Presurgical predictors of early cognitive outcome after brain tumor resection in glioma patients

Andrea Zangrossi a, b, *, Erica Silvestri b, c, Marta Bisio b, d, Alessandra Bertoldo b, c, Serena De Pellegrin e, Antonino Vallesi a, Alessandro Della Puppa f, Domenico D’Avella g, Luca Denaro g, Renato Scienza g, Sara Mondini h, Carlo Semenza b, 1, Maurizio Corbetta a, b, c, e, i

a Department of Neuroscience, University of Padova, Italy
b Padova Neuroscience Center (PNC), University of Padova, Italy
c Department of Information Engineering, University of Padova, Italy
d Department of Biomedical Sciences, University of Padova, Italy
e Neurology Clinical Unit, University Hospital of Padova, Padova, Italy
f Neurosurgery Clinical Unit, Department of Neuroscience, Psychology, Pharmacology and Child Health, Careggi University Hospital and University of Florence, Florence, Italy
g Academic Neurosurgery, Department of Neuroscience, University of Padova, Italy
h Department of Philosophy, Sociology, Pedagogy and Applied Psychology, University of Padova, Padova, Italy
i Venetian Institute of Molecular Medicine, VIMM, Foundation for Advanced Biomedical Research, Padova, Italy

ARTICLE INFO

Keywords:
Cognition
Glioma
Neurosurgery
Precision medicine
Neuropsychological testing

ABSTRACT

Gliomas are commonly characterized by neurocognitive deficits that strongly impact patients’ and caregivers’ quality of life. Surgical resection is the mainstay of therapy, and it can also cause cognitive impairment. An important clinical problem is whether patients who undergo surgery will show post-surgical cognitive impairment above and beyond that present before surgery. The relevant prognostic factors are largely unknown.

This study aims to quantify the cognitive impairment in glioma patients 1-week after surgery and to compare different pre-surgical information (i.e., cognitive performance, tumor volume, grading, and lesion topography) towards predicting early post-surgical cognitive outcome.

We retrospectively recruited a sample of N = 47 patients affected by high-grade and low-grade glioma undergoing brain surgery for tumor resection. Cognitive performance was assessed before and immediately after (~1 week) surgery with an extensive neurocognitive battery. Multivariate linear regression models highlighted the combination of predictors that best explained post-surgical cognitive impairment.

The impact of surgery on cognitive functioning was relatively small (i.e., 85% of test scores across the whole sample indicated no decline), and pre-operative cognitive performance was the main predictor of early post-surgical cognitive outcome above and beyond information from tumor topography and volume. In fact, structural lesion information did not significantly improve the accuracy of prediction made from cognitive data before surgery.

Our findings suggest that post-surgery neurocognitive deficits are only partially explained by preoperative brain damage. The present results suggest the possibility to make reliable, individualized, and clinically relevant predictions from relatively easy-to-obtain information.

1. Introduction

Gliomas are the most frequent primary malignant intracranial tumors in adults, and represent 81% of all primary central nervous system tumors (Ostrom et al., 2014). Neurocognitive deficits are commonly experienced and strongly impact the quality of life of both patients and caregivers (Bergo et al., 2016). An important question in the clinical routine is whether cognitive performance will deteriorate after brain surgery. Surgical resection is the mainstay of therapy, and it can also cause cognitive impairment. An important clinical problem is whether patients who undergo surgery will show post-surgical cognitive impairment above and beyond that present before surgery. The relevant prognostic factors are largely unknown.

This study aims to quantify the cognitive impairment in glioma patients 1-week after surgery and to compare different pre-surgical information (i.e., cognitive performance, tumor volume, grading, and lesion topography) towards predicting early post-surgical cognitive outcome.

We retrospectively recruited a sample of N = 47 patients affected by high-grade and low-grade glioma undergoing brain surgery for tumor resection. Cognitive performance was assessed before and immediately after (~1 week) surgery with an extensive neurocognitive battery. Multivariate linear regression models highlighted the combination of predictors that best explained post-surgical cognitive impairment.

The impact of surgery on cognitive functioning was relatively small (i.e., 85% of test scores across the whole sample indicated no decline), and pre-operative cognitive performance was the main predictor of early post-surgical cognitive outcome above and beyond information from tumor topography and volume. In fact, structural lesion information did not significantly improve the accuracy of prediction made from cognitive data before surgery.

Our findings suggest that post-surgery neurocognitive deficits are only partially explained by preoperative brain damage. The present results suggest the possibility to make reliable, individualized, and clinically relevant predictions from relatively easy-to-obtain information.
surgery, the mainstay of glioma therapy, above and beyond the clinical status pre-surgery.

The literature on cognitive impairment in gliomas has mainly focused on tumor malignancy (Bosma et al., 2007; Desmurget et al., 2006), volume and topography (Arbulu et al., 2021; Banerjee et al., 2015). Tumor malignancy is taken into account by categorizing gliomas as low-grade (LGG; grade I and II) vs high-grade (HGG; grade III and IV) according to the WHO grading system, with the latter being more fast-growing, aggressive, and more frequently associated with cognitive deficits (Bosma et al., 2007). In contrast, LGG preserve patient’s cognitive performance for years (Desmurget et al., 2006). Tumor volume is also correlated with cognitive decline (e.g., executive functions, verbal fluency and perceptual speed; Hendriks et al., 2017; van Kessel et al., 2020).

