Assessment of neuropsychological function during early treatment of diffuse glioma

Mallory Owen[1*], Rafael Romero-Garcia[2*], Alexa McDonald[3], Emma Woodberry[3], Moataz Assem[4], Pedro Coelho[5], Rob C Morris[6], Stephen J Price[6], Tom Santarius[6,7], John Suckling[2,8,9], Yaara Erez[4], Michael G Hart[6]

1. University of Cambridge School of Clinical Medicine, 2. Department of Psychiatry, University of Cambridge, 3. Department of Neuropsychology, Cambridge University Hospitals NHS Foundation Trust, 4. MRC Cognition and Brain Sciences Unit, University of Cambridge, 5. Neurophys Limited, 6. Department of Neurosurgery, Addenbrooke's hospital, Cambridge, 7. Physiology, Development and Neuroscience, University of Cambridge, 8. Behavioural and Clinical Neuroscience Institute, University of Cambridge; 9. Cambridge and Peterborough NHS Foundation Trust.

Correspondence to:
Rafael Romero-Garcia, rr480@cam.ac.uk
Herchel Smith Building for Brain and Mind Sciences
Robinson Way
Cambridge CB2 0SZ
Tel 1223 336581

Funding

RRG is funded by a Guarantors of Brain Post-Doctoral Fellowship award.
YE is funded by a Royal Society Dorothy Hodgkin Research Fellowship.
MA is funded by the Cambridge Trust – Yousef Jameel Scholarship.
SJP is supported by the National Institute for Health Career Development Fellowship (CDF-2018-11-ST2-003). This report is independent research supported by the National Institute of Research (NIHR Career Development Fellowship, Mr Stephen Price, CDF-2018-11-ST2-003). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care.
MGH received an award from The Brain Tumour Charity (ref: RG86218) to fund this work.

Conflict of Interest – There are no known conflicts of interest for any author of this manuscript

Authorship

Study design: MGH, RRG, TS, JS, YE, MO, MA, PC, SJP

Data Acquisition: AM, KC, PC, YE, MA, EW, MGH

Data Analysis: MO, RRG, MGH

Manuscript Writing: MO, MGH, RRG

*****These authors contributed equally

Background

Cognitive function in patients with diffuse glioma is frequently impaired and can have a profound impact on quality of life. Accurate, reproducible and accessible tools to assess cognition are mandatory to understand the effects of the tumour and treatment. Our hypothesis was that an app-based assessment would be complementary to traditional neuropsychological testing, thereby aiding in defining cognitive profiles and trajectories during early treatment of diffuse glioma.

Methods

Seventeen subjects with diffuse low-grade gliomas completed a traditional neuropsychological assessment battery before and after surgery. In addition an app-based tablet assessment (OCS-Bridge) was administered pre- and post-operatively as well as longitudinally at 3- and 12-month follow-up. Deficit rates, mean performance, and changes over time were compared using standardized z-scores between the two testing methods. Unsupervised k-means clustering was performed on individual cognitive tests in each battery.

Results

Preoperative testing showed an average of 2.88 deficits and 1.18 deficits per patient on neuropsychological testing and the tablet-based OCS-Bridge assessment, respectively. Digit span testing demonstrated agreement between testing modalities, but otherwise there was no significant correlation (Pearson's correlation: $p=0.7723$, r-value = 0.0758, df = 16). Longitudinal assessment revealed dynamic changes in attention and nonverbal skills. Traditional assessment was more sensitive to memory deficits, showing 22 preoperative deficits within the cohort vs. 1 for the app-based assessment, while app-based assessment was more sensitive to nonverbal skills, showing 8 deficits preoperatively vs. none in the traditional assessment. Clustering analysis did not create clusters along the predetermined domains, indicating that certain individual tests may test more than one cognitive function.

Conclusions

These data suggest app-based assessment is reliable and complementary to data obtained from traditional neuropsychological testing. Advantages include efficiency, facilitation of longitudinal testing, and increased sensitivity in domains of non-verbal skills and attention. Patients with diffuse glioma show subtle neuropsychological impairments, unique cognitive profiles, and discrete trajectories during early treatment, therefore judicious assessment is imperative.

Keywords: glioma, neuropsychology, neuropsychiatry, neurosurgery, cognitive function

Key Points:

- Patients with glioma have subtle cognitive changes requiring judicious testing
- App-based assessment is effective in understanding neuropsychological function
- Distinct cognitive profiles and trajectories are demonstrated during treatment

Importance of the Study

Understanding cognitive function over time is an essential marker of quality of life and treatment outcomes in patients with diffuse glioma. Because diffuse gliomas are associated with few impairments, careful choice of individual neurocognitive testing is paramount to ensure sufficient sensitivity. Additionally, individuals with similar presentations or seemingly innocuous tumour locations can have profound differences in cognitive outcomes (Romero-Garcia et al., 2020).

Studies have demonstrated that traditional neuropsychological testing can be resource-intensive and unsuitable for administration at multiple time points. New app-based testing modalities aim to address these issues, and theoretically offer advantages in terms of novel interactive paradigms, ease of administration, and efficiency of repeated testing in longitudinal studies. However, they have not been extensively validated for use in this population, nor compared with traditional neuropsychology testing.

This study compares two neurocognitive testing modalities in a single population with diffuse glioma. Our study demonstrated that patients with diffuse glioma had subtle deficits, and the two modalities were complementary in their sensitivity to deficits in different domains of cognitive function. Traditional neuropsychology testing was more sensitive to memory deficits, while app-based assessment was more sensitive nonverbal skills and attention deficits. In addition, the app-based assessment allowed identification of longitudinal cognitive trajectories, while combined analysis identified discrete cognitive profiles specific to patients with diffuse glioma. Overall these data demonstrate the importance of cognitive testing in this population and highlight the complementary role of combined testing.

INTRODUCTION

Patients with diffuse glioma can develop impairments in multiple cognitive domains either before or after surgery that may correspond with variables such as tumour morphology, extent of resection, and post-surgical treatment (Duffau, 2005; Lang et al., 2017; Taphoorn and Klein, 2004). These impairments can have a profound impact on quality of life and recovery trajectory (Ek et al., 2017; Reijneveld et al., 2016), and preoperative neurocognitive functioning may predict future functional status (Noll et al., 2018) and prognosis (Klein et al., 2003). Adjuvant chemotherapy and radiotherapy may also independently be associated with neurocognitive impairment, adding to the burden from the primary lesion and surgical resection (Wefel et al., 2011; Lawrie et al., 2019; Hilverda et al., 2010). Neurocognitive functioning may be affected by individual characteristics such as age (Kaleita et al., 2004), as well as psychological distress (Anderson et al., 1999). Additionally, tumor recurrence is correlated with decreasing neurocognitive function (Wefel et al., 2011; Meyers and Hess, 2003). Therefore, understanding cognitive function over time is an essential marker of quality of life and treatment outcomes in patients with diffuse glioma. It is vital to develop accurate, easy-to-use, convenient tools which can establish a detailed, comprehensive understanding of neurocognitive functioning at several points along the treatment pathway.

