Biopsy versus resection in the management of malignant gliomas: a systematic review and meta-analysis

A review

ABRAHAM TSITLAKIDIS, M.D., M.SC., NICOLAS FOROGLOU, M.D., PH.D., CHRISTOS A. VENETIS, M.D., M.SC., IOANNIS PATSALAS, M.D., PH.D., ATHANASIOS HATZISOTIRIOU, M.D., PH.D., and PANAGIOTIS SELVIARIDIS, M.D., PH.D.

1First Department of Neurosurgery, AHEPA University Hospital; 2Unit for Human Reproduction, First Department of Obstetrics and Gynaecology, Papageorgiou General Hospital, Aristotle University of Thessaloniki; and 3Agios Loukas Clinic, Thessaloniki, Greece

Object. The aim of this study was to answer the question whether quality of life and progression-free and overall survival are increased in adults with supratentorial malignant glioma who are treated with cytoreductive resection as compared with those who only undergo biopsy.

Methods. A literature search of the electronic databases MEDLINE, EMBASE, and CENTRAL was performed to identify relevant studies published before May 2008. Hand-searching of reference lists of the identified studies and relevant review articles was also performed. A study was considered eligible, regardless of study design (prospective or retrospective), if: 1) quality of life and/or progression-free and/or overall survival was compared among adult patients undergoing biopsy or resection, and 2) patient age and Karnofsky Performance Scale scores were not significantly different among the 2 groups compared.

Results. One randomized controlled trial and 4 retrospective studies (involving a total of 1111 patients) were found eligible for this systematic review. A meta-analysis of the eligible studies demonstrated a significant increase in overall survival in the patients treated with resection instead of biopsy (hazard ratio 0.61, 95% CI 0.52–0.71, p < 0.0001, fixed-effect model). Although statistical pooling was not feasible, the available data suggest that quality of life was increased in patients treated with resection rather than biopsy, while there did not seem to be any significant difference in progression-free survival between the 2 groups.

Conclusions. Based on the best available evidence, it appears that cytoreductive resection in adults with supratentorial malignant glioma is associated with improved overall survival as compared with biopsy. However, well-designed prospective studies are needed for more solid conclusions to be drawn. (DOI: 10.3171/2009.7.JNS09758)
Biopsy versus resection in the management of malignant gliomas

Methods

Search Strategy

A literature search was conducted by 2 of the authors (A.T. and N.F.) in the electronic databases MEDLINE, EMBASE and CENTRAL in May 2008, using a search strategy that included terms that described the disease (for example, glioma), the treatment (for example, surgery), and the type of studies of interest (for example, comparative study) and facilitated search techniques like expanding of terms to identify studies answering the research question of interest. No language limitations were applied.

The reference lists of all studies considered for inclusion in the systematic review and of all relevant review articles were hand-searched. All studies identified by this process were subsequently screened by 2 independent reviewers (A.T. and N.F.).

Finally, the authors of the studies included in the systematic review were contacted to obtain any available unpublished data related to the subject of the systematic review.

Study Selection

Initially, the titles of the articles were screened to exclude articles that did not refer to the disease and the therapeutic intervention under examination.

Subsequently, the abstracts of the articles were checked to exclude articles that did not provide information related to the research question of interest. We included studies in which 1) QOL and/or progression-free and/or overall survival was compared among adult patients with supratentorial malignant gliomas (regardless of specific histological type) undergoing biopsy or resection; and 2) an effort was made to prevent selection bias by controlling the confounding variables, especially age and KPS score—by randomization, matching, or stratification—or, as a last resort, by post hoc verification that there were no differences between the groups under study. These 2 variables (age and KPS score) are considered significant prognostic factors3,12–14,18,19,40,66,75,80,84,89,94,101,103,104 known to influence the selection of the type of surgical intervention.2,5,16,19,24,34,35,46,51,75,81–83

No limitation was posed with respect to study design (cohort or case-control study), data collection process (prospective or retrospective), sample size, particular surgical techniques used, or the way the outcomes were described, as long as they were reported separately for the 2 study groups.

In every phase of this process, the selection was conservative, so that, whenever doubts about whether a study conformed to the inclusion criteria remained, they would be elucidated in a succeeding phase. Whenever doubts remained after the retrieval of the full text, we attempted to communicate with the researchers in order to retrieve the missing information.

Bibliographical reviews and letters to the editor with subjects relevant to the study according to the keywords were identified among the references that were retrieved from the databases. Their full text was retrieved in order to identify further studies from their reference list.