Tumor topography also plays a role in cognitive outcome, especially when the dominant hemisphere is affected (Hendriks et al., 2017; Taphoorn and Klein, 2004). An attempt to reach a more fine-grained comprehension of the link between lesion anatomy and cognitive deficits in gliomas has been made through the voxel-based lesion symptom mapping (VLSM) approach (Bates et al., 2003), a popular technique in the stroke literature. However, these studies (Banerjee et al., 2015; Shalllice et al., 2010) suggest that lesion topography accurately predicts deficits in some domains (e.g., language; Banerjee et al., 2015; visuospatial; Herbet and Duffau, 2022), but not in others (e.g., visuospatial memory; Campanella et al., 2018). These findings could be explained by considering that some neuropsychological tests are not optimized for anatomical damage localization (Campanella et al., 2018). Another possible explanation is that VLSM may not be sensitive in gliomas given their slow infiltration into healthy tissue which may lead to functional reorganization (Aerts et al., 2016; Desmurget et al., 2006; Duffau, 2003; Hendriks et al., 2018), weakening the link between lesion localization and cognitive deficits. This idea fits with both LGG (Rijnen et al., 2020) and HGG (Fandino et al., 1999; Majos et al., 2017), and it is particularly true for cognitive functions that are more distributed across different brain regions. Accordingly, when a brain lesion is characterized by a sudden blood supply interruption (such as in stroke) these functions are better predicted by functional networks alterations than lesion location per se (Siegel et al., 2016). Moreover, the edema surrounding glioma lesion also contributes to neurological deficits (Butterbrod et al., 2019).

Taken together, these findings suggest that focusing only on tumor-related information might lead to inaccurate prediction of cognitive deficits in glioma patients (Aerts et al., 2016; Rijnen et al., 2020; Taphoorn and Klein, 2004). This explains why a key challenge in neuro- oncology is to predict cognitive disability following tumor resection (for a recent review see Sinha et al., 2019), the first mainstay of glioma treatment. Since one of the main goals of tumor resection is to maintain the highest quality of life (QoL), of which cognition is a major contributor (Klein et al., 2012; Taphoorn and Klein, 2004), a critical issue in counselling patients and caregivers on the opportunity of brain surgery is whether surgery will cause cognitive impairment (Hendriks et al., 2018). This issue is only partly understood. Some studies (Campanella et al., 2018, 2015) have investigated the acute effects of glioma surgery, focusing either on specific tasks (Charras et al., 2015; Herbet and Duffau, 2022; Wilson et al., 2015) or on left-lateralized tumors (Satoer et al., 2012). Immediately after surgery, sensory, motor, and language deficits have been reported mainly in LGG patients with recovery within three months (e.g., Bello et al., 2007; Teixidor et al., 2007). Conversely, HGG patients show milder changes in cognition post-surgery (Ng et al., 2015; Santini et al., 2012; Talacchi et al., 2011). A recent meta-analysis (Ng et al., 2019) including both LGG and HGG patients reported a positive impact of surgery on measures of attention, language, learning and memory, but a negative impact on executive functions. In clinical practice the pre-surgical cognitive status is considered a strong prognostic indicator for survival (Johnson et al., 2012) and a reliable index of tumor progression (Butterbrod et al., 2019), and several authors have pointed out that a pre-surgical neuropsychological assessment in glioma patients shall be a goal of future research (Lemaitre et al., 2022; Ng et al., 2019).

In this study we aimed to measure whether surgery worsens cognitive functions trying to answer the question “will my dad be worse after surgery?”. A second aim was to measure the value of pre-surgery cognitive measures as predictors of early (1 week) post-surgery cognitive function as compared to other prognostic factors, specifically age, education, tumor location, grading and volume.

2. Materials and methods

2.1. Participants

A sample of N = 49 glioma patients undergoing brain surgery for tumor resection were retrospectively recruited from the University Hospital of Padova (Italy) between March 2011 and March 2017. All patients underwent awake surgery using the sleep-awake-sleep anaesthetic procedure (Della Puppa et al., 2015) with the aid of MRI navigation and intraoperative neurophysiological monitoring. Pre-surgical mapping involved functional MRI in 12.8 % (N = 6) and tractography in 6.4 % (N = 3). Moreover, 5-ALA fluorescence guidance was adopted in 29.7 % patients (N = 14) as in (Della Puppa et al., 2013).

To be included in the study patients had to be (I) aged 18 years or older (no upper age limit), (II) free from other neurological or psychiatric disorders and (III) having been administered an extensive neuropsychological battery (see below) both before (T0) and shortly after the surgery (~1 week; T1). Moreover, we excluded patients with multiple lesions. Two patients were excluded because imaging data were not available, thus the final sample was composed of N = 47 patients (mean age = 51.6y, age standard deviation (SD) = 12.7y; 25F; Table 1). Patients were diagnosed both LGG (49 %) and HGG (51 %), more specifically WHO grade I was diagnosed in 12 patients (26%), grade II in 11 (23 %), grade III in 14 (30%) and grade IV in 10 (21 %). All patients gave their written informed consent to use their data for research when they were hospitalized, and our retrospective study was approved by the Bioethical Committee of the University Hospital of Padova.

2.2. MRI data acquisition and lesion segmentation

T1-weighted (both with and without contrast agent, i.e., gadolinium) and FLAIR images were preoperatively acquired on a 3 T Philips Ingenia scanner for 26 out of 47 patients. For the remaining patients, analogous clinical protocols have been acquired on a 1.5 T Siemens Avanto scanner (13 patients) and on a 1 T Philips Polaris scanner (8 patients).