Despite recognition of the importance of neurocognitive function both for quality of life and clinical outcome in patients with diffuse glioma, several issues remain with current neurocognitive testing. First, they are time and resource-intensive. Traditional testing may lead to fatigue in patients with brain tumors (Taphoorn and Klein, 2004). They also require administration by a trained neuropsychologist in a hospital or clinica setting, which limits repeated use during long-term follow-up of patients with diffuse glioma.

Second, no consensus has been reached on the optimal battery of tests, timing, or setting of cognitive testing in patients with diffuse glioma. Although recent strides have been made towards defining a standardized cognitive assessment tool (Rofes et al., 2017), a diverse range of different assessments are still used clinically to assess cognitive function (Ng et al., 2019; Satoer et al., 2016). Patients with diffuse glioma show relatively few cognitive impairments overall, making many traditional testing batteries developed for other neurological conditions, such as the Mini Mental State Examination, unsuitable in this population (Brown et al., 2006; Taphoorn and Klein, 2004; Wefel et al., 2016; Wang et al., 2018).

Third, it remains difficult to predict cognitive outcomes of surgical or medical intervention in glioma patients (Taphoorn and Klein, 2004) and the impact of specific anatomical tumor location on quality of life and cognitive impairment remains poorly understood (Sagberg et al., 2019). While recent studies have demonstrated a relationship between tumor location and specific functional domains (Noll et al., 2016), it has also been shown that diffuse gliomas can have long-ranging effects on networks throughout the brain (Romero-Garcia et al., 2019; Hart et al., 2019; Harris et al., 2014; van Dellen et al., 2012), making accurate prediction of function based on anatomy difficult.

Several new testing modalities have become available in the past decade that have the potential to enhance the accuracy and ease-of-use of cognitive testing for both research and patient care purposes. App-based assessments theoretically offer advantages in terms of novel interactive paradigms, ease of administration and efficiency of repeated testing in longitudinal studies (Savulich et al., 2017; Freedman et al. 2018). Examples of these tools include computerized batteries like the Cambridge Neuropsychological Test Automated Battery (CANTAB, https://www.cambridgecognition.com/cantab/), and the OCS-Bridge assessment (https://ocs-bridge.com/) used in the current study. Although some of these tests have shown only moderate correlation with results of traditional neuropsychological testing (Smith et al., 2013), they are increasingly used in both research contexts and in routine patient care. Automated batteries allow for more rapid testing that is less resource-intense, as they do not require administration by a trained neuropsychologist. They also make it easier to assess certain aspects of cognitive function that are difficult to assess with traditional methods, such as reaction time. Finally, automated computerized or tablet-based tests may in the future present the potential for use independently by the patient, which would provide more granular data during surgical follow-up.

Despite the potential benefits of app-based cognitive testing to complement traditional neuropsychology testing, these tools have not been validated in this population. This study compares two cognitive testing modalities in a single prospective cohort of 17 patients with diffuse glioma who underwent awake surgical resection. The first approach was traditional neuropsychology testing which was performed both before and after surgery as per routine clinical practice. The second approach was a novel, app-based cognitive assessment, which was performed pre- and postoperatively and additionally at 3 and 12-month follow-up. Our hypothesis was that app-based assessment would be consistent with and complementary to traditional neuropsychological testing, thereby aiding in defining cognitive profiles and trajectories during early treatment of diffuse glioma. Specifically, we hypothesized that the latter offers

advantages in terms of novel interactive paradigms, ease of administration, and efficiency of repeated testing in longitudinal studies.

MATERIALS AND METHODS

Subjects

This study was approved by the Cambridge Central Research Ethics Committee (protocol number 16/EE/0151), and all participants provided written informed consent. Participants were selected on the basis of having pre-operative MRI scans consistent with a non-enhancing diffuse glioma and who were planned to undergo awake brain surgery. Full inclusion and exclusion criteria as well as imaging protocols have been previously described (Romero-Garcia et al., 2020).

Cognitive Assessment

Cognitive assessment was carried out using two distinct methods at various time points before and after surgery. The OCS-Bridge Assessment (https://ocs-bridge.com/) is a novel app-based screening tool administered via tablet that is designed to be used for the assessment of cognitive function in a wide range of neurological conditions. It consists of three parts: 1) the Oxford Cognitive Screen (OCS), 2) the Cambridge Attention, Memory and Perception Screen, and 3) the PHQ-9 and GAD-7 for the assessment of mood disturbance. The OCS assesses function across five cognitive domains of attention, perception, memory, language, praxis and numeracy skills, and has mainly been used previously as a cognitive screening tool in stroke (Mancuso et al., 2018; Demeyere et al., 2015). The Cambridge Attention, Memory and Perception Screen was specifically designed to extend the Oxford Cognitive Screen, and assesses perception, attention, verbal working memory and prospective memory via a variety of previously validated cognitive screening tests. For the purposes of this study, the PHQ-9 and GAD-7 portions of the OCS-Bridge were not performed, although mood disturbance was assessed clinically during the neuropsychological assessments. Assessment was performed pre-operatively, post-operatively prior to discharge, as well as at 3-month and 12-month follow-up, and took between 20-35 minutes to complete.

This was compared with traditional neuropsychological assessment that was performed before and after surgery at an average of 2 weeks postoperatively as per routine clinical practice. The battery comprised 26 independent measures of cognitive function across eight domains: verbal memory, nonverbal memory, verbal skills, nonverbal skills, attention, executive function, and mood disturbance

using a variety of previously validated tests (Meyers et al., 2013; Quental et al., 2013; Vlaar et al., 2003) The battery included elements of both the Weschler Adult Intelligence Scale IV, which has previously been used to assess cognitive functioning in glioma patients (Coalson et al., 2010), and the Brain Injury Rehabilitation Trust Memory and Information Processing Battery, which was specifically designed for patients with neurological injuries (Crawford and Garthwaite, 2007). Testing took approximately one hour to complete and was administered by a registered neuropsychologist in a clinical setting. A full list of the tests included on both the OCS-Bridge and the neuropsychological assessments can be found in **Supplementary Table 1**. A deficit in either modality was defined as performance two standard deviations below the mean of a reference control population on any particular test or any measurement taken from a test. The total deficits in each domain were defined as the number of individual tests within that domain on which a participant demonstrated a deficit.

Individual domains from the neuropsychological battery and the OCS-Bridge assessment were combined into four generalized functional domains for the purposes of direct comparison: attention, memory, verbal skills, and nonverbal skills. Rationalization is shown in **Table 1**. The neuropsychological battery also included two tests of executive function that provided four independent measures, as well as two assessments of mood, the Beck Anxiety Inventory and the Beck Depression Inventory. These were included in the analysis as their own domains that only included tests from the neuropsychological battery.

Statistical Analysis

All statistical analysis was carried out in R (R Core Team, 2014), and figures were produced using the package ggplot2 (Wickham, 2009). Z-scores were computed for preoperative data only for each individual task for both the neuropsychological battery and the OCS-Bridge assessment by subtracting the mean and dividing by the standard deviation of each participant's score for each task. Z-scores were multiplied by -1 for any task that used a scale in which functional improvement was characterized by a lower score in order to maintain comparability. For each combined domain, an average Z-score for each participant was computed by taking the average of the scaled scores for each task. The number of tasks that contributed to each domain differed and full details are available in **Supplementary Table 1.**

Unsupervised clustering analysis was carried out using the K-means algorithm. Individual tests that did not show any variation in score between subjects were omitted from the clustering analysis. Out of the original 39 OCS-Bridge tests, 25 were included in the clustering analysis and 14 were excluded due

to lack of variability. Out of 26 of the original Neuropsychological tests, 25 were included, as well as two neuropsychiatric tests and two additional measures of digit span. Clustering was performed on values that had been centered to a mean value of 0 with a standard deviation of 1. K-means clustering was done with a k value of 4 based on the within groups sum of squares. Analysis was perfomed using the kmeans function in R version 3.6.1. Clusters were visualized by performing a principal components analysis (PCA) using the fviz function in the factoextra package and plotting the individual cognitive tests against the first and second principal components.

