Comparing Glioblastoma Surgery Decisions Between Teams Using Brain Maps of Tumor Locations, Biopsies, and Resections

Domenique M.J. Müller, MD1; Pierre A.J.T. Robe, MD, PhD2; Roelant S. Eijgelaar, Msc3; Manix G. Witte, PhD3; Martin Visser, Msc2; Jan C. de Munck, PhD2; Marieke L.D. Broekman, MD, PhD, LLM2; Tatjana Seute, MD, PhD2; Jeroen Hendrikse, MD, PhD2; David P. Noske, MD, PhD2; William P. Vantendorp, MD, PhD2; Frederik Barkhof, MD, PhD2,4; Mathilde C.M. Kouwenhoven, MD, PhD1; Emmanuel Mandonnet, MD, PhD5; Mitchel S. Berger, MD6; and Philip C. De Witt Hamer, MD, PhD1

PURPOSE The aim of glioblastoma surgery is to maximize the extent of resection while preserving functional integrity, which depends on the location within the brain. A standard to compare these decisions is lacking. We present a volumetric voxel-wise method for direct comparison between two multidisciplinary teams of glioblastoma surgery decisions throughout the brain.

METHODS Adults undergoing first-time glioblastoma surgery from 2012 to 2013 performed by two neuro-oncologic teams were included. Patients had had a diagnostic biopsy or resection. Preoperative tumors and postoperative residues were segmented on magnetic resonance imaging in three dimensions and registered to standard brain space. Voxel-wise probability maps of tumor location, biopsy, and resection were constructed for each team to compare patient referral bias, indication variation, and treatment variation. To evaluate the quality of care, subgroups of differentially resected brain regions were analyzed for survival and functional outcome.

RESULTS One team included 101 patients, and the other included 174; 91 tumors were biopsied, and 181 were resected. Patient characteristics were largely comparable between teams. Distributions of tumor locations were similar, suggesting absence of indication variation. Differentially resected regions were identified in the anterior limb of the right internal capsule and the right caudate nucleus, indicating treatment variation. Patients with (n = 12) and without (n = 6) surgical removal in these regions had similar overall survival and similar permanent neurologic deficits.

CONCLUSION Probability maps of tumor location, biopsy, and resection provide additional information that can inform surgical decision making across multidisciplinary teams for patients with glioblastoma.

INTRODUCTION Standard treatment of glioblastoma consists of surgery and radiotherapy with concomitant and adjuvant chemotherapy. Nevertheless, the prognosis remains dismal, with a 1-year survival rate of 39%.1 Initial neuro-oncologic treatment decisions include whether to undergo resective surgery or diagnostic biopsy or refrain from surgery, if the patient is unfit or unmotivated or the tumor is considered nonresectable.2,3 More extensive resections are associated with longer survival in glioblastoma, often recognized as a causal relation.4,7 The aim of neurosurgery is therefore to maximize tumor removal while preserving functional integrity. The surgical dilemma is that patient outcome depends on the decision of where to stop tumor removal. If stopped too early, more residual disease remains, with the risk of earlier tumor recurrence and shorter survival. If stopped too late, more tumor-infiltrated functional brain is removed, leading to brain function deficits and reducing the patient’s quality of life.8-11

Neurosurgical decision making depends on several factors before surgery: the condition, age, and comorbidity of the patient; neurologic signs and symptoms; brain location of the tumor and derived resectability; expectations of the neurosurgeon regarding benefit and risk for the patient; and motivation of the patient after discussion of these expectations.5,12-25 Intraoperatively, neurosurgeons use various techniques to distinguish tumoral tissue from functional brain tissue, such as image guidance with neuronavigation, magnetic resonance imaging (MRI), ultrasound or fluorescence,16 and intraoperative stimulation mapping.17 These factors and techniques add up to practice variation in neuro-oncologic decision making, which was captured by population-based cohort studies evaluating patterns of care in patients with glioblastoma.18-22 Biopsy percentages in...
these reports range widely: 5%,22 12%,20 24%,18 34%,19 and 44%.21 The quality of glioblastoma resection is usually reported as the percentage of patients with a gross total resection, which also ranges considerably: 28%,21 32%,18 46%,20 and 47%.22 Quality registries are being developed to evaluate the quality of neuro-oncologic care23,24; however, no standards exist for the evaluation of neuro-oncologic and neurosurgical decision making among multidisciplinary teams. For comparison of teams, bias resulting from heterogeneity of tumor locations within the brain should be addressed; we proposed using a voxel-wise approach in nonenhancing glioma.25