FLAIR-hyper-intense areas were manually drawn in the patient’s native space by an expert neuroradiologist using ITK-SNAP software (https://www.itksnap.org; Yushkevich et al., 2006). Then, the advanced normalization tools (ANTS; Avants et al., 2008) was used to non-linearly register individual anatomical images to the symmetric MNI atlas (1 mm³) and each binary tumor mask was then normalized to this space. The tumor mask included both lesion boundaries and edema, which is known to impact cognition (Campanella et al., 2015; Tucha et al., 2000). Tumor volume was computed as the extent of the entire mask (in mm³). Lesion locations scarcely represented within our sample (i.e., voxels lesioned only in few patients) were discarded from the subsequent statistical analysis by retaining only voxels lesioned in at least 10 % of the sample (5 out of 47).

2.3. Neurocognitive data

We used the Esame Neuropsicologico Breve 2 (ENB2; Mondini et al., 2011), a neuropsychological battery validated on the Italian population and employed in previous investigations on glioma patients (Della Puppa et al., 2015). This battery is composed of 16 tests, i.e., digit span, immediate and delayed prose memory, interference memory (10 and 30 secs), trail making test parts A (TMT-A) and B (TMT-B), token test (five
items), phonemic fluency, abstract reasoning, a brief version of the cognitive estimation, overlapping figures, spontaneous drawing, copy drawing, clock drawing, ideational and ideomotor praxis test. Cognitive scores were first Z-scored on data from N = 370 age-matched healthy controls. Then, to study how cognitive functions were affected by surgery, we calculated a ΔZ score (i.e., T1 — one week post-surgery—minus T0 —pre-surgery) for each test and we investigated the post-surgical decline by means of a series of one-sample t-tests. Moreover, the proportion of patients showing worsening of performance between T0 and T1 (ΔZ < -1) was statistically compared to the percentage of patients showing stable or improved performance (ΔZ > -1) after the surgical procedure (van Kessel et al., 2020) by means of a series of chi-square tests. Finally, to rule out the possibility that the resulting cognitive changes were biased by practice effect, we computed the Reliable Change Index (RCI; Jacobson and Truax, 1991; Maassen et al., 2009).

Finally, a Principal Component Analysis (PCA) was run on T0 cognitive data to reduce their dimensionality (Corbetta et al., 2015; Ramsey et al., 2017). Components were retained only if their eigenvalue was > 1 (Kaiser’s criterion; Kaiser, 1960), moreover the cumulative variance explained by the set of retained components had to be ≥ 80 %.

This procedure resulted in N = 8 retained components (Supplementary Fig. 2), whose individual scores were included in the statistical models described in paragraph 2.5. The first component (PC1) mainly loaded on verbal comprehension; PC2 loaded on verbal comprehension, visuospatial analysis and planning; PC3 and PC4 related to praxis; PC5 and PC8 loaded on working memory and language. Tumor grading was classified according to the WHO classification system (Truax, 1991) for individual patients and each test. Notably, since the ENB2 battery provides test–retest reliability values only for 10 tests (digit span, immediate and delayed prose memory, interference memory at 10 and 30 secs, TMT-A, TMT-B, phonemic fluency, overlapping figures, clock drawing test) the RCI was computed only for these tests. Values of RCI < -1.96 or > 1.96 indicate significant worsening or improvement, respectively, while values in between suggest a stable performance (Jacobson and Truax, 1991; Maassen et al., 2009).

2.4. Tumor features and topography

We took into consideration tumor grading and volume as features potentially related to cognition. Tumor grading was classified according