Studies Identified

The electronic search, after the exclusion of duplicates, resulted in the retrieval of 4889 publications. The titles of these articles were examined to exclude irrelevant studies, resulting in 435 potentially eligible publications. The abstracts of these studies were scrutinized, and eventually 12 studies that could provide data to answer our study question were identified, as well as 25 bibliographic reviews and letters to the editor.8–11,15,21,37,39,42,43,62,64,65,67,72–75,79,80,92,95,97,105,106

Subsequently, the reference lists of all the bibliographic reviews, the letters, and the candidate studies were retrieved and examined. One potentially eligible study28 was identified by this process. Thus there were 13 candidates for inclusion in the systematic review.

The full text of these studies was examined thoroughly, resulting in the exclusion of 8 studies. Specifically, studies were excluded because age and/or KPS were not reported as comparable22,68,69,78,109 (5 studies), the question of interest was not answered1,93 (2 studies), or children were included among the patients examined and no data regarding only the adult patients were available28 (1 study).

To ascertain whether an effort to control the confounding variables was made, whenever there was doubt, communication with the researchers was pursued (Fig. 1).

Data Extraction

Data extraction was conducted independently by 2 of the authors (A.T. and C.A.V.). The following data were recorded: study characteristics (study design, power analysis, recruitment interval, observation duration, sample size, outcomes, and association with the intervention); patient data (inclusion and exclusion criteria, age, KPS score, and initial symptoms); tumor data (histological characteristics, location, preoperative size, and midline shift); procedural data (distribution criteria, surgical techniques description, postoperative morbidity and mortality description, estimation of degree of resection, and number of surgeons); adjuvant treatment data (for radiotherapy: protocol, number of patients treated, number of patients who completed treatment, dose given, and time until initiation of radiotherapy; for chemotherapy: chemotherapeutic agent and number of patients who participated in a chemotherapy protocol); and outcome data (definition and median value of overall survival and progression-free survival; definition and estimation of a measure of QOL). Whenever the extraction of important data was not possible, communication with the researchers was attempted.

Outcomes

The main outcome measure of the meta-analysis was overall survival (time interval from the outset of the measurement to death, in months). Progression-free survival (time interval from the outset of the measurement to diagnosis of recurrence, in months) and quantitative expression of QOL were the secondary outcome measures of this meta-analysis.

The most suitable effect measure for variables that...
are expressed as time-to-event, like overall survival and progression-free survival, is the HR, which is similar to risk ratio. It describes how likely (or unlikely) is the occurrence of the event in the participant at a specific moment, if he is given the treatment evaluated instead of the control treatment. It is notable that when variables of this type are examined in a clinical study or a meta-analysis, it is assumed that the HR remains constant during the observation (proportional hazards assumption).²⁶

In the meta-analysis of survival studies, the log hazard ratio estimate (LnHR) and the standard error of log hazard ratio (SE[LnHR]) are used. However, their values are rarely given directly in the published article and they should be calculated from other data:²⁷,¹⁰⁸ 1) from the value of the log-rank statistic and the log-rank statistic variance; 2) from the limits of the CI of HR that results from log-rank test or Cox regression; 3) from the value of the significance level (p value) of the log-rank test or Cox regression and the total number of patients and deaths in every group; or 4) as a last resort, directly from the Kaplan-Meier curves, if the total number of patients is known.

These calculations were performed with a personal computer using Microsoft Excel 2003 (Microsoft Corp.),⁶⁴ according to the methodology described by Parmar et al.⁷⁷ and Whitehead,²⁰,¹⁰⁸ while the isolation of the Kaplan-Meier curves and the extraction of the percentages of cumulative survival were performed with the program ImageJ 1.40g (National Institutes of Health, 1997–2008),¹ in the platform Java 1.5.0.7. Whenever independent patient data were provided, calculations were performed with the statistical package SPSS 15.0 (SPSS Inc., 2005). Quality of life, as a distinctive value in a scale, can be analyzed as a continuous variable, if enough data exist. The weighted mean difference of the postoperative QOL or its change before and after the operation was used as effect measure. However, whenever the scales used were not similar between studies, statistical pooling was performed with the use of standardized mean difference.

**Quantitative Data Synthesis**

The quantitative synthesis of the time-to-event effect measures as well as the QOL effect measures was planned to be carried out using the inverse variance method.

