0117 923 0000 Minicom 0117 94 9869 www.uhbt.nhs.uk



University Hospitals Bristol NHS Foundation Trust
Tel 0117 923 0000 Minicom 0117 94 9869 www.uhbt.nhs.uk

I can confirm the following agreements are in place:

Project Lead

By completing this section:

- I agree to ensure that this project is completed, the results disseminated, and a report is provided to: Richard Hancock, Research and Development Team.
- I understand that non-anonymised (staff/patient) data must not be taken outside the Trust and will act in accordance with the Data Protection Act / Caldicott Principles at all times.
- I understand that the results belong to the Trust and that a report may be made available to anyone who requests it.

Alex Marsh

Senior Clinician / Manager

By completing this section:

- I confirm that this project has been agreed as part of the speciality programme and that I will give my full support to it.
- I will ensure the dissemination of results and lead on the development and implementation of an action plan (as necessary) in order to obtain improvements in the quality of care provided.

Ingram Wright

Clinical Project Convener

- I approve the project described above and confirm that it has been appropriately reviewed for methodological quality, resource implication and importance to the Trust.

Richard Hancock

If you have any questions, please do not hesitate to contact me.

Kind regards,

Professor Ingram Wright
Clinical Lead & Head of Psychology, Women's and Children's Division
University Hospitals Bristol & Weston NHS



University Hospitals Bristol NHS Foundation Trust
Tel 0117 923 0000 Minicom 0117 94 9869 www.uhbt.nhs.uk



University Hospitals Bristol NHS Foundation Trust
Tel 0117 923 0000 Minicom 0117 94 9869 www.uhbt.nhs.uk

Quality Control

BRCH Quality Control Guide

The aim of this guide is to check the quality of the fMRI data. You may highlight minor flaws, which can be fixed (to an extent) with pre-processing steps. You may also highlight major issues, which may require a repeat scan or indicate a need to check the scanner set up. To make these steps easier, save your scans in the Brain Imaging Data Structure (BIDS) standard format. Please see further down for a list of MRI Brain artefacts with exemplar images. Any findings must be documented within the fMRI report.

1. Inspect the Anatomical and Functional Images for artefacts (e.g. scanner spikes incorrect orientation, poor contrast, etc.).

1.1. Open the anatomical image in SPM using the display button on your SPM GUI

- The anatomical image will be displayed in the SPM viewer in axial, sagittal, and coronal views. You can close any of the windows if you only want to focus on a subset of the views.

1.2. Inspect for Gibbs Ringing Artefacts (GRA)

- These are lines that look like ripples in a pond. They may indicate an error in the reconstruction of the MR signal from the scanner. These ripples may also be caused by the patient moving too much during the scan.

→ In either case, if the ripples are large enough, they may cause pre-processing steps like BET or normalisation to fail.

→ Example (see Figure S1):