Combined Domain	**Neuropsychology**	**OCS-Bridge**
Attention	Attention	Attention
Memory	Nonverbal Memory Verbal Memory	Memory Prospective Memory Verbal Working Memory
Verbal Skills	Verbal Skills	Language
Nonverbal Skills	Nonverbal Skills	Praxis Number Perception
Executive Function	Executive Function	-
Mood	Mood	-

Table 1. Combined domains across neuropsychological testing and OCS-Bridge testing

RESULTS

Demographics and Data Completeness

Seventeen subjects between the ages of 22 and 56 were recruited. Demographic data is presented in **Table 2**. No significant difference was found in total deficits reported preoperatively between right and left handed participants (t=0.547, p=0.6183, df = 3.34, 95% CI -3.73 – 5.406), gender (t=01.175, p=0.258, df = 14.77, 95% CI -4.106-1.189), or between right- and left-sided tumours (t=0.845, p=0.411, df = 14.41, 95% CI -1.53-3.53). There was no significant correlation between number of deficits and age of participant (Pearson correlation: p=0.39, df = 15, r-value = -0.22, 95% CI -0.65 – 0.29). All participants completed pre-operative traditional neuropsychology and OCS-Bridge assessments. Fourteen participants completed traditional neuropsychology after surgery, and eight completed postoperative OCS-Bridge assessments. Eleven participants completed an OCS-Bridge assessment at 3-month follow-up and at 12-month follow-up. Sixteen of seventeen participants had at least one OCS-Bridge assessment after surgery.

Patient	Age	Gender	Handedness	Presentation	Hemisphere	Location	Tumour Grade / Pathology
1	40	Female	Left	Seizures	Left	Frontal	Grade II Oligodendroglioma
2	30	Male	Right	Seizures	Right	Insula	Grade II Astrocytoma
3	30	Male	Right	Seizures	Left	Temporal / Insula	Grade IV Glioblastoma
4	50	Female	Right	Incidental	Right	Insula	Grade II Oligodengroglioma
5	60	Female	Right	Recurrence	Left	Frontal / SFG / frontal pole	Grade II Oligodendroglioma
6	20	Female	Left	Seizures	Right	Frontal / IFG	Grade I Ganglioglioma
7	30	Male	Right	Seizures	Right	Frontal / SFG & MFG	Grade III Astrocytoma
8	30	Male	Right	Seizures	Right	Frontal / MFG	Grade III Astrocytoma
9	50	Male	Left	Seizures	Left	Temporal / ITG	Grade IV Glioblastoma
10	40	Female	Right	Seizures	Right	Frontal / MFG	Grade II Oligodendroglioma
11	30	Male	Right	Seizures	Left	Frontal / SFG / frontal pole	Grade II Astrocytoma
12	30	Female	Right	Headaches	Left	Temporal / MTG	Grade III Astrocytoma
13	30	Female	Right	Seizures	Left	Superior Temporal Gyrus	Grade I Ganglioglioma
14	60	Female	Right	Seizures	Left	Superior Temporal Gyrus	Grade II Astrocytoma
15	30	Male	Right	Seizures	Left	Superior Temporal Gyrus	Grade III Astrocytoma
16	30	Male	Right	Seizures	Left	SFG/SMA & Pre-central	Grade IV Glioblastoma
17	30	Male	Right	Seizures	Left	Inferior frontal	Grade III Astrocytoma

SFG, Superior Frontal Gyrus; MFG, Middle Frontal Gyrus; IFG, Inferior Frontal Gyrus; ITG, Inferior Temporal Gyrus; MTG, Middle Temporal Gyrus; SMA, Supplementary Motor Area. Ages have been rounded to off for anonymity.

Table 2. Demographic Information

Neuropsychiatric Function

There was no significant association between the change in either the Beck Anxiety Index (BAI) (t=1.2789, p=0.22, 95% CI -2.4 – 9.6) or Beck Depression Index (BDI) score (t=0.87, p=0.4, 95% CI -3.8-8.95) and total deficits pre- and postoperatively. Of the 17 participants who completed preoperative neuropsychological testing, 17.6% (3/17) scored in the range of mild anxiety on the BAI, and 11.7% (2/17) scored in the moderate anxiety range. The rest scored within the normal range. Of the 14 participants who completed postoperative testing, 11.7% (2/17) tested in the mild anxiety range and 11.7% (2/17) tested within the moderate anxiety range. Three participants showed a significant decrease

in their BAI score while two showed a significant increase in BAI before and after surgery. Before surgery, 17.6% (3/17) of participants scored in the mild depression range on the Beck Depression Index (BDI), while 1 scored in the moderate range and 1 scored in the severe depression range. After surgery, 2 scored within the moderate range, 1 within the severe and 1 within the mild range. 17.6% (3/17) of participants showed significant improvement in their BDI score while 5.8% (1/17) showed a decline. All three participants lost to follow-up on neuropsychological testing were within the normal range on their BAI and BDI on preoperative testing.

Formal Neuropsychological Testing

Out of the 26 measures on the neuropsychological battery, 10 tests showed no variation within the cohort, with no patient showing a deficit on these tests. The other 16 tests demonstrated an impairment at least once across the cohort.

Before surgery, 79% (14/17) of participants had a deficit in at least one domain, with a mean of 2.88 (SD = 2.47) deficits per participant. Three participants did not show any deficits on testing before surgery. Deficits occurred in the domains of attention (7), verbal memory (7), verbal skills (6), nonverbal memory (4), and executive function (3). Given the presentations were either with seizures or as an incidental finding, it is likely that these cognitive profiles were under-appreciated by both clinicians and patients.

After surgery, all but one participant tested had a deficit in at least one domain, with a mean of 4.50 deficits per participant (SD = 3.40). Compared with testing before surgery, total cognitive deficits reduced in 43% (6/14) participants, remained unchanged in 7% (1/14), and increased in 50% (7/14), with an average increase of 4.14 deficits in those 7 participants (SD = 3.18). Changes in deficits for each participant are shown in **Figure 1.** Across participants who had testing both before and after surgery, three domains (verbal memory, nonverbal memory, and verbal skills) demonstrated an increase in deficits after surgery, two domains (nonverbal skills and attention) remained unchanged, and one domain (executive function) improved. No participants demonstrated a deficit in nonverbal skills at any time point.

OCS-Bridge Testing

Of the 39 individual cognitive tests included in the OCS-Bridge battery, 33% (13/39) did not show any deficits before or after surgery, or on follow-up testing. Before surgery, 59%% (10/17) of participants had a deficit in at least one domain, with a mean of 0.94 deficits (SD = 1.08) deficits per

participant. Deficits before surgery occurred in the domains of number (4), perception (3), attention (3), language (1), praxis (1), and verbal working memory (1). No participants showed a deficit in memory or prospective memory before surgery.