Voxel-wise methodologies were introduced for low-grade gliomas to evaluate surgical decision making with volumetric analysis of preoperative tumor and residual disease throughout the brain.26,27 The results of voxel-wise analyses can be plotted in standard brain space, which is called a probability map. Surgical decision making in nonenhancing gliomas has been compared among neurosurgical teams using resection probability maps (RPMs).25

Ultimately, the best decision for a patient with glioblastoma regarding biopsy or resection, and in case of resection, the extent of tumor removal, should be based on a collaborative body of objective data rather than on the intuition of one or a few surgeons. A step in that direction is to compare and discuss the patterns of current care. In this study, we address the proof of concept for the use of probability maps to assess and compare the patterns of (intra)operative decision making between multidisciplinary teams in patients with glioblastoma.

**METHODS**

**Patient Inclusion and Data Collection**

Consecutive patients older than 17 years with a histopathologic diagnosis of supratentorial glioblastoma and first-time brain surgery in 2012 or 2013 were included from two brain tumor treatment centers: the Vrije Universiteit Medical Center in Amsterdam and the University Medical Center in Utrecht, the Netherlands. Patients underwent either resection with glioblastoma removal or biopsy for diagnosis. We defined tumor removal for histopathologic diagnosis only by an open or stereotactic approach as biopsy and more contrast-enhanced tumor removal than necessary for a histopathologic diagnosis by craniotomy as resection. The histopathologic diagnosis was based on WHO 2007 criteria. Written informed consent was obtained from all patients.

The two multidisciplinary neuro-oncologic teams had a similar approach to decision making in tumor board meetings on surgical indications for new patients by consensus of neurosurgeons based on patient and tumor characteristics. Intraoperative decision making was performed by neurosurgical teams under the supervision of a neuro-oncologic surgeon from a pool of two or three per center, respectively. Image guidance with neuronavigation was routinely applied. Functional MRI, DTI tractography, and intraoperative stimulation mapping for language and motor functions were used by indication. The Utrecht team had started using fluorescence-guided microscopy in 2013; the Amsterdam team had not.

Baseline characteristics of these patients were extracted from digital department databases, as well as from pre- and postoperative MRI. MRI acquisition protocols were comparable between the two teams (Data Supplement). A complication was defined as a new intervention or intensive care admission. A functional decline was defined as a drop in Karnofsky performance score of > 20 between admission and 6 weeks after surgery.

**Brain Map Construction**

The construction of brain maps is described in detail in the Data Supplement. Briefly, the enhancing portions of tumors were segmented on pre- and postoperative T1-weighted gadolinium-enhanced images. For patients who underwent biopsy, postoperative MRI was usually omitted, and we therefore considered the residual tumor to be identical to the preoperative tumor. Next, tumor volumes were non-linearly registered to the Montreal Neurological Institute (MNI) 152 1-mm brain template to compare tumor locations between patients and teams.