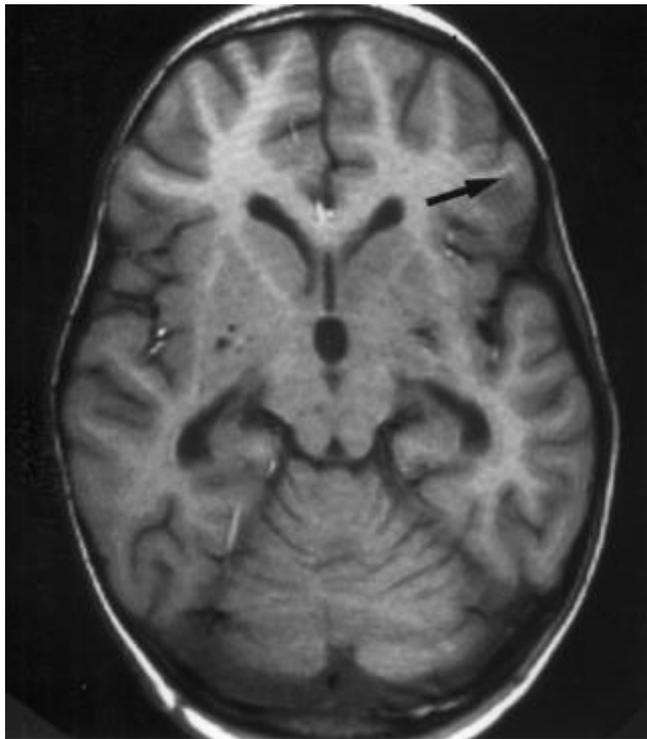


Figure S1. Example of Gibbs Ring Artifacts

- Report any artefacts
- 1.3. Inspect each plane for abnormal intensity differences within the grey or white matter.
- These may indicate pathologies, such as aneurysms or cavernomas, and they should be reported to the CESS radiologist (Dr Marcus Likeman) right away if not already reported on previous scans. Do not include these findings in the report but do document in the neuropsychology notes that you have requested a scan review.
- 1.4. When you are done looking at the anatomical image, click on the Display button again, navigate to the patient's fMRI folder, and select the functional image.
- A new image will be displayed in the orthogonal viewing windows. This image also looks like a brain, but it is not as clearly defined as the anatomical image.

This is because the resolution is lower. The functional scans are lower resolution in part because they are collected at a very fast rate.

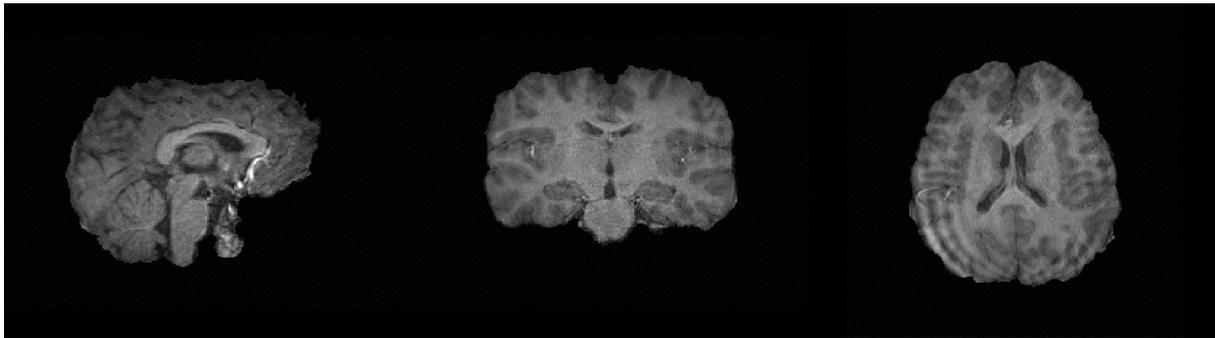
1.5. Inspect the image for extremely bright or extremely dark spots in the grey or white matter, as well as for image distortions such as abnormal stretching or warping.