This quantitative synthesis was conducted with the program MIX 1.7,⁷ using the computational machine of

---

**Fig. 1.** Flowchart illustrating the steps of the systematic review and meta-analysis. n = number of studies.
Biopsy versus resection in the management of malignant gliomas

Microsoft Excel. Statistical heterogeneity was estimated with the Cochran Q statistic, based on the null hypothesis that all the studies derive from the same population. Since this specific statistic often has little power to identify true heterogeneity between 2 studies, wherever it did not take statistically significant values, the value of the I² statistic was taken under consideration. This statistic describes the percentage of differentiation between the studies that is due to heterogeneity rather than to chance, and values greater than or equal to 50% were considered indicative of significant heterogeneity. Wherever statistically significant heterogeneity was not identified, the calculation of the summary effect measure was carried out with the inverse variance fixed-effect model, while, wherever statistically significant heterogeneity was identified, it was carried out with the DerSimonian-Laird inverse variance random-effects model. Moreover, Egger regression was used to estimate publication bias, and the trim-and-fill method was used to estimate the sensitivity of the results against possible publication bias.

Sensitivity analysis was carried out by excluding studies the methodological quality of which was considered marginal for inclusion in the quantitative data synthesis. Furthermore, it was considered appropriate that sensitivity analysis should also be done by excluding studies for which the only outcome data were the Kaplan-Meier curves, as the data extraction process used in this case could be inaccurate.

Subgroup analyses were performed based on the grade of the gliomas (studies that were limited to Grade IV gliomas and studies that included Grade III gliomas) and on the age of the patients (studies that were limited to the elderly and studies with a broad spectrum of age). The possible differentiation between the subgroups was explored with the Q\textsubscript{int} statistic, which is compared with the chi-square distribution for df = n – 1 degrees of freedom.

Results

Systematic Review

The primary characteristics of the 5 studies that were included in the systematic review are presented in detail (Tables 1–6).

Eligible studies were published between 1993 and 2003. Of the 5 reviewed studies, 4 were retrospective cohort studies, and 1 was a randomized controlled trial. There was substantial heterogeneity with respect to KPS scores as well, with mean values ranging from 55 to 84.2. These 2 confounding factors were handled as continuous variables in 4 studies and as stratified variables in 1 study. So, for the control of between-groups similarity, post hoc tests that compared the mean values of the variables in the 2 groups were used in 3 studies and tests that compared the distribution of the groups among

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Country</th>
<th>Study Type (power analysis)</th>
<th>Recruitment Time Int, Max Obs Dur</th>
<th>Sample Size—No. of Pts, No. of Deaths (M/F)</th>
<th>Outcome</th>
<th>Assoc w/ Resection</th>
<th>Commun w/ Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simpson et al., 1993</td>
<td>US</td>
<td>retrosp cohort (NA)</td>
<td>1974–1989,† 33–36 mos</td>
<td>645 (NR), 615‡ (NR)</td>
<td>median o/all survival</td>
<td>positive</td>
<td>yes</td>
</tr>
<tr>
<td>Kelly &amp; Hunt, 1994</td>
<td>US</td>
<td>retrosp cohort (NA)</td>
<td>July 1984–Jun 1992, ~2 yrs</td>
<td>128 (67:55), NR</td>
<td>mean o/all survival, mean time to neurol deter</td>
<td>positive</td>
<td>yes</td>
</tr>
<tr>
<td>Kiwit et al., 1996</td>
<td>Germany</td>
<td>retrosp cohort, matching pr analysis (NA)</td>
<td>Jan 1986–Dec 1990, ~50 wks</td>
<td>80 (51:29)‡, 80 (51:29)‡</td>
<td>median o/all survival, KPS score on discharge</td>
<td>positive</td>
<td>yes</td>
</tr>
<tr>
<td>Kreth et al., 1999</td>
<td>Germany</td>
<td>retrosp cohort (NA)</td>
<td>1991–1994, ~2 yrs</td>
<td>228 (NR), 225 (121:104)‡</td>
<td>median o/all survival (a), change in KPS score after op (b) &amp; RT (c)</td>
<td>positive only for (b)</td>
<td>yes</td>
</tr>
<tr>
<td>Vuorinen et al., 2003</td>
<td>Finland</td>
<td>RCT (no)</td>
<td>1993–1996, 576 days</td>
<td>30 (NR), 23 (11:12)</td>
<td>median o/all survival (a), median time to deter (b)</td>
<td>positive only for (a)</td>
<td>yes</td>
</tr>
</tbody>
</table>

* Assoc = association; Commun = communication; deter = deterioration; Dur = duration; Int = interval; NA = not applicable; neurol = neurological; NR = not reported; Obs = observation; o/all = overall; pr = pair; Pt = patient; RCT = randomized clinical trial; retrosp = retrospective; RT = radiotherapy.
† Duration of recruitment obtained from Curran et al.
‡ Calculated based on data from the article.
strata were used in 1 study,\textsuperscript{91} while in the study of Kiwit et al.\textsuperscript{57} matching was applied. We considered all 3 ways of control acceptable (Table 2).