Three types of probability maps were constructed by voxel-wise aggregations of the registered tumor volumes, each
reflecting clinical decision making. First, a tumor probability map (TPM) representing tumor burden before treatment was constructed by summation of preoperative tumors divided by the total number of patients treated by a team. A TPM difference between two teams would indicate referral or recruitment bias given a tumor location. Second, a biopsy probability map (BPM) representing the likelihood of a biopsy procedure over resection was constructed for each team by dividing the number of biopsies per location by the number of resections. A BPM difference between two teams would indicate treatment variation in indications for biopsy given a tumor location. This can be considered a measure of multidisciplinary neuro-oncologic team decision making. Third, an RPM representing the likelihood of residual disease after surgery was constructed. A difference between RPMs of patients undergoing resection would indicate treatment variation in tumor removal given an indication for resection. This can be considered a measure of neurosurgical team decision making during surgery. In addition, a difference between RPMs based on all patients in a cohort, including those undergoing biopsy or resection, would then indicate treatment variation in tumor removal for all patients with newly diagnosed glioblastoma. This can be considered an overall measure of multidisciplinary neuro-oncologic decision making after first-time surgery for newly diagnosed glioblastoma.

Subgroup Analysis
As an example of quality evaluation from treatment variation, we determined the subgroup of patients involved in the largest cluster of differentially resected voxels. Therefore, we extracted clinical follow-up and radiologic data from medical records to explore differences in outcome potentially related to treatment variation.

**RESULTS**

**Patient and Treatment Characteristics**

A total of 279 patients were treated, 102 by the Amsterdam team and 177 by the Utrecht team. For the probability map analysis, 101 (99%) and 174 patients (98%) were available, respectively. In two patients, the preoperative MR scan was unavailable from the referring hospital, and in two patients, the glioblastoma was nonenhancing. Patients were mostly comparable between teams in baseline characteristics (Table 1).

An MR scan within 3 days after resection was obtained in 77 patients (100%) who underwent resection in Amsterdam and in 68 (65%) of 104 in Utrecht. To adhere to current imaging standards,28 we based the resection probabilities for and in 68 (65%) of 104 in Utrecht. To adhere to current imaging standards,28 we based the resection probabilities for patients undergoing biopsy or resection, would then indicate treatment variation in tumor removal for all patients with newly diagnosed glioblastoma. This can be considered an overall measure of multidisciplinary neuro-oncologic team decision making after first-time surgery for newly diagnosed glioblastoma.

**Subgroup Analysis**

As an example of quality evaluation from treatment variation, we determined the subgroup of patients involved in the largest cluster of differentially resected voxels. Therefore, we extracted clinical follow-up and radiologic data from medical records to explore differences in outcome potentially related to treatment variation.

**RESULTS**

**Patient and Treatment Characteristics**

A total of 279 patients were treated, 102 by the Amsterdam team and 177 by the Utrecht team. For the probability map analysis, 101 (99%) and 174 patients (98%) were available, respectively. In two patients, the preoperative MR scan was unavailable from the referring hospital, and in two patients, the glioblastoma was nonenhancing. Patients were mostly comparable between teams in baseline characteristics (Table 1).

An MR scan within 3 days after resection was obtained in 77 patients (100%) who underwent resection in Amsterdam and in 68 (65%) of 104 in Utrecht. To adhere to current imaging standards,28 we based the resection probabilities for evaluation of neurosurgical decisions on the patients with postoperative scans within 3 days. The quality of neurosurgical team decisions in patients undergoing resection was not significantly different between teams in terms of residual volume, extent of resection, or outcome (Table 1).

**Referral Bias**

A significant difference was indicated in relative tumor incidence by the TPM comparison between both teams (Fig 1). In several regions in the left hemisphere, regions with differential tumor incidence were identified, such as the caudate nucleus, anterior limb of the internal capsule, and temporal stem. These observations seemed to be based on systematic differences, such as referral or recruitment bias.