1.6. Inspect the image for excessive motion.

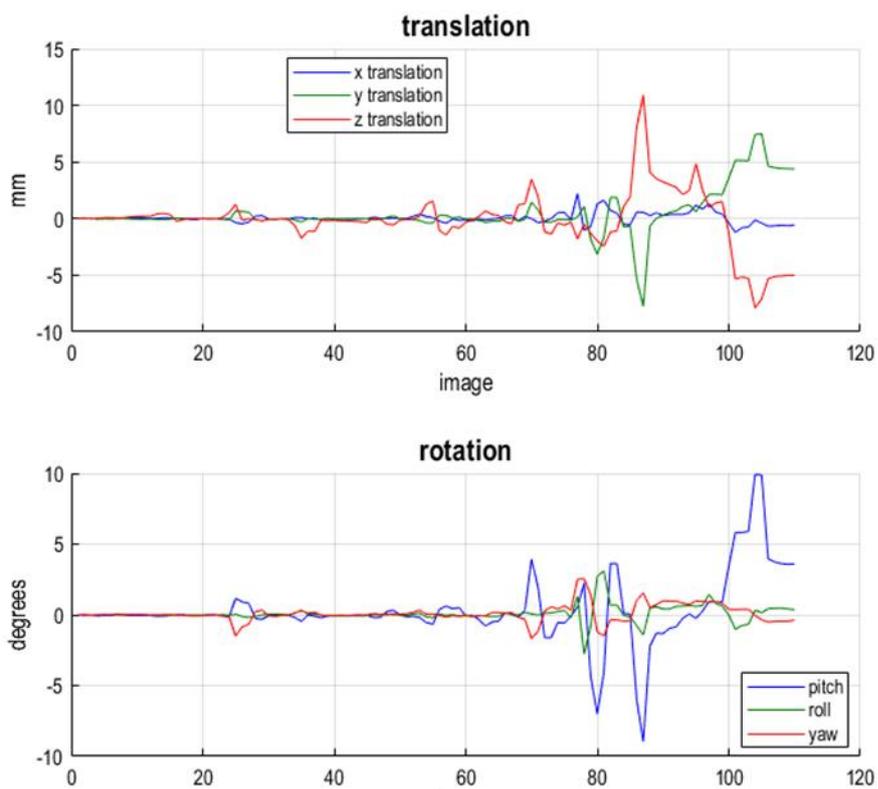
- Load the 4D series and play as a movie. Look for signs of significant movement (you can see this as the brain will wiggle as the movie progresses).

Examples of Brain MRI Artefacts: Somasundaram & Kalavathi, Orient. J. Comp. Sci. & Technol., Vol. 5(1), 135-141 (2012)

Quality Control Metrics for Task Based Scan



Supplementary Figure S2. T1 inspection demonstrates significant Gibbs ring artefacts



Supplementary Figure S3. Motion parameters demonstrate movement of >1 voxel (i.e. 3mm). Significant movement is introduced after 70th volume.

Supplementary Table S1. Cluster levels for Whole-Brain Analysis.

The top three clusters show significant activation in the occipital fusiform gyrus, angular gyrus, and cerebellar regions.

Statistics: *p*-values adjusted for search volume

set-level		cluster-level				peak-level					mm mm mm		
<i>p</i>	<i>c</i>	<i>p</i> _{FWE-corr}	<i>q</i> _{FDR-corr}	<i>k</i> _E	<i>p</i> _{uncorr}	<i>p</i> _{FWE-corr}	<i>q</i> _{FDR-corr}	<i>T</i>	(<i>Z</i> _E)	<i>p</i> _{uncorr}			
0.000	11	0.000	0.001	216	0.000	0.000	0.072	6.31	5.78	0.000	-14	-82	-22
						0.001	0.113	6.03	5.57	0.000	-26	-80	-28
						0.003	0.283	5.63	5.24	0.000	-14	-76	-32
		0.008	0.499	17	0.166	0.002	0.248	5.74	5.33	0.000	-38	-52	32
		0.016	0.499	9	0.309	0.005	0.366	5.51	5.14	0.000	16	-74	-28
		0.005	0.499	25	0.098	0.012	0.670	5.29	4.96	0.000	60	-36	-4
		0.021	0.499	6	0.408	0.018	0.670	5.18	4.87	0.000	-28	-62	-50
		0.014	0.499	10	0.284	0.023	0.670	5.11	4.81	0.000	-64	-22	0
		0.026	0.555	4	0.505	0.023	0.670	5.11	4.81	0.000	14	-14	50
		0.021	0.499	6	0.408	0.024	0.670	5.09	4.80	0.000	-46	2	50
		0.019	0.499	7	0.371	0.026	0.670	5.07	4.78	0.000	-36	0	40
		0.021	0.499	6	0.408	0.031	0.731	5.03	4.74	0.000	-38	-66	-28
		0.038	0.762	1	0.762	0.050	0.990	4.89	4.62	0.000	-46	14	40