A parameter that differed between the studies was the histological type of the tumors, as either solely glioma Grade 4/IV patients (3 studies)\textsuperscript{55,63,91} or patients with glioma Grade 3/III or 4/IV (2 studies)\textsuperscript{57,100} were included. The WHO grading system was used in 3 studies\textsuperscript{58,60,110} and the scale of Daumas-Duport in 1 study;\textsuperscript{23} the authors of the fifth study, Simpson et al.,\textsuperscript{91} did not report what scale was used. Only the study of Kelly and Hunt\textsuperscript{55} referred to the cell line from which the tumor was considered to be derived. The anatomical site of the tumor was specified in 4 studies\textsuperscript{55,63,91,100} (Table 3).

The surgical techniques used were analyzed in detail in 4 studies.\textsuperscript{55,57,63,100} The operations were performed either by 1 surgeon\textsuperscript{55,100} in 2 studies and by many\textsuperscript{57,63,91} in 3 (Table 4). The description of the radiotherapeutical protocols that were used varied from detailed\textsuperscript{57,63,91} (3 studies) to nonexistent\textsuperscript{55} (1 study), while in the study of Vuorinen et al.,\textsuperscript{100} only the total dose of radiation was reported. The 2 groups (patients undergoing no surgery other than biopsy and those undergoing resection) were comparable in terms of participation in radiotherapy and/or total dose of radiation in 3 studies,\textsuperscript{57,63,100} although in one of these studies, that of Kreth et al.,\textsuperscript{63} only patients that completed the radiotherapeutical protocol were included.

In the study of Kelly and Hunt,\textsuperscript{55} significant differ-

### Table 2: Patient characteristics by study\textsuperscript{*}

<table>
<thead>
<tr>
<th>Study</th>
<th>Inclusion Criteria</th>
<th>Age—Mean (range, ctrl meth), Strat</th>
<th>KPS Score—Mean (range, ctrl meth), Strat</th>
<th>Initial Sx &amp; No. of Pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simpson et al., 1993</td>
<td>NR</td>
<td>NR (post hoc eval), ≤39 yrs 76,\textsuperscript{†}</td>
<td>NR (post hoc eval), 80–100 291,\textsuperscript{†}</td>
<td>NR</td>
</tr>
<tr>
<td>Kelly &amp; Hunt, 1994</td>
<td>NR</td>
<td>71.14 yrs (65–63 yrs, post hoc eval), NR</td>
<td>84.2 (60–100, post hoc eval), NR</td>
<td>epileptic crises 36, incr ICP 6, neural deficit 86</td>
</tr>
<tr>
<td>Kiwit et al., 1996</td>
<td>NR</td>
<td>59 yrs (matching)</td>
<td>55 (matching)</td>
<td>NR</td>
</tr>
<tr>
<td>Kreth et al., 1999</td>
<td>no Hx of low-grade glioma, tumor recur, deep loc, chemo, implants, or absent/incompl RT</td>
<td>NR (NR, post hoc eval), NR</td>
<td>NR (NR, post hoc eval), NR</td>
<td>epileptic crises 63, other 162\textsuperscript{†}</td>
</tr>
<tr>
<td>Vuorinen et al., 2003</td>
<td>radiol diagnosed high-grade glioma, KPS score ≥60, age &gt;65 yrs, consent</td>
<td>71.48 ± 4.1 yrs (66–80 yrs, post hoc eval)</td>
<td>73.48 ± 11.1 yrs (60–90, post hoc eval)</td>
<td>epileptic crises 1, incr ICP 5, neural deficit 18</td>
</tr>
</tbody>
</table>

\* chemo = chemotherapy; ctrl = control; eval = evaluation; Hx = history; incompl = incomplete; incr = increased; loc = location; meth = method; radiol = radiologically; recur = recurrence; Strat = stratification.

\† Calculated based on data from the article.

### Table 3: Tumor characteristics by study\textsuperscript{*}

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Histological Characteristics</th>
<th>Location</th>
<th>Side</th>
<th>Preop Tumor Size</th>
<th>Midline Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simpson et al., 1993</td>
<td>IV: 645 (NR)</td>
<td>NR</td>
<td>fr 280, par 161, temp 178, occ 26</td>
<td>&lt;5 cm 245, 5–10 cm 359, ≥10 cm 41</td>
<td>NR</td>
</tr>
<tr>
<td>Kiwit et al., 1996</td>
<td>III: 26, IV: 54 (WHO)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Kreth et al., 1999</td>
<td>IV: 228 (WHO)</td>
<td>NR</td>
<td>57, other lobe 168</td>
<td>rt 112, lt 113</td>
<td>NR 77</td>
</tr>
<tr>
<td>Vuorinen et al., 2003</td>
<td>III: 5, IV: 18 (WHO)</td>
<td>NR</td>
<td>fr 7, par 1, temp 7, occ 2, fr-par 1, fr-temp 1, temp-par 1, par-occ 2, temp-occ 1</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

\* Anat = anatomical; CC = corpus callosum; fr = frontal; par = parietal; fr-par = frontoparietal; fr-temp = frontotemporal; histol = histological; occ = occipital; par = parietal; par-occ = parietooccipital; temp = temporal; temp-occ = temporooccipital; temp-par = temporoparietal.