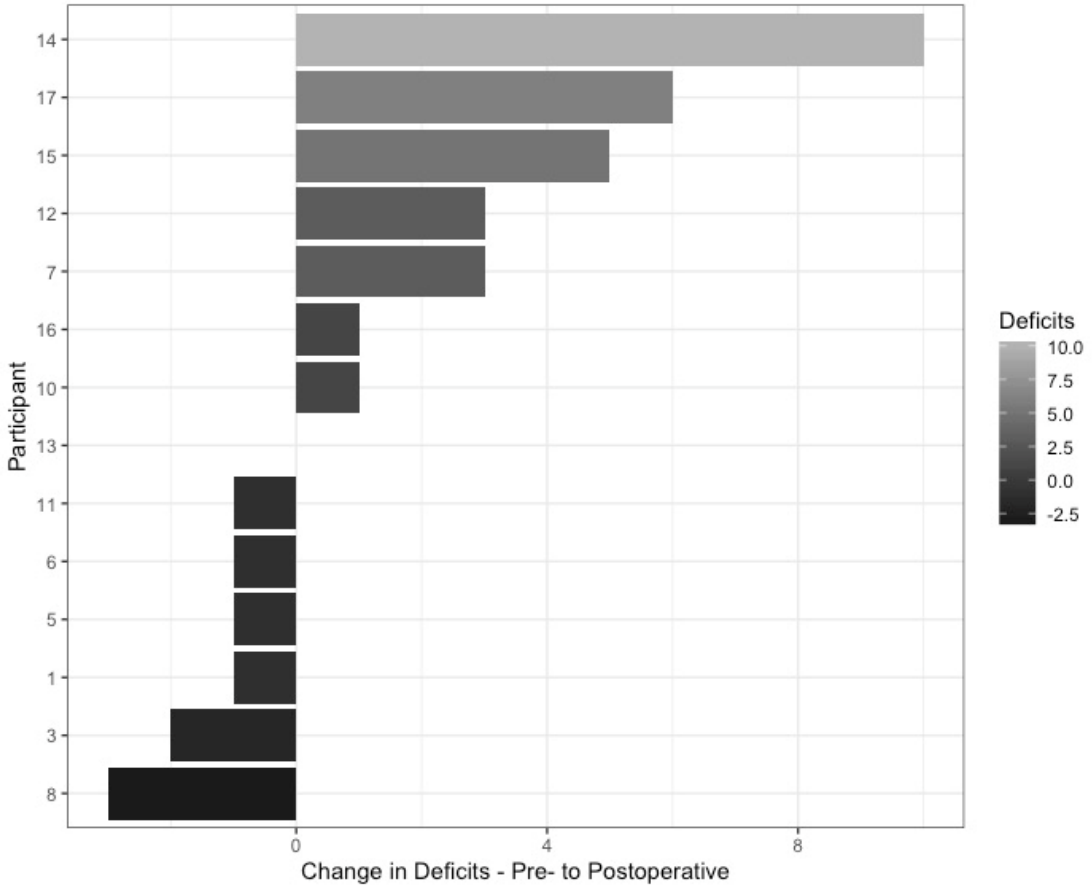

Figure 1. Change in total number of deficits for each patient between pre- and postoperative traditional neuropsychology testing.

OCS-Bridge Longitudinal changes

An advantage of OCS-Bridge was that it facilitated multiple longitudinal assessments in order to ascertain cognitive trajectories. The number of deficits over time as assessed by OCS-Bridge is shown in **Figure 2**. The greatest number of deficits were seen in attention and nonverbal skills. Of the 16 participants who had at least one postoperative assessment, 44% (7/16) showed an improvement in the number of deficits by their last assessment, 25% (4/16) remained stable, and 31% (5/16) showed an increase in the number of deficits by their last assessment. Four participants who had multiple follow-ups showed a pattern of increased deficits on either postoperative or 3-month testing which resolved by their last follow-up date. Three of these four participants showed a transient increase of at least one deficit in

perception, and two showed a transient increase in memory deficits. Ten of the sixteen participants who had at least one follow-up OCS-Bridge assessment ended the study with no deficits on formal testing.

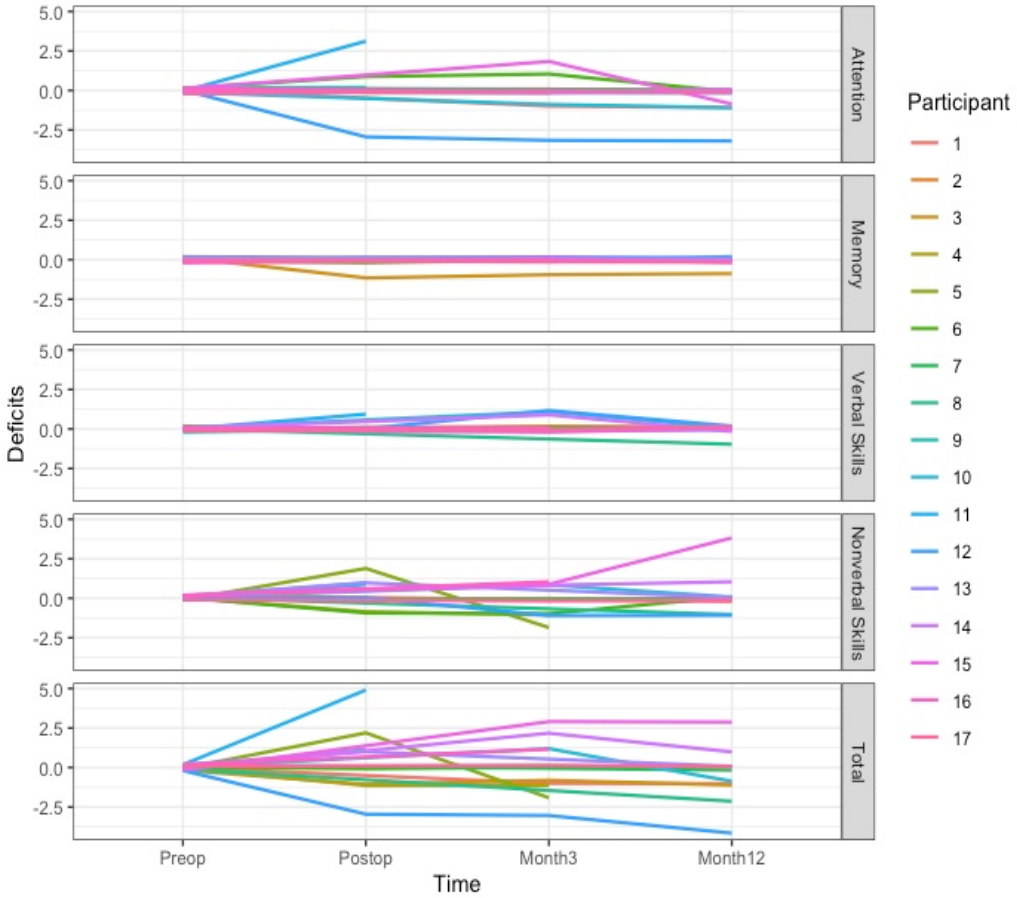

Figure 2. Number of cognitive deficits over time based on OCS-Bridge assessment preoperatively, postoperatively and at 3 and 12 month follow-ups.