| Patient ID | Tumor type     | Surgery date (year) | Grade | Hemisphere affected | Tumor volume, ml | Gender | Age   | Education |
|-----------|----------------|---------------------|-------|---------------------|------------------|--------|-------|-----------|
| 1         | Glioma         | 2016                | II    | R                   | 18.42            | Male   | 65    | 11        |
| 2         | Glioblastoma   | 2012                | IV    | R                   | 98.82            | Male   | 45    | 8         |
| 3         | Glioma         | 2016                | I     | L                   | 3.28             | Female | 54    | 8         |
| 4         | Astrocytoma    | 2011                | IV    | L                   | 45.95            | Female | 58    | 8         |
| 5         | Glioblastoma   | 2013                | III   | L                   | 19.24            | Male   | 68    | 8         |
| 6         | Glioma         | 2014                | III   | L                   | 13.35            | Female | 29    | 18        |
| 7         | Astrocytoma    | 2014                | IV    | L                   | 38.92            | Female | 43    | 18        |
| 8         | Glioma         | 2013                | II    | L                   | 33.50            | Female | 49    | 8         |
| 9         | Glioma         | 2014                | II    | L                   | 11.67            | Male   | 64    | 8         |
| 10        | Astrocytoma    | 2015                | IV    | L                   | 99.26            | Male   | 70    | 13        |
| 11        | Glioma         | 2013                | I     | R                   | 11.85            | Female | 23    | 18        |
| 12        | Glioblastoma   | 2014                | III   | L                   | 371.36           | Male   | 50    | 16        |
| 13        | Glioma         | 2017                | I     | R                   | 192.11           | Male   | 37    | 13        |
| 14        | Glioblastoma   | 2014                | IV    | R                   | 12.19            | Female | 69    | 8         |
| 15        | Glioma         | 2012                | II    | L                   | 23.97            | Female | 64    | 8         |
| 16        | Oligodendroglioma | 2013          | III   | L                   | 56.35            | Female | 65    | 8         |
| 17        | Glioma         | 2012                | III   | L                   | 20.39            | Male   | 59    | 13        |
| 18        | Glioma         | 2012                | I     | L                   | 94.46            | Female | 46    | 13        |
| 19        | Glioma         | 2017                | III   | L                   | 34.47            | Male   | 71    | 5         |
| 20        | Astrocytoma    | 2016                | IV    | R                   | 27.52            | Male   | 50    | 8         |
| 21        | Glioma         | 2011                | III   | R                   | 125.20           | Male   | 41    | 8         |
| 22        | Glioma         | 2013                | II    | L                   | 44.46            | Male   | 66    | 5         |
| 23        | Astrocytoma    | 2017                | IV    | L                   | 89.21            | Male   | 59    | 4         |
| 24        | Glioblastoma   | 2013                | III   | L                   | 8.72             | Female | 52    | 12        |
| 25        | Glioma         | 2012                | II    | R                   | 23.93            | Female | 52    | 5         |
| 26        | Glioma         | 2012                | II    | R                   | 150.87           | Female | 47    | 13        |
| 27        | Glioma         | 2012                | III   | L                   | 272.77           | Male   | 65    | 8         |
| 28        | Glioma         | 2013                | I     | L                   | 37.52            | Male   | 32    | 13        |
| 29        | Glioma         | 2015                | I     | L                   | 5.59             | Female | 34    | 11        |
| 30        | Glioblastoma   | 2013                | III   | L                   | 30.57            | Male   | 72    | 5         |
| 31        | Glioma         | 2014                | I     | L                   | 31.57            | Female | 37    | 13        |
| 32        | Glioma         | 2011                | II    | L                   | 36.12            | Male   | 26    | 13        |
| 33        | Glioma         | 2014                | I     | L                   | 8.36             | Male   | 42    | 8         |
| 34        | Astrocytoma    | 2015                | III   | L                   | 30.17            | Female | 43    | 13        |
| 35        | Glioblastoma   | 2013                | IV    | R                   | 34.13            | Male   | 49    | 13        |
| 36        | Glioma         | 2013                | I     | L                   | 3.99             | Female | 44    | 18        |
| 37        | Glioma         | 2012                | III   | L                   | 55.98            | Female | 46    | 8         |
| 38        | Astrocytoma    | 2016                | IV    | R                   | 12.47            | Female | 67    | 11        |
| 39        | Glioma         | 2012                | II    | L                   | 31.24            | Female | 57    | 18        |
| 40        | Glioma         | 2013                | III   | R                   | 13.08            | Female | 51    | 18        |
| 41        | Glioma         | 2014                | I     | L                   | 107.50           | Female | 45    | 18        |
| 42        | Glioblastoma   | 2014                | IV    | L                   | 32.47            | Female | 39    | 18        |
| 43        | Glioma         | 2011                | I     | L                   | 26.82            | Male   | 45    | 9         |
| 44        | Glioblastoma   | 2012                | III   | L                   | 25.77            | Female | 63    | 8         |
| 45        | Glioma         | 2011                | I     | L                   | 15.79            | Male   | 63    | 18        |
| 46        | Glioma         | 2015                | II    | L                   | 14.36            | Male   | 52    | 13        |
| 47        | Glioma         | 2012                | II    | L                   | 40.30            | Male   | 56    | 13        |
to the WHO (grade I to IV), while volume was calculated on the MNI normalized lesion masks (in cm$^3$), as specified above. Furthermore, we included tumor topography in the statistical analysis. To this end, we ran a PCA on lesion locations (in MNI space) across all patients to reduce data dimensionality (first row Fig. 1; see Fig. 2 for the first three PCs). Then, like for the neurocognitive data, components with associated eigenvalues $>1$ and showing a cumulative variance explained $>80\%$ were retained for further analysis ($N=10$; Supplementary Fig. 3). Individual scores were computed for each component and used as predictors in the regression models.

2.5. Models’ comparison and prediction of early post-surgical cognitive outcome

Our main aim was to compare different sets of information available preoperatively (T0) in predicting cognitive test scores at T1. More specifically, we identified four sets of predictors (each one including several variables): demographics (age and years of education), cognition at T0 (8 variables representing the first eight PCs), tumor features (volume and grading) and tumor topography (10 PCs). Then, we tested which combination of these sets of variables best predicted cognitive scores 1-week after the surgery. To this end, we built a series of multivariate linear regression models with the full set of T1 cognitive scores as dependent variables, with the aim of predicting the cognitive profile as described by the neuropsychological battery employed. In other words, each model (Mi) included a specific combination of sets of predictors (Table 2), spanning from the baseline model (M1) including only demographics to the full model (M4) including all predictors (Fig. 1 shows a representation of the method).

The nested models were compared by means of three independent Likelihood Ratio Tests (LRTs) testing whether the addition of a set of predictors to a simpler model significantly improved model fit. This procedure identified three models that were then compared in a higher-order LRT to find the best model to describe T1 cognitive data. Furthermore, we also compared the multivariate models by means of the corrected Akaike Information Criterion (AICc; Akaike, 1987; Hurvich and Tsai, 1989). The AICc is a widely recommended index (Burnham and Anderson, 2002) to select the model (with the lowest AICc value) with the best balance between likelihood and parsimony (i.e., number of predictors), hence accounting for the risk of overfitting. Moreover, it allows to compare non-nested models. Since this measure was designed for univariate models, we built all models independently for each dependent variable (i.e., T1 cognitive scores). Then, for each univariate model, we computed the AICc and we averaged AICcs within each model to obtain a single model-specific value.