**Variation in Biopsy Indications**

No differentially biopsied regions were identified by the BPM comparison between teams (Fig 2). The BPMs

---

**Table 1** Baseline Patient and Treatment Characteristics and Outcomes

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Amsterdam Team (n = 101)</th>
<th>Utrecht Team (n = 174)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, years</td>
<td>62.3 (37)</td>
<td>66.1 (38)</td>
<td>.798</td>
</tr>
<tr>
<td>Female sex</td>
<td>37 (37)</td>
<td>68 (38)</td>
<td></td>
</tr>
<tr>
<td>Karnofsky performance score</td>
<td></td>
<td></td>
<td>.160</td>
</tr>
<tr>
<td>100</td>
<td>9 (9)</td>
<td>5 (3)</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>22 (22)</td>
<td>54 (31)</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>23 (23)</td>
<td>40 (23)</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>28 (28)</td>
<td>34 (20)</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>12 (12)</td>
<td>19 (11)</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>7 (7)</td>
<td>12 (7)</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>3 (2)</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Year of surgery</td>
<td></td>
<td></td>
<td>.617</td>
</tr>
<tr>
<td>2012</td>
<td>49 (49)</td>
<td>77 (44)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>52 (52)</td>
<td>97 (56)</td>
<td></td>
</tr>
<tr>
<td>Tumor in left brain hemisphere</td>
<td>39 (39)</td>
<td>95 (55)</td>
<td>.012</td>
</tr>
<tr>
<td>Median preoperative tumor volume, mL</td>
<td>28.7 (17)</td>
<td>37.9 (17)</td>
<td></td>
</tr>
<tr>
<td>MRI timing</td>
<td></td>
<td></td>
<td>.082</td>
</tr>
<tr>
<td>Days between preoperative MRI and surgery</td>
<td>−4</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>Days between postoperative MRI and surgery</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Surgical strategy</td>
<td></td>
<td></td>
<td>.001</td>
</tr>
<tr>
<td>Biopsy</td>
<td>21 (21)</td>
<td>70 (40)</td>
<td></td>
</tr>
<tr>
<td>Resection</td>
<td>80 (79)</td>
<td>104 (60)</td>
<td></td>
</tr>
<tr>
<td>Median residual tumor volume, mL</td>
<td>1.0</td>
<td>0.3</td>
<td>.438</td>
</tr>
<tr>
<td>No. of patients undergoing resection with residue &lt; 3 mL</td>
<td>59 (74)</td>
<td>44 (42)</td>
<td></td>
</tr>
<tr>
<td>Median extent of resection, %</td>
<td>96.3</td>
<td>96.8</td>
<td>.206</td>
</tr>
<tr>
<td>No. of patients undergoing resection with extent of resection &gt; 95%</td>
<td>45 (56)</td>
<td>43 (41)</td>
<td>.406</td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
<td>.083</td>
</tr>
<tr>
<td>Median OS, months</td>
<td>12</td>
<td>9</td>
<td>.187</td>
</tr>
<tr>
<td>Complications</td>
<td>3 (3)</td>
<td>7 (4)</td>
<td>.750</td>
</tr>
<tr>
<td>Functional decline</td>
<td>5 (5)</td>
<td>20 (11)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: MRI, magnetic resonance imaging; OS, overall survival.
demonstrated that consensus existed regarding in which brain regions biopsy was preferred over resection, such as the thalamus, internal capsule, splenium of the corpus callosum, and periventricular region around the occipital horns, and conversely in which brain regions resection was preferred over biopsy, such as the anterior temporal lobe, right prefrontal cortex, and right inferior parietal lobe. The absence of differences in the BPMs suggests that brain location was not an important factor contributing to multidisciplinary decision making on performing a biopsy procedure. Nevertheless, the overall biopsy percentage differed, with 21 (21%) and 70 patients (40%) undergoing biopsy in Amsterdam and Utrecht, respectively ($P = .002$). Patients undergoing biopsy did not differ with regard to age (median, 63.1 and 65.9 years), condition (Karnofsky score $< 70$ in six [29%] and 16 [24%]), or tumor volume (median, 27 and 29 mL), respectively.