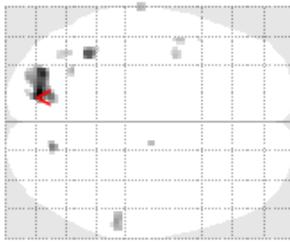
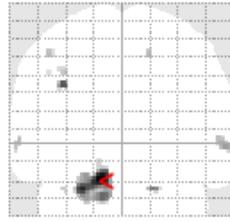
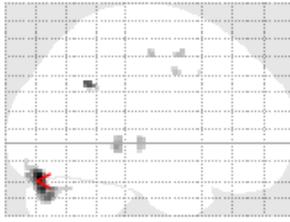
table shows 3 local maxima more than 8.0mm apart

Height threshold: $T = 4.89$, $p = 0.000$ (0.050)
 Extent threshold: $k = 0$ voxels
 Expected voxels per cluster, $\langle k \rangle = 9.410$
 Expected number of clusters, $\langle c \rangle = 0.05$
 FWE_p: 4.887, FDR_p: Inf, FWE_c: 1, FDR_c: 216

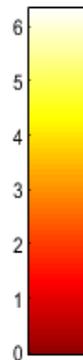
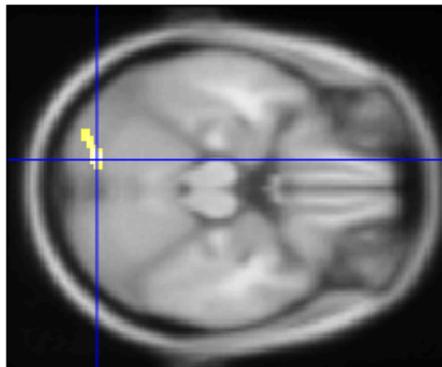
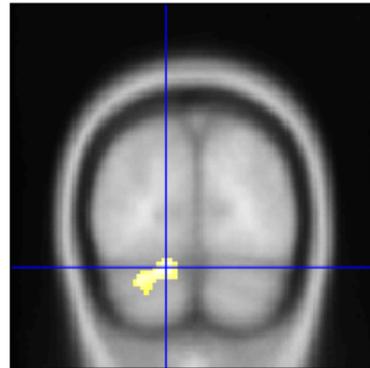
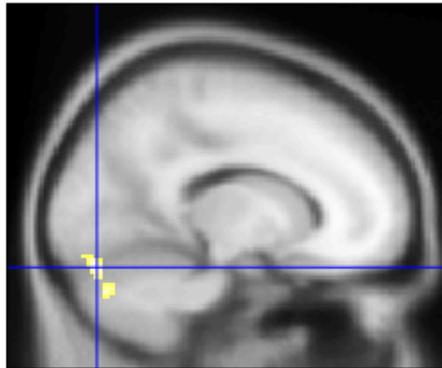
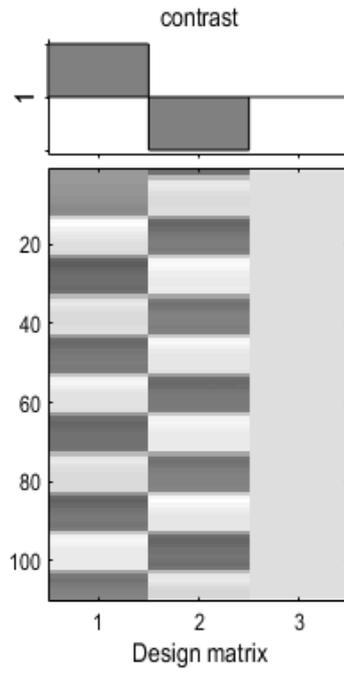
Degrees of freedom = [1.0, 102.0]
 FWHM = 14.1 14.5 12.4 mm mm mm; 7.1 7.3 6.2 {voxels}
 Volume: 1864536 = 233067 voxels = 684.1 resels
 Voxel size: 2.0 2.0 2.0 mm mm mm; (resel) = 317.73 voxels

language

SPM_{minip}
[-14, -82, -22]