Comparison of OCS-Bridge with Traditional Neuropsychological Testing

The total number of deficits detected in different domains varied between the neuropsychological battery and the OCS-Bridge assessment, and overall the neuropsychology battery detected more deficits than OCS-Bridge (**Figure 3A**). Preliminary comparison of the OCS-Bridge app was carried out using two tests that were duplicated in the traditional neuropsychological testing: forward and backward digit span (**Figure 3B)**. No significant difference was found between forward digit span from the traditional testing (mean = 6.29, SD = 1.40) and the OCS-Bridge test (mean = 6.76, SD = 0.90) on paired t-test (t = -1.367, p = 0.1905). Additionally, no significant difference was found between backward digit span on traditional testing (mean = 4.47, SD = 1.28) and on OCS-Bridge testing (mean = 4.82, SD = 1.55) on paired t-test (t=

-0.972, p = 0.3456). One subject had a difference of three or more points on forward digit span testing, and two subjects showed a difference of three or more on backward digit span testing**.**

Overall, the traditional neuropsychological testing battery detected more deficits in cognitive function than OCS-Bridge. The neuropsychological tests detected 44 preoperative deficits among the 17 subjects in the four combined domains of Attention, Memory, Verbal Skills and Nonverbal Skills, while OCS-Bridge detected 13 overall deficits and 28 possible deficits preoperatively (**Figure 3C**). The largest differences in detection were in the domains of memory (Neuropsychological = 22 deficits, OCS-Bridge = 1 deficit and 6 possible deficits), and Nonverbal Skills (Neuropsychological = 0 deficits, OCS-Bridge = 5 deficits and 12 possible deficits).

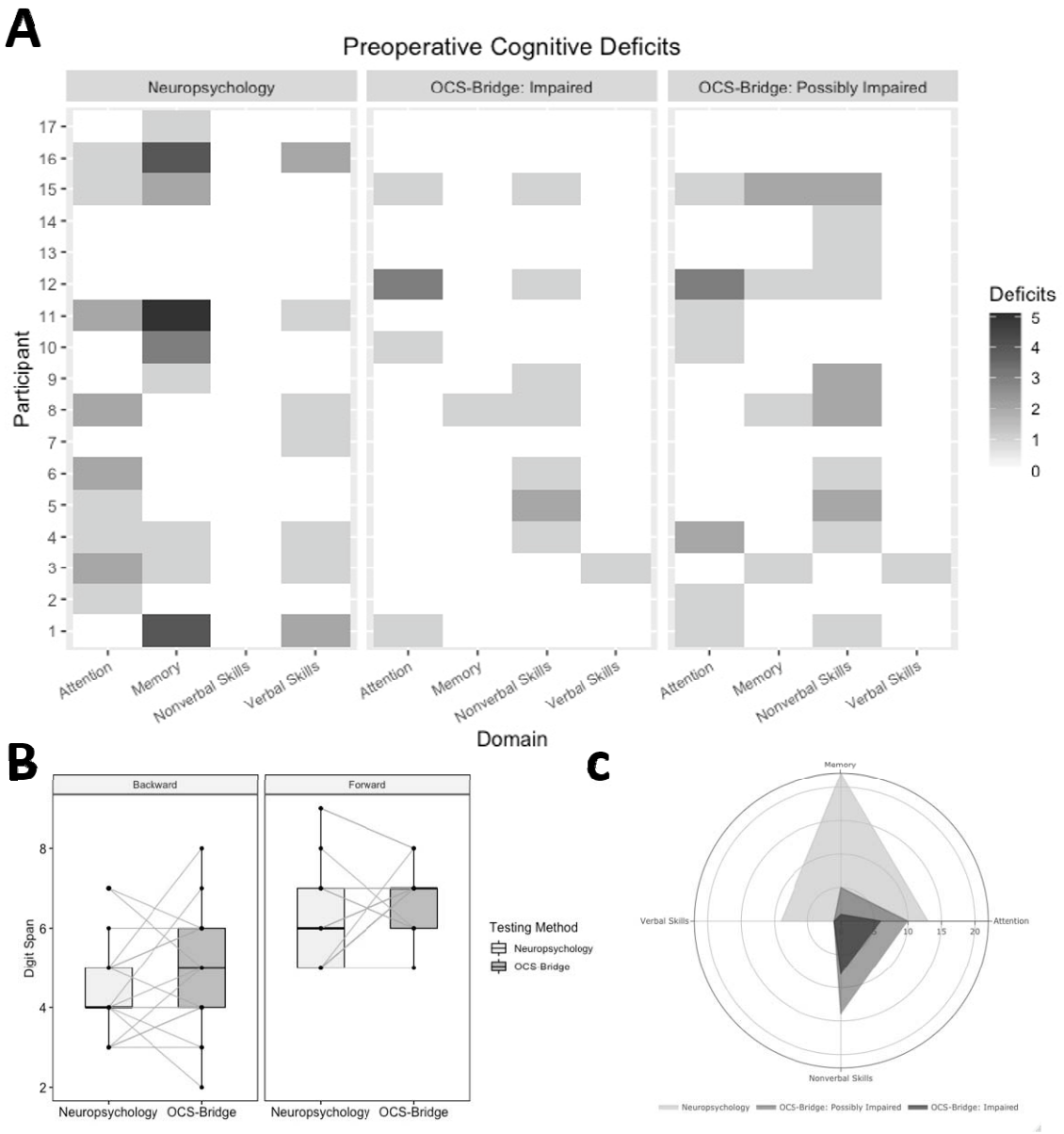

Figure 3. A. Cognitive deficits by domain from preoperative neuropsychological testing and OCS-Bridge assessments, B. Comparison of digit span analysis between neuropsychological testing and OCS-Bridge, C. Total preoperative deficits shown by neuropsychological testing and OCS-Bridge

Z-scores were calculated from both the neuropsychological testing and the OCS-Bridge data for each individual test and then averaged over each domain. There was no significant correlation between the total preoperative z-score for neuropsychological testing and OCS-Bridge using Pearson's product-moment correlation (p=0.7723, df = 15, r-value = 0.0758, 95% CI -0.42 – 0.53). There was significant correlation between OCS-Bridge and traditional neuropsychological testing in the domain of Attention (p=0.0413, df=15, r-value =-0.499, 95% CI -0.7902 - -0.0245), which showed a negative correlation in z-scores, but this correlation did not survive correction for multiple comparisons. None of the other three domains showed significant correlations between the two testing methods (**Figure 4**).

