Intra-Operative Language Assessment for Adult Brain Tumor Survivors: A Systematic Review

Carissa Jolley
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Communication Sciences and Disorders Commons, and the Education Commons

Recommended Citation

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
INTRA-OPERATIVE LANGUAGE ASSESSMENT FOR ADULT BRAIN TUMOR PATIENTS: A SYSTEMATIC REVIEW

by

Carissa Jolley

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Communicative Disorders and Deaf Education

(Communication Sciences)

Lisa Milman, Ph.D., CCC-SLP
Committee Chair

Joann Tschanz, Ph.D.
Committee Member

Ronald Gillam, Ph.D.
Committee Member

Randy Jensen, M.D., Ph.D.
Committee Member

D. Richard Cutler, Ph.D.
Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah
2022
ABSTRACT

Intra-operative Language Assessment for Adult Brain Tumor Patients:
A Systematic Review

by
Carissa Johnston, Master of Science
Utah State University, 2022

Major Professor: Dr. Lisa Milman
Department: Communicative Disorders and Deaf Education

Awake craniotomy and language assessment procedures for brain tumor survivors can drastically affect patient outcomes and quality of life. The goal of intra-operative language mapping is to help preserve vital communicative functions. However, there is currently no standardized set of measures for intra-operative language assessment. This systematic review identified behavioral language measures used during intra-operative procedures and the resulting outcomes for adult brain tumor patients, with the aim of helping clinicians and researchers select intra-operative language assessments supported by the highest levels of evidence.

PRISMA guidelines were used to systematically identify articles. Patient demographic and medical information, neuroimaging data, intra-operative language assessment measures, language outcomes, and quality of evidence data was obtained using a data extraction manual and form. Descriptive methods were used for synthesis of the information.

The nineteen identified articles included a total of 471 patients with a variety of tumor types and tumor severity. Direct electrical stimulation (DES) was used across all
studies, with a few utilizing supplemental neuroimaging methods. Across studies, 37 language measures were used peri-operatively to assess 21 different language domains, and 26 language measures were used intra-operatively to assess 18 different language domains. The majority of patients (88%) reported as working prior to resection were able to return to work. Eight studies were classified as American Academy of Neurology (AAN) evidence level 3, ten studies as evidence level 2, and one study as evidence level 1. Discussion: This review supports awake language mapping during tumor resection and identifies the most widely used measures and associated language outcomes during post-operative stages. Further research is needed to establish best practice for selecting intra-operative assessments best suited for individual patient needs.
ACKNOWLEDGMENTS

A special thank you to the family, friends, LANR lab members, professors, and colleagues for their support and patience as I worked from proposal to end of this final document. I could not have completed it without your support and assistance.

Carissa Johnston
**CONTENTS**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Methods</td>
<td>6</td>
</tr>
<tr>
<td>Inclusion/Exclusion Criteria</td>
<td>6</td>
</tr>
<tr>
<td>Search Strategies</td>
<td>7</td>
</tr>
<tr>
<td>Electronic Search Databases and Search Terms</td>
<td>7</td>
</tr>
<tr>
<td>Ancestral and Descendant Search</td>
<td>9</td>
</tr>
<tr>
<td>Study Selection</td>
<td>9</td>
</tr>
<tr>
<td>Data Extraction</td>
<td>10</td>
</tr>
<tr>
<td>Data Synthesis and Interpretation</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td>Participant Demographics</td>
<td>12</td>
</tr>
<tr>
<td>Brain Tumor Characteristics</td>
<td>12</td>
</tr>
<tr>
<td>Neurosurgical Procedures Associated with Awake Language Mapping</td>
<td>18</td>
</tr>
<tr>
<td>Language Assessment/Tasks Associated with Awake Language Mapping</td>
<td>20</td>
</tr>
<tr>
<td>Pre-Post-Operative Assessment</td>
<td>20</td>
</tr>
<tr>
<td>Intraoperative Assessment</td>
<td>20</td>
</tr>
<tr>
<td>Language Outcomes Following Awake Mapping</td>
<td>26</td>
</tr>
<tr>
<td>Quality of Evidence</td>
<td>30</td>
</tr>
<tr>
<td>Discussion</td>
<td>31</td>
</tr>
<tr>
<td>Patient Demographics and Brain Tumor Characteristics</td>
<td>33</td>
</tr>
<tr>
<td>Neurosurgical Procedures Associated with Awake Language Mapping</td>
<td>34</td>
</tr>
<tr>
<td>Language Assessment/Tasks Associated with Awake Language Mapping</td>
<td>35</td>
</tr>
<tr>
<td>Recommendations for Clinical Application</td>
<td>36</td>
</tr>
<tr>
<td>Language Outcomes Following Awake Mapping</td>
<td>37</td>
</tr>
<tr>
<td>Quality of Evidence</td>
<td>39</td>
</tr>
<tr>
<td>Contribution</td>
<td>40</td>
</tr>
</tbody>
</table>
Limitations .......................................................................................................................... 41
Conclusion .......................................................................................................................... 41
References .......................................................................................................................... 43
Appendices ......................................................................................................................... 60
  Appendix A. Study Inclusion Form .................................................................................. 61
  Appendix B. PsychInfo Search Terms ............................................................................. 66
  Appendix C. Web of Science Search Terms .................................................................... 68
  Appendix D. References of Included Studies ................................................................... 70
  Appendix E. Manual for Rating Intraoperative Brain Tumor Survivor (BTS)
  Language Assessment Studies ....................................................................................... 74
LIST OF TABLES

Table 1 Participant Demographics ............................................................... 14
Table 2 Tumor Characteristics ..................................................................... 16
Table 3 Neurosurgical Procedures Associated with Awake Language Mapping .... 19
Table 4 Pre- and post-operative Language Assessment .................................. 22
Table 5 Intra-operative Language Assessment ............................................... 24
Table 6 Language Outcomes Following Awake Mapping ............................... 28
Table 7 Quality Indicator Scores .................................................................. 30
Table 8 Persistent Language Impairments Compared to Most Commonly Assessed Domains ............................................................... 39
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. Study Inclusion Flowchart</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>
Intraoperative Language Assessment for Adult Brain Tumor Patients: A Systematic Review

Introduction

More than 308,000 new cases of central nervous system (CNS) primary tumor are diagnosed each year (Central Brain Tumor Registry of the United States, CBTRUS, Ostrom et al., 2019; Globocan, 2020). Although brain tumor (BT) patients are frequently combined with stroke/traumatic brain injury patients in discussions of recovery trajectory, BT patients present differently than other patient populations with respect to language and other behavioral characteristics (Anderson et al., 1990; Davie et al., 2008). Surgical resection of the tumor, the most common initial medical treatment (McFaline-Figueroa & Lee, 2018; Meng et al., 2015; Ritaccio et al., 2018), can further impact cognitive-communicative functions, and subsequently the survivor’s quality of life. It is also widely recognized that various surgical procedures, including intra-operative cognitive assessment, can differentially affect outcomes in a variety of patient populations and tumor types (Dziedzic & Bernstein, 2014; Duffau 2018; Kanno & Mikuni, 2015; McGirt et al., 2009; Saito et al., 2015). Currently, however, there is no standard of care for intra-operative language assessment of brain tumor survivors undergoing tumor resection.

Surgical procedures for brain tumors include biopsy, craniotomy, endoscopy, and laser ablation (John Hopkins Medicine, 2021). Awake craniotomy is considered by many to be the gold standard for resection of supratentorial tumors, particularly for lesions in motor or language areas (Dilmen et al., 2016; Duffau, 2018; Leal et al., 2017; Raffa et al., 2019; Ritaccio et al., 2018). A critical issue in awake craniotomy is optimizing onco-functional balance (i.e., maximizing the extent of tumor resection while preserving vital
functions). Gross total resection is associated with longer survival rates (Brown et al.,
2016; Chang et al., 2011; Chikenzie et al., 2017; Duffau, 2018; Ritaccio et al., 2018;
Sanai & Berger, 2010; Xia et al., 2018) but can be associated with greater morbidity
affecting functions such as language and related cognitive domains that affect quality of
life (De Witt Hamer et al., 2012; Duffau, 2013; McGirt et al., 2009).

A variety of neuroimaging methods to maximize tumor resection without
compromising vital functions have been developed (Tarapore et al., 2012). Recent
imaging methods of interest include intra-operative fluorescent imaging to augment
visual separation of cancerous margins from healthy tissue (Lakomkin & Hadjipanayis,
2018; Nagaya et al., 2017); and intra-operative MRI (iMRI) to monitor intra-operative
brain shift (Gandhe & Bhave, 2018; Caras et al., 2020). Awake procedures such as direct
electrical/electrocortical stimulation (DES) allow neurosurgeons to define boundaries of
functional neural tissue and thereby make decisions about what tissue can be excised.
Historically, DES is the standard for intra-operative mapping of cerebral tumors in
eloquent areas. Though some debate exists about the relative effectiveness of awake
craniotomy and intra-operative DES relative to other surgical procedures (Buchfelder &
Zhao, 2018; Nossek et al., 2013; Rutten et al., 2002), it has remained the most widely
used and researched approach to achieving onco-functional balance. Evidence shows that
the procedure results in higher extent of resection, longer survival time, and better post-
operative neurological outcomes than when surgery is done under general anesthesia (Bu
et al., 2020; De Witt Hamer et al., 2012; Dziedzic & Bernstein, 2014; Groshev et al.,
2017; Sacko et al., 2011). Studies have found that this technique is the most sensitive in
identifying essential language areas (De Witte et al., 2013; Mandonnet et al., 2010),
allows mapping of subcortical tissue (Duffau, 2018; Mandonnet et al., 2010), and provides direct, immediate feedback (Duffau, 2015).

During the DES procedure, electrodes are used to stimulate specific sites in the brain with an electric current. During the electric pulse, which usually lasts 2-8 seconds (Ritaccio et al., 2018; Saito et al., 2015; Szelényi et al., 2010), the patient is asked to complete various assessment tasks. Disruption of vital functions such as speech, language, or motor activity during DES indicates that the stimulated tissue is part of an essential circuit associated with that function and that removal is likely to have negative effects on patient outcomes. Since language is vulnerable to both tumor pathology and neurosurgery, it is almost always assessed as part of the DES procedure (Duffau, 2018; Sanai et al., 2008). Selection and/or administration of the language assessment tasks is primarily the responsibility of a speech-language pathologist (O’Neill et al., 2020; Sefcikova et al., 2019) or neuropsychologist (Kelm et al., 2017; Sanai et al., 2008; Sefcikova et al., 2019).

Due to the nature of awake craniotomy, intra-operative imaging, and DES, it is essential that all assessment tasks are completed within a relatively short time interval. A patient may be awake from 45 minutes to a couple hours, depending on the size and location of the tumor (Raeke, 2018). Tasks used to intra-operatively evaluate language vary widely due to differences associated with tumor histology and presentation (Ostrom et al., 2019), procedural differences across clinicians and/or institutions (Mandonnet et al., 2017; Sefcikova et al., 2020) and/or when batteries are tailored to individual patient language profiles based on preoperative assessments (Chan et al., 2019; Hall et al. 2021). Many studies report using brief screening tests (Satoer et al, 2016). The most basic
battery includes simple movement and speech/language tasks such as systematic counting, reading, and naming (Ottenhausen et al., 2015; Sanai et al., 2008; Zanin et al., 2018). Other tasks used to assess language might include spontaneous speech, semantic associations, and repetitions. Occasionally tasks probing additional cognitive functions, such as memory or simple calculations, are also included (O’Neill et al., 2020; Ottenhausen et al., 2015; Rofes, Mandonnet, et al., 2017; Ruis, 2018; Talacchi et al., 2012).

The importance of intra-operative language mapping cannot be overstated; by assessing language functions intra-operatively, surgeons can determine which specific microregions of the cortex are still functioning in language processing and thus preserve critical areas necessary for post-surgical performance. Ultimately, the goal is to maximally preserve abilities that support post-surgical language use and participation across a range of social and vocational contexts. Given the importance of awake craniotomy to patient outcomes and quality of life, it is imperative that the intraoperative language assessments used during these procedures are adequate for determining “vital” tissue (Mandonnet et al., 2020; Miceli et al., 2012).

In 2013, De Witte et al. evaluated common standardized intraoperative language tasks and the scientific theory behind them. They concluded that most intraoperative tasks (such as those described above) were not standardized and had limited psychometric data; sensitivity of various language tasks needed further research; and that most literature lacked valid scientific basis. Since then, a number of individual protocols specifically for intraoperative language assessment have been developed, such as the Dutch Linguistic Intra-operative Protocol battery (DuLIP, Alves et al., 2020; De Witte et
al., 2015); the Russian Naming Test for mapping noun and verb production of Russian speakers (Dragoy et al., 2016; Dragoy et al., 2020); the verb and noun test for peri-operative testing with DES or nTMS mapping (VAN-POP, Ohlerth et al., 2020); and strategies for simultaneous assessment of speech and motor function (Mandonnet et al., 2017). However, the available literature supporting these protocols consists mainly of descriptive clinical reports and single cases studies. Furthermore, many studies focus primarily on medical aspects, such as extent of resection or anesthesia protocols, rather than language assessment and post-operative language outcomes.

Other related reviews have focused on pre-operative and post-operative (vs. intra-operative) neuropsychological assessment of brain tumor patients (Noll et al., 2019), a broad survey of intra-operative assessment tasks employed across diverse clinical groups (Ruis, 2018), and a critical review focusing specifically on intra-operative assessment of syntax, also for a mixed clinical group (Zanin et al., 2017). The general consensus from these reviews (which collectively evaluated literature up to 2016) was that there was a need to balance the special requirements of intra-operative language assessment with the need for higher psychometric standards. Specific recommendations included: 1) Increased use of standardized core language protocols, 2) more comprehensive assessment of language (and other cognitive domains), and 3) longitudinal post-surgical language outcome assessment to validate the selection of specific intra-operative language measures.

This systematic review follows up on these recommendations and evaluates the evidence supporting language assessment of adult brain tumor patients during awake craniotomy. Following up on previous recommendations is invaluable in understanding
current state-of-the-art procedures and in planning future research. Additionally, evaluation of evidence supporting current protocols has the potential to guide future decisions made by clinicians for selecting intra-operative language assessments supported by the most current and highest levels of evidence, thereby directly and positively impacting brain tumor survivor quality of life.

Specific research questions included: 1) What are the demographic and medical characteristics of the brain tumor patients included in these publications? 2) What neurosurgical procedures are associated with awake language mapping? 3) What language assessment tasks/measures are associated with awake mapping? 4) What language outcomes are reported following awake mapping? and 5) What is the quality of evidence for this literature?

Methods

Inclusion/Exclusion Criteria

Since the goal of this paper was to identify and evaluate cutting-edge procedures currently being used in practice, inclusion criteria were crafted to ensure that only studies with the highest levels of evidence would be included. Studies selected for inclusion met all of the following criteria: 1) available in English; 2) published in a peer-reviewed journal and contained original research; 3) published between 1976-June 2021 (electronic search 2014-2021); 4) was an experimental or quasi-experimental study, or correlational study comparing relation between pre-/intra- and post-operative treatment variables; 5) included at least one adult patient or participant (over 18 years of age) with a diagnosis of brain tumor; 6) identified the intra-operative neuroimaging technique(s) used; 7) included intra-operative assessment of at least 3 language domains; 8) at least one of the three
intra-operative language domains was assessed using a standardized and/or experimental task that is replicable (i.e., a published measure and/or adequate description of administration, cueing, stimuli, scoring, and interpretation procedures allowing for replication); 9) reported post-operative language outcome data using at least one standardized language assessment. For the full form used to screen articles, see Appendix A.

Search Strategies

Search strategies included an electronic search, an ancestral search, and a citation search. PRISMA guidelines (Schlosser, Ralph, & Sigafos, 2007; Shamshee et al., 2015) were followed for each of the strategies.

Studies were identified by: 1) searching for relevant papers in previous related reviews covering the time period between 1976-2016 (DeWitte & Marien, 2013; Martin-Monzon et al., 2020; Rofes & Miceli, 2014; Ruis, 2018; Zanin et al., 2017); 2) an electronic search covering 2015-to June 2021; and 3) ancestral and descendent searches of all identified articles.