The best model resulting from this comparison was then evaluated in terms of accuracy in the prediction of T1 cognitive scores by means of a leave-one-out design (i.e., the model was trained on N-1 observations and its predictive ability was tested on the left-out observation). All statistical analyses were performed by means of R software (R Core Team, 2019) and Matlab version R2017b (Mathworks Inc.).

3. Results

A sample of glioma patients with cognitive impairments undergoing surgery for tumor resection was retrospectively recruited at the University Hospital of Padova (Italy), with N = 47 being included in the final analysis. All patients were MRI scanned and underwent a neurocognitive assessment both before and after surgery. Lesion volume was highly variable across patients (range = 4.95–442.27 cm$^3$; median = 37.65 cm$^3$, MAD = 30.76 cm$^3$) and the lesions were more frequently located in the left hemisphere specifically in the frontal and temporal...
lobes (Fig. 2, line 1). The first three PCs explained ~ 55 % of lesions variance and captured temporo-parietal (PC1), frontal (PC2) and temporal (PC3) components of lesions. Fig. 2 (line 2 to 4) shows the topographical distribution of the first three components. The individual scores of the first 10 PCs (Supplementary Fig. 4) were used in the statistical analysis.

3.1. Small impact of surgery on cognition

The representation of the patients showing declined, stable, or improved performance (in terms of ΔZ) are shown in Fig. 3. Taking into account the cognitive scores across the whole sample (i.e., 47 patients × 16 scores = 752 scores), 85 % of scores showed a stable or improved performance (70 %: ΔZ ≤ 0; 15 %: ΔZ > 0), with only 15 % showing a clinically relevant worsening, quantified as a within-patient decline of at least 1 SD (ΔZ ≤ -1) as in previous studies (e.g., van Kessel et al., 2020).

At the group-level, we found ΔZ scores indicating a significant decline (after Bonferroni correction) only for TMT-A (t(46) = -3.73, 95 %, CI = [-0.58, -0.17], p = .004), verbal fluency (t(46) = -4.37, CI =...
[−0.80, −0.29], p < .001) and overlapping figures (t(46) = −3.48, CI = [−0.72, −0.19], p = .009). Conversely, no significant differences emerged when comparing the proportion of patients with worsened (ΔZ ≤ −1) vs stable/improved (ΔZ > −1) performance over time (all chi-square ps > 0.55) in any of the neurocognitive tests. These findings were confirmed (all chi-square ps > 0.47) also when running the same chi-square analysis on the percentage of patients with RCI indicating worsened (RCI < −1.96) or stable/improved performance (RCI > −1.96), thus ruling out practice effect. The patients per each RCI category (i.e., worsened, stable, improved) are depicted in Supplementary Fig. 5. Importantly, we did not find a significant relation between the proportion of worsened vs stable/improved patients and tumor grade (HGG vs LGG) both considering ΔZ (all chi-square ps > 0.127) and RCI (all chi-square ps > 0.19).

Overall, these results suggest that most within-patient changes in cognitive performance following surgery for tumor removal were not significant (see also Supplementary figure 7), also considering practice effect.

Aside from highlighting changes in cognitive performance due to surgery, more clinically relevant information is that related to the pre-surgical cognition PCs significantly improved a demographics only model (Pillai = 5.48, F[128,104] = 1.77, p = .001). Conversely, the further addition of tumor features (Pillai = 3.56, F[80,50] = 1.54, p = .051) and tumor topography data (Pillai = 5.45, F[160,150] = 1.12, p = .24) did not significantly improve model likelihood. In LRT 2 tumor topography significantly improved a demographics-only model (Pillai = 5.98, F[160,150] = 1.39, p = .02), however the model was significantly improved after adding pre-surgical cognitive components (Pillai = 5.29, F[128,104] = 1.59, p = .007), and did not further improve with the inclusion of tumor features. LRT3 showed that adding tumor features significantly improved a model with only demographics (Pillai = 3.82, F[80,50] = 2.03, p = .004), but did not improve after adding tumor topography PCs (Pillai = 5.81, F[160,150] = 1.30, p = .051). Conversely, a full model including also T0 cognitive data was significantly more accurate (Pillai = 5.28, F[128,104] = 1.57, p = .008).

These results suggest that tumor features and topography slightly improved a simpler model based only on demographics, but the best prediction was obtained by inserting pre-surgical cognition (Fig. 4). To formally test this, a higher order LRT was run including the best models from each LRT (M2a, M3b and M4, respectively) to identify the final best model. This comparison confirmed that neither the addition of tumor topography (M3b; Pillai = 5.68, F[160,150] = 1.23, p = .097) nor the combination of tumor features and topography (M4; Pillai = 3.44, F[80,50] = 1.38, p = .11) significantly improved a model that included only demographics and pre-surgical cognition (M2a). The AICc computation showed that model M2a was that with the lowest value, thus indicating the best balance between model likelihood and parsimony (Supplementary Fig. 6).

Taken together, these results show that adding the pre-surgical cognitive components as predictors significantly improved simpler models. Furthermore, a model including only demographics and pre-surgical cognitive components emerged as a more parsimonious choice which allows to keep a high level of predictive ability while including less variables.

The same procedure was repeated using the percentage of worsened tests as dependent variables and the results were confirmed

---

### 3.2. Early post-surgical cognitive outcome is better predicted by pre-surgical cognition than by tumor features and topography

To test which set of predictors best explained cognition after surgery we built three sets of models with different combinations of predictors, which were compared by means of three LRTs (Fig. 4). The statistical assumptions for multivariate regression were met. Specifically, we controlled all models for multicollinearity by means of the Variance Inflation Factors (VIF) which should be no potentially harmful collinearity (Bowerman and O’Connell, 1990; Myers, 1990). Finally, we checked for homoscedasticity (Goldfeld-Quandt test; Goldfeld and Quandt, 1965) and for autocorrelation of residuals (Durbin-Watson test; Durbin and Watson, 1971).