**Variation in Tumor Removal**

Tumor removal between teams was comparable in many brain regions. However, the Amsterdam team more often resected tumor in the right caudate nucleus and anterior limb of the right internal capsule than the Utrecht team (Fig 3). Common regions where tumor was generally removed included the temporal lobes, prefrontal cortices, and superior parietal lobules. Other tumor-infiltrated regions were commonly not removed, such as the anterior perforating substance and corticospinal tract. The precision of these probabilities varied by location depending on the number of patients with information (Fig 1).

**Subgroup Analysis of Patients With Different Tumor Removal**

The brain region with the largest differential resection probability was the right caudate nucleus and surrounding white matter pathways, in which the Amsterdam team decided to remove the tumor, whereas the Utrecht team did not. This would probably have remained unnoticed without a probability map comparison. We identified 18 patients who underwent resection. No differences were noted in neurologic outcome, survival, postsurgical hydrocephalus, or ventricular dissemination of glioblastoma for patients with ($n = 12$) and without ($n = 6$) tumor removal (Data Supplement).

**Variation in Oncologic Quality of Glioblastoma Surgery**

The decisions for the whole cohort on whether to perform biopsy or resection and the observed extents of resection were similar between the two teams in both hemispheres (Fig 4). Therefore, overall decisions on biopsy and extent of resection do not depend on brain location.

**DISCUSSION**

The main finding of this study is that probability maps of tumor location, biopsy, and resection enable the evaluation...
of surgical decision making among multidisciplinary teams in patients with glioblastoma.

Its potential use is illustrated by some examples. First, we identified a significant difference in tumor lateralization between two teams, and with TPM comparison, a more detailed specification of the brain regions involved could be made. This bias could be related to regional referral patterns of which we are as yet unaware. Because each center is considered to serve designated geographic regions, neurologists seldom cross-refer patients to other hospitals. Second, we identified a significant difference in biopsy percentage between two teams. This could not be attributed to specific brain locations according to the BPM analysis, nor to patient age, condition, or tumor volume. The difference in fluorescence guidance and application of stimulation mapping is not an explanation, because these techniques were used by the team performing more biopsies and fewer extensive resections. It seems other factors contributed to the differential overall biopsy preference, possibly related to expectations between the multidisciplinary teams regarding the benefits and risks of resection.

In addition, several of the observed differences in neurosurgical decision making were related to unresolved clinical dilemmas. Several plausible explanations were collected as feedback from both neuro-oncologic teams for the difference in resection probability lateral of the frontal horn at the level of the caudate nucleus in the right hemisphere. As another example of the potential use of the probability map comparison, our findings highlight these dilemmas. First, opinions differed on the justification of surgical removal of the nondominant caudate nucleus. Arguments against removal included the functional anatomy of the caudate nucleus, which has been associated with behavior, sometimes resulting in severe cognitive impairment, such as abulia, as learned from observations in stroke and other focal lesions. In addition, the caudate nucleus is part of the negative motor network, connected to the supplementary motor area through the frontostriatal tract. For supracomplete glioma resection, the caudate nucleus is considered a strict functional boundary by neurosurgeons who push the limits of resection. In contrast, deficits were absent after caudate removal in low-grade glioma surgery. In our data, no argument was raised against removal of the caudate to preserve the patient's functional integrity, although no systematic neuropsychological examinations were available for these patients. Second, the opinions also differed on justification of ventricular entry at the anterior horn of the lateral ventricle between the two teams. One team favored more extensive tumor removal, whereas the other team primarily intended to avoid adverse events resulting...
from ventricular entry, such as hydrocephalus or tumor seeding. Hydrocephalus is an uncommon complication after glioma surgery, but it has been reported more frequently after ventricular entry. Tumor seeding can be facilitated by wide ventricular opening. Although early publications did not find support for this, later publications have indicated more frequent leptomeningeal glioblastoma dissemination after ventricular entry. In our data, tumor seeding only occurred in cases with opening of the ventricle, although this difference was not statistically significant with this limited number of patients. Altogether, our data do not support avoidance of ventricular entry to prevent these complications. Third, the opinions differed on the use of intraoperative stimulation mapping with cognitive tasks in the right frontal lobe. One team argued that cognitive impairment is an uncommon outcome and that cognitive mapping has not been validated and could impede tumor removal. The other team had started to develop a cognitive mapping protocol. More recently, awake mapping of nonlanguage functions has been proposed for low-grade glioma surgery, particularly in the right hemisphere. On the basis of these arguments and a small number of patients in our study, the dilemma over removal of a tumor-infiltrated caudate nucleus remains.