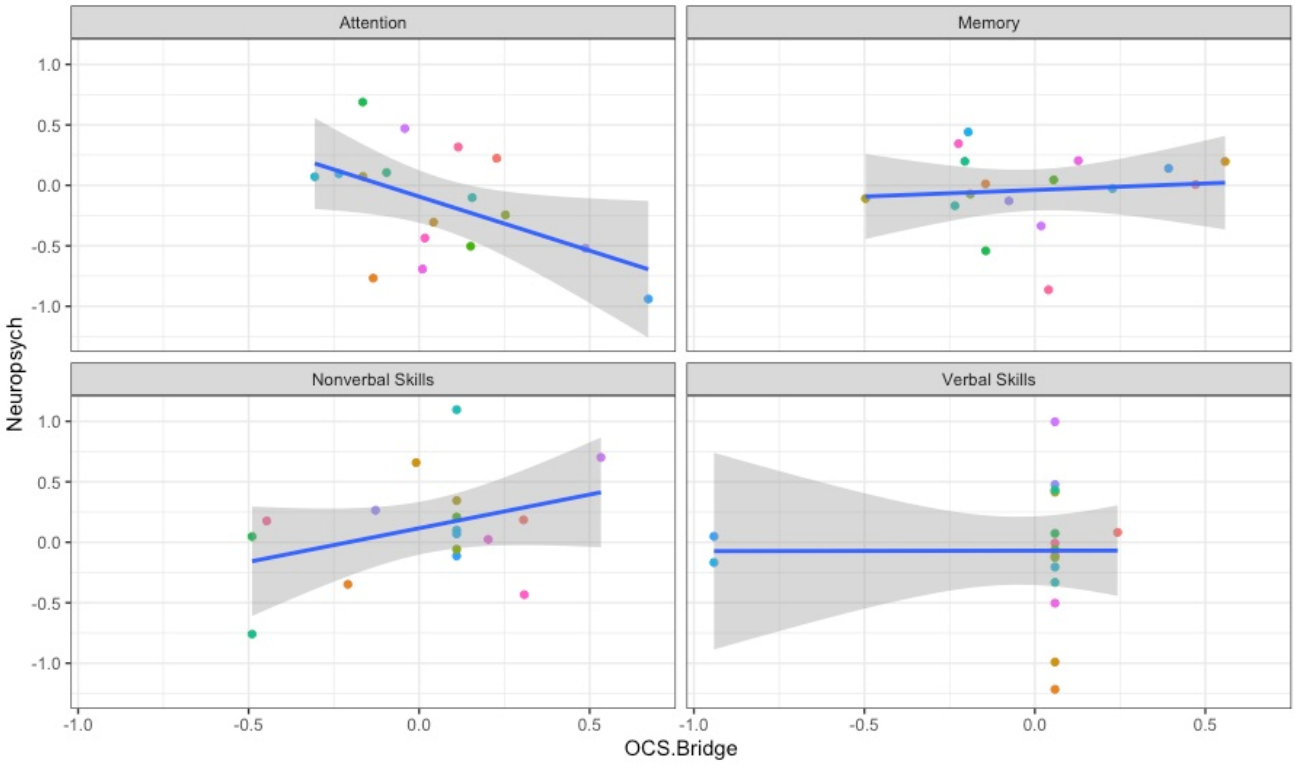

Figure 4. Total preoperative z-scores for OCS-Bridge and neuropsychological testing separated into combined domains.

Unsupervised Clustering

Because the deficits in the given domains did not correlate between the two testing modalities, unsupervised clustering was carried out to assess whether individual tests would cluster into groups

different from the original domains. Results of unsupervised k-means clustering using the preoperative Z-scores from both the OCS-Bridge and the Neuropsychology Battery for each participant is shown in **Figure 5.** Clusters were named for their dominant test type. Cluster 1 included 10 tests, predominantly Memory tests from the Neuropsychology battery and included all but one test from the AMIPB and BMIPB groups of tests. Cluster 2 included 10 tests in total, 8 from the Neuropsychology battery and 2 from OCS-Bridge, and was mostly composed of Attention and Verbal Skills tests. Cluster 3 included only 4 tests, 3 of which were from the SALT task within the OCS-Bridge battery, a test of Attention. Finally, Cluster 4 was the largest cluster, with 29 of the total tests, and included all but one of the tests of Nonverbal Skills, as well as 10 out of 14 of the OCS-Bridge Attention tests. Tasks which included several individual tests, such as the OCS-Bridge Hearts task or the OCS-Bridge SALT task, clustered together, in Cluster 4 and Cluster 3 respectively. The Beck Depression Inventory fell into Cluster 1, while the Beck Anxiety Inventory fell into Cluster 4.

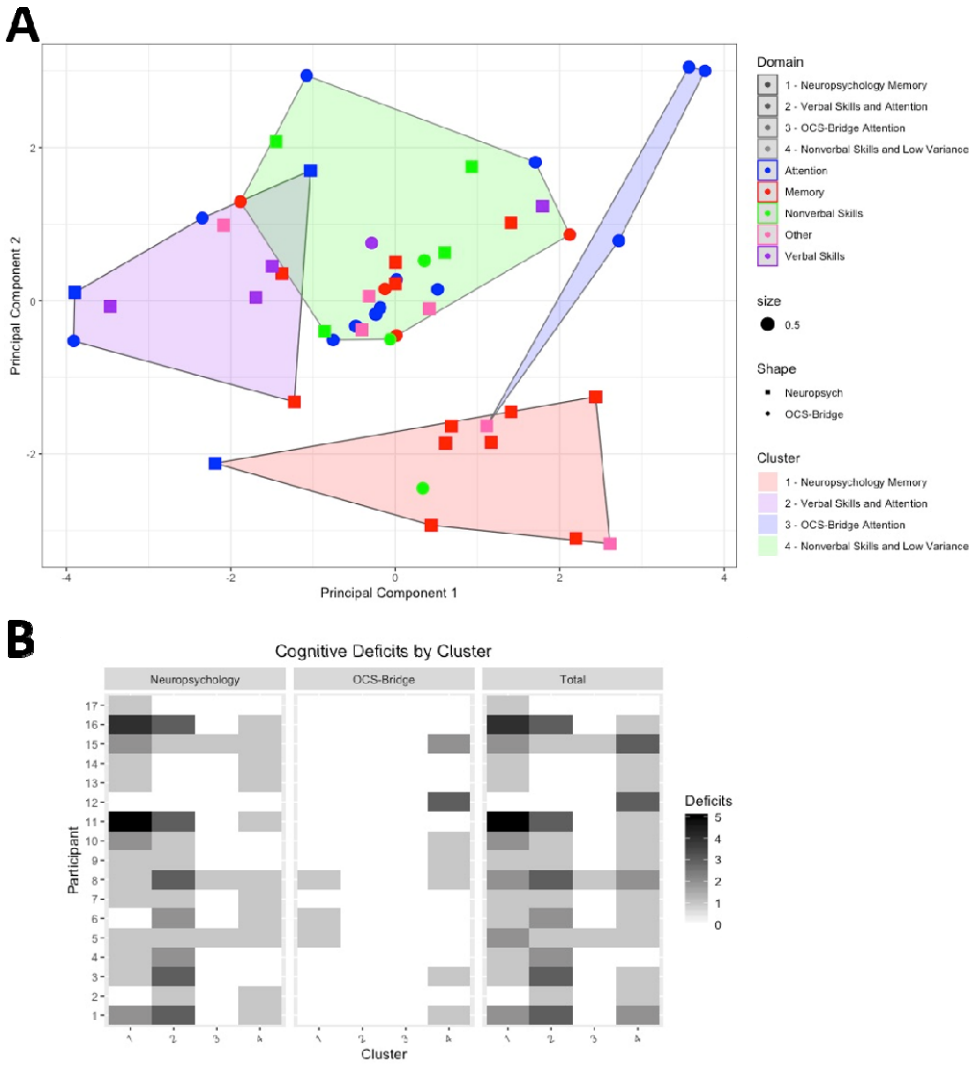

Figure 5. K-means clustering of individual cognitive tests and validation against deficit numbers. A. Principal components analysis plotted versus the k-means clustering results, B. Cognitive deficits for each participant by cluster, both before and after surgery.