The first model comparison (LRT 1) showed that the addition of cognition PCs significantly improved a demographics only model (Pillai = 5.48, F[128,104] = 1.77, p = .001). Conversely, the further addition of tumor features (Pillai = 3.56, F[80,50] = 1.54, p = .051) and tumor topography data (Pillai = 5.45, F[160,150] = 1.12, p = .24) did not significantly improve model likelihood. In LRT 2 tumor topography significantly improved a demographics-only model (Pillai = 5.98, F[160,150] = 1.39, p = .02), however the model was significantly improved after adding pre-surgical cognitive components (Pillai = 5.29, F[128,104] = 1.59, p = .007), and did not further improve with the inclusion of tumor features. LRT3 showed that adding tumor features significantly improved a model with only demographics (Pillai = 3.82, F[80,50] = 2.03, p = .004), but did not improve after adding tumor topography PCs (Pillai = 5.81, F[160,150] = 1.30, p = .051). Conversely, a full model including also T0 cognitive data was significantly more accurate (Pillai = 5.28, F[128,104] = 1.57, p = .008).

These results suggest that tumor features and topography slightly improved a simpler model based only on demographics, but the best prediction was obtained by inserting pre-surgical cognition (Fig. 4). To formally test this, a higher order LRT was run including the best models from each LRT (M2a, M3b and M4, respectively) to identify the final best model. This comparison confirmed that neither the addition of tumor topography (M3b; Pillai = 5.68, F[160,150] = 1.23, p = .097) nor the combination of tumor features and topography (M4; Pillai = 3.44, F[80,50] = 1.38, p = .11) significantly improved a model that included only demographics and pre-surgical cognition (M2a). The AICc computation showed that model M2a was that with the lowest value, thus indicating the best balance between model likelihood and parsimony (Supplementary Fig. 6).

Taken together, these results show that adding the pre-surgical cognitive components as predictors significantly improved simpler models. Furthermore, a model including only demographics and pre-surgical cognitive components emerged as a more parsimonious choice which allows to keep a high level of predictive ability while including less variables.

The same procedure was repeated using the percentage of worsened tests as dependent variables and the results were confirmed

---

**Fig. 4.** Model comparison by Likelihood Ratio Tests (LRT). Each LRT included nested models, i.e., each model included the previous and added a set of predictors. **M2a** p < .05; (*) = p < .06.
(Supplementary Figure 8). Then, to rule out the possibility that the results were driven by specific individuals (e.g., outliers), we ran a sensitivity analysis by repeating the LRTs for N times, each time on data from N-1 patients, confirming M2a as the most frequently selected model across all iterations (Supplementary Figure 9). Finally, as a further control analysis, we ran a subject-specific model selection by testing the accuracy of each model in predicting T1 cognitive scores in a leave-one-out design, and by checking which model was the best in explaining each patient data (Supplementary Figure 10). The results showed that M2a was the most frequently selected model at the individual level.

3.3. Glioma lesion topography and early post-surgical cognitive worsening

The previous results highlight that glioma topography did not significantly improve a model including pre-surgical cognitive data in predicting individual post-surgical cognition. This is surprising given the traditional emphasis on localization of function in neurology. However, in the models the cognitive status was summarized by PCs that sum up variability across different tasks and de-emphasize the importance of more specific lesion topography information as overall changes in cognition may be less localized than more specific deficits. We could observe that lesions associated with worsening or improved/stable cognitive performance were largely overlapped. This suggests that lesion topography is poorly informative about surgery-based cognitive worsening, and lesion per se do not seem to localize well cognitive functions, thus confirming previous results. This is also supported by the maps of worsening probability shown in Fig. 5.

In line with previous literature, lesion topography explains early post-surgical worsening only in few cognitive domains, e.g., verbal working memory (Interference Memory) was more likely worsened in patients with left frontal and temporo-parietal lesions (Emch et al., 2019), analogical reasoning (Abstract Thinking) worsened in patients with left frontal pole lesions (Urbanski et al., 2016), memory retrieval (Prose Memory) by tumors involving prefrontal cortex (Barredo et al., 2015), and cognitive flexibility (Overlapping Figures) was related to bilateral fronto-parietal lesions (Uddin, 2021).

3.4. Prediction of cognitive outcome 1 week after surgery

The best model resulting from the previous analyses (including demographics and T0 cognitive scores) was then used to predict the pattern of cognitive scores at T1 with a leave-one-out design. Prediction accuracy was quantified by computing Pearson’s correlation between actual and predicted T1 cognitive scores (see Fig. 6; prediction error distribution is shown in Supplementary Figure 11). The multivariate
prediction showed an overall mean actual-predicted correlation $r = 0.57$ ($p < .001$, $SD = 0.19$, range 0.09-0.79), with all significant correlations but those related to TMT-B ($r = 0.28$) and Praxis ($r = 0.09$) test scores. This indicates that the model including demographics and PCs of T0 cognitive scores accurately predicted 14/16 T1 scores at the individual level.

4. Discussion

The present study aimed to quantify and predict the short-time cognitive sequelae of surgery for glioma resection, by comparing lesion topography, volume, tumor grading, and cognitive functioning prior to surgery.