Electronic Search Databases and Terms. The electronic search was completed using the PubMed, Web of Science, and PsychInfo databases. Each database search was completed separately, with key terms being coded according to the specifications of each individual database. The search terms below were entered into PubMed as given; the same terms as they were coded for the other databases (as recommended by USU and U of U librarians) can be found in Appendices B (PsychInfo) and C (Web of Science). In order to maximize consistency across the three databases, no limiting terms or filters were used other than publication year.
Search terms as entered into PubMed:

language OR speech OR apraxia OR aphasia OR communication [Mesh] OR "auditory comprehension" OR listening OR reading OR writing OR conversation OR discourse OR naming OR "word retrieval" OR "sentence production" OR "sentence comprehension" OR grammar OR linguistics OR semantics OR spelling OR linguistic OR "action naming" OR "verbs and sentences" OR “spontaneous speech” OR "language disorders"[Mesh] OR eloquent OR dyslexia OR “reading comprehension” OR alexia AND resection OR "awake craniotomy" OR "awake surgery" OR neuronavigation OR "intra-operative imaging" OR neurosurgery OR craniotomy OR "direct electrocortical stimulation" OR "intra-operative magnetic resonance imaging" OR "direct electrical stimulation" OR "intra-operative stimulation mapping" OR "cortical stimulation" OR "intra-operative monitoring" OR "diagnostic imaging" [Mesh] OR "aphasia/surgery" [Mesh] OR "intra-operative neurophysiological monitoring" [Mesh]) AND (glioma OR tumor OR cancer OR astrocytoma OR ependymoma OR glial OR glioblastoma OR malignant OR meningioma OR medulloblastoma OR metastasis OR oligodendroglioma OR oncology OR chondrosarcoma OR craniopharyngioma OR cyst OR hemangioma OR hemangioblastoma OR hemangiopericytoma OR lipoma OR lymphoma OR neurofibromatosis OR oligoastrocytoma OR schwannoma AND evaluation OR assessment OR test OR task OR "standardized test battery" OR analysis OR tool OR protocol OR screening OR measure OR "mapping diagnostic" OR monitoring OR examine OR investigate OR probe OR questionnaire OR standardized OR "non standardized" OR appraising OR "neuropsychological tests" [Mesh] OR "language tests" [Mesh] OR “mental status test” AND brain OR cerebral OR cortical OR subcortical OR
frontal lobe OR parietal lobe OR temporal lobe OR occipital OR limbic OR thalamic OR cranial OR “central nervous system” OR “arcuate fasciculus” OR “longitudinal fasciculus” OR “posterior cingulate cortex” OR “uncinate fasciculus” OR “fronto-occipital fasciculus” OR “parietopontine” OR “cingulate cortex”.

The electronic search yielded a total of 1,898 original articles for screening (see Figure 1). A total of 15 articles from this search method met the inclusion criteria.

**Ancestral and Descendant Search**

Both the ancestral and descendant searches were completed for all articles included in the review (See Figure 1).

An ancestral search was completed by screening all references within the articles that met inclusion criteria. For the descendant search, all included articles were entered into Google Scholar and citations checked for additional relevant articles. Relevant articles were screened to see if they met the inclusion criteria. An additional four articles were identified through these methods, with ancestral and descendant searches also being completed for each new addition.

**Study Selection**

Initial screening of all articles, regardless of the search strategy by which they were identified, consisted of reviewing article titles and abstracts for relevance. If it was unclear from the title and abstract if the article should be included/excluded, each reviewer independently accessed the entire article to determine inclusion/exclusion. To ensure reliability, a second screener independently reviewed 50% of the records. All reviewers used a standardized form developed specifically for this study (Appendix A) to
increase intra-coder agreement. Inclusion reliability was 98.3%, and any disagreements were discussed until a consensus was reached.

The search process identified a total of 19 papers for inclusion. See Figure 1 below. For the list of references for all included articles, see Appendix D.

**Figure 1**

*Study Inclusion Flowchart*

**Data Extraction**

An adapted (from Milman et al., 2017; Watt, 2018) data extraction manual (Appendix E) and spreadsheet were used to extract information from the included articles. The form focused on five primary areas: 1) participant information such as basic demographics (e.g., age, gender, language(s) spoken) and medical status (e.g., tumor laterality, type, severity), 2) neuroimaging data related to intra-operative language mapping (e.g., candidacy restrictions, type(s) of imaging employed, extent of resection),
3) pre-, post-, and intra-operative language assessment data (e.g., tasks, language domains assessed, standardization); 4) post-operative language outcomes (e.g., timing of follow-ups, language impairments) and 5) key indicators of research quality (e.g., design, blinding, intent to treat) of included studies using quality indicator scales (e.g., ASHA QI, RoBiNT, PEDro).

A second coder independently extracted information from 57.9% (11/19) of the included articles. All coding was conducted by SLP graduate students who successfully completed reliability coding training (measured by 80% reliability accuracy on at least one training paper previously coded by both CJ and LM with 100% agreement). Interrater reliability was derived by dividing the total number of items that the coders agreed on by the total number of items coded. Interrater reliability was calculated as 95%. All disagreement was discussed until there was complete consensus.

Data Synthesis and Interpretation

A reference list of included articles was generated (see Appendix D). Tables and figures display data regarding study inclusion, patient demographics and medical characteristics, intra-operative imaging, language assessment, outcomes, and research quality indicators. Due to variability across studies (protocols and tasks, demographic characteristics, tumor location, study design), descriptive/qualitative synthesis was used, following procedures modeled by previous reviews (Murray et al., 2018; Noll et al., 2019; Ruis, 2018; Watt, 2018; Zanin et al., 2018).

Results
Data regarding patient demographics and medical characteristics, intra-operative imaging, language assessment, outcomes, and research quality are presented in the text and tables below.

**Participant Demographics**

Demographic data for all participants are summarized in Table 1. Of 586 participants from the included studies, 471 (80%) were brain tumor survivors that underwent awake craniotomy. Of the 471 brain tumor participants undergoing awake craniotomy (plus 3 additional participants whose gender information was inseparable from the reported brain tumor survivor (BTS) data), 278 were male and 196 were female. The weighted average age of the brain tumor survivors was 44.7 and ranged from 15 to 75. Participants were recruited from various nations in Europe (n= 16 studies) and Asia (n= 3 studies). Of the available data, severity of pre-existing language impairments was classified as mild (n= 31), moderate (n=2), severe (n= 3), or unspecified (n= 91). The study by Leote et al. (2020) excluded patients with aphasia (as determined by an error rate greater than 50% on any of the given language tests) and Zigiotto et al. (2020) only considered patients for awake craniotomy if the individual had “no language deficit, confusion and/or anxiety disorder” (p. 98). Reported comorbidities included seizures (n= 45), hemiparesis (n=11), intracranial hypertension (n=8), unspecified motor deficit (n=6), central facial palsy (n=1), astereognosia (n=1), and dysphagia (n=unspecified).

**Brain Tumor Characteristics**

Brain tumor characteristics are summarized in Table 2. Tumors were localized to the left hemisphere (n= 372), right hemisphere (n=81) or were unspecified (n=18). In most cases tumors were restricted to a single cortical lobe or region, including tumors
restricted to the frontal lobe (n= 182), the temporal lobe (n=99), the parietal lobe (n= 40), the occipital lobe (n=4), insula (n=20), paralimbic (n=10), and subcortical areas (n=5). In addition, tumors were reported in diffuse areas (spanning more than one lobe) (n=80) or were unspecified/described without quantitative data (n=31). Diffuse tumors spanned two (n= 71) or three (n=9) lobes. Tumor subtypes were also identified as astrocytoma (n=76), glioblastoma (n= 76), oligodendroglioma (n= 47), oligoastrocytoma (n= 11), glioma (n= 247), oligodendroglioma astrocytoma (n= 1), neuroglial (n=1), gliosarcoma (n=1), ependymoma (n=1), cavernoma (n=4), DNET (n=2), and metastasized (n= 4). Tumors were further classified (WHO, Louis et al., 2016) as grade I (n=2), as grade II (n= 130), between grades II and III (n= 4), grade III (n= 47), and grade IV (n= 36); or simply described as “low grade” (n= 88), “high grade” (n= 106), “other” (n= 7), or grade was not reported (n= 51).
<table>
<thead>
<tr>
<th>Study</th>
<th>BTS underwent AC/Total participants</th>
<th>Age Mean (Range)</th>
<th>Gender (M/F)</th>
<th>Language(s) Spoken (n) or country where study conducted</th>
<th>Description of Baseline (pre-operative) Language Impairments (n)</th>
<th>Comorbidities (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019)</td>
<td>23/23</td>
<td>48 (21-72)</td>
<td>13/10</td>
<td>- Italy</td>
<td>Speech (7), general cognition (7)</td>
<td>Hemiparesis (2), Dysphagia (unspecified)</td>
</tr>
<tr>
<td>Bartha et al. (2000)</td>
<td>5/5</td>
<td>35 (±5.1)</td>
<td>3/2</td>
<td>- Austria</td>
<td>Mild articulation (1), Mild writing (1), Moderate fluency (2), Severe fluency (2)</td>
<td>None</td>
</tr>
<tr>
<td>Bello et al. (2007)</td>
<td>88/88</td>
<td>43 (32-68)</td>
<td>43/45</td>
<td>- Italy</td>
<td>Mild impairment (23)</td>
<td>“Mild motor deficit” (6)</td>
</tr>
<tr>
<td>Chang et al. (2018)</td>
<td>17/19</td>
<td>41 (20-65)</td>
<td>9/10b</td>
<td>- Taiwan</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>De Witte et al. (2015)</td>
<td>3/3</td>
<td>46 (41-53)</td>
<td>3/0</td>
<td>- Netherlands</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Herbet et al. (2019)</td>
<td>11/11</td>
<td>44 (27-58)</td>
<td>7/4</td>
<td>French (10); French-Arabic Bilingual (1)</td>
<td>Mild naming (1)</td>
<td>-</td>
</tr>
<tr>
<td>Leote et al. (2020)</td>
<td>19/19</td>
<td>50 (31-68)</td>
<td>11/8</td>
<td>Portuguese (19)</td>
<td>General language (6); Aphasia (3)</td>
<td>Contralateral hemiparesis (1), Central facial palsy (1), Seizures (12)</td>
</tr>
<tr>
<td>Mandonnet et al. (2019)</td>
<td>10/12</td>
<td>40 (21-72)</td>
<td>6/5 b</td>
<td>- France</td>
<td>-</td>
<td>Seizures (11)</td>
</tr>
<tr>
<td>Martino et al. (2018)</td>
<td>32/32</td>
<td>44 (36-40), No VM: 46 (25-62)</td>
<td>20/12</td>
<td>- Spain</td>
<td>General language (3)</td>
<td>-</td>
</tr>
<tr>
<td>Moritz-Gasser et al. (2013)</td>
<td>8/8</td>
<td>42 (21-52)</td>
<td>6/2</td>
<td>- France</td>
<td>Naming (2); Semantics (4); semantic fluency (6)</td>
<td>-</td>
</tr>
<tr>
<td>Motomura et al. (2020)</td>
<td>50/50</td>
<td>44 (17-69)</td>
<td>33/17</td>
<td>- Japan</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nakajima et al. (2019)</td>
<td>66/66</td>
<td>43 (16-73)</td>
<td>39/27</td>
<td>- Japan</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rofes et al. (2017)</td>
<td>6/6</td>
<td>49 (37-70)</td>
<td>4/2</td>
<td>Italian (6)</td>
<td>General language Mild (4), Severe (1)</td>
<td>-</td>
</tr>
<tr>
<td>Satoer et al. (2018)</td>
<td>18/39</td>
<td>BTS: 38 (19-62) NHC: 39 (19-62)</td>
<td>15/3</td>
<td>- Netherlands</td>
<td>Mild deficits in incomplete sentences, naming, and categorical fluency (BTS group)</td>
<td>-</td>
</tr>
<tr>
<td>Study</td>
<td>BTS underwent AC/Total study participants</td>
<td>Age Mean (Range)</td>
<td>Gender (M/F)</td>
<td>Language(s) Spoken (n) and/or country where study conducted</td>
<td>Baseline (pre-operative) Language Impairments (n)</td>
<td>Comorbidities (n)</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Signorelli et al. (2007)</td>
<td>28/101</td>
<td>53 (28-74)</td>
<td>17/11</td>
<td>- France and Italy</td>
<td>Mixed aphasia (5), Nonfluent aphasia (7), Memory (5)</td>
<td>Seizures (15), Intracranial hypertension (8), Facial-brachial hemiparesis (8), Astereognosia (1)</td>
</tr>
<tr>
<td>Starowicz-Filip et al. (2020)</td>
<td>15/15</td>
<td>48 (24-63)</td>
<td>10/5</td>
<td>- Poland</td>
<td>3/9 ACE III score below 82, 5/9 ACE III below 88</td>
<td>-</td>
</tr>
<tr>
<td>Tomasinio et al. (2020)</td>
<td>49/49</td>
<td>42 (15-67)</td>
<td>24/25</td>
<td>Italian (49)</td>
<td>Reading (7), verbal short-term memory (2), object naming (8) action naming (12), verbal fluency (7)</td>
<td>-</td>
</tr>
<tr>
<td>Zemmoura et al. (2015)</td>
<td>7/7</td>
<td>42 (23-59)</td>
<td>5/2</td>
<td>French-English Bilingual (1), French (4), English (1), Not specified (1)</td>
<td>Mild reading (1)</td>
<td>Seizures (7)</td>
</tr>
<tr>
<td>Zigiotto et al. (2020)</td>
<td>16/33</td>
<td>51 (32-74)</td>
<td>10/6</td>
<td>Italian (16)</td>
<td>None d</td>
<td>None</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>471/586</strong></td>
<td><strong>15-74</strong></td>
<td><strong>278/196</strong></td>
<td><strong>Europe (16), Asia (3)</strong></td>
<td><strong>-</strong></td>
<td><strong>Seizures (45), Hemiparesis (11), Intracranial hypertension (8), Motor (6), Central facial palsy (1), Astereognosia (1), Dysphagia (unspecified)</strong></td>
</tr>
</tbody>
</table>

NHC=Neurologically Healthy Controls; BTS= brain tumor survivors, VM= DES with verbal memory mapping; No VM= did not do verbal memory during intra-operative DES; ACE III = Addenbrooke’s Cognitive Examination