\† Calculated based on data from the article.
Biopsy versus resection in the management of malignant gliomas

**TABLE 4: Operation characteristics by study***

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Allocation Criteria</th>
<th>Bx Method, Allocated No. of Pts (postop morb), (postop mort)</th>
<th>Resect Method, Allocated No. of Pts (postop morb), (postop mort)</th>
<th>Degree of Resect (est meth, time of est)</th>
<th>No. of Surgeons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simpson et al., 1993</td>
<td>NR</td>
<td>NR, 107 (NR), (NR)</td>
<td>NR, 538† (NR), (NR)</td>
<td>partial resect 413, total resect 125 (NR, NR)</td>
<td>many†</td>
</tr>
<tr>
<td>Kelly &amp; Hunt, 1994</td>
<td>a) enhancing mass on images ≥1/3 of tumor, b) midline shift ≥15 mm, c) loc not in midline, d) pt consent</td>
<td>stereo Bx, 88 (hematoma 3), (hematoma 2, tumor progr 2)</td>
<td>stereo vol resect, 40 (perm neural deter 2), (pulm embolism 1)</td>
<td>total resect 40 (CT, NR)</td>
<td>1‡</td>
</tr>
<tr>
<td>Kiwit et al., 1996</td>
<td>informed pt’s choice</td>
<td>CT-guided stereo Bx in series, 40 (NR), (NR)</td>
<td>craniot &amp; microsurg resect or conv macrosurg excision, 40 (NR), (NR)</td>
<td>(surgeon’s eval &amp; CT, 1–3rd postop day)</td>
<td>many</td>
</tr>
<tr>
<td>Kreth et al., 1999</td>
<td>NR</td>
<td>CT-guided stereo Bx, 100 (NR 1†), (NR 1)</td>
<td>microsurg as rad as poss resec, 128 (NR 7†), (NR 2)</td>
<td>(surgeon’s eval, NA) difft grps in the 2 ops</td>
<td></td>
</tr>
<tr>
<td>Vuorinen et al., 2003</td>
<td>randomization (closed env method)‡</td>
<td>stereo Bx, 16 (0), (NR)</td>
<td>open tumor resect w/ craniot, 14 (hematoma 1), (NR)</td>
<td>(CT/MRI, 1–3rd postop day)</td>
<td>1</td>
</tr>
</tbody>
</table>

* Bx = biopsy; conv = conventional; craniot = craniotomy; difft = different; env = envelope; est = estimation; grp = group; macrosurg = macrosurgical; microsur = microsurgical; morb = morbidity; mort = mortality; perm = permanent; poss = possible; progr = progression; pulm = pulmonary; rad = radical; resect = resection; stereo = stereotactic; vol = volumetric.  
† Calculated based on data from the article.  
‡ From communication with researchers.

ences in adjuvant therapy were observed. Furthermore, in the study of Simpson et al.,91 there was no control for differences in the adjuvant treatment, while, in relation to the outcomes, only the Kaplan-Meier curves were available. For this reason, a sensitivity analysis was performed with the 2 above studies excluded to check the robustness of the results obtained. The study of Kreth et al.63 clarified that no patient was subjected to other forms of treatment, while in the other studies the chemotherapy protocols and numbers of patients participating in them were not reported (Table 5).

**TABLE 5: Adjuvant therapy in the biopsy/resection groups***

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Total Dose (protocol)</th>
<th>No. of Pts Who Participated</th>
<th>No. of Pts Who Completed RT</th>
<th>Waiting Time</th>
<th>Total Dose Received</th>
<th>Chemo Agent (no. of pts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simpson et al., 1993</td>
<td>60 Gy, 64.8 Gy, 72 Gy, 76.8 Gy, 81.6 Gy.</td>
<td>all rec’d RT (107/538†), NR for each protocol separately</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>none (NR), BCNU (NR), MeCCNU + DTIC (NR)</td>
</tr>
<tr>
<td>Kelly &amp; Hunt, 1994</td>
<td>NR</td>
<td>105 (71†/34‡)</td>
<td>96 (62/34†)‡</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Kiwit et al., 1996</td>
<td>60 Gy (40 Gy WBRT + 20 Gy)</td>
<td>59† (29/30)</td>
<td>NR</td>
<td>≥2 wks</td>
<td>38 Gy/42 Gy</td>
<td>NR</td>
</tr>
<tr>
<td>Kreth et al., 1999</td>
<td>60 Gy (2 Gy/day × 5 days/ wk × 6 wks)</td>
<td>228 (100/128)</td>
<td>228 (100/128)</td>
<td>1 wk/3 wks</td>
<td>60 Gy/60 Gy</td>
<td>none</td>
</tr>
<tr>
<td>Vuorinen et al., 2003</td>
<td>60 Gy</td>
<td>19 (10/9)</td>
<td>8 (4/4)</td>
<td>median time 25 ± 13 days/42 ± 24 days</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