As a first result, the analysis on cognitive changes between T0 and T1 using ΔZ scores revealed that only TMT-A, overlapping figures and verbal fluency showed a significant decline one week after surgery. These results are in line with the literature on cognitive deficits occurring early after surgery for gliomas which shows the main involvement of executive functions (for a review see Ng et al., 2019). On the other hand, the decline in verbal fluency could be explained by the unbalance of our sample towards left hemisphere lesions. However, when considering RCI and ΔZ levels, we did not find significant differences in any of the neurocognitive tests, with most scores across patients remaining relatively stable post-surgery (e.g., ΔZ > -1). These two findings are not in contradiction: a consistent weak worsening from pre- to post-surgery at the group level does not deny that clinically most patients remained stable within one standard deviation from normal across time points. Taken together, these results are consistent with studies showing that post-surgical cognitive deficits were not caused by surgery nor by perioperative causes in LGG (Klein et al., 2002; Schei et al., 2022) and HGG (Klein et al., 2001; Ng et al., 2019; Santini et al., 2012; Schei et al., 2022; Talacchi et al., 2011). Our results are in line with these findings and suggest that cognitive deficits occurring shortly after surgery (~1 week) mainly depend on cognitive deficits already

Fig. 6. Prediction of T1 cognitive scores. The figure shows the correlation between actual and predicted T1 cognitive scores using model M2a (including demographics and cognition before surgery as predictors) with a leave-one-out design. n.s. = not significant.
explain the differential pattern of cognitive impairment in HGG and LGG cases, since the tumor growth permits neuroplasticity processes to compensate for the cognitive decline after surgery. Our findings could be explained considering the functional brain organization taking place during insidious tumor growth, such as in the case of gliomas (Cargnelutti et al., 2020) and relates to the so-called lesion map (Aerts et al., 2016). Neurocognitive assessment and did not consider the variability of tests' scores in the healthy population (i.e., Z-scoring of patients’ scores on controls), which can help to better quantify the possible cognitive decline after surgery.

The second aim of our study was to directly compare lesion topography, tumor features and preoperative cognitive status towards predicting cognitive deficits 1-week after surgery. Our results suggest that cognitive status prior to surgery predict postsurgical cognition better than tumor features and lesion topography. This is in line with previous studies showing that cognitive deficits are only partially explained by tumor grade, size or location (Taphoorn and Klein, 2004). One possible explanation is that glioma progression impacts functional networks organization (Aerts et al., 2016; Cargnelutti et al., 2020) and structural connectivity causing both local and long-distance cognitive dysfunctions. In other words, cognitive deficits depend only partially on tumor location per se and presumably reflect networks disconnection (Aerts et al., 2016). For these reasons it is plausible that cognitive deficits emerging from a neurocognitive assessment are more informative than structural damage alone, as they reflect both local (i.e., damaged tissue) and long-distance (e.g., altered structural and functional network) effects of glioma proliferation. This could also account for the differential neurocognitive effects between intrinsic tumor and other kinds of physiological events (e.g. stroke). Similarly as in gliomas, stroke can also cause functional alterations in brain regions that appear structurally normal (Carter et al., 2012, 2016; He et al., 2007), and it has been highlighted that stroke focal brain lesions have a widespread impact on both structural and functional connectivity (Griffis et al., 2009; Salvaglio et al., 2020). For these reasons, symptom-mapping can benefit from the combination of connectivity- and lesion-based approaches to explain neurocognitive deficits both in stroke (Yourganov et al., 2016) and glioma patients (De Baene et al., 2019; Nakajima et al., 2022).

Given the focal nature of structural damage caused by stroke, we could speculate that the correspondence between structural damage and neurobehavioral deficits is stronger in stroke than in glioma patients. Accordingly, our results suggest that presurgical cognitive data can better predict postoperative cognitive deficits compared to lesion topography. It is important to note though, that we are not claiming that cognition explains a higher proportion of behavioral variance than lesion or tumor features (or their combination), but rather that a model including only cognition shows a higher likelihood given postsurgical cognitive data. Importantly, our findings apply both to LGG and HGG patients, since models including grading were not comparatively selected, and since we did not find grade-related differences in our best model.

Taken together, our results have important implications for the estimation of the cognitive risk of surgical resection in terms of presurgical mapping. Indeed, we suggest that it is critical to consider the brain networks potentially affected by a lesion and not only its topographical location, and that an indirect way to do it is to quantify pre-surgical neurocognitive impairments. Moreover, it should be noticed that, although prediction of long-term cognitive outcome (e.g., 3–6 months) in glioma patients is an important indicator of good survival QoL, short-term prediction is also desirable since it can help caregivers and family members in knowing what to expect immediately after surgery. This also supports a recent study highlighting cognitive assessment into the clinical management of glioma patients as a crucial tool to inform patients, caregivers, and clinicians on the cognitive functioning they should expect (Rijnen et al., 2020). An early post-surgical cognitive assessment may also help in settings a prompt cognitive treatment and in understanding the effects of surgery without the confounding of postsurgical adjuvant therapy.

The present study suffers from some limitations. First, our prediction may be enhanced by information about resection volume. Unfortunately, this information was not available. Second, given the retrospective nature of the present work we could not control for practice effect in cognitive retesting by using parallel forms. Thus, we made use of test–retest data available for the ENB2 battery, that regarded 10 tests out of 16, and with retesting at 1-month. Third, we had to deal with structural images acquired with different scanners and this could affect the generalizability of our results. Moreover, the cognitive battery adopted includes brief versions of well-known cognitive tests which could have limited application for the accurate cognitive profiling of patients. Finally, further limitations of the present investigation are represented by small sample size and by the topographical representation of tumors in this sample (mainly left-lateralized). Given these limitations, further studies are required to confirm our findings on independent and larger samples.