a Standard deviation; range not available

b Sex of patient(s) that did not undergo AC remained unspecified

c Median data only/ mean not available

d awake criteria included no lang impairment
Table 2  
*Tumor Characteristics*

<table>
<thead>
<tr>
<th>Study (n)</th>
<th>Tumor Laterality (L/R)</th>
<th>Anatomic Area of Tumor</th>
<th>Tumor Type</th>
<th>Tumor Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019) (n=23)</td>
<td>13/10</td>
<td>Fronto-temporo-insular (3), Frontal (8), Parietal (2) Fronto-parietal (2), Fronto-insular (1), Parieto-insular (1) Parieto-temporal (6)</td>
<td>DNET (1), Astrocytoma (4), Oligodendroglioma (6), Anaplastic astrocytoma (3), Glioblastoma (9)</td>
<td>I (1), II (10), III (3), IV (9)</td>
</tr>
<tr>
<td>Bartha et al. (2000) (n=5)</td>
<td>5/0</td>
<td>Involvement of Hippocampus, Amygdala, and Basal ganglia, Insula, pre-post central gyrus, superior temporal gyrus, temporal pole, frontal operculum, inferior frontal and supramarginal gyrus</td>
<td>Astrocytoma (4), Oligodendroglioma (1)</td>
<td>I (1), II (3), III (1)</td>
</tr>
<tr>
<td>Bello et al. (2007) (n=88)</td>
<td>69/1</td>
<td>Fronto (39), Paralimbic (10), Temporal (32), Parietal (7)</td>
<td>Glioblastoma (36), Anaplastic glioma (8), Glioma (44)</td>
<td>Low grade (44), High grade (44)</td>
</tr>
<tr>
<td>Chang et al. (2018) (n=17)</td>
<td>17/0</td>
<td>Frontal lobe (12), Parietal (1), Fronto-temporal (3), Fronto-parietal (1), Fronto-insular (1), Insula (1)</td>
<td>Glioblastoma (2), Anaplastic astrocytoma (4), Anaplastic oligoastrocytoma (4), Diffuse Astrocytoma (3), Oligodendroglioma (4)</td>
<td>II (7), III (8), IV (2)</td>
</tr>
<tr>
<td>De Witte et al. (2015) (n=3)</td>
<td>3/0</td>
<td>Temporo-parietal (1), Fronto-temporal (1), Parieto-occipital (1)</td>
<td>Astrocytoma (2), Oligoastrocytoma (1)</td>
<td>II/III (3)</td>
</tr>
<tr>
<td>Herbet et al. (2019) (n=11)</td>
<td>11/0</td>
<td>Temporal (10), Temporo-occipital (1)</td>
<td>Diffuse glioma (11)</td>
<td>II (11)</td>
</tr>
<tr>
<td>Leote et al. (2020) (n=19)</td>
<td>19/0</td>
<td>Temporal (6), Frontal (8), Parietal (2), Insula (3)</td>
<td>Anaplastic astrocytoma (2), Astrocytoma (4), Glioblastoma (5), Oligodendroglioma astrocytoma (1), Oligodendroglioma (6), Neuroglial (1)</td>
<td>High grade (8)</td>
</tr>
<tr>
<td>Mandonnet et al. (2019) (n=10)</td>
<td>7/3</td>
<td>Insula (10)</td>
<td>IDH-mutated glioma (10)</td>
<td>-</td>
</tr>
<tr>
<td>Martino et al. (2018) (n=32)</td>
<td>20/12</td>
<td>Supplementary motor area (12), Lateral premotor/frontal operculum (6), Superior parietal/postcentral gyrus (1), Insula (6), Temporal (7)</td>
<td>Glioma (22), Anaplastic glioma (10)</td>
<td>II (22)</td>
</tr>
<tr>
<td>Moritz-Gasser et al. (2013) (n=8)</td>
<td>8/0</td>
<td>- (Temporal or temporo-occipital regions)</td>
<td>Glioma (8)</td>
<td>II (8)</td>
</tr>
<tr>
<td>Motomura et al. (2020) (n=50)</td>
<td>34/16</td>
<td>Superior frontal (31), Middle frontal gyrus (10), Inferior frontal gyrus (1), Precentral gyrus (5), Cingulate gyrus (3)</td>
<td>Diffuse astrocytoma (21), Oligodendroglioma (13), Oligoastrocytoma (4), Anaplastic astrocytoma (5), Anaplastic oligodendroglioma (7)</td>
<td>II (38), III (12)</td>
</tr>
<tr>
<td>Study</td>
<td>Tumor Laterality (L/R)</td>
<td>Anatomic Area of Tumor</td>
<td>Tumor Type</td>
<td>Tumor Grade (n)</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Nakajima et al. (2019) (n=66)</td>
<td>36/30</td>
<td>Frontal (31), Temporal (18), Parietal (14), Occipital (3)</td>
<td>Glioma (66)</td>
<td>II (24), III (19), IV (23)</td>
</tr>
<tr>
<td>Rofes et al. (2017) (n=6)</td>
<td>6/0</td>
<td>Prefrontal (5), Parietal (1)</td>
<td>Oligodendroglia (2), Oligoastrocytoma (1), Glioblastoma (1), Astrocytoma (1), Anaplastic astrocytoma (1)</td>
<td>II (2), III (3), IV (1)</td>
</tr>
<tr>
<td>Satoer et al. (2018) (n=18)</td>
<td>16/2</td>
<td>“Language areas” (Inferior frontal gyrus/supramarginal gyrus/angular gyrus/inferior temporal gyrus/middle temporal gyrus/superior temporal gyrus/Subcentral Gyrus) (11); “Non-language areas” (Precentral gyrus/middle frontal gyrus/superior frontal gyrus) (5), Unspecified (2)</td>
<td>Glioma (18)</td>
<td>Low grade (11), High grade (7)</td>
</tr>
<tr>
<td>Signorelli et al. (2007) (n=28)</td>
<td>27/1</td>
<td>Temporal (10), Parietal (4), Frontal (3), Temporo-parietal (3), Fronto-parietal (1), Paracentral lobule and supplementary motor area (2), Temporo-insular (3), Fronto-temporo-insular (2)</td>
<td>Mixed glioma (2), Oligodendroglia (7), Anaplastic oligodendroglia (1), Gliosarcoma (1), Metastasis of epidermoid carcinoma (1), Pilocytic astrocytoma (2), Recurrent ependymoma (1), Glioblastoma (13)</td>
<td>II (3), Low grade (5)</td>
</tr>
<tr>
<td>Starowicz-Filip et al. (2020) (n=15)</td>
<td>15/0</td>
<td>Temporal (3), Parietal (1), Frontal (8) Temporo-insular (3)</td>
<td>Astrocytoma diffusum (11), Astrocytoma malignum (1), Glioblastoma (1), Glioma miutum paritum anaplasticum (1), Oligoastrocytoma (1)</td>
<td>II (12), II/III (1), III (1), IV (1)</td>
</tr>
<tr>
<td>Tomasino et al. (2020) (n=49)</td>
<td>49/0</td>
<td>Temporal (6), Frontal (13), Parietal (5), Fronto-temporo-insular (4), Temporo-insular (3), Temporo-parietal (10), Fronto-temporal (3), Temporo-occipital (1), Frontal-insular (2), Frontal-medial orbital (1), Temporo-hippocampal (1)</td>
<td>Metastases (3), Cavernoma (4), Pilocytic astrocytoma (1), Glioma (40), DNET (1)</td>
<td>Low grade (28), High grade (14), Other (7)</td>
</tr>
<tr>
<td>Zemmoura et al. (2015) (n=7)</td>
<td>7/0</td>
<td>Hippocampus (2), Temporal (3), Occipital (1), Left fusiform/occipital lobe (1)</td>
<td>Glioma (7)</td>
<td>II (7)</td>
</tr>
<tr>
<td>Zigiotto et al. (2020) (n=16)</td>
<td>10/6</td>
<td>Temporal (4), Frontal (8), Parietal (2), Temporo-insular (2)</td>
<td>Anaplastic astrocytoma (7), Glioblastoma (9)</td>
<td>High grade (16)</td>
</tr>
<tr>
<td>Totals (n=471)</td>
<td>372/81 Not reported= 18</td>
<td>Frontal (182), Temporal (99), Parietal (40), Occipital (4), Insula (20), Diffuse (80), Paralimbic (10), Subcortical (5), Unspecified (31)</td>
<td>Astrocytoma (76), Glioblastoma (76), Oligodendroglia (47), Oligoastrocytoma (11), Glioma (247), Oligodendroglia astrocytoma (1), Neuroglial (1), Gliosarcoma (1), Metastasized (4), Ependymoma (1), Cavernoma (4), DNET (2)</td>
<td>I (2), II (130), II/III (4), III (47), IV (36), Low grade (88), High grade (106), Other (7), Not reported (51)</td>
</tr>
</tbody>
</table>

a Rest not reported
b 2 additional participants not with BT included; unable to specify
Neurosurgical Procedures Associated with Awake Language Mapping

Neurosurgical procedures related to language mapping during awake craniotomy are summarized in Table 3. Six studies provided criteria for awake craniotomy candidacy, including characteristics such as age (n=1 study), tumor location (n= 2 studies), presence of cardiac pacemakers (n=1 study), anesthesiologic contraindications (n=1 study), indications of intracranial pressure (n= 1 study), behavioral abilities (n=1 study), psychological profiles (n=2 studies), and presurgical aphasia or language deficits (n= 2 studies). DES was used in all of the included studies. Imaging techniques paired with DES included: intraoperative ultrasound (iUS) (n=1 study), navigated fMRI (nfMRI) (n=1 study); intraoperative MRI (iMRI) (n=1 study); and monitoring of electrical discharge through electrocorticography (eCoG) (n= 2 studies). For DES electrical stimulation, all studies used a bipolar probe stimulator, with frequency widths 50-60 Hz. Amperage varied, and the duration of stimulations ranged from one millisecond to four seconds. Extent of Resection (EOR) was reported as gross total (n= 77), subtotal/partial (n= 48), supratotal (n= 46), and four studies (Altieri et al., 2020; Martino et al., 2018; Satoer et al. 2018; Tomasino et al., 2020), including a total of 122 participants, reported group data with reported ranges of 12.6-202%. 
<table>
<thead>
<tr>
<th>Study (n)</th>
<th>I/E Criteria of AC Candidates</th>
<th>Type of Intraoperative Imaging (n)</th>
<th>Imaging Specifications</th>
<th>EOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019) (n=23)</td>
<td>- <em>tumor not in high-functional area</em></td>
<td>DES (23), eCoG (23)</td>
<td>DES: bipolar electrode eCoG: stimulation intensity of 2.5 mA to 10 mA</td>
<td>Median 82.42% (R: 12.6-100%)</td>
</tr>
<tr>
<td>Bartha et al. (2000) (n=5)</td>
<td>-</td>
<td>DES (5)</td>
<td></td>
<td>GT (2), ST (3)</td>
</tr>
<tr>
<td>Bello et al. (2007) (n=88)</td>
<td>No symptoms of high intracranial pressure or preexisting language impairment</td>
<td>DES (88)</td>
<td>Bipolar electrode (60 Hz, pulse duration 1 ms to 4 s), amplitude began at 2 mA and increased by 0.5 mA</td>
<td>-</td>
</tr>
<tr>
<td>Chang et al. (2018) (n=17)</td>
<td>-</td>
<td>DES (17)</td>
<td>Bipolar electrode (50 Hz, pulse duration 1 ms), amplitude ranged from 2 to 10 mA</td>
<td>GT (11), ST (3), “partial” (3)</td>
</tr>
<tr>
<td>De Witte et al. (2015) (n=3)</td>
<td>-</td>
<td>DES (3)</td>
<td></td>
<td>ST (3)</td>
</tr>
<tr>
<td>Herbet et al. (2019) (n=11)</td>
<td>-</td>
<td>iUS (11); DES (11)</td>
<td>Bipolar electrode (60-Hz, pulse width 1 ms); amperage ranged from 1 to 4 mA</td>
<td>-</td>
</tr>
<tr>
<td>Leote et al. (2020) (n=19)</td>
<td>* lesion less than 20 mm from language brain regions or pathways, less than 50% on presurgical neuropsychological tests, no cardiac DES pacemakers</td>
<td>DES (19), nMRI (13)</td>
<td>DES: bipolar electrode (60 Hz, pulse duration 1000ms), amperage ranged from 1 to 5 mA, nMRI: used to inform and modulate the effort dispended (time and number of stimuli) for DCS at a specific brain region</td>
<td>-</td>
</tr>
<tr>
<td>Mandonnet et al. (2019) (n=10)</td>
<td>“Behavioral traits judged incompatible with procedure; sub optimal participation during awake period expected” due to age</td>
<td>DES (10)</td>
<td>(60 Hz, pulse duration 500 microseconds) amperage ranged from 1 to 3 mA</td>
<td>GT (1), ST (7), partial (4)</td>
</tr>
<tr>
<td>Martino et al. (2018) (n=32)</td>
<td>-</td>
<td>DES (32)</td>
<td>- *“bipolar electrode (60 Hz, pulse duration 1 msec), amplitude ranged from 2 to 8 mA”</td>
<td>VM group Mean: 82.2% (20-100%); No VM group Mean: 79% (51-100%)</td>
</tr>
<tr>
<td>Moritz-Gasser et al. (2013) (n=8)</td>
<td>-</td>
<td>DES (8)</td>
<td>Amplitude began at 1.5 mA and increased by 0.5 mA until reproducible response obtained</td>
<td>-</td>
</tr>
<tr>
<td>Motomura et al. (2020) (n=50)</td>
<td>-</td>
<td>DES (50), iMRI (50)</td>
<td>DES: bipolar electrode (60 Hz; pulse duration 0.5 msec) amperage ranged from 1 to 2 mA, iMRI: “0.4 T vertical field MR scanner”</td>
<td>GT (14), ST (6), SpT (13)</td>
</tr>
<tr>
<td>Nakajima et al. (2019) (n=66)</td>
<td>-</td>
<td>DES (66)</td>
<td></td>
<td>GT (27), ST (11), SpT (28)</td>
</tr>
<tr>
<td>Rofes et al. (2017) (n=6)</td>
<td>-</td>
<td>DES (6)</td>
<td>Bipolar electrode (50 ~ 60 Hz, pulse duration 0.2 ms); amperage ranged from 2.5 to 10 mA, stimulus for 4 seconds</td>
<td>-</td>
</tr>
<tr>
<td>Sato et al. (2018) (n=18)</td>
<td>-</td>
<td>DES (18)</td>
<td>Bipolar electrode (50 Hz, pulse duration 1 ms); amperage ranged from 6 to 12 mA</td>
<td>Mean 69% (3.12 - 91.33%)</td>
</tr>
<tr>
<td>Signorelli et al. (2007) (n=28)</td>
<td>-</td>
<td>DES (28); eCoG (28)</td>
<td>Bipolar electrode (60 Hz, pulse duration 1 ms), amplitude began at 1 mA and increased 0.5- to 1 mA until the threshold for after discharges were established by eCoG</td>
<td>GT (22), ST (6), fatality (1)</td>
</tr>
<tr>
<td>Starowicz-Filip et al. (2020) (n=15)</td>
<td>“Able to tolerate intraoperative functional monitoring until terminal stage of resection”</td>
<td>DES (15)</td>
<td>Amperage ranged from 2mA to 25mA</td>
<td>-</td>
</tr>
<tr>
<td>Tomasino et al. (2020) (n=49)</td>
<td>-</td>
<td>DES/real time neuropsychological testing (RTNT) (49)</td>
<td></td>
<td>Mean 93.16% (R:45-100%)</td>
</tr>
<tr>
<td>Zemmoura et al. (2015) (n=7)</td>
<td>-</td>
<td>DES (7)</td>
<td>Bipolar electrode (60 Hz, pulse length 1 ms) amplitude ranged from 1.5 to 4 mA, stimulation of each site lasted maximum 3 s,</td>
<td>-</td>
</tr>
<tr>
<td>Zigiotto et al. (2020) (n=16)</td>
<td>“Adequate psyche, profile and attitude”; no language deficit, confusion, and/or anxiety disorders; no anesthesiologic contra-indication</td>
<td>DES (16)</td>
<td>(60 Hz, pulse duration 1 ms), amplitude ranged from 2 to 4 mA</td>
<td>SpT (33)</td>
</tr>
</tbody>
</table>

DES= Direct electrical stimulation, iUS= intraoperative ultrasound, nMRI= navigated functional magnetic resonance imaging, iMRI= intraoperative magnetic resonance imaging, eCoG= electrocorticography; VM= group with verbal memory assessed during AC; No VM= group not assessed with verbal memory during AC; GT = gross total; ST = subtotal; SpT = supratotal

* Study criterion, not specific to participation in AC

*b 28 received language mapping; 73 received motor mapping only

* Rest not reported

* From citation (Martino et al., 2015)
Language Assessments/Tasks Associated with Awake Mapping

**Pre- and Post-Operative Assessment**

Pre- and post-operative language measures are summarized in Table 4. Of 37 language *assessments*, the most frequently used were verbal fluency tasks (n=9 studies), the DO-80 (n=7 studies), the Boston Diagnostic Aphasia Test/short form (BDAE/sfBDAE, n=4 studies), the Pyramids and Palm Trees Test (PPTT, n=4 studies), the Wechsler Adult Intelligence Scale/3rd Ed (WAIS/WAIS-III, n=4 studies), and unspecified naming tests (n=4 studies). One study did not specify the language assessment measures used but did use a well-known rating scale (National Outcomes Measurement System). Of the 21 *language domains* assessed peri-operatively, the most common were naming (n=17 studies), verbal fluency (n=12 studies), reading (n=10 studies), auditory comprehension (n=10 studies), and semantics (n=10 studies). Of the 37 language measures, 29 (78%) were either a published test measure or replicable from the report, 7 (19%) were deemed not replicable due to missing information, and one naming test (05%) was inferred to use a published test but this was not explicitly stated or reported. All studies administered at least one test both pre- and post-operatively; the majority (n=16) of studies administered multiple measures pre-/post- operatively, and 13 studies administered at least one measure at all three time points (pre-, post-, and intra-operatively). Only one study (Zemmoura et al., 2015) administered all measures used at all three time points.