* Numbers of patients are given as biopsy group/resection group. Abbreviations: rec’d = received; WBRT = whole-body RT.  
† Calculated based on data from the article.  
‡ p < 0.05.
no reason to consider that there was significant deviation in the measurements. Overall survival was the only outcome studied by all the researchers. A positive association between resection of the tumor and survival was observed in the majority of the studies (4 of 5), whereas Kreth et al.63 did not observe a statistically significant association (Table 6). Quantitative expression of QOL was explored in 3 studies. Kiwit et al.57 recorded KPS scores on discharge, which differed significantly in favor of the resection group, and Kreth et al.63 reported the direction of change of KPS scores (relative to preoperative values) at 2 time points, after the operation and after radiotherapy. Although there was a significant difference in favor of the resection group after surgery, after completion of radiotherapy the changes in KPS were similar in the 2 groups. Moreover, Vuorinen et al.100 recorded discharge KPS scores, although they did not present summary values for the 2 groups. Further analysis of the individual patient data reported did not detect a statistically significant difference in the median values and distribution of postoperative KPS scores between the 2 groups (p = 0.15, Mann-Whitney U test; p = 0.11, Kolmogorov-Smirnov Z test) or in the median values and distribution of the change of KPS scores in relation to the preoperative one (p = 0.69, Mann-Whitney U test; p = 0.81, Kolmogorov-Smirnov Z test) (Table 6).

![Fig. 2. Forest plot of quantitative synthesis of HRs with all the studies that were included in the meta-analysis. The position of the rectangles corresponds to the value of the effect measure for each study, their size to its respective weight, and the horizontal lines correspond to the 95% CI. The position of the diamond (dashed line) corresponds to the value of the resultant effect measure and its left and right angles to its 95% CI. An HR < 1 favors resection. An HR > 1 favors biopsy (inverse variance fixed-effect model).](image-url)
Biopsy versus resection in the management of malignant gliomas

Meta-Analysis

Overall Survival. The HR for overall survival between the biopsy and resection groups was significantly reduced in favor of the resection group (1111 patients, HR 0.61, 95% CI 0.52–0.71, p < 0.0001, inverse variance fixed-effect model) (Fig. 2), while statistically significant heterogeneity was not identified (Cochran Q statistic 7.32, p = 0.12, $\tau^2 = 45.4\%$).

No publication bias was traced (Egger regression, intercept $-1.42$, 95% CI $-7.14$ to $4.30$, p = 0.49) (Fig. 3), while the estimation of the influence of probable publication bias in the results with the trim-and-fill method (Fig. 4) did not note significant alteration (HR 0.62, 95% CI 0.53–0.72).

Furthermore, the repetition of the quantitative synthesis, after exclusion of the studies in which the influence of performance bias was dubious or the outcome data were reported only via Kaplan-Meier curves, did not result in substantial alteration of the results (HR 0.72, 95% CI 0.58–0.90, p = 0.003, inverse variance fixed-effect model) (Fig. 5A). Subgroup analysis based on the grade of the gliomas did not show a significant difference ($Q_{int} = 0.10$, p = 0.75) in the effectiveness of the resection between the studies of patients with Grade IV glioma (3 studies, heterogeneity $Q = 6.15$, p = 0.046, $I^2 = 67.5\%$ and HR 0.61, 95% CI 0.45–0.82, p = 0.001, inverse variance random effects model) (Fig. 5B) and those of patients with Grade III or IV glioma (2 studies, HR 0.57, 95% CI 0.39–0.85, p = 0.006, inverse variance fixed-effect model, heterogeneity $Q = 1.07$, p = 0.3, $I^2 = 6.9\%$) (Fig. 5C).

Subgroup analysis based on the age of the patients did not trace statistically significant differentiation ($Q_{int} = 1.29$, p = 0.26) of the result of the resection between the studies that focused on elderly participants ($\geq 65$ years) (2 studies, HR 0.50, 95% CI 0.35–0.71, p < 0.0001, inverse variance fixed-effect model, heterogeneity $Q = 0.5$, p = 0.48, $I^2 = 0$) (Fig. 5D) and the studies that included patients with a wide age range (3 studies, heterogeneity $Q = 5.35$, p = 0.07, $I^2 = 62.6\%$ and HR 0.64, 95% CI 0.48–0.85, p = 0.0025, inverse variance random-effects model) (Fig. 5E).

Progression-Free Survival

The only study of the systematic review that recorded progression-free survival was that of Vuorinen et al., and thus meta-analysis was not feasible with respect to this measure. In that study, duration of progression-free survival was not significantly different between the 2 groups compared.

Quality of Life

Although 3 studies provided data on QOL of the patients with high-grade glioma based on KPS scores, quantitative synthesis of effect measures could not be performed in a meaningful way, as Kiwit et al. did not provide a measure of variation for the postoperative KPS scores in 2 groups and Kreth et al. provided only the direction of the postoperative and postradiation change in KPS scores. On the other hand, only Vuorinen et al. provided enough data to extract an effect measure. It should be noted, though, that 2 of the 3 studies provided evidence of a beneficial effect of resection as compared with biopsy on QOL.

Discussion

The results of our systematic review suggest that cytoreductive resection instead of biopsy of high-grade gliomas is associated with prolongation of overall survival of the patients. This result persisted in all the sensitivity analyses performed. Furthermore, we did not note any significant alteration of the association during subgroup
analysis based in the range of the ages of the patients or the grades of the gliomas included in the studies.

In particular, no substantial alteration of the results was found when the quantitative synthesis was repeated without the 2 studies in which the methodological quality was considered marginal and/or the only outcome data were in the Kaplan-Meier curves. Inverse variance fixed-effect model (A, C, and E), and inverse variance random effects model (B and D).

**FIG. 5.** Forest plots of quantitative synthesis of HR with the exclusion of the studies in which the influence of performance bias was dubious and/or the only outcome data were the Kaplan-Meier curves (A), and of HRs of studies with Grade 4/IV gliomas (B), Grade III and IV gliomas (inverse variance fixed effect model) (C), elderly patients (D), and wide range of ages (E). Inverse variance fixed-effect model (A, C, and E), and inverse variance random effects model (B and D).

Furthermore, the efficacy of the various treatment modalities, including surgery, in elderly patients with high-grade glioma still remains unclear. In the present systematic review, a subgroup analysis based on the age of the patients suggested that an improved clinical outcome with resection as compared with biopsy is also present in the elderly.

Overall survival is the most frequently recorded outcome in the clinical studies of high-grade gliomas and the most accurately measured. On the other hand, the measurement of progression-free survival depends heavily on the definition of progression and the means by which it is measured. Practically, of course, progression-free survival is an important outcome measure, but the only study reporting this outcome measure did not show a significant difference between the 2 groups compared. It should be noted, however, that this absence of effect might be attributed to the small number of patients included in that study (27 patients). Moreover, the authors reported that, when adjustment for grade was applied, the HR was significantly different in favor of the resection as compared with the biopsy group. Nevertheless, it should be emphasized that this difference was only marginally significant (HR = 2.757, 95% CI 1.004–7.568, p = 0.0491), and for this reason the finding should be interpreted with caution.

Similarly important is the measurement of QOL; many scales, including the KPS, have been proposed for the quantitative measurement of this variable. In the set of studies included in this systematic review, there was incompatibility of the measures of QOL. As a consequence, no meta-analysis could be performed, although 2 of the 3 studies that recorded a measure of this variable provided evidence of a beneficial effect of resection as compared with biopsy on QOL.

The design of the systematic review had to take into consideration parameters that influence prognosis or selection of treatment and should be controlled. Age and preoperative performance status, usually expressed by KPS score, are significant prognostic factors known preoperatively and considered to influence the selection of the type of surgical treatment. Other prognostic parameters, such as the histological type, grade, and genetic profile of the tumor are unknown until surgery and can be controlled only by randomization. Mass effects and especially raised ICP, when confronted, render resection of the tumor the only valid choice. The proximity of the tumor to an eloquent area of the brain limits the percentage of total resection, although extended resection may still be possible. Among types of adjuvant treatment, radiotherapy is a significant prognostic factor for overall survival and should be controlled to study the performance of surgery, and chemotherapy bears a slight advantage for overall survival.

The studies included in the present systematic review make up the most trustworthy available evidence with re-
Biopsy versus resection in the management of malignant gliomas

spect to the question of whether biopsy or resection of the supratentorial malignant gliomas is preferable in adults. Although the intention of many researchers was to study the relationship between the extent of resection and the outcome of surgery, extent of resection can be known fully only after an operation and it may vary substantially. In consequence, the aim of the present meta-analysis was not to examine the relationship between the extent of resection and surgical outcome. Instead, our aim in this meta-analysis was to provide information to help in making the choice between biopsy and resection in planning an operation.

In the present systematic review and meta-analysis, firm inclusion criteria were adopted in order for solid conclusions to be drawn. However, most of the studies included were observational and were characterized by certain methodological flaws and sources of potential systematic error that should be noted. Specifically, with regard to the inclusion and exclusion criteria, the majority of the studies did not take into account the existence of mass effects and the location of the tumor in relation to critical areas, factors that limit the choice of operation. In particular, we observed that the 2 studies that included analysis of patients’ initial symptoms did not exclude patients with raised ICP, while at least 2 studies did not exclude patients with midline shift. On the other hand, the study of Kelly and Hunt included patients with brainstem tumor location in relatively low numbers in relation to the total. No study included information about the relation of the tumors to eloquent areas of the brain.

Despite the effort of the authors of the eligible studies, the possibility of selection bias (systematic differences in comparison groups) cannot be fully excluded. Because of the lack of a sufficient number of randomized clinical trials that study the question of performing biopsy or resection in adult patients with supratentorial high-grade glioma, it was decided a priori that nonrandomized studies would also be included in the systematic review, to explore the best available evidence on the subject. One study applied stratification and another matching, but they were both retrospective and thus this methodology might not be adequate.

In relation to performance bias (systematic differences in care provided apart from the intervention being evaluated), it should be taken into account that the studies are surgical series, a characteristic of which is that there is no concealment of allocation from the moment of the operation on. Furthermore, in some studies the operations were performed by many surgeons, while in 1 study the 2 different types of operation were performed by 2 different surgical groups. However, the repetition of the quantitative synthesis without the 2 studies in which the influence of performance bias was dubious did not result in substantial alteration of the results.

Although no evidence of a differential effect of modern chemotherapeutic agents, such as temozolomide and bevacizumab, on the overall survival of patients after resection or biopsy has been presented in the literature, such a scenario cannot be excluded. This hypothesis could not be examined in the current review and meta-analysis due to the fact that these agents were not used in any of the eligible studies.

With regard to attrition bias (systematic differences in withdrawals from the trial), the only prospective study did not apply an intention-to-treat analysis, but an on-treatment analysis instead. In that study, however, the exclusion of the cases that did not have the histological diagnosis of high-grade glioma reflects the real-time conditions in which the clinical trial was conducted and one of the expected limitations of this kind of study.

Furthermore, in relation to detection bias (systematic differences in outcome assessment), some studies did not provide data about postoperative morbidity and mortality. Accurate data regarding cause of death were not reported in any study, and comorbidity can play a significant role in mortality—although in one study an effort was made to control for comorbidities.

The proportional hazards assumption is included among the limitations of the present study, but it is necessary in survival analysis. It should also be noted that, as independent patient data or estimations of log-hazard ratio and its variance were not available for the majority of the studies, the effect measures were calculated indirectly.

An individual patient data meta-analysis could probably address to some extent many of the limitations of this meta-analysis. However, such an approach was not planned a priori and it is doubtful that it would be feasible given the retrospective nature of most of the eligible studies.

A well-designed prospective randomized trial comparing biopsy versus resection might be able to answer the research question of interest with more confidence. It has been proposed that, assuming a median survival of 100 days in the biopsy arm, a sample size of 500 patients (250 in each arm) would be required to detect a 30-day increase in the patient survival in the resection arm with a power of 80%, an accrual period of 2 years, and a follow-up period of 2 years. These numbers are not likely to be feasible for a single-center study, and thus a coordinated multicenter study is probably needed. Such a trial might also allow the assessment of the relative value of resection versus biopsy in combination with modern chemotherapeutic agents available for the treatment of malignant gliomas, such as temozolomide and bevacizumab.

Conclusions

In conclusion, based on the best available evidence, it seems that cytoreductive resection in adults with supratentorial malignant glioma is associated with improved overall survival as compared with biopsy. However, well-designed prospective studies assessing the comparative efficacy of resection versus biopsy in these patients in terms of overall and progression-free survival as well as QOL are warranted. Data originating from such studies will allow conclusions to be drawn with greater confidence and could facilitate the evidence-based provision of the optimal treatment to these patients.
Disclaimer

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

References

40. Green SB, Byar DP, Walker MD, Pistenmaa DA, Alexander E Jr, Batzdorf U, et al: Comparisons of carmustine, procarba -zine, and high-dose methylprednisolone as additions to sur-

Please include this information when citing this paper: published online September 11, 2009; DOI: 10.3171/2009.7.JNS09758.