In conclusion, since human cognition arises from a complex interplay between networks of functionally connected brain areas, it is plausible that neurocognitive deficits are only partially explained by preoperative brain damage. Importantly, these deficits may also be linked to altered long-distance connectivity in critical functional networks. Neurocognitive assessment tools are designed to capture cognitive deficits which depend on both structural and functional brain alterations. In the present work, we demonstrated that cognitive performance before surgery can reliably estimate post-surgical cognitive deficits and does not benefit from the addition of neuroimaging data. In the precision medicine framework, the present results support the systematic inclusion of neurocognitive testing in the clinical routine of glioma patients to help making clinically relevant individualized outcome predictions.
NeuroImage: Clinical 36 (2022) 103219

References

MC was supported by FLAG-ERA JTC 2017 (grant ANR-17-HBPR-0001); Italian Ministry of Research - Departments of Excellence (MAR-T,ECCCELLENZ1A81_01); Fondazione Cassa di Risparmio di Padova e Rovigo (Grant Agreement number 55403); Italian Ministry of Health (RF-2008-12366899); Celegin Foundation Padova (CUP C94I2000420007); BIAL foundation (grant n. 361/18); H2020-SC5-2019-2, Grant Agreement number 869505); Italian Ministry of Health - C94I20000420007; European Community H2020 (H2020-SC5-2019-2, Grant Agreement number 869505); European School of Network Neuroscience-euSNN, H2020-SC5-2019-2, (Grant Agreement number 869505); European Community H2020 (H2020-SC5-2019-2, Grant Agreement number 869505); Italian Ministry of Health (RF-2019-12369300).

Funding

Celegin Foundation Padova (CUP C94I2000420007)

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2022.103219.

Data curation. Carlo Semenza: Conceptualization, Supervision, Writing – review & editing. Maurizio Corbetta: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available upon reasonable request.

Acknowledgements

Bosma, I., Vos, M.J., Heimans, J.J., Taphoorn, M.J.B., Aaronson, N.K., Postma, T.J., van der Ploeg, H.M., Muller, M., Vanderport, W.P., Slotman, B.J., Klein, M., 2007. The consequences of neurocognitive dysfunction in high-grade glioma patients.1. Neuro. Oncol. 9, 53–62. https://doi.org/10.1215/15228177-2006-012.

Bowerman, B.L., O’Connell, R.T., 1990. Linear statistical models: an applied approach. Pacific Grove, Dunbury.

Burnham, K.P., Anderson, D.R., 2002. Information and Likelihood Theory: A Basis for Model Selection and Inference, in: Model Selection and Multimodal Inference. A Practical Information-Theoretic Approach. Sprenger, pp. 49–97.

Carbetta, M., Ramsey, L., Corbetta, M. 2012. Why use a connectivity-based approach to study stroke and recovery of function? Neuroimage 62, 2271–2280. https://doi.org/10.1016/j.neuroimage.2012.02.070.

Charras, P., Herbet, G., Deverdun, J., de Champlain, N.M., Dufau, H., Bartolomeo, P., Bonnetbouc, F., 2015. Functional reorganization of the attentional networks in low-grade glioma patients: a longitudinal study. Cortex 63, 27–41. https://doi.org/10.1016/j.cortex.2014.08.010.

Corbetta, M., Ramsey, L., Callejas, A., Baldassarre, A., Hacker, C.D., Siegel, J.S., Astafiev, S.V., Rengachary, J., Zinn, K., Lang, C.E., Connor, L.T., Puce, T., Strube, M., Carter, A.R., Shulman, G.L., Corbetta, M., 2010. Resting interhemispheric functional magnetic resonance imaging connectivity predicts performance after stroke. Ann. Neurol. 67, 365–375. https://doi.org/10.1002/ana.21905.

Corti, A.R., Shulman, G.I., Corbetta, M., 2018. Why use a connectivity-based approach to study stroke and recovery of function? Neuroimage 62, 2271–2280. https://doi.org/10.1016/j.neuroimage.2012.02.070.

Desmurget, M., Bonnefond, B., Dufau, H., 2006. Contrasting acute and slow-growing lesions: a new door to brain plasticity. Brain 130, 898–914. https://doi.org/10.1093/brain/woo301.

Dufau, H., 2003. Functional recovery after surgical resection of low grade gliomas in eloquent brain: hypothesis of brain compensation. J. Neurol. Neurosurg. Psychiatry 74, 901–907. https://doi.org/10.1136/jnnp.74.7.901.

Durbin, J., Watson, G.S., 1971. Testing for serial correlation in least squares regression. III. Biometrika 58, 1. https://doi.org/10.2307/2334313.

Emch, M., van Bastian, C.C., Koch, K., 2019. Neural correlates of verbal working memory: an fMRI meta-analysis. Front. Hum. Neurosci. 13 https://doi.org/10.3389/fnhum.2019.00180.

Fandino, J., Kollias, S.S., Wieser, H.G., Valavanis, A., Yonekawa, Y., 1999. Intraoperative subcortical languagetract mapping guides surgical removal of gliomas involving speech areas. Neurosurgery 60, 67–82. https://doi.org/10.1227/01.NEU.0000054906.58601.e2.

Funding

Celegin Foundation Padova (CUP C94I2000420007)

Identification of potential brain plasticity. Neurology 210, 116589. https://doi.org/10.1016/j.neuroimage.2020.116589.