**Intra-operative Assessment**

Intra-operative language assessment data is summarized in Table 5. Eight studies pre-operatively trained stimuli used in the intra-operative language assessments, while 11 did not specify whether pre-operative training was provided. Of the 26 intra-operative
language assessment measures identified across studies, the most frequently reported were counting (n= 12 studies), Pyramids and Palm Trees Test (n= 8 studies), the DO-80 (n=6 studies), unspecified naming tasks (n=9 studies), verbal fluency tasks (n=4 studies), discourse (n=4 studies), and unspecified language comprehension tasks (n=4 studies). Of the 18 language domains measured intra-operatively, the most commonly evaluated were naming (n= 19 studies), automatic speech (n= 14 studies), semantics (n= 12 studies), spontaneous speech (n= 6 studies), and reading (n= 6 studies). Of the 26 measures identified across studies, the majority used published or replicable tests. However, for these measures, five studies did not report replicable counting, six studies did not report replicable naming tasks, one study did not report a replicable reading task, one study did not report a replicable repetitions task, one study did not report a replicable verbal fluency task, one study did not report a replicable discourse task, one study did not report a replicable memory span task, and one study did not report a replicable “sentence construction” task. Intra-operative language assessment stimuli were presented using a computer or laptop (n=9 studies), “slides” (n=1 study), or presentation modality was not reported (n= 9 studies).
<table>
<thead>
<tr>
<th>Study</th>
<th>Language Measures</th>
<th>Language Domains Assessed as Designated by Author(s)</th>
<th>Standardized/Replicable (Y/No)</th>
<th>Time Points Administered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019) (n=23)</td>
<td>“Writing, motor speech, comprehension, expression, reading, pragmatics, attention, memory, problem solving and visuo-perceptive functions… scored with the NOMS scale”</td>
<td>W, AC, R, O-attention, VM, O-problem solving, O-visuoperception</td>
<td>No (but rated with NOMS)</td>
<td>Pre, Post</td>
</tr>
<tr>
<td>Bartha et al. (2000) (n=5)</td>
<td>Boston Naming Test (BNT) (German) Aachen Aphasia Test (AAT) Rivermead Behavioral Memory Test Verbal fluency (categorical)</td>
<td>N AC, W, Rep, R VM, AC VF</td>
<td>Y</td>
<td>Pre, Post, Intra-operative Pre, Post Pre, Post Pre, Post Pre, Post</td>
</tr>
<tr>
<td>Chang et al. (2018) (n=17)</td>
<td>DO-80 Pyramids and Palm Trees (PPTT) Short form Boston Diagnostic Aphasia Exam- Chinese (sBDAE)</td>
<td>N Sem VR, Rep, AC, P</td>
<td>Y</td>
<td>Pre, Post, Intra-operative Pre, Post Pre, Post Pre, Post Pre, Post</td>
</tr>
<tr>
<td>Herbet et al. (2020) (n=11)</td>
<td>Naming test *unspecified; has 80 items</td>
<td>N - *inferred Y</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td>Leote et al. (2020) (n=19)</td>
<td>during fMRI: Semantic judgment (PALPA-Portuguese) Syntactic judgment Verbal fluency (categorical verb) Neurological examination: Picture naming Pyramids and Palm Trees Test (PPTT)</td>
<td>Sem SynG VF N Sem VF</td>
<td>Y</td>
<td>Pre Pre Pre, Post Pre, Post Pre, Post Pre, Post</td>
</tr>
<tr>
<td>Mandonnet et al. (2019) (n=10)</td>
<td>DO-80 Verbal fluency (categorical) Pyramids and Palm Trees Test (PPTT) Mathematics The Free and Cued Selective Reminding Test (FCSRT) Stroop Color and Word Test (SCWT) Pictures of Facial Affect</td>
<td>N VF Sem Calc VM O-attention, O-executive function O-emotion recognition</td>
<td>Y</td>
<td>Pre, Post, Intra-operative Pre, Post Pre, Post Pre, Post Pre, Post Pre, Post Pre, Post</td>
</tr>
<tr>
<td>Martino et al. (2018) (n=32)</td>
<td>DO-80 Boston Diagnostic Aphasia Exam (BDAE)- Commands Boston Diagnostic Aphasia Exam (BDAE)- Reading Rey Auditory Verbal Learning Test (RAVLT) WAIS-III Digit Span tests Symbol Digit Modality Test (SDMT) Stroop Color and Word Test (SCWT)</td>
<td>N AC R VM VM O-mental flexibility O-mental flexibility</td>
<td>Y</td>
<td>Pre, Post, Intra-operative Pre, Post Pre, Post Pre, Post Pre, Post Pre, Post Pre, Post</td>
</tr>
<tr>
<td>Study</td>
<td>Language Measures</td>
<td>Language Domains Assessed</td>
<td>Standardized/Replicable (Y/No)</td>
<td>Time Points Administered</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Moritz-Gasser et al. (2013) (n=8)</td>
<td>DO-80 Pyramids and Palm Trees Test (PPTT) Verbal fluency (semantic)</td>
<td>N</td>
<td>Y</td>
<td>Pre, Post, Intra-operative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sem</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VF (Sem)</td>
<td>Y</td>
<td>Pre</td>
</tr>
<tr>
<td>Motomura et al. (2020) (n=50)</td>
<td>Wechsler Adult Intelligence Scale 3rd Ed. (WAIS-III) Wechsler Memory Scale-Revised (WMS-R) O-Verbal comprehension, Sem VM, O-attention</td>
<td>O</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal fluency</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td>Nakajima et al. (2019) (n=66)</td>
<td>Picture naming WAIS III Picture arrangement task 2-back test Expression recognition test</td>
<td>N</td>
<td>No</td>
<td>Pre, Post, Intra-operative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sem</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VM</td>
<td>Y</td>
<td>Pre</td>
</tr>
<tr>
<td>Rofes et al. (2017) (n=6)</td>
<td>Battery for Analysis of Aphasic Deficits (BADA) Naming tasks Mental Deterioration Battery (MDB) Stroop Color and Word Test (SCWT) Verbal (and Spatial) Immediate Memory Span</td>
<td>N, VF, Rep, R N VM, AC, O-inhibition R, O-attention VM</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td>Satoe et al. (2018) (n=18)</td>
<td>Boston Naming Test (BNT) Verbal fluency (semantic) Spontaneous speech variables test</td>
<td>N</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VF</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SS-corrections, incomplete sentences, MLUw; Rep</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td>Tomasino et al. (2020) (n=49)</td>
<td>Naming task Verbal fluency Batteria per l’Analisi dei Deficit Afascici (BADA) reading Tests of oral and ideomotor apraxia WAIS Digit Span tests Token Test</td>
<td>N</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VF</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P, VF</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VM</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AC</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td>Zemmoura et al. (2015) (n=7)</td>
<td>DO-80 Verbal fluency (Phonological and semantic) Montreal-Toulouse Aphasia Battery (MT-86) Reading</td>
<td>N</td>
<td>Y</td>
<td>Pre, Post, Intra-operative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VF</td>
<td>Y</td>
<td>Pre, Post, Intra-operative*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R, W</td>
<td>Y</td>
<td>Pre, Post, Intra-operative</td>
</tr>
<tr>
<td>Zigliotto et al. (2020) (n=16)</td>
<td>Dissocizioni semantiche intercategoriali Tre test clinici di recerca e produzione lessicale Mental Deterioration Battery (MDB) DO-80 Verbal fluency (semantic)</td>
<td>N</td>
<td>Y</td>
<td>Pre, Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AC</td>
<td>Y</td>
<td>Pre</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VF</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VF</td>
<td>Y</td>
<td>Pre, post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>No</td>
<td>Pre, post</td>
</tr>
</tbody>
</table>

*inferred

\* indicates missing or insufficient information

AS = automatic speech, N = naming, Rep = repetitions, P = phonology, SynG = syntax/grammar, VF = verbal fluency, AC= auditory comprehension, R = reading, W = writing, VM = verbal memory, VR = verbal reasoning/judgement, Calc= calculation, SS = spontaneous speech, O = other, DuLIP= Dutch Linguistic Intraoperative Protocol,
<table>
<thead>
<tr>
<th>Study</th>
<th>Pre-operative Training of Intraoperative Items</th>
<th>Intra-operative Assessment</th>
<th>Language Domains Assessed</th>
<th>Standardized/Replicable (Y/No)</th>
<th>Presentation Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019)</td>
<td>-</td>
<td>Counting DO-80 PPTT</td>
<td>AS</td>
<td>No Y</td>
<td>C</td>
</tr>
<tr>
<td>Mandonnet et al. (2019)</td>
<td>-</td>
<td>Counting * Naming * (DO-80)</td>
<td>AS</td>
<td>No Y</td>
<td>-</td>
</tr>
<tr>
<td>Bartha et al. (2000)</td>
<td>-</td>
<td>Automatic speech AAT Naming</td>
<td>AS</td>
<td>Y Y</td>
<td>-</td>
</tr>
<tr>
<td>Leote et al. (2020)</td>
<td>Y</td>
<td>Counting DO-80 PPTT</td>
<td>AS</td>
<td>Y Y</td>
<td>-</td>
</tr>
<tr>
<td>Chang et al. (2018)</td>
<td>Y</td>
<td>Counting DO-80 PPTT</td>
<td>AS</td>
<td>No Y</td>
<td>-</td>
</tr>
<tr>
<td>De Witte et al. (2015)</td>
<td>Y</td>
<td>Counting Discourse</td>
<td>AS</td>
<td>Y Y</td>
<td>C</td>
</tr>
<tr>
<td>Herbet et al. (2019)</td>
<td>-</td>
<td>Counting DO-80 PPTT</td>
<td>AS</td>
<td>Y Y</td>
<td>C</td>
</tr>
<tr>
<td>Martino et al. (2018)</td>
<td>-</td>
<td>Counting Naming * (DO-80)</td>
<td>AS</td>
<td>No Y</td>
<td>- “slides”</td>
</tr>
<tr>
<td>Motomura et al. (2020)</td>
<td>-</td>
<td>Naming Digit span</td>
<td>N</td>
<td>Y Y</td>
<td>C</td>
</tr>
<tr>
<td>Nakajima et al. (2019)</td>
<td>-</td>
<td>Naming WAIS-III Pic arranging Spatial 2-back test</td>
<td>N VR; “mentaling”</td>
<td>Y Y</td>
<td>-</td>
</tr>
<tr>
<td>Rofes et al. (2017)</td>
<td>Y</td>
<td>Counting Naming</td>
<td>AS</td>
<td>Y Y</td>
<td>C</td>
</tr>
<tr>
<td>Satoer et al. (2018)</td>
<td>-</td>
<td>Naming Repetitions Discourse</td>
<td>N Rep (P, Sem, SynG)</td>
<td>No Y</td>
<td>-</td>
</tr>
<tr>
<td>Signorelli et al. (2007)</td>
<td>Y</td>
<td>Naming (BDAE?) Word fluency</td>
<td>N</td>
<td>Y Y</td>
<td>C</td>
</tr>
<tr>
<td>Study</td>
<td>Pre-operative Training of Intraoperative Items</td>
<td>Intra-operative Assessment</td>
<td>Language Domains Assessed</td>
<td>Standardized/ Replicable (Y/N)</td>
<td>Presentation Modality</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------</td>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Starowicz-Filip et al. (2020) (n=15)</td>
<td>-</td>
<td>Automatic speech Naming (ACE III?) Verbal fluency Speech comprehension (ACE III?) Discourse</td>
<td>AS N VF AC SS</td>
<td>Y</td>
<td>-</td>
</tr>
<tr>
<td>Tomasino et al. (2020) (n=49)</td>
<td>Y</td>
<td>BADA naming Repetition P &amp;P discrimination Digit span BADA reading Picture description</td>
<td>N Rep P VM R SS</td>
<td>Y No</td>
<td>C</td>
</tr>
<tr>
<td>Zigiotto et al. (2020) (n=16)</td>
<td>-</td>
<td>Counting Laiacona-Capitani PPTT Verb generation Reading the Mind in the Eyes Stroop test Reading test</td>
<td>AS N N Sem N O-motion recognition O-attention R</td>
<td>Y Y Y Y</td>
<td>-</td>
</tr>
</tbody>
</table>

*Performed for selective patients

‘-’ indicates insufficient information; was counted as unspecified task and non-replicable

P&P Discrimination=phonemic and phonological discrimination; AAT = Aachen Aphasia Test; BADA = Batteria per l’Analisi del Deficit Afasico; ACE-III = Addenbrooke’s Cognitive Examination III; DO-80 = Test de denomination oracle d’images; PPTT = Pyramids and Palm Trees Test; DuLIP = Dutch Linguistic Intraoperative Protocol; WAIS-III = Wechsler Adult Intelligence Scale Third Edition; MT-86 = Montreal- Toulouse Aphasia Battery; BDAE = Boston Diagnostic Aphasia Examination

AS = automatic speech, N = naming, Rep = repetitions, P = phonology, SynG = syntax/grammar, VF = verbal fluency, AC= auditory comprehension, R = reading, W = writing, VM = verbal memory, VR = verbal reasoning/judgement, Calc= calculation, SS = spontaneous speech, O= other(-specified)

C = computer/laptop
Language Outcomes Following Awake Mapping

Language outcomes are summarized in Table 6. From the available data, during the pre-operative period language impairments were reported in domains of naming (n=55), phonology (n=8), repetition (n=3), semantics (n=5), verbal fluency (n=16), auditory comprehension (n=3), verbal memory (n=16), verbal reasoning (n=1), attention (n=1), reading (n=8), and general language (n=61). During the acute post-operative period (immediate through 3 months post), there was a general increase of participants with impairments in the domains of phonology (n=12), repetition (n=4), semantics (n=19), verbal fluency (n=27), auditory comprehension (n=4), syntax and grammar (n=1), reading (n=35), writing (n=2), attention (n=7), visuospatial perception (n=9), verbal reasoning (n=2), and higher cognition (inhibition, “mentalizing”) (n=13); a decrease compared to pre-operative baselines in the number of participants with impairments in domains of naming (n=34) and general language (n=113); and no overall change in verbal memory (n=21). Of the available data, long-term (three months post-operative and beyond) outcomes compared to baseline showed a greater number of persons with documented impairments in the domains of cognitive mentalizing (n=1), semantics (n=7), syntax and grammar (n=1), reading (n=12), writing (n=1), and visuospatial perception (n=6); a decrease in number of individuals with impairment in the domains of naming (n=10), phonology (n=3), verbal fluency (n=9), auditory comprehension (n=2), verbal memory (n=8), and general language (n=27); and no change in the domain of repetitions (n=3) and attention (n=1).
Six studies indicated that language therapy took place during the post-operative stage. Of the 104 patients reported as working professionally prior to tumor resection, 92 (88%) were able to return to work.
## Table 6
**Language Outcomes Following Awake Mapping**

<table>
<thead>
<tr>
<th>Study</th>
<th>Language Domain</th>
<th>Pre-operative (Total Impaired /Total BTS)</th>
<th>Acute Post-operative (1D-up to 3M) (Total Impaired /Total BTS)</th>
<th>Long-term (3M+) (Total Impaired /Total BTS)</th>
<th>Post-operative Language Therapy Reported (Y/No)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019)</td>
<td>AC, VM, W, R, O-attention, O-problem-solving, O-visuo perception</td>
<td>7 (avg group NOMS score)</td>
<td>6 (avg NOMS)</td>
<td>7 (avg NOMS)</td>
<td>No</td>
</tr>
<tr>
<td>Bartha et al. (2000)</td>
<td>N (BNT), Rep, P, VF, AC, W, R, VM</td>
<td>1/5</td>
<td>4/5</td>
<td>1/5</td>
<td>Y-reported patients received none</td>
</tr>
<tr>
<td>Bello et al. (2007)</td>
<td>General language</td>
<td>23/88 &quot;mild language deficit&quot;</td>
<td>59/88 new or worsened</td>
<td>16/88 mild, 2/88 moderate or severe</td>
<td>No</td>
</tr>
<tr>
<td>Leote et al. (2020)</td>
<td>General language</td>
<td>6/19</td>
<td>4/19</td>
<td>1/3</td>
<td>No</td>
</tr>
<tr>
<td>Mandonnet et al. (2019)</td>
<td>N, VF, Sem, VM, O-attention, O-emotion recognition</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>Moritz-Gasser et al. (2013)</td>
<td>N, Sem, VM, VF, P</td>
<td>2/8</td>
<td>1/2</td>
<td>5/12</td>
<td>No</td>
</tr>
<tr>
<td>Motomura et al. (2020)</td>
<td>General language</td>
<td>-</td>
<td>15/50 &quot;speech disturbance&quot;</td>
<td>50/50 improved from pre-operative</td>
<td>No</td>
</tr>
<tr>
<td>Nakajima et al., (2019)</td>
<td>N, VM, O-visuospatial, O-mentoringalizing</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>Rofes et al. (2017)</td>
<td>N, Rep, N/SynG (finite verbs), VM, AC, W, O-attention, O-inhibition/VR</td>
<td>1/6</td>
<td>2/6</td>
<td>1/6</td>
<td>No</td>
</tr>
<tr>
<td>Study</td>
<td>Language Domain</td>
<td>Pre-operative (Total Impaired /Total BTS)</td>
<td>Acute Post-operative (1D-up to 3M) (Total Impaired /Total BTS)</td>
<td>Long-term (4M+) (Total Impaired /Total BTS)</td>
<td>Post-operative Language Therapy Reported (Y/No)</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------</td>
<td>------------------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Satoer et al. (2018) (n=18)</td>
<td>N VF Incomplete sentences MLUw</td>
<td>BTS group mean BTS group mean BTS group mean -</td>
<td>BTS group mean BTS group mean BTS group mean BTS group mean</td>
<td>BTS group mean BTS group mean</td>
<td>No</td>
</tr>
<tr>
<td>Starowicz-Filip et al. (2020) (n=15)</td>
<td>N</td>
<td>-</td>
<td>-</td>
<td>2/15</td>
<td>No</td>
</tr>
<tr>
<td>Tomasino et al. (2020) (n=49)</td>
<td>N VF R VM</td>
<td>8/49 (object), 12/49 (verb) 7/49 7/49 2/49</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zigiotto et al. (2020) (n=7)</td>
<td>VM</td>
<td>-- (39.3% mean accuracy on REY) -23.8% mean accuracy on REY</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zemmoura et al. (2015) (n=7)</td>
<td>N P Sem VF R</td>
<td>-</td>
<td>-</td>
<td>4/7 3/7 4/7 4/7 5/7 5/7 5/7</td>
<td>5/7</td>
</tr>
</tbody>
</table>

W = weeks, M = months, NOMS= National Outcomes Measurement System, AS = automatic speech, N = naming, Rep = repetitions, P = phonology, SynG = syntax/grammar, Sem = semantics, VF = verbal fluency, AC = auditory comprehension, R = reading, W = writing, VM = verbal memory, VR = verbal reasoning/judgement, Calc = calculation, SS = spontaneous speech; O= other; MLUw= mean length of utterance, number of words; ‘-‘ indicates all patents normal functioning or not reported
REY= rey-osterith memory tests (complex figure, 15-word recall, 15-word repetitions list)
*Rest not reported

a Only group data reported; not individualized results
Quality of Evidence

The quality of evidence for individual studies is reported in Table 7. According to Academy of Neurologic Communication Disorders and Sciences (ANCDS, 2006) classifications (see Appendix E), study designs were controlled group design (CGD, n= 6 studies), case series (uncontrolled with data not averaged across cases) (n=3), uncontrolled group (UG, n= 4), randomized control trial (RCT, n=1), correlational study (Corr, n=1), and single participant design (SPD, n=4). According to the American Academy of Neurology (AAN, 2007) level of evidence scale (corresponding with the research design type), studies included evidence level 3 (n=8 studies), level 2 (n=10 studies), and level 1 (n=1 study). The mean ASHA QI score (max = 9) was 4.68 (SD= 1.003, range= 3-6). The mean RoBINT score (max= 30) was 8.86 (SD=1.77, range= 6-11), and the mean PEDro score (max= 10) was 5.67 (SD= 0.78, range= 4-7).

Table 7

Quality Indicator Scores

<table>
<thead>
<tr>
<th>Study (n)</th>
<th>AAN* Research Design Type</th>
<th>AAN* Level of Evidence</th>
<th>ASHA QI Score (max = 9)</th>
<th>RoBINT (max = 30)</th>
<th>PEDro (max = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altieri et al. (2019)</td>
<td>CGD</td>
<td>2</td>
<td>4</td>
<td>NA</td>
<td>6</td>
</tr>
<tr>
<td>Bello et al. (2007)</td>
<td>CGD</td>
<td>2</td>
<td>6</td>
<td>NA</td>
<td>6</td>
</tr>
<tr>
<td>Bello et al. (2007)</td>
<td>Case Series</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>NA</td>
</tr>
<tr>
<td>Chang et al. (2018)</td>
<td>UG</td>
<td>3</td>
<td>3</td>
<td>NA</td>
<td>6</td>
</tr>
<tr>
<td>De Witte et al. (2015)</td>
<td>Case Series</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>NA</td>
</tr>
<tr>
<td>Herbet et al. (2019)</td>
<td>CGD</td>
<td>2</td>
<td>6</td>
<td>NA</td>
<td>6</td>
</tr>
<tr>
<td>Leste et al. (2020)</td>
<td>RCT</td>
<td>1</td>
<td>6</td>
<td>NA</td>
<td>7</td>
</tr>
<tr>
<td>Mandonnet et al. (2020)</td>
<td>Case Series</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>NA</td>
</tr>
</tbody>
</table>
### Discussion

To be included in this review, articles were required to include at least one patient older than 18 with a diagnosis of brain tumor; identify the neuroimaging technique(s) used during awake craniotomy; include intra-operative assessment of at least three language domains using at least one standardized and/or replicable measure; and report post-operative language outcome data using at least one standardized language assessment. 19 studies met these criteria. This discussion begins by summarizing the primary research findings and then reviews demographic characteristics, intra-operative neuroimaging techniques, language tests and domains assessed pre-, post-, and intra-
operatively, short- and long-term post-surgical language outcomes, as well as the overall quality of this research.

Patient demographic information varied widely, as did tumor characteristics, though the majority of tumors were located in left peri-sylvian areas. The majority of participants had mild-moderate language impairments prior to resection; this and the fact that some studies excluded participants with pre-existing language impairments is likely due to the fact that the purpose of language mapping during AC is to preserve existing language abilities.

Direct electrical stimulation (DES) was used in all of the included studies, and five studies included additional supplemental neuroimaging methods. Peri-operatively, the majority of measures were published and/or replicable and most assessed the language domains of naming, verbal fluency, reading, auditory comprehension, and semantics. The measures used to assess these and other language domains varied widely. Intra-operatively, a slightly less wide range of language assessments were used than those used peri-operatively, but more of these measures were not published/replicable. The most common domains assessed intra-operatively were naming, automatic speech, semantics, spontaneous speech, and reading. Language outcomes showed an overall increase in language impairment during the acute (up to three months post-operative) recovery phase compared to baseline (pre-operative), followed by a general lessening of deficits. The most common deficits identified long term (beyond three months post-operative), were in areas of general language, naming, verbal fluency, verbal memory, and semantics. The majority of the articles included in this review were designated as low
levels of evidence (AAN classifications two and three); however, one study was designated as a randomized control trial (AAN classification level one).

**Patient Demographics and Brain Tumor Characteristics**

One interesting finding is the geographic areas of recruitment for the included studies. The majority of studies were completed in Europe (n=16), with the remaining studies originating from Asia (n=3). This is not to say that DES and intraoperative language mapping assessments are not used in other parts of the world (e.g., North or South America), but rather that research from these countries has not focused on language assessment/outcomes (e.g., Almeida et al., 2014; Brown et al., 2020), did not meet study inclusion criteria (e.g., Shen et al., 2019; Shields & Choucair, 2014), and/or may not yet be published.

In keeping with the fact that 16 of the included articles recruited patients from Europe, most participants spoke Italian, French, or other European languages. Most of the participants did not use English as their first language. Only one participant in one study is explicitly stated as being a monolingual English speaker (Zemmoura et al., 2015). A second participant, from the same study, was described as French-English bilingual. Of the six articles that reported languages spoken by patients (n=121), 71 (58.7%) spoke Italian. This is particularly interesting given that English, then Mandarin Chinese, are the most frequently-spoken languages in the world (Szmigiera, 2021).

As expected, the majority of patients had left hemisphere and/or peri-sylvian tumors; but not all. In addition, relative to other studies, participants had a higher-than-expected number of tumors restricted to the frontal lobe (38.6%), compared to parietal lobe (8.49%) or temporal lobe (21%). Though many tumors classified as diffuse (n=80)
also extended into the parietal lobe (28/80, 35%). The findings reported here are largely consistent with a recent report by Hall et al. (2021) of 469 patients treated with awake craniotomy. In their study, tumor location was identified as frontal lobe (47%), parietal lobe (22.8%), and temporal lobe (17.9%) (Hall et al., 2021).

Compiling data for tumor grade and malignancy proved difficult due to variance in level of detail reported across studies. Four studies reported tumor grade as only “high” or “low” grade rather than the more detailed WHO grade classifications, while some of this data was selectively reported or not reported at all. However, given that low grade (grades I and II) tumors are generally slower-growing and more benign than the more malignant higher grade (grades III and IV) tumors, it could be stated that this study included 220 “more benign” tumors and 189 “more malignant” tumors, with 62 either not reported or between stages II and III.

**Neurosurgical Procedures Associated with Awake Language Mapping**

Only four studies stated conditions determining a patient’s candidacy for awake craniotomy, as opposed to other treatment (e.g., surgery under general anesthesia or “asleep” surgery). Two additional articles had study inclusion criteria that provided insight into the conditions that might prevent patients from undergoing awake resection. Of the reported conditions from these four studies, eight considerations centered on mental or behavioral factors (e.g., excluding individuals that did not meet criteria on psychological or language measures) and seven centered on physical or medical factors (e.g., excluding young or very elderly individuals or patients with intracranial pressure). However, the fact that the majority of articles included in this review did not specify
factors limiting patient candidacy for AC decreases the usefulness of these studies for clinicians in determining candidacy for their own patients.

Though DES was used for language mapping in all of the included studies, a wide range of complementary imaging techniques (intra-operative ultrasound, navigated functional magnetic resonance imaging, intra-operative magnetic resonance imaging, electrocorticography) were included in this sample. None of these supplemental techniques were favored over another by study count. This attests to the variety of imaging methods available and the unique data and information each technique provides (Gandhe & Bhave, 2018; Caras et al., 2020; Cirillo et al., 2019).

One of the most consistent aspects of the data collected was the DES stimulation process. Though there were some small variances in amplitude and duration of stimulations, many have converged on the extensive, well-known work of Wilder Penfield and George Ojemann (Berger et al., 1990; Ojemann et al., 1989; Penfield & Boldrey, 1937; Tarapore et al., 2012). The stimulation procedure is fairly standardized.

Language Assessments/Tasks Associated with Awake Mapping

In contrast to the well-standardized DES procedures described above, intra-operative procedures for behavioral language assessment are still in development. There is large variability in the number of measures used and the language domains assessed, with little justification for selecting particular measures or domains.

Naming, automatic speech (e.g., counting), and verbal fluency were among the most common domains of language assessed intra-operatively. This overlaps in part with findings by Martin-Mendoza et al. (2020), who found that naming, reading aloud, and generating verbs were the most common intra-operative tasks. The difference in reported
findings could be attributed to differences between classifications/terminology used in the
two reviews. This study classified all confrontational naming, including generation of
verbs from pictures, within the general domain of “naming” rather than separating object
or action naming. In addition, our criteria of requiring a minimum of three intra-operative
measures excluded a significant number of studies.

**Recommendations for Clinical Application**

The articles included in this review had several clinical recommendations for
selecting intra-operative language measures. Rofes et al. (2017) recommended a naming
task that includes both object naming and finite verb production over either naming task
alone, as their study found the combined naming to be more thorough. They also
recommended using standardized measures with normative data when possible, though
warned that normative data may be based on clinical populations that do not include brain
tumor patients (Rofes et al., 2017). Chang et al. (2018) found that the Pyramids and Palm
Trees Test had greater sensitivity and specificity in predicting post-operative language
deficits, compared to the DO-80, though intra-operative performance on both intra-
operative measures correlated with accuracy in the post-operative language measure (the
sfBDAE). Zemmoura et al. (2015) found that phonological and reading tasks were critical
for mapping areas in specific locations such as the posterior portion of the inferior
longitudinal fascicule. This coincides with the recommendation by De Witte et al. (2015)
that tumor location should be taken into account when selecting intra-operative language
assessments. Inclusion of measures assessing other neurocognitive functions, such as
digit span tests (Motomura et al., 2020) and verbal memory tests (Martino et al., 2018)
was also recommended to prevent significant decline of post-surgical abilities in these
areas. Mapping of subcortical areas was also recommended (Bello et al., 2007), as this was associated with a lower incidence of persistent long-term language impairments, albeit with a higher incidence of transient impairment in the acute post-operative phase. Satoer et al. (2018) recommended evaluating multiple linguistic levels during intra- and peri-operative assessment. It appears that the best intra-operative measures are personalized to tumor location, assess a wide range of language domains, and assess language ability at multiple linguistic levels. Selection of language measures that best meet these criteria requires further research and development.

These recommendations come from studies rated as level 2 according to the AAN (2017) levels of evidence, with the exception of the studies by Satoer et al. (2018) and Chang et al. (2018), which were level 3 studies.

**Language Outcomes Following Awake Mapping**

As expected, language outcomes showed a general pattern of increased impairment during the acute phase (immediate post-operative through three months post), followed by general improvement of language functions in long-term assessments compared to baseline (pre-operative measures) The exception to this was the study by Satoer et al. (2018), which showed the greatest decline in the average ability of the brain tumor survivor group between three months post-operative and the one-year follow up. However, as noted by the authors, this “is not in line with most cognitive outcome studies” (p. 115). This result was attributed to possible tumor recurrence. The overall long-term language impairments reported in Satoer et al.’s (2018) study were still considered to be mild.
The most common long-term (persisting longer than three months) language impairments were in the areas of general language (n=17), reading (n=12), naming (n=10), verbal fluency (n=9), and verbal memory (n=8). This is largely in agreement with previous reports (Brownsett et al., 2019; Chang et al., 2015; Papagno et al. 2016; Wilson et al., 2015). Though naming has been called the most frequently occurring post-operative impairment (Chang et al., 2015; Wilson et al., 2015), presenting deficits are likely dependent upon location of the tumor. This may partly explain our findings of higher occurrence of long-term impairment in reading over naming. Specifically, long-term deficit in naming has been associated with resection of temporal lobe areas (Krauss et al., 1996; Wilson et al., 2015). Our findings showed a higher occurrence of tumor in frontal and parietal rather than temporal areas.

There was considerable overlap between the language domains assessed (peri- and intra-operatively) and the domains of language with long-term impairments. This finding, summarized in Table 8 below, provides some support for the content validity of the selected language measures. Greater overlap was found between peri-operative (vs. intra-operative) measures and persistent language impairments. This is to be expected because it is the post-operative measures that detect and score post-operative language deficits. However, the domains of naming and reading are both frequently assessed intra-operatively and both are among the most common long-term impairments.
Table 8
Persistent Language Impairments Compared to Most Commonly Assessed Domains

<table>
<thead>
<tr>
<th>Language Domains</th>
<th>Language Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Commonly Assessed Peri-Operatively</td>
<td>Naming, Verbal Fluency, Reading, Auditory Comprehension, Semantics</td>
</tr>
<tr>
<td>Most Commonly Assessed Intra-Operatively</td>
<td>Naming, Automatic Speech, Semantics, Spontaneous Speech, Reading</td>
</tr>
<tr>
<td>Most Common Persistent (longer than 3 months post-operative) Language Impairments</td>
<td>General language, Reading, Naming, Verbal Fluency, Verbal Memory</td>
</tr>
</tbody>
</table>

An additional item of note concerning language outcomes is the sensitivity of measures used to detect possible impairments. Satoer et al. (2018) observed, “Most linguistic variables did not correlate with standardized tests usually administered in glioma patients. A spontaneous speech analysis seems to be more sensitive than standardized language tests for the detection of (surgery induced) language deficits until long-term” (p. 115). A higher incidence of post-operative language deficits has been found to be correlated with the existence of pre-surgical language deficits (Clavreul et al., 2021; Rolston et al., 2015) and possibly how extensive the post-operative language assessments are, as suggested by Brownsett et al. (2019).

Future studies investigating the extent to which identification of long-term language impairments are linked to variables such as those above (e.g., tumor location, pre-surgical language impairment, psychometric properties of language measures, such as content validity and test sensitivity) are still needed to further guide clinicians in selecting appropriate measures.

Quality of Evidence

With the exception of one randomized control trial (RCT), the majority of the studies were classified as intermediate or low-level evidence. That said, the studies included in this review represent the highest available evidence for intra-operative
language assessment of brain tumor survivors. Relative to other studies, strengths included reporting pre-, post, and intra-operative language status with at least one standardized and/or replicable measure (as per the inclusion criteria), reporting statistical significance for primary study findings, and describing awake craniotomy/language mapping (treatment) procedures. Generally, randomization and blinding of participants/assessors were weak points of the included studies, again with the exception of the randomized control trial. Many group studies did not have control groups or clearly defined control conditions.

**Contribution**

Previous related reviews covered literature up to 2016. This review shows that most recent literature is responding to recommendations made by previous reviews evaluating intra-operative language assessment practices (DeWitte & Marien, 2013; Martin-Monzon et al., 2020; Rofes & Miceli, 2014; Ruis, 2018; Zanin et al., 2017). This study compiles updated information on state-of-the-art practices for patient candidacy, intra-operative neuro-imaging procedures, and clinical recommendations for intra-operative language assessment, with the majority of included studies being published after 2016. In addition, this review not only looked at the most commonly-used assessments (Martin-Monzon et al., 2020; Ruis, 2018) but also evaluated the language domains assessed and the language outcomes following surgical procedure. Our findings found that more extensive language assessment is occurring and is even being encouraged by clinicians implementing this in practice, but that there may still be a need for using more standardized assessments intra-operatively.
Interest in this topic continues to expand, as shown in the descendent search. Though almost all of the articles identified through this search method did not meet our inclusion criteria, many of them from most recent years focused on language assessments specific to brain tumor or intra-operative language assessment and/or norming of new experimental protocols. As research continues to develop, more research may include standardized protocols more appropriate for this population.

**Limitations**

This study is limited in that the electronic search was only for the years 2014-2021. However, this limitation was partially offset by including articles identified in previously published reviews (DeWitte & Marien, 2013; Martin-Monzon et al., 2020; Rofes & Miceli, 2014; Ruis, 2018; Zanin et al., 2017) covering the literature back to 1976. Moreover, neuroimaging, and specifically intra-operative language assessment, have evolved rapidly since the first intraoperative review/protocol for language assessment was published. In addition, this review only reflects what is done in the reported literature and not what is done in all treatment clinics. As more language assessment batteries specifically for intra-operative use are being developed and tested, follow-up reviews will provide further guidance and insight.

**Conclusion**

This review provides a broad overview of state-of-the-art intra-operative language assessment practices for BT patients. A range of neuroimaging techniques and intra-operative language assessments were utilized across a broad range of patients with diverse tumor characteristics. Value lies in looking at trends in current assessment practices, analyzing the evidence, and considering the relations between demographic and
tumor characteristics, intra-operative procedures, and language outcomes. This information is invaluable to both neurosurgeons and administrators of intra-operative language assessments for making clinical decisions that will optimize best-of-care and quality of life for BT patients.
References


Appendix A. Study Inclusion Form
(adapted from Rajinder & Koul, 2015; and Watt, 2018)

Abbreviated Reference: Coder: Date:

STEP 1. Initial screen for inclusion: Please read through the list of publication titles (and abstracts if needed) in Rayyan. If the answer is clearly ‘No’ to ANY of the questions listed below, enter your rationale for exclusion (only enter the first reason for exclusion) in Rayyan using the number and letter codes listed below. If the answer to ALL questions could be ‘Yes’ (i.e., you are undecided and need more information), then continue with step 2.

STEP 2. Second screen for inclusion: Retrieve the full article and read as much of the article as necessary to make a decision. If the answer is ‘No’ to ANY of the questions listed below, enter your rationale for exclusion using the number and letter codes listed below.

<table>
<thead>
<tr>
<th>Criteria for Inclusion</th>
<th>Y or N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Is the full publication printed/available in English?</td>
<td></td>
</tr>
<tr>
<td>If ‘No,’ enter J In Rayyan for Reason excluded. For Label enter the language (e.g., “Spanish”)</td>
<td></td>
</tr>
<tr>
<td>2. Was the article published between 2014-June 2021?</td>
<td></td>
</tr>
<tr>
<td>Notes:</td>
<td></td>
</tr>
<tr>
<td>● Articles will be included if any version of the publication was published after 2014.</td>
<td></td>
</tr>
<tr>
<td>Reasons for Exclusion:</td>
<td></td>
</tr>
<tr>
<td>2a) The publication was not published within the specified time frame.</td>
<td></td>
</tr>
<tr>
<td>In Rayyan for Reason excluded enter 2a. For Label enter the year of publication.</td>
<td></td>
</tr>
</tbody>
</table>
3. Was the article published in a peer-reviewed journal and does it contain original research?

Notes:

- Systematic reviews will be excluded. In Rayyan, mark relevant reviews as ‘maybe’ so that they can be referenced in the introduction.

Reasons for Exclusion:

3a) The publication was not published in a peer-reviewed journal.

3b) Publication is a review of previously published research (relevant publications will be referenced in the introduction/discussion and checked for references).

3c) Publication is a duplicate (e.g., exact duplicate or online/hard copy reference of a publication that has already been referenced).

In Rayyan for *Reason excluded* enter 3a, 3b, or 3c.

4. Was the publication an experimental or quasi-experimental study (e.g., small n or single participant design including pre-/post treatment designs, controlled group design, randomized control trials, or meta-analysis) or a correlational study comparing the relation between pre-/intra- and post-treatment variables?

Notes:

- Refer to page 11 of the manual for more information regarding single participant designs (e.g., multiple baseline design across participants, behaviors, contexts; alternating treatment designs; multiple probe designs; withdrawal/reversal designs; changing criterion designs), controlled group designs, randomized control trials, and meta-analyses.

- The following non-experimental (no control group, condition, or phase) studies will be excluded: Single case studies/case reports; position papers; expert opinion; best practice recommendations.

- Retrospective studies of patient records can be included.
5. Participants: Does the study include at least one adult patient or participant (over 18 years of age) with diagnosis of brain tumor?

Notes:
- The tumor(s) may be anywhere in the brain (cerebrum, cerebellum, or brainstem).
- Cases with co-morbidities (stroke/seizures) and/or cases receiving concurrent medical treatments (e.g., chemotherapy/radiation) may be included.
- Metastases are to be included.
- Stroke subsequent to brain tumor excision may be included if there was at least one follow-up of language between resection and stroke.

Reasons for Exclusion:

5a) Participant(s) are in the wrong age group (<18 years), and/or this is a cadaver/autopsy, immunohistology, or genetic tissue study.

5b) None of the participants have a diagnosis of brain tumor (e.g., there is no diagnosis of tumor and/or the tumor is in the spinal cord and/or outside the CNS).

5c) Brain tumor is not the primary diagnosis. The study focuses on another pathology. (e.g., one or more participants have an additional diagnosis of brain tumor, but little information is given about the tumor and/or focus is on another primary condition)

In Rayyan for *Reason excluded*, enter 5a, 5b, or 5c.

6. Intervention: Does the publication identify the intra-operative neuroimaging technique(s) used?

Notes:
- This includes publications that do not exclusively focus on intra-operative imaging.

Reasons for Exclusion:
6a) The publication does not state that intra-operative imaging was used OR imaging was not intra-operative.

6b) No specific intra-operative imaging method (e.g., direct electrocortical stimulation/direct electrical stimulation (DES), electrocorticography (ECoG), fluoroscopy, intraoperative MRI (iMRI)) is given in the publication.

In Rayyan for *Reason excluded* enter 6a or 6b.

7. **Intervention:** Does the publication include intra-operative assessment of **naming** (e.g., confrontation or picture object and/or action naming) and at least two other tasks assessing the following:
   - Automatic speech (e.g., counting)
   - Reading
   - Writing
   - Spontaneous speech
   - Auditory comprehension
   - Syntax/grammar
   - Phonology
   - Semantics (e.g., sorting, picture-word matching)
   - Repetitions
   - Timed fluency tasks (generative naming and/or speech efficiency)
   - Bilingual speakers (e.g., naming in multiple language/code switching)
   - Calculation
   - Verbal memory
   - Verbal reasoning/judgement

**Notes:**
- Publications can focus on neuroimaging and/or medical/behavioral variables.

**Reasons for Exclusion:**

7a) The publication does not include any description of intra-operative assessment, or the assessment(s) are not of language (e.g., the study includes assessment of motor functions but not language).

7b) Only one or two domains of language processing is assessed (e.g., naming).
In Rayan for *Reason excluded* enter 7a or 7b.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>8.</strong> Intervention: Is at least <strong>one</strong> of the intra-operative language assessments/tasks (see examples listed above) a standardized test and/or an experimental task that is replicable (i.e., a published measure and/or adequate description in the study of administration, cueing, stimuli, scoring, and interpretation procedures allowing for replication)?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If ‘No,’ enter 8 for <em>Reason excluded</em> in Rayyan.</td>
</tr>
<tr>
<td><strong>9.</strong> Outcomes: Does the study report post-operative language outcome data using at least one standardized language assessment (e.g., WAB-R, BDAE-3, SCCAN, BNT, MoCA, etc.)?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If ‘No,’ enter 9 in Rayan for <em>Reason excluded</em>.</td>
</tr>
</tbody>
</table>
Appendix B. PsychInfo Search Terms

(language OR speech OR apraxia OR aphasia OR communication OR "auditory comprehension" OR listening OR reading OR writing OR conversation OR discourse OR naming OR "word retrieval" OR "sentence production" OR "sentence comprehension" OR grammar OR linguistics OR semantics OR spelling OR linguistic OR "action naming" OR "verbs and sentences" OR “spontaneous speech” OR eloquent OR dyslexia OR “reading comprehension” OR alexia)

AND (resection OR "awake craniotomy" OR "awake surgery" OR neuronavigation OR "intra-operative imaging" OR neurosurgery OR craniotomy OR "direct electrocortical stimulation" OR " intra-operative magnetic resonance imaging" OR "direct electrical stimulation" OR "intra-operative stimulation mapping" OR "cortical stimulation" OR "intra-operative monitoring" OR "diagnostic imaging" OR "intra-operative neurophysiological monitoring")

AND (glioma OR tumor OR cancer OR astrocytoma OR ependymoma OR glial OR glioblastoma OR malignant OR meningioma OR medulloblastoma OR metastasis OR oligodendroglioma OR oncology OR chondrosarcoma OR craniopharyngioma OR cyst OR hemangioma OR hemangioblastoma OR hemangiopericytoma OR lipoma OR lymphoma OR neurofibromatosis OR oligoastrocytoma OR schwannoma)

AND (evaluation OR assessment OR test OR task OR "standardized test battery" OR analysis OR tool OR protocol OR screening OR measure OR "mapping diagnostic" OR monitoring OR examine OR investigate OR probe OR questionnaire OR standardized OR "non standardized" OR appraising OR "neuropsychological tests" OR “mental status test”)
AND (brain OR cerebral OR cortical OR subcortical OR “frontal lobe” OR “parietal lobe” OR “temporal lobe” OR occipital OR limbic OR thalamic OR cranial OR “central nervous system” OR “arcuate fasciculus” OR “longitudinal fasciculus” OR “posterior cingulate cortex” OR “uncinate fasciculus” OR “fronto-occipital fasciculus” OR “parietopontine” OR “cingulate cortex”)}
Appendix C. Web of Science Search Terms

TS=(language OR speech OR apraxia OR aphasia OR communicat* OR "auditory comprehension" OR listening OR reading OR writing OR conversation OR discourse OR naming OR "word retrieval" OR "sentence production" OR "sentence comprehension" OR grammar OR linguistics* OR semantics* OR spelling OR "action naming" OR "verbs and sentences" OR “spontaneous speech” OR eloquent OR dyslexia OR “reading comprehension” OR alexia)

AND (resection OR "awake craniotomy" OR "awake surgery" OR neuronavigation* OR "intra-operative imaging" OR neurosurgery OR craniotomy OR "direct electrocortical stimulation" OR " intra-operative magnetic resonance imaging" OR "direct electrical stimulation" OR "intra-operative stimulation mapping" OR "cortical stimulation" OR "intra-operative monitoring" OR "diagnostic imaging" OR "intra-operative neurophysiological monitoring")

AND (glioma OR tumor* OR cancer OR astrocytoma OR ependymoma OR glial OR glioblastoma OR malignant OR meningioma OR medulloblastoma OR metastasis OR oligodendroglioma OR oncology OR chondrosarcoma OR craniopharyngioma OR cyst OR hemangioma OR hemangioblastoma OR hemangiopericytoma OR lipoma OR lymphoma OR neurofibromatosis OR oligoastrocytoma OR schwannoma)

AND (evaluation OR assessment OR test OR task OR "standardized test battery" OR analysis OR tool OR protocol OR screening OR measure OR "mapping diagnostic" OR monitoring OR examine OR investigate* OR probe OR questionnaire OR standardized* OR "non-standardized" OR appraising* OR "neuropsychological tests" OR “mental status test”)


AND (brain OR cerebral OR cortical OR subcortical OR “frontal lobe” OR “parietal lobe” OR “temporal lobe” OR occipital OR limbic OR thalamic OR cranial OR “central nervous system” OR “arcuate fasciculus” OR “longitudinal fasciculus” OR “posterior cingulate cortex” OR “uncinate fasciculus” OR “fronto-occipital fasciculus” OR “parietopontine” OR “cingulate cortex”)
Appendix D. References of Included Studies


Appendix E: Manual for Rating Intraoperative Brain Tumor Survivor (BTS) Language Assessment Studies

**Inclusion Criteria:** To be included in the evidence table, publications must:

1. Be written in **English**
2. Be published in a peer-reviewed journal and contain original research. Reviews will be excluded.
3. Be an experimental, quasi-experimental, or cor relational pre/intra- and post-craniotomy study.
4. Include at least one adult (over 18 years) patient or participant with a diagnosis of brain tumor.
5. Identify the intra-operative neuroimaging technique used during awake craniotomy.
6. Include Intraoperative assessment of at least three of the following language domains: Naming, automatic speech, reading, writing, spontaneous speech, auditory comprehension, syntax/grammar, phonology, semantics, repetitions, timed fluency tasks, bilingual tasks, verbal memory, verbal reasoning/judgement. Studies that focus on pre-surgical assessment and/or incidentally mention speech/language assessment following the surgery without intraoperative evaluation and/or comment on its effectiveness but provide no evidence regarding what assessment took place should be excluded.
7. Assess intraoperative language using (at least) one standardized test and/or experimental task that is replicable.

*Year criterion (2014-2021) was only used for electronic search

See Appendix A (Study Inclusion Form) for more detailed information about inclusion/exclusion criteria.

**Instructions for entering information from treatment studies into the Table of Evidence**

**Introduction**

This manual is designed to make data extraction easier and to facilitate inter-coder reliability. First, data extraction roughly follows the order in which information is presented in a typical research article. Specifically, the coding manual and accompanying excel spreadsheet are divided into 13 broad sections:

I. General coding of publication and raters (Columns A-C)
II. Citation/Reference (Columns D-H)
III. Study Overview - Abstract (Columns I-L)
IV. Language/Communication Assessment Research Questions or Objectives - Introduction (Column M)
V. Participant Population— General—Methods/Results (Columns N-P)
VI. Participant Demographics: Brain Tumor Survivors only (Columns Q-AK)
VII. Pre-operative Language Assessment (Columns AL-AT)
VIII. Intraoperative Mapping and Resection: Technical - Methods (Columns AU-BA)
IX. Intraoperative Language Ax for BTS (Columns BB-BQ)
Second, since some study information may be presented in multiple locations in a publication (with variable levels of detail/consistency), guidelines are provided for each section/variable, indicating where to retrieve specific study information. For instance, at the start of section \textbf{IV. Research Questions or Objectives}, the guidelines specify that this information should be extracted (when possible) from the Introduction section of the article, or if not available in the Introduction, the information should be extracted from other sections of the publication that provide the most relevant information. This is summarized in the guidelines below by a ‘go to’ arrow followed by the preferred information source/article section (→INTRODUCTION).

In addition, codes are used to make data extraction easier, facilitate inter-coder reliability, and support future data analyses. As detailed below, letters (usually the first letter of a key word) are used to code categorical data, such as gender (e.g. M = male; F = female) or outcomes. Codes can also have suffixes that provide more detailed information about the initial code. For example, with respect to outcome codes, an improvement (I) in the target behavior that was statistically significant (SS), would be transcribed as: I-SS. For single subject designs with more than one participant, a second suffix is used to indicate the number of participants associated with each category (e.g. D-NS-1; I-SS-1; I-NS-1) (averages are used for group studies).

The following general codes are used for all variables:

\textbf{O-specify} Use capital letter ‘O’ to indicate ‘Other’ i.e. information relevant to a particular variable that is not captured by the available codes and/or if quantitative information is requested (e.g. age) but only descriptive information is provided (e.g. ‘older adults’). The “O” should be followed by a dash and explanation. Examples

i. With respect to aphasia classification, there is no code for isolation aphasia, so this would be entered as “\textit{O-isolation aphasia}”.

ii. With respect to years of education, if the only information provided was the inclusionary criteria that participants had a minimum of 9 years of education, you would enter as “\textit{O-inclusionary criteria: participants had at least 9 yrs of education}”.

\textbf{NA} Use NA to indicate that information is not relevant for a particular study (e.g., if you are evaluating a group study, you would enter ‘NA’ in column CA (which asks for a quality rating of single subject designs).

Use the period symbol to denote missing data (information that is not reported in the publication under review). Do NOT leave any cells empty. The ‘.’ code serves two functions. First, it clarifies for
you and others that the entry was actually coded (vs. overlooked/skipped accidentally). Second, it allows Excel to read this information as missing when computing any statistics (e.g., means, SD, etc.).

* If you have a question/comment about any information that is entered that you would like to discuss with other reviewers, put an asterisk in the column and then enter your question/comment in that column.

? If you have a question about something that is ambiguous in the publication AND you think that this could be clarified by contacting study authors, add the question mark symbol at the end of the entry, and then enter your question in column BW (Question for study authors).

When entering descriptive information, be as specific as possible with respect to terminology e.g., enter ‘recurrent medulloblastoma’ (vs. ‘tumor’).

Medical terminology: If you are uncertain about terminology in the publication, go to the following reference cite for clarification:
https://www.merriam-webster.com/medical

I. General coding of publication and raters (columns A-C)

A. **Study number** (assigned in advance by project lead)
B. **Rater 1** (assigned in advance by project lead)
C. **Rater 2** (assigned in advance by project lead)

**Citation/Reference (Columns D-H)**

D. **Citation** – enter full reference using APA style without indenting
E. **First author** – enter last name of first author
F. **Year** – enter year of publication
G. **Title of publication** – enter title of publication
H. **Journal** – enter name of journal

III. Study Overview (Columns I-L) → ABSTRACT

This section is meant to capture broad information about the study. In most cases it will correspond to information about the surgical procedure, but in some cases, the purpose, participants, results, and/or interpretation will be broader than the language assessment. For example, the purpose may be to assess the effectiveness of a general procedure to maximize extent of resection for brain tumor survivors (that includes information about the intraoperative speech/language assessment).

I. **Purpose/rationale** – provide a 1 sentence summary of the authors’ stated purpose/rationale for the study. For example: *To examine the effects of direct electrical stimulation on the left frontal aslant* (variable x) *on articulation* (target language/communication assessment/behavior y).

J. **Participant description** – enter the number of participants, age range, and the diagnoses (e.g., *5 adults with a mean age of 49 yrs and dx of aphasia*).
K. Results – Summarize (in 1-2 sentences) the primary outcome(s) as reported by study authors (e.g., Treatment x resulted in statistically significant improvement in extent of resection for awake craniotomy over general anesthesia).

L. Author’s Interpretation – Summarize (in 1-2 sentences) the authors’ conclusions (e.g., Treatment x may be a feasible assessment protocol for awake craniotomy)

IV. Language/Communication Research Questions or Objectives (Column M) → INTRODUCTION
This section covers specifics related to the language and/or speech assessment (as opposed to other medical/general rehabilitation).

M. Research Questions/Objectives specific to language/communication tx – Briefly list up to three most important research questions/objectives related to intraoperative imaging or intraoperative language assessment (1 sentence/question). Example: Was intraoperative assessment X effective in preserving language outcomes?
V. Participant Population (Columns N-P) →METHODS/RESULTS
This section covers specifics related to the patients that participated in awake craniotomy language mapping (as opposed to other medical/general treatments). If no information is provided, code as missing data (i.e., enter “.”)

Most of this information will be described in detail in the METHODS section. However, if there is a discrepancy between what was planned (METHODS section) and what was actually done (RESULTS section) record what was actually done (usually this will be in the RESULTS section).

N. Number of participants total - Enter the total number of treated participants in the study (regardless of participant group (e.g., awake craniotomy vs. general anesthesia).

O. Number of participants with dx other than brain tumor – Enter the number of treated participants who did NOT have a diagnosis of brain tumor. If all treated participants had a brain tumor enter the number “0”.

P. Diagnoses other than brain tumor – List the primary medical diagnoses of participants and (if available) the number of participants with each diagnosis (e.g., epilepsy-3). If all participants had a brain tumor enter ‘NA’.

VI. Participant Demographics: Brain Tumor Survivors only (Columns Q-AL) →METHODS/RESULTS
This section covers specifics related to participants who received awake craniotomy speech/language assessment (as opposed to other medical/general rehabilitation). If no information is provided, code as missing data (i.e., enter “.”).

Most of this information should be provided in the Methods section of the article. However, if there is a discrepancy between information in the Methods and Results section (e.g. a discrepancy between inclusionary/exclusionary criteria in the Methods and characteristics of the final sample as presented in the Results) enter the information that you think most accurately reflects the characteristics of the sample that

Q. Number of BT Survivors – Enter the number of participants with a history of brain tumor.

R. BTS/Total – Enter the proportion of BTS participants relative to other participants (e.g., 8/12).

S. Age – Enter the age (in years) of the participant at the start of the study. If a study involves multiple BTS participants, enter the average age of these participants.

T. Age Range – Enter the minimum and maximum ages (e.g., 18-76). If min-max is not available, enter the most closely related descriptive information (e.g., Inclusionary criteria: Participants were between 21-85 years of age).

U. Gender – Specify gender using the codes:
- M = male
- F = female

Specify the number of participants in each category (e.g., M-2; F-1).

V. Education – Enter the average years of education at the start of the study.

W. Education Range – For studies with more than one BTS participant, enter the minimum and maximum years of education (e.g., 12-23). If only descriptive data is provided about education (for individuals/groups), the descriptive information should be
entered in this column (e.g., *minimum high school education; all participants were literate native-speakers of English*).

**X. Employment Status** - Enter the employment status of participants at the start of the study using the codes:
- W = working
- S = student
- D = currently on disability leave
- R = Retired
- O = Other - specify

If available, provide more specific information about employment (e.g., W - teacher)

For studies with more than one BTS participant, specify the number of participants in each category (e.g., W-1; D-3).

**Y. Time Post dx (months)** – Enter the number of months from the time of initial brain tumor diagnosis to surgical resection. For studies with more than one BTS, enter the average TPO. If only descriptive information is available enter the descriptive information using the code “O-other” (e.g., O- ‘newly diagnosed tumor’).

**Z. Surgical Tx Timing in relation to other medical treatment**: Briefly describe the timing of the awake craniotomy in relation to more general tumor-related medical treatments (Examples: *Chemotherapy began 1 week following surgical resection.*).

**AA. Living Situation** – Summarize living situation using these codes:
- A=Lives alone
- F = Lives with family/friends
- M = Lives in a medical/assisted living facility and receives professional care
- O = Other - specify

For studies with more than one BTS participant, enter the number of participants in each category (e.g., A-3; F-5).

**AB. Hand** – Indicate handedness using these codes:
- R=Right-handed
- L= Left-handed
- A=Ambidextrous
- O = Other - specify

For group studies, enter the number of participants in each category (e.g., R-3; L-2). If numeric data is not available enter any available descriptive data (e.g., *R&L participants were included*).

**AC. Ethnicity/Language Spoken** – Use the following codes to indicate ethnicity:
- Aml=American Indian or Alaska Native
- A=Asian
- BAAm=Black or African American
- NH=Native Hawaiian or another Pacific Islander
- W=White

Enter the number of participants in each category (e.g., W-2; BAAm-1). If numeric data is not available enter any available descriptive data (e.g., *White, BAAm, and Asian participants; or Mixed ethnicity sample*).
If a language other than English is spoken by the participant(s), include any data on the language spoken by the participant(s) (e.g., O-3 Italian speaking, 3 French speaking or A-6-3 Mandarin, 3 Cantonese).

**AD. Location of Tumor: Laterality** – Use the following codes:

- R = right hemisphere
- L = left hemisphere
- B = bilateral

Enter the number of BTS participants in each category (e.g., L-2; B-2). If numeric data is not available enter any available descriptive data (e.g., *Tumors were located in L, R, and/or bilateral structures*).

**AE. Location of Tumor: Anatomic areas** – List the anatomic structures affected by the tumor. Note that the location of the tumor is sometimes indicated by the name of the tumor (e.g., ‘medullosblastoma’ indicates that the tumor includes tissue in the cerebellum. If you are unsure if the name provides information about tumor location, use the reference search engines listed at the top of pg. 3 of this document. Use the following codes to denote tumor location:

- AF = Arcuate Fasciculus
- BG = basal ganglia
- BS = brainstem
- Ce = cerebellar
- FL = frontal lobe
- HPC = hippocampus
- Hypo = hypothalamus
- I = insula
- L = limbic
- Me = medulla
- Mi = midbrain
- MLF = middle/medial longitudinal fasciculus
- OL = occipital lobe
- Po = pons
- PL = parietal lobe
- Th = thalamus
- TL = temporal lobe
- UF = uncinate fasciculus

O-specify = other (specify)

If more than one structure is identified (which will often be the case), list all structures (e.g., FL, I, PL, TL). If more specific information is provided (in addition to the basic anatomic structure) follow the code with a dash and include the area affected (e.g., *FL–Broca’s area, PL-post central & supramarginal gyrus*). For small n studies (≤ 5 BTS participants) list the regions for individual participants (e.g., *P1: FL-Brocas area; P2-PL-angular gyrus*). For larger group studies > 5 BTS participants) enter the tumor locations included in the study as a single list.

**AF. Tumor type** - Identify the type of tumor (e.g., astrocytoma). For studies with more than one BTS, enter the number of participants with each tumor type (e.g., astrocytoma-3; glioma-2). If numeric data is not available enter any available descriptive data (e.g.,
Participants had various types of tumors including: astrocytomas, gliomas, and meningiomas. See appendix B for a list of common tumor types.

**AG. Tumor grade** – Use the following (WHO/American Brain Tumor Society codes to enter the tumor grade as specified by the study author(s):

- I = Grade 1
- II = Grade 2
- III = Grade 3
- IV = Grade 4
- O = Other – specify (e.g., O-tumors varied with respect to complexity)

For studies with more than one BTS participant, enter the number of participants with each tumor grade (e.g., I-3; II-1). If numeric data is not available enter any available descriptive data (e.g., varied, including: Grades II-IV). If tumor grade is not explicitly stated, simply code as missing information (i.e., enter “.”)

**AH. Tumor Diagnostic Criteria** – Enter the criteria used to diagnose the brain tumor using the following codes:

- **PE** = Physical examination of brain and spinal cord function (reflexes, muscle strength, vision, coordination, balance, alertness) by a medical professional – specify professional (e.g., PE – Physician or PE – neurologist or PE – not specified). If multiple medical professionals were involved in the diagnosis, list all professionals.
- **I** = Imaging – specify test (e.g., I-MRI; I-CT, etc.). If multiple tests were used, list all tests.
- **B** = Biopsy – Specify type (e.g., B - Stereostatic or B- Open/craniotomy).

If more than one criterion was used, list all criteria that were used (e.g., PE-not specified, I-PET, B-Open).

**AI. Comorbidities: Speech/Swallowing** – List any comorbidities associated with speech/swallowing using the following codes:

- A = Apraxia
- Dysarth = Dysarthria
- Speech = general speech motor impairment (type not specified)
- Dysphag = swallowing impairment
- O = Other (specify)

If more than one comorbidity is identified, list all (e.g., A, Dysarth, Dysphag). If specified, indicate whether co-morbidity is associated with tumor (t) or unrelated (u) or not specified (NR) (e.g., A-t). For studies with more than one BTS, enter the number of participants with each co-morbidity (e.g., A-2, Dysarth-2). If numeric data is not available enter any available descriptive data (e.g., O-varied, including: Speech & Dysphag).

**AJ. Comorbidities: Other** - List other comorbidities using the codes:

- C = Cognitive-specify
- Psy = Psychiatric-specify
- S = Sensory-specify
- M = Motor-specify
- Med = medical-specify

For studies with more than one BTS, enter the number of participants with each co-morbidity (e.g., C-executive function-2; Psy-depression-2). If numeric data is not
available enter any available descriptive data (e.g., *O-varied, including: depression, headache*).

**AK. Preoperative Neuroimaging** - List any additional neuroimaging data that was provided to support treatment planning using the following codes:

- **CT** = cat scan
- **MRI/fMRI** = magnetic resonance imaging/functional magnetic resonance imaging
- **PET** = positron emission topography
- **EEG** = electrocorticography
- **F** = fluoroscopy
- **TMS/nTMS** = transcranial magnetic stimulation/navigated transcranial magnetic stimulation
- **Other**- specify

For studies with more than one BTS, enter the number of participants for which each neuroimaging technique was used (e.g., *CT-2, MRI-8*).

**VII. Pre-operative Language Assessment (Columns AL-AT)**

→ METHODS/RESULTS

**AL. Name of Pre-operative Language Assessment(s)** - List any pre-operative language assessments used (e.g., *BNT, DO80, etc.*). If specific data is not available list/enter any available descriptive data (e.g., *speech and language tasks such as naming or battery used to assess patient’s cognitive and communicative abilities*).

**AM. Timing of Pre-operative Language Assessments(s) in Relation to Awake Craniotomy** - Describe the timing of pre-operative language assessments, as given by the author (e.g., *All pre-operative language assessments were completed 1 month to 1 week before surgery*; or *Language baselines were collected the day prior to surgery*).

**AN. Standardization** – Indicate whether *at least one* of the pre-operative language assessments was standardized (i.e., replicable) using the following codes:

- **Y** = Yes
- **N** = No

If yes, also re-enter the abbreviated name of the assessment/task (e.g., *Y- WAB-2*) or the general description as given by author (e.g., *experimental dynamic assessment task*), followed by a subcode for intended population of the assessment. Use the following codes (e.g., *Y-WAB-2- G*):

- **BT** = Standardized specifically for brain tumor populations
- **G** = Test for more general populations (e.g., general aphasia test or general cognition test)

**AO. Pre-operative Training/Trials of Intraoperative Items** - Indicate whether the intraoperative language assessment items were trained or trialed (e.g., excluding incorrect items from test stimuli) prior to the awake craniotomy using the following codes:

- **Y** = Yes
- **N** = No
If Yes (Y), also indicate whether the pre-operative training/trials did (Y) or did not (N) impact the subsequent intraoperative language assessment protocol (e.g., Y-N, or Y-Y-five images were removed from the items used intraoperatively).

**AP. Language Domains Assessed Pre-Operatively** – List all the language domains assessed during the pre-operative language assessments using the following codes:

- **N** = Naming (confrontational or convergent naming, noun or verb)
- **AS** = Automatic speech (e.g., counting, days of the week)
- **R** = Reading
- **W** = Writing
- **SS** = Spontaneous Speech
- **AC** = Auditory Comprehension
- **SynG** = Syntax/Grammar
- **P** = Phonology
- **Sem** = Semantics (e.g., sorting, picture-word matching)
- **Rep** = Repetitions
- **F** = Timed fluency tasks (e.g., generative naming and/or speech efficiency)
- **BL** = Bilingual speakers (e.g., naming in secondary language/code switching)
- **Cal** = Calculation
- **VM** = Verbal memory
- **VR** = Verbal reasoning/judgement
- **O** = other -specify

**AQ. Description of Pre-operative Language Assessments/Tasks** – Include the authors description of the pre-operative language assessment tasks. (e.g., “The test battery included Swedish standard tests to detect aphasia, and instruments for detecting more subtle language disorders. The latter included tests of lexical retrieval (naming and word fluency) and a test (BeSS) of high-level language functions” (Antonsson et al., 2017)).

Depending on the level of description included by the authors, this may be the same as column AM.

**AR. Presentation Modality** - List all of the methods used to present pre-operative language assessment stimuli using the following codes:

- **C** = laptop computer/computer monitor
- **ED** = another electronic device (e.g., iPad)
- **Flash** = physical flash cards
- **B** = binder or standardized assessment stimuli booklet
- **VR** = virtual reality
- **O** = Other-specify

If further information is given, use a secondary coding (e.g., L- powerpoint).

**AS. Duration of Pre-operative Language Assessment(s)** – Enter the amount of time used for all pre-operative language assessments, in minutes. If numerical is not available or only partially calculable (from a standardized manual), enter the estimated data and any descriptive data (e.g., WAB-2 approximately 45, plus additional tasks).
**AT. Existing Language Impairment- Pre-operative**- Enter the level of severity (as specified by the study authors) during the pre-operative period based off the pre-operative language assessments or report the assessment score(s). Use the following codes:

- **N** = Normal Function
- **Mi** = Mildly Impaired
- **Mo** = Moderately Impaired
- **S** = Severely Impaired
- **O** = Other-specify (e.g., ‘O-borderline normal’)

If there is information on which language domains are impaired, indicate them with a subc coding (see AQ): (e.g., *Mi-N*).

If no terms are used to describe the severity level of the aphasia, but scores are provided for a standardized aphasia test (e.g., WAB-R), enter the name of the test followed by a dash and then the summary score (e.g.,

‘WAB-R AQ: 92’).

For studies with more than one BTS participant, enter the number of participants with each level of aphasia severity (e.g., ‘Mi-2, S-2’). If numeric data is not available, enter any available descriptive data (e.g., ‘O-varied, including Mi & Mo’ or ‘O-WAB-R AQ: 47-92’)

**VIII. Intraoperative Mapping and Resection: Medical/Technical (columns AU-BB)** → METHODS/RESULTS

This section covers specifics related to the technical and medical aspects of intraoperative mapping and resection (as opposed to aspects of the language assessment/mapping). If no information is provided about the procedure, code as missing data (i.e., enter “.”)

**AU. Inclusion/Exclusion Criteria of Awake Craniotomy Participant Selection**- Enter any inclusion/exclusion criteria considered when evaluating patient eligibility for awake craniotomy procedure, as the author(s) described them (e.g., “Patients with major neurological deficits that preclude testing of function (pronounced aphasia or motor), poor neuropsychological scores (Mini-Mental State Examination, Aphasia Quotient Score for language function), significant deficits in concentration and sustained attention, patients suffering with severe anxiety or depression, claustrophobia, disinhibition, apathy and disorganized behavior were not candidates for awake
craniotomy” (Chowdury et al. 2018)). If there is no information for patient eligibility, enter ‘.’.

**AV. Prior Surgeries** – Indicate whether any of the BTS participants previously underwent surgery using the following codes:

- **Y** = Yes
- **N** = No

If yes, specify the number of BTS participants that underwent surgical treatment prior to the report (e.g., Y-4).

**AW. Ratio Underwent Awake Craniotomy**- Enter the number of BTS participants that underwent awake craniotomy treatment in ratio to the total number of BTS participants in the study (e.g., 11/20).

**AX. Type of Intraoperative Imaging**- Indicate the mapping technique used with the following codes. List all intraoperative imaging techniques used in the study.

- DES = direct electrocortical stimulation/direct cortical stimulation, etc.
- iMRI = intraoperative MRI
- eCoG = electrocorticography
- F = fluoroscopy
- O = other (specify)

**AY. Imaging Specifications**- Give a brief description of the imaging specifications as described by the author. If more than one intraoperative imaging technique was used, enter the intraoperative imaging type code followed by the description (e.g., ‘DES-bipolar stimulator with tips spaced 5 mm apart, biphasic current with pulse frequency 60 Hz, stimulation 4 sec’).

**AZ. Identification of Language Sites**- Give a brief description of how language (positive and/or negative) sites were defined intraoperatively by study authors (e.g., “A site was considered eloquent (language positive) if disturbances were observed in more than two trials”).

**BA. Extent of Resection (EOR)**- Enter the extent of resection using the following codes, followed by the number of patients that received the level of resection (e.g., GT-22, ST-30, SpT-2):

- **GT** = Gross Total (complete removal of all tumors visible in MRI)
- **ST** = Subtotal (some tumorous tissue is left)
- **SpT** = Supratotal (complete removal of all tumorous tissue visible in MRI and some additional unaffected brain tissue beyond the tumor border)

If this data is unavailable, enter any description given by the author (e.g., mean % of resection) OR calculate the mean extent of resection if possible.

**IX. Intraoperative Language Ax for BTS (columns BB-BQ)** →

**METHODS/RESULTS**

This section covers specifics related to the *intraoperative language assessment* (as opposed to pre- or post-surgical assessments/language therapies). If no information is
provided about the language and/or communication assessment, code as missing data (i.e., enter “.”)
If the description or information on the assessment task comes from a secondary (citated) source, enter “.” in the column followed by a “*” or “**” and the description from the secondary source. Be sure to include the reference.

**BB. Presentation Modality**- List all of the methods used to present intraoperative assessment stimuli using the following codes:
- L= laptop computer
- ED= another electronic device (e.g., iPad)
- Flash= physical flash cards
- B= binder or standardized assessment stimuli booklet
- VR= virtual reality
- O= Other-
   - specify
   If further information is given, use a secondary coding (e.g., L- powerpoint).

**BC. Total Duration of Assessment (minutes)**- Enter the total time the patient was awake for intraoperative language mapping assessment, in minutes (e.g., 90). For studies with more than one BTS participants, enter the average duration of assessment (e.g., mean 100 or range 45-90).

**BD. Administrator of Intraoperative Ax** – Specify who administered the intraoperative assessment using the following codes:
- SLP = Speech-language pathologist/therapist
- NP = Neuropsychologist
- O = Other (Specify)

*Before completing this section (BE-BP), review IEF Items 7 and 8*

Summarize information for *at least 3 primary tasks/measures* (M). The following should be directly related to and further detail the tasks identified in inclusion/exclusion screening (see IEF, items 7 and 8)

**BE (M1), BJ (M2), & BM (M3). Intraoperative Language Assessment (M) (subtest or task): Name**– Enter the subtest/task name or a description of the task, as provided by the author (e.g., modified BNT or nonverbal semantic associations task).

**BF (M1), BJ(M2), & BN (M3). Task Standardization**- Enter whether the measure (task) follows a standardized (formal) procedure using the following codes:
- Y = Yes
- N= No

If yes, follow up with two subcodings. Use the following codes:
- First:
  - BT = Standardized specifically for brain tumor populations
  - G= Test for more general populations.
- Second:
  - IO = Standardized specifically for use during intraoperative context
  - Gen= Standardized for general or other (not intraoperative context)
**BG** (M1), **BK** (M2), & **BO** (M3) **Language Domain Assessed**- Enter the language domain(s) assessed by the intraoperative language assessment task using the following codes:

- N= Naming (confrontational or convergent naming, noun or verb)
- AS = Automatic speech (e.g., counting, days of the week)
- R = Reading
- W = Writing
- SS = Spontaneous Speech
- AC = Auditory Comprehension
- SynG = Syntax/Grammar
- P = Phonology
- Sem= Semantics (e.g., sorting, picture-word matching)
- Rep = Repetitions
- F= Timed fluency tasks (e.g., generative naming and/or speech efficiency)
- BL = Bilingual speakers (e.g., naming in secondary language/code switching)
- Cal = Calculation
- VM = Verbal memory
- VR = Verbal reasoning/judgement
- O = other -specify

**BH** (M1), **BL** (M2), & **BP** (M3) **Description of Task**- Give a brief description of the task as given by the author (e.g., “Patients are given a black-and-white drawing or a video of an action, and they are asked to answer with a verb in the infinitive (e.g., 'to jump') or in the progressive form ('jumping’)” (Bogka et al., 2003); or Patients counted backward from 20).

**BQ. Additional Tasks**- If more than three intraoperative language assessments were completed, enter the domain and corresponding assessment/task name here (e.g., SS-conversation between tasks or O-Line Bisection Task).

**X. Post-Operative Language Outcomes (columns BR-CF) → RESULTS**

This section covers specifics related to any language outcomes following surgical treatment. If no information is provided about the language and/or communication intervention, code as missing data (“.”).

Post-Operative Language Assessment (columns BR-BY)

**BR. Timing of Follow-up(s)** – Enter the number of follow-ups and timing in relation to awake craniotomy, in weeks (W) or months (M). For example, ‘3-1W, 3M, 6M.’

**BS. Name of Post-operative Language Assessment(s)** - List all of the post-operative language assessment(s) used (e.g., BNT, DO80, experimental dynamic assessment task, etc.) to assess language outcomes.

**BT. Standardization** – Enter the formal standardized test that was used to assess post-operative language outcome and diagnosis of potential aphasia. (e.g., WAB-R, BDAE), followed by a subcoding for intended population. Use the following codes (e.g., WAB-2-G)

- BT = Standardized specifically for brain tumor populations
G= Test for more general populations.

**BU. Language Domains Assessed Post-Operatively** – List all the language domains assessed during the pre-operative language assessments using the following codes:
- **N** = Naming (confrontational or convergent naming, noun or verb)
- **AS** = Automatic speech (e.g., counting, days of the week)
- **R** = Reading
- **W** = Writing
- **SS** = Spontaneous Speech
- **AC** = Auditory Comprehension
- **SynG** = Syntax/Grammar
- **P** = Phonology
- **Sem** = Semantics (e.g., sorting, picture-word matching)
- **Rep** = Repetitions
- **F** = Timed fluency tasks (e.g., generative naming and/or speech efficiency)
- **BL** = Bilingual speakers (e.g., naming in secondary language/code switching)
- **Cal** = Calculation
- **VM** = Verbal memory
- **VR** = Verbal reasoning/judgement
- **O** = Other -specify

**BV. Description of Post-operative Language Assessments/Tasks** – Give a brief description of the task as given by the author (e.g., a more complete version of the BNT than was administered intraoperatively was used to assess naming or Patients counted backward from 20).

**BW. Presentation Modality**- List all of the methods used to present pre-operative language assessment stimuli using the following codes:
- **C** = laptop computer/computer monitor
- **ED**= another electronic device (e.g., iPad)
- **Flash**= physical flash cards
- **B**= binder or standardized assessment stimuli booklet
- **VR**= virtual reality
- **O**=Other -specify

If further information is given, use a secondary coding (e.g., *L- powerpoint*).

**BX. Duration of Post-operative Language Assessment(s)** – Enter the amount of time used for all post-operative language assessments, *in minutes*. If numerical is not available or only partially calculable (from a standardized manual), enter the estimated data and any descriptive data (e.g., WAB-2 approximately 45, plus additional tasks).

**BY. Existing Language Impairment- Post-operative**- There are several parts to this question.

*First*, enter the level of severity (as specified by the study authors) during the post-operative period based off the last post-operative language assessments, or report the assessment score(s). Use the following codes:
- **N** = Normal Function
- **Mi** = Mildly Impaired
- **Mo** = Moderately Impaired
S = Severely Impaired
O = Other-specify (e.g., ‘O-borderline normal’)

If no terms are used to describe the severity level of the aphasia, but scores are provided for a standardized aphasia test (e.g., WAB-R), enter the name of the test followed by a dash and then the summary score (e.g., ‘WAB-R AQ-92’). For studies with more than one BTS participant, enter the number of participants with each level of aphasia severity (e.g., ‘Mi-2, S-2’). If numeric data is not available, enter any available descriptive data (e.g., ‘O-varied, including: Mi & Mo’ or ‘O-WAB-R AQ: 47-92’)

Second, indicate the trajectory of language outcomes using the following format: - participant ID/group-timing(#W/M)- Descriptive/statistically significant (D/SS)- Increase/Decrease (I/D)- domains (list). E.g., Mi-CRA-1W & 3M-SS-I-N.

BZ. Patient/Caregiver-Reported Language Outcomes – Report any language outcomes as reported by the patient or caregiver (e.g., ‘caregiver reported word finding difficulties’).

Post-Operative Language Therapy Treatment (columns CA-CG)

CA. Post-Operative Language Therapy Treatment Name- Enter the name of the Treatment (e.g., ‘Semantic Feature Analysis’) and/or the treatment approach (e.g., ‘Life Participation’). If no information is provided other than “language treatment,” enter ‘language treatment.’

CB. Duration of Language Therapy Treatment Sessions- Enter the length of treatment sessions in minutes (e.g., 50).

CC. Frequency of Language Treatment Sessions – Enter the frequency that treatment was administered using the following codes:
D = Daily-specify frequency (e.g., D-x1 to indicate sessions occurred once daily).
W= Weekly-specify frequency (e.g., W-x3 to indicate 3 sessions/week)
M= Monthly-specify frequency (e.g., M-x2 to indicate 2 sessions/month)
A = Annually-specify frequency (e.g., A-x1 to indicate 1 session annually)
L = Limited-specify frequency (e.g., L-2 sessions administered post-surgery)
O = Other-specify frequency (e.g., O-sessions administered at patient request)

CD. Number of Language Therapy Treatment Sessions – Indicate the total number of treatment sessions that were actually administered. For studies with more than one BTS participant, enter the average number of treatment sessions.
**CE. Duration of Treatment** – indicate in weeks the actual duration of treatment (e.g., 6).

**CF. Language Domains Addressed in Post-Operative Language Treatment** – List all the language domains addressed during the post-operative language treatments/therapy using the following codes:
- N = Naming (confrontational or convergent naming, noun or verb)
- AS = Automatic speech (e.g., counting, days of the week)
- R = Reading
- W = Writing
- SS = Spontaneous Speech
- AC = Auditory Comprehension
- SynG = Syntax/Grammar
- P = Phonology
- Sem = Semantics (e.g., sorting, picture-word matching)
- Rep = Repetitions
- F = Timed fluency tasks (e.g., generative naming and/or speech efficiency)
- BL = Bilingual speakers (e.g., naming in secondary language/code switching)
- Cal = Calculation
- VM = Verbal memory
- VR = Verbal reasoning/judgement
- O = other - specify

**XI. Quality of Evidence for Study (columns CG-CN)** – This section should be completed based on information provided throughout the publication. If no information is provided about the language and/or communication assessment, code as missing data (i.e. enter “.”)

**CG. Design Description (By Author)** – provide a brief description of the (language/communication) tx study design as described by the author (e.g., ‘Case report’ or ‘Clinical recommendations’ or ‘Group study’).

**CH. Design Type** – Enter your assessment of the design type based on the authors’ description and other related information provided in the abstract and results section of the document. Use the following codes (Tate et al., 2013; see also ANCDs criteria):
- **Uncontrolled non-experimental studies** (no control group, condition, or phase):
  - **Corr** = Correlational study comparing the relation between two variables
  - **Case** = Case series – an uncontrolled clinical description of treatment of one or more persons (data are not averaged across cases), includes B (data collected during tx phase only), AB (data collected before and during treatment); and Pre/Post (data collected before and after tx, but not during tx)
  - **UG** = Uncontrolled (no control group/condition) descriptive group (data are averaged across participants) study

- **Controlled experimental studies** (either a repeated control phase in single participant designs, or a control condition/group in group studies):
  - **SPD** = Single participant design (data are not averaged across individuals) – an experimentally controlled treatment in which i).the individual with
tumor(s) serves as his/her own control by prospective and intensive study during multiple (minimum of two) treatment phases (usually A&B); ii). There is a specific operationally defined behavior targeted for treatment that is measured *repeatedly* during all phases; includes classic withdrawal/reversal e.g. ABA designs and extensions (e.g., ABCB), multiple baseline design (MBD) across participants, MBD across behaviors, MBD across contexts, multiple probe design, alternating treatments design (ATD), changing criterion design, and other combinations.

**CGD** = Controlled (there is a control condition or group) design. Includes within and between group designs.

**RCT** = CGD with random assignment of participants to different experimental conditions/treatments

**MA** = Meta-Analysis providing a statistical summation of multiple controlled experimental studies

**O** = Other-specify

Three quality rating scales should be completed for all publications. First, for all publications, complete the ASHA Quality Indicators Scale (ASHA QI in column CI), developed for all study designs. Second, complete the quality rating scale that corresponds to the specific study design. Specifically, complete the RoBiNT scale (column CJ) for small n studies (case studies and/or single subject designs, results not averaged across individuals); OR the PEDro scale (column CK) for group studies (results averaged across individuals) OR the RQR scale (column CL) for qualitative studies. Lastly, complete the [NitkoMurrayRohde] scale (column CM) developed for this review. *All worksheets used to compute ratings are listed as separate tabs in the excel file.*

**CI. ASHA Quality Indicators (ALL studies)** – After reading the full publication, complete the ASHA QI worksheet (Cherney et al., 2008) as it pertains to the intraoperative language/communication assessment and enter the total ASHA QI score in this column.

**CJ. RoBiNT (Single subject/small n studies only)** – After reading the full publication, complete the RoBiNT (Tate et al., 2013) worksheet as it pertains to the intraoperative language/communication assessment and enter the total RoBiNT score in this column.

**CK. PEDro (Group studies only)** – After reading the full publication, complete the PEDro worksheet (Murray et al., 2013) as it pertains to the intraoperative language/communication assessment and enter the total PEDro score in this column.

**CL. RQR (Qualitative studies only)** - After reading the full publication, complete the RQR (Simmons-Mackie et al., 2016) worksheet and enter the total RQR score in this column.

**CM. NitkoMurrayRohde (All studies)** – After reading the full publication, complete the NitkoMurrayRohde worksheet as it pertains to intraoperative language/communication assessment and enter the final score in this column.

**CN. Level of Evidence (AAN)** – Classify the design type using the American Academy of Neurology (AAN) classification codes:

3 = Corr, Case, or UG designs
2 = SPD or CGD
1 = RCT or MA

XII. Study Conclusions (columns CO-CU) → DISCUSSION
If no information is provided about the language and/or communication assessment, code as missing data (i.e. enter “.”). Before completing this section, review columns I, K, & M which summarize study purpose & results (from abstract) and research questions (from introduction).

CO. Results: Main Study Findings- From the Results section, summarize the primary study findings corresponding to the research questions listed in columns K and/or M using the following coding format
(Note: Complete for only 1 finding/RQ for up to 3 RQs):
RQ#-specify analysis test- primary finding including # participants-test statistic-probability level
[e.g., RQ1-Barnard’s test- 3/5 participants had more naming errors during DES vs. no DES condition- (WALD r:4.062, 2.646, 3. 821; nuisance .06, .11, .14)-p <.01]

CP. Descriptive Summary of Main Findings – Based on information provided in the Discussion section:

a. Provide a single sentence that best summarizes the authors’ main conclusion for each research question (these should correspond to the research questions specified in column M). Specify i. the research question and ii. whether the conclusions apply to all study participants OR specifically to participants with intraoperative mapping for brain tumor, and iii. The primary study conclusion. Examples:
RQ1 Ax Validity: Results of this study suggest that the protocol is valid for intraoperative language mapping in individuals with parieto-temporal brain tumor.
RQ2 DV Effectiveness: Furthermore, results for Ax using virtual reality compared to traditional presentations indicated that virtual reality assessments are viable for expanding intraoperative assessment stimulus.

b. Additional information/commentary can be included (use grey font) following this initial summary sentence. Only the initial sentence will be counted towards reliability. Example: RQ2 Comparison of 2 Assessments: For automated speech tasks there was no SS difference between virtual reality or physical cards; Pts who were younger were less likely to have visual adjustment time in virtual reality presentation.

CQ. Conclusions: BTS vs other populations – Based on information provided in the Discussion section, provide a one sentence summary of any conclusions made by the authors that directly compare outcomes for participants with brain tumor with participants who have other etiologies (e.g., There was no difference in the performance of participants on the task regardless of whether surgical treatment was for brain tumor resection or epileptic seizures).
**CR. Author Recommendations:** List any clinical/research recommendations made by the authors with respect to intraoperative language assessment for BTS. Use the following codes:

- MR = More research needed–specify
- MDI = More demographic information needed about participants with BT–specify
- MLTI = More language treatment information needed–specify
- MMTI = More medical treatment information needed–specify
- O = Other–specify

**CS. Interesting Discussion Items:** There will be no reliability on this column. Enter anything from the discussion that you found interesting or worth noting. This could include comments on the discussion, any discrepancies, interesting findings that were not captured by the RQs, etc. (e.g., differences in activation/DES response between cortical/subcortical, or other neuroanatomical areas and/or participant populations).

**CT. Reviewer Comments/Questions** – Enter any general comments/questions that you have about the publication/coding here. For questions that are specific to information in a particular column, specify the column.

**CU. Questions for Study Authors** – Specify any questions for clarification that you would like to ask the study authors.

---

**XII. Ancestral Search (column CV)** → REFERENCE LIST

**CV. # References** – After reading the study references and completing the References worksheet, enter the total number of references that you think have a strong probability of addressing intraoperative speech/language assessment for adult brain tumor survivors. List the references of interest in the “Ancestral Search” tab on the data extraction spreadsheet.