SPM{T₁₀₂}



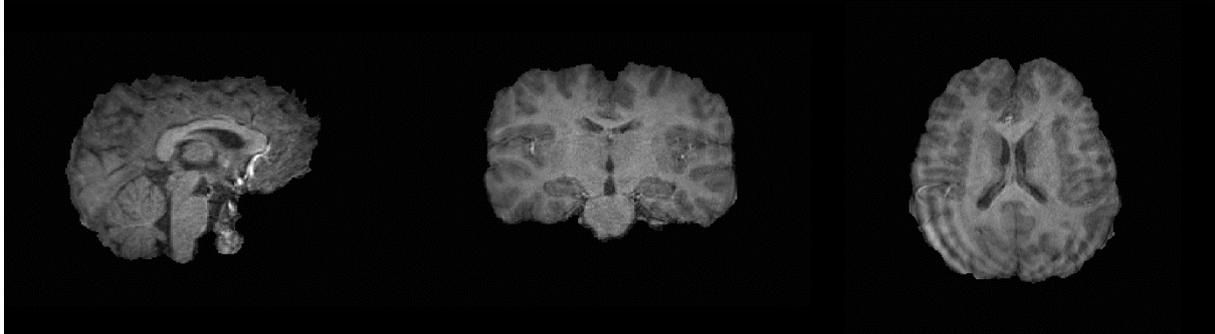
Supplementary Figure S4. The highest-level whole-brain analysis cluster rendered onto a glass brain (top) and a standard T1 brain (bottom; SPM Canonical avg152T1.nii).

MR Behavioural Report: Patient struggled to engage with task in scanner.

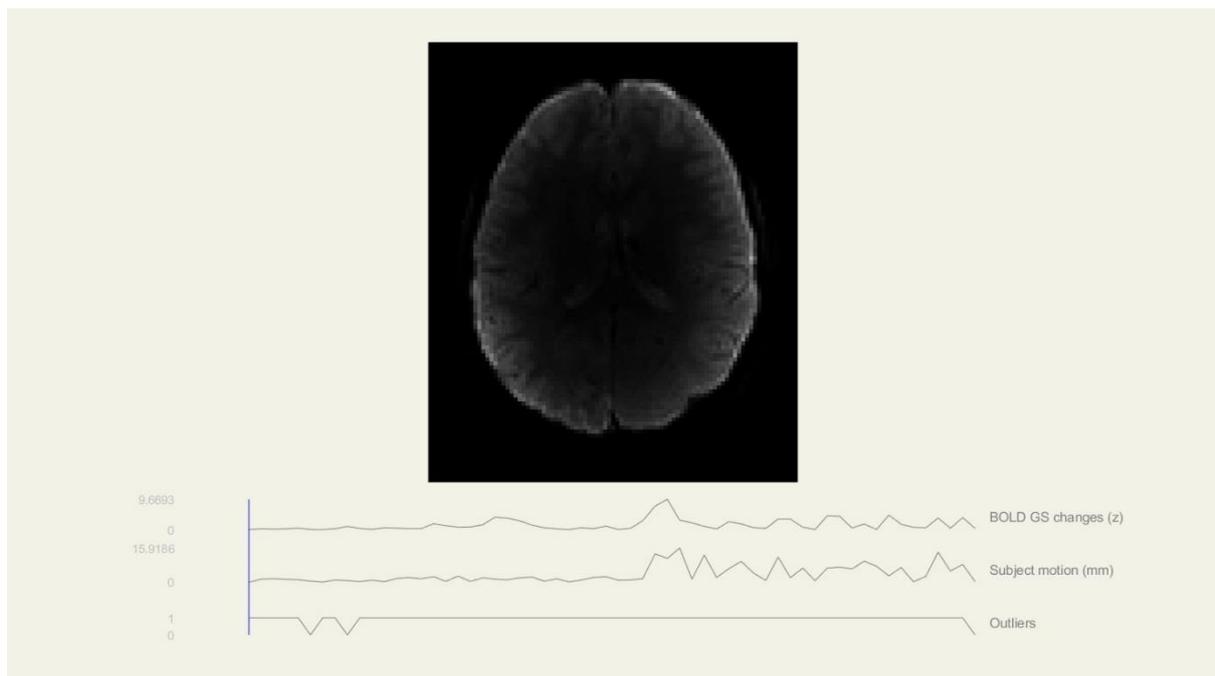
Conclusion: The presence of gibbs ring artefacts indicate motion during T1 acquisition. Significant motion was detected for >c.35% of the functional scan. Presence of motion in both scans may indicate MR is poorly ~~tolerated~~ ~~accepted~~. Whole brain analysis demonstrated significant clusters outside expected areas (e.g. inferior frontal gyrus, superior temporal gyrus). Patient report demonstrates poor compliance with task. Together, these indicate compromised quality of scan and recommendation for exclusion from study.