DISCUSSION

Summary

Current traditional neuropsychological testing can be inefficient and resource intensive, making it difficult to perform at multiple time points during the patient journey (Taphoorn and Klein, 2004). Tablet-based assessments tools such as OCS-Bridge provide rapid, easy-to-use alternatives to traditional pen-and-paper testing, yet little work has been done to demonstrate the validity of these tests, or to assess their relative strengths and weaknesses when compared with traditional neuropsychological batteries. This paper presents data from a cohort of 17 patients with diffuse glioma who underwent neuropsychological testing with both modalities before and after surgery, and at several follow-ups, to address these questions. When combined, these two approaches may present a complementary picture of a patient's cognitive and emotional status over time and aid in the understanding of recovery trajectories after surgical resection of low-grade glioma.

Benefits of OCS-Bridge

There are several benefits of using a tablet-based system in conjunction with traditional neuropsychological testing in glioma patients. First, a tablet-based system is easy to transport and can be administered anywhere, versus traditional testing which is usually done in a hospital setting. Second, the OCS-Bridge assessment is easy to administer and does not necessarily require the presence of a trained neuropsychologist, making it more practical for use at multiple time points both pre- and post-surgery. Ease of use also presents the possibility that these assessments could be modified to be performed by the patient independently, furthering their potential for tracking subtle changes in cognitive functioning over time during recovery. Third, a tablet-based system may pick up subtle differences in certain areas of performance that traditional testing cannot, such as precise reaction times and accurate visual acuity scores. These results are in keeping with previous studies which have shown moderate validity of automated testing when compared with traditional neuropsychological batteries (Smith et al., 2013).

Complementary Role of Traditional and App-based Assessment

Traditional neuropsychological testing exhibited higher sensitivity to deficits in memory, picking up 22 total deficits, when compared with the tablet-based testing, which picked up only one deficit and six possible deficits. This may be partially accounted for by the difference in the number of independent measures of memory – the pen and paper neuropsychological tests used twelve independent measures of memory, whereas the OCS-Bridge assessment used eight. The OCS-Bridge testing, on the other hand, showed higher sensitivity to nonverbal deficits, showing five deficits and 12 possible deficits when compared with the six deficits picked up with the neuropsychological pen and paper tests. These results indicate potential complementarity of the two modalities. Due to its automated, tablet-based system, the OCS-Bridge assessment is able to assess aspects of perception, reaction time and other nonverbal skills that are difficult to test with pen-and-paper assessments.

The differences between the two modalities used here also highlight the importance of careful individual test selection. Many of the individual tests used on both the neuropsychological battery and the OCS-Bridge assessment demonstrated potential ceiling effects, with no patients scoring less than perfectly. While this may be due to lack of deficits in these tests within the cohort, there is also the potential that these tests were not sensitive enough to pick up subtle deficits within this patient population. This is especially likely given the different deficits picked up by the two tests – there were many memory deficits detected on traditional neuropsychological testing that were not detected using OCS-Bridge, and vice-versa for nonverbal deficits. This may be due to the cut-offs for “deficit” vs. “no deficit” used being too stringent, or the tests themselves being insensitive to deficits. Both modalities were initially designed to cover a wide range of neurological pathology, and therefore they may need modification to be appropriate for use in low-grade glioma patients, who are overall mostly young and relatively well, with few obvious neuropsychological deficits. These differences between tests also raise the possibility that binary classification into “deficit” vs. “no deficit” may not be as relevant as measurement of continuous functioning and comparison within a given patient over time. Clinically, a drastic change in functional ability, even if it does not meet the objective criteria for a neuropsychological deficit, is likely to be distressing to patients and relevant to their recovery trajectory.

The results presented above indicate potentially complementary roles of different modalities of cognitive testing in the assessment and postoperative follow-up of glioma patients. Traditional testing administered in a hospital or clinical setting by a qualified neuropsychologist demonstrated high sensitivity to cognitive deficits, particularly in the domains of memory and verbal skills, when compared

with the OCS-Bridge tablet-based system. The tablet-based system showed higher sensitivity to deficits in nonverbal skills including deficits in perception, and took less time, expertise and resources to administer and analyze. These differences indicate that there may be a role for a combined assessment schedule using both modalities for glioma patients in future to gain a more comprehensive understanding of their functional abilities pre- and postoperatively and as they recover in the longer term.

Study Limitations

Limitations of this study include incomplete outcome data and attrition that are inevitable when performing multiple assessments in a longitudinal manner. To ameliorate this risk care needs to be taken in timing assessments and coordinating them with routine clinical care. In particular the immediate post-operative assessment was not infrequently limited by fatigue while follow-up data was on occasion precluded by the participants geographical distance from centralized assessment. In a related manner, there were logistical and technical limitations to the assessments, which is not entirely unexpected during the beta-testing phase and with such an ambitious study protocol. Overall, the density of cognitive data is remarkable in this population, which is a testament to the study protocol. There was also heterogeneity introduced into the data between subjects who received only surgical intervention and those who had follow-up with chemotherapy or radiotherapy. Treatment with radiotherapy has previously been shown to be independently correlated with increasing cognitive deficits (Wefel et al., 2011; Lawrie et al., 2019; Hilverda et al., 2010). Unfortunately, due to the limited sample size it is impossible to comment on the differential contributions of treatment regimens to cognitive functioning in our cohort. Nevertheless these data present a ‘real-world’ view of treatment pathways that will be better reflective of outcomes encountered in series outside of clinical research studies. Finally, there are inevitable issues surrounding practice effects when using the same cognitive assessment tools over time in the same population (Cacciamani et al., 2018; Barron et al. 2015). There is a tradeoff between the direct comparability of assessments at different time points and the potential for practice effects. There is potential for future research aiming to address this problem in the design of app-based cognitive assessment tools.

Conclusions

Cognitive function in patients with diffuse glioma during early treatment is a complex and dynamic interplay between multiple factors. Both traditional and app-based assessment have complementary roles in assessment. This work provides the framework for robust, objective, and

accessible assessment across multiple centres that will aid in generating the necessary large datasets suitable so that we may better understand individual patient outcomes.

Acknowledgements

We thank all patients for generous involvement in the study. We also thank to Luca Villa, Rohit Sinha and Jessica Ingham for their contribution to the study. This research was supported by The Brain Tumour Charity. RRG was funded by the Guarantors of Brain.

References

Anderson, S. I., Taylor, R. & Whittle, I. R. Mood disorders in patients after treatment for primary intracranial tumours. *Br J Neurosurg* **13**, 480–485 (1999).

Barron, A. B. *et al.* Embracing multiple definitions of learning. *Trends in Neurosciences* **38**, 405–407 (2015).

Brown, P. D. *et al.* Detrimental Effects of Tumor Progression on Cognitive Function of Patients With High-Grade Glioma. *JCO* **24**, 5427–5433 (2006).