No standards exist for quality measurement of neuro-oncologic care. Often, for a patient cohort, a variety of outcome measures are reported. Examples of oncologic outcome measures include the percentage of patients undergoing biopsy, median extent of resection, percentage of patients undergoing gross total resection, median residual disease volume, time to progression, and overall survival time. Examples of functional outcome measures include the percentage of patients with neurologic deficits, average cognitive performance scores, and health-related quality-of-life indices. We postulate that in addition to these outcome measures, probability map comparison of tumors, biopsies, and resections aids in evaluating oncologic outcome without bias from brain location. Probability maps thus enable outcome comparison among hospitals treating patients with diverging complexity of tumor locations. Another advantage is that probability map information is
available early in the course of disease, so quality assessment is not delayed until progression or survival. To optimize decision making and improve quality of care, a balance is required between oncologic and functional goals.\textsuperscript{52} That is, quality evaluation should not be narrowed to “less red is more quality” in BPMs and RPMs.

Several strengths of our study can be discerned. Volumetric analysis of residual disease was an important innovation compared with surgeon impression of resection completeness for an individual patient; now probability map comparison extends this by aggregating results for many patients per multidisciplinary team. New patterns of disease and treatment can be discovered, as demonstrated. By applying these methods at high resolution, we avoid assumptions on eloquence and interpretation of brain region classifications. Furthermore, an open-minded approach toward quality evaluation, learning, and improvement by changing current opinions was required from both teams. The well-documented patient cohorts were sizeable and seemed to represent clinical practice, which encourages extrapolation of our results. Limitations to our findings include technical and inclusion issues. The technical observations on segmentation and registration are described in detail in the Data Supplement. Regarding inclusion, patients not undergoing surgery, and consequently without histopathologic diagnosis, were not included in our study. The impact of this potential selection bias is small, because we identified only three patients in Amsterdam with a radiologically presumed glioblastoma without any surgery.

Our results have a number of practical implications. First, BPMs and RPMs can identify patients who will benefit most from intraoperative stimulation mapping and guide intraoperative stimulation mapping to the tumor regions with most function variability to potentially improve decision making. Instead of a static concept of functional and nonfunctional brain territories (ie, brain eloquence\textsuperscript{14}), we now have a dynamic quantitative measurement per brain voxel of resectability that applies to a patient cohort treated by a team during a period of time. Second, neuro-oncologic registries have been initiated for the collection of patient data for quality evaluation.\textsuperscript{21,24,53} Our findings support the inclusion of standardized imaging as part of this collection. Third, our findings may serve patient counseling regarding decision making for biopsy or resection indications within multidisciplinary teams and for optimization of surgical strategies within neurosurgical teams. Fourth, probability maps may speed up knowledge transfer between neuro-oncologic teams. Fifth, the use of probability maps may enhance education and training of young neurologists, neurosurgeons, and radiologists by putting treatment decisions for individual patients into a broader perspective of aggregated patient cohorts and by providing intuitive and objective brain map visualization.
Future work could focus on larger patient populations to resolve the pictured clinical dilemmas and identify more patterns of treatment variation. The impact on patient outcome of identifying and resolving dilemmas resulting from treatment variation remains to be determined. Furthermore, RPMs could be used for individualized patient care to predict resection results, contribute to image-guided surgical plans, and evaluate resection results. In conclusion, probability maps of tumor location, biopsy, and resection provide additional information that can inform surgical decision making across multidisciplinary teams for patients with glioblastoma.

REFERENCES