Quality Control Metrics for Resting state

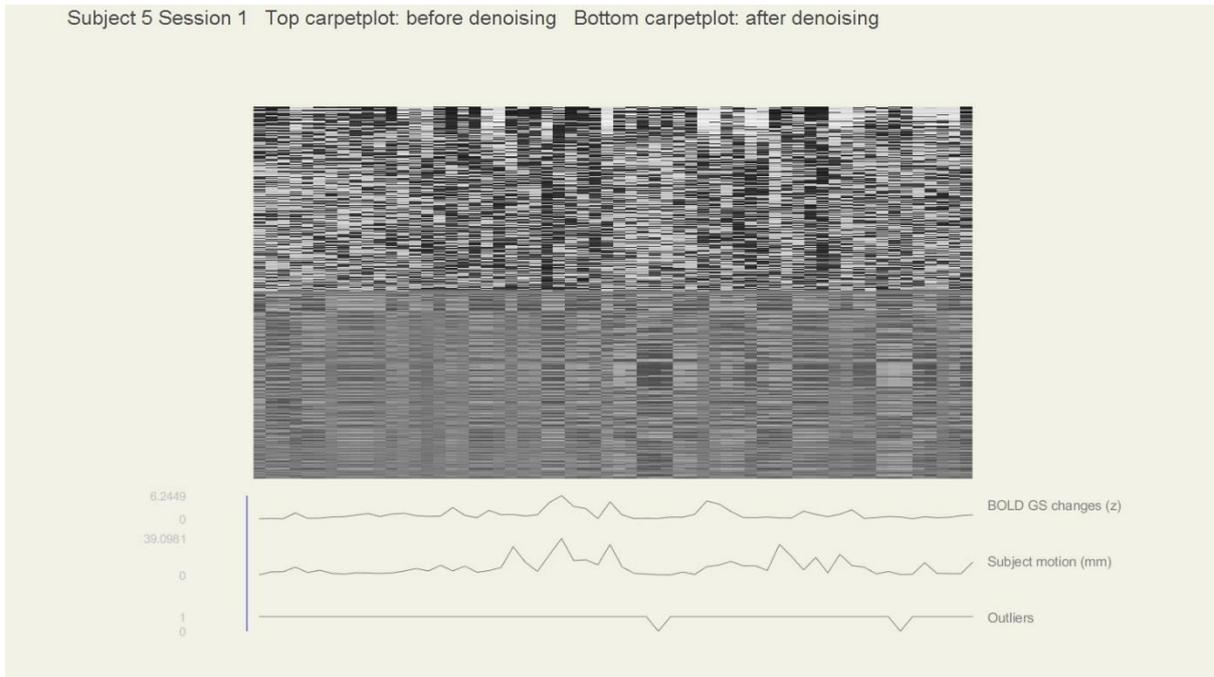
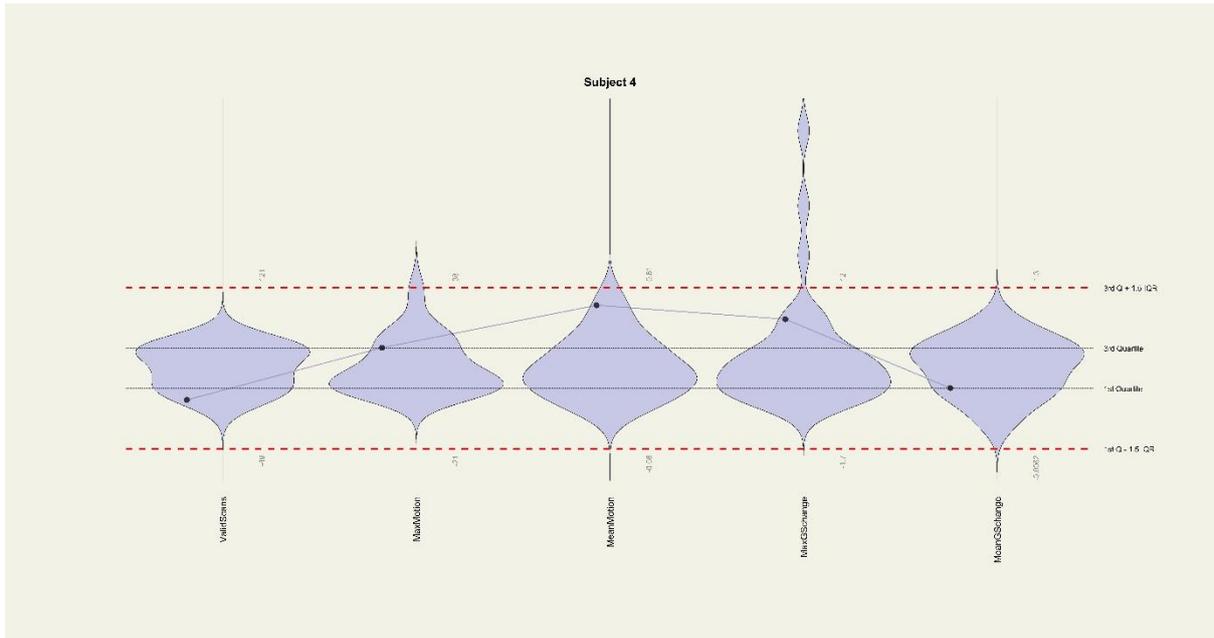
Example of Case Excluded for Motion



Supplementary Figure S5. T1 inspection demonstrates significant Gibbs ring artefacts



Supplementary Figure S6. Significant outliers have been identified. Acquisitions with framewise displacement above 0.9mm or global BOLD signal changes above 5 s.d. are flagged as potential outliers.



Supplementary Figure S7. Bottom carpetplot demonstrates notable residual noise after denoising processing and motion trace demonstrates significant levels of motion at multiple timepoints. Outliers persist despite denoising.

Conclusion: The presence of gibbs ring artefacts indicate motion during T1 acquisition. Significant motion was detected for functional scan. Presence of motion in both scans may indicate MR is poorly tolerated. Subject motion is concordant with global BOLD signal change. Significant outliers were detected in QC analysis. Significant levels of noise remain following denoising processing. Together, these indicate compromised quality of scan and recommendation for exclusion from study.

Plans for Dissemination

Journal Publication

We plan to submit our publication to *Epilepsia* in an abbreviated format, editing down the introduction, results and discussion to comply with journal limits.

NHS Dissemination

A formal report will be compiled and presented to the NHS Research and Development department at University Hospitals Bristol and Weston NHS Trust. In addition, the findings of this research will be presented to the Children's Epilepsy Surgery Service, Bristol Royal Hospital for Sick Children.