Abstract. Large phase III studies have identified limited survival benefits with chemotherapy in high-grade glioma. However, numerous clinical trials have been published previously with smaller patient numbers and no control groups. A small positive effect could be missed this way, resulting in premature rejection of possible beneficial treatment. In order to perform a treatment arm summarizing analysis (TASA), a database was created summarizing treatments published between January 1976 and June 2002. In this database, one record represents a cohort of patients treated in the same way. Various patient cohort characteristics such as median age, and outcome measures including median overall survival times (mOS), were documented. Two-hundred and twenty publications were documented with a total number of 17213 cases treated in 337 treatment groups. There was a statistically significant relationship between the distribution of histological grades (p<0.001) and the outcome, and a better outcome in younger patient populations (p<0.01). However, the known influence of the median Karnowsky performance scale could not be confirmed in this database. The extent of surgery showed a positive influence only when excluding relapse studies, while the positive effect of radiation was clear in all subgroups (p<0.05). Clinical studies that included nitrosourea in the treatment had a significantly better outcome than those with platinum drugs or without chemotherapy. We conclude that TASA, representing a novel way to perform a meta-analysis, is valid since it confirms the known treatment effects and, therefore, has the potential to provide new insights by combining information from different clinical treatment studies.

Correspondence to: Johannes Wolff, MD, Dept. Pediatrics, MD Anderson Cancer Center, University of Texas, 1515 Holcombe Blvd, Unit 87, Houston, TX 77030, U.S.A. Tel: (+1) 713 792 3366, Fax: (+1) 713 794 5042, e-mail: jwolff@mdanderson.org

Key Words: Malignant glioma, nitrosourea.
nitrosurea, the overall gain in life expectancy is small. This poses the question that maybe another small benefit has been overlooked among all the other trials with drugs less extensively studied, which could add further benefit to survival. The major problems with most of these studies were small patient numbers and often the lack of control groups. In addition, many of the previous studies had not been as stringent with respect to patient eligibility criteria, often including various histological diagnoses (11, 12), age groups (13) and Karnofsky Performance Score (KPS) (14). Finally, the end-points differed, including overall survival, event-free survival and response, with hardly any study reporting all of them. This makes comparisons between different studies difficult and small benefits are easily overlooked.

Thus, we conducted a literature study documenting the characteristics of the study populations (patient cohorts), the treatment characteristics and the reported outcome parameters from published clinical trials in a database, and analyzed how the characteristics of the study population influenced the outcome. The goal was to take this influence into consideration when comparing in the outcomes in the search for possible beneficial treatments. In this database, one record represents one treatment arm of any given study. A typical phase II study will be represented by one record; a typical randomized phase III study will be represented by two or more records. We named this approach a "Treatment Arms Summarizing Analysis" (TASA). This approach gives away the advantage of control groups within the same studies, but it thereby allows the combining of information available from phase II and phase III studies. We hypothesized that improved survival by nitrosurea would be confirmed by this model, even when the data of those randomized trials were excluded from the analysis.

Materials and Methods

The first step of this analysis was a literature search using the Medline and Cancerlit databases, including publications from January 1976 up to June 2002. The following languages were included: English, French, German, Italian and Spanish. The search was conducted exploring search terms "glioma" and "brain neoplasms". Subheadings applied to these terms included "clinical trial", "phase I trial", "phase II trial", "phase III trial", "therapy", "drug therapy", "radiation therapy" and "surgery". These searches were extended by hand searching the reference lists of identified trials and bibliographies of relevant books and review articles. Three different reviewers screened all titles identified by the first searches. Abstracts were downloaded for all titles of potential relevance and/or child patients with primary intracranial AA or GBM; or ii) phase II studies enrolling adult