Cacciamani, F. *et al.* Evidence of practice effect in CANTAB spatial working memory test in a cohort of patients with mild cognitive impairment. *Applied Neuropsychology: Adult* **25**, 237–248 (2018).

Coalson, D. L., Raiford, S. E., Saklofske, D. H. & Weiss, L. G. WAIS-IV. in *WAIS-IV Clinical Use and Interpretation* 3–23 (Elsevier, 2010). doi:10.1016/B978-0-12-375035-8.10001-1.

Crawford, J. R. & Garthwaite, P. H. Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology* **21**, 611–620 (2007).

Demeyere, N., Riddoch, M. J., Slavkova, E. D., Bickerton, W.-L. & Humphreys, G. W. The Oxford Cognitive Screen (OCS): Validation of a stroke-specific short cognitive screening tool. *Psychological Assessment* **27**, 883–894 (2015).

Duffau, H. Lessons from brain mapping in surgery for low-grade glioma: insights into associations between tumour and brain plasticity. *The Lancet Neurology* **4**, 476–486 (2005).

Ek, L., Kristoffersen Wiberg, M. & Vestberg, S. Decline in executive functions and speed in suspected low-grade gliomas: A 3-year follow-up of a clinical cohort. *Applied Neuropsychology: Adult* **25**, 376–384 (2017).

Freedman, M. *et al.* The Toronto Cognitive Assessment (TorCA): normative data and validation to detect amnestic mild cognitive impairment. *Alz Res Therapy* **10**, 65 (2018).

Harris, R. J. *et al.* Altered functional connectivity of the default mode network in diffuse gliomas measured with pseudo-resting state fMRI. *J Neurooncol* **116**, 373–379 (2014).

Kaleita, T. A. *et al.* Prediction of neurocognitive outcome in adult brain tumor patients. *J. Neurooncol.* **67**, 245–253 (2004).

Lang, S. *et al.* Assessment of Cognitive, Emotional, and Motor Domains in Patients with Diffuse Gliomas Using the National Institutes of Health Toolbox Battery. *World Neurosurgery* **99**, 448–456 (2017).

Lang, S. *et al.* Functional Connectivity in Frontoparietal Network: Indicator of Preoperative Cognitive Function and Cognitive Outcome Following Surgery in Patients with Glioma. *World Neurosurgery* **105**, 913-922.e2 (2017).

Lawrie, T. A. *et al.* Long-term neurocognitive and other side effects of radiotherapy, with or without chemotherapy, for glioma. *Cochrane Database of Systematic Reviews* (2019) doi:10.1002/14651858.CD013047.pub2.

Mancuso, M. *et al.* Using the Oxford Cognitive Screen to Detect Cognitive Impairment in Stroke Patients: A Comparison with the Mini-Mental State Examination. *Frontiers in Neurology* **9**, (2018).

Meyers, C. A. & Hess, K. R. Multifaceted end points in brain tumor clinical trials: Cognitivedeterioration precedes MRI progression. *Neuro-Oncology* **5**, 89–95 (2003).

Meyers, J. E., Zellinger, M. M., Kockler, T., Wagner, M. & Miller, R. M. A Validated Seven-Subtest Short Form for the WAIS-IV. *Applied Neuropsychology: Adult* **20**, 249–256 (2013).

Motomura, K. *et al.* Supratotal Resection of Diffuse Frontal Lower Grade Gliomas with Awake Brain Mapping, Preserving Motor, Language, and Neurocognitive Functions. *World Neurosurgery* **119**, 30–39 (2018).

Motomura, K. *et al.* Neurocognitive and functional outcomes in patients with diffuse frontal lower-grade gliomas undergoing intraoperative awake brain mapping. *Journal of Neurosurgery* 1–9 (2019) doi:10.3171/2019.3.JNS19211.

Ng, J. C. H. *et al.* Effects of surgery on neurocognitive function in patients with glioma: a meta-analysis of immediate post-operative and long-term follow-up neurocognitive outcomes. *J Neurooncol* **141**, 167–182 (2019).

Noll, K. R., Bradshaw, M. E., Weinberg, J. S. & Wefel, J. S. Neurocognitive functioning is associated with functional independence in newly diagnosed patients with temporal lobe glioma. *Neuro-Oncology Practice* **5**, 184–193 (2018).

Noll, K. R., Ziu, M., Weinberg, J. S. & Wefel, J. S. Neurocognitive functioning in patients with glioma of the left and right temporal lobes. *J Neurooncol* **128**, 323–331 (2016).

Quental, N. B. M., Brucki, S. M. D. & Bueno, O. F. A. Visuospatial Function in Early Alzheimer's Disease—The Use of the Visual Object and Space Perception (VOSP) Battery. *PLoS ONE* **8**, e68398 (2013).

Reijneveld, J. C. *et al.* Health-related quality of life in patients with high-risk low-grade glioma (EORTC 22033-26033): a randomised, open-label, phase 3 intergroup study. *The Lancet Oncology* **17**, 1533–1542 (2016).

Rofes, A. *et al.* Survey on current cognitive practices within the European Low-Grade Glioma Network: towards a European assessment protocol. *Acta Neurochir* **159**, 1167–1178 (2017).

Sagberg, L. M. *et al.* Brain atlas for assessing the impact of tumor location on perioperative quality of life in patients with high-grade glioma: A prospective population-based cohort study. *NeuroImage: Clinical* **21**, 101658 (2019).

28.

Satoer, D., Visch-Brink, E., Dirven, C. & Vincent, A. Glioma surgery in eloquent areas: can we preserve cognition? *Acta Neurochir* **158**, 35–50 (2016).

Savulich, G. *et al.* Cognitive Training Using a Novel Memory Game on an iPad in Patients with Amnestic Mild Cognitive Impairment (aMCI). *International Journal of Neuropsychopharmacology* **20**, 624–633 (2017).

Smith, P. J., Need, A. C., Cirulli, E. T., Chiba-Falek, O. & Attix, D. K. A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with "traditional" neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology* **35**, 319–328 (2013).

Taphoorn, M. J. & Klein, M. Cognitive deficits in adult patients with brain tumours. *The Lancet Neurology* **3**, 159–168 (2004).

van Dellen, E. *et al.* MEG Network Differences between Low- and High-Grade Glioma Related to Epilepsy and Cognition. *PLoS ONE* **7**, e50122 (2012).

Vlaar, A. M. M. & Wade, D. T. The Adult Memory and Information Processing Battery (AMIPB) test of information-processing speed: a study of its reliability and feasibility in patients with multiple sclerosis. *Clinical Rehabilitation* **17**, 386–393 (2003).

Wang, Q. *et al.* A prospective longitudinal evaluation of cognition and depression in postoperative patients with high-grade glioma following radiotherapy and chemotherapy. *J Can Res Ther* **14**, 1048 (2018).

Wefel, J. S. *et al.* Neurocognitive function in patients with recurrent glioblastoma treated with bevacizumab. *Neuro-Oncology* **13**, 660–668 (2011).

Wefel, J. S., Noll, K. R., Rao, G. & Cahill, D. P. Neurocognitive function varies by IDH1 genetic mutation status in patients with malignant glioma prior to surgical resection. *NEUONC* **18**, 1656–1663 (2016).

Wickham, H. *ggplot2*. (Springer International Publishing, 2016). doi:10.1007/978-3-319-24277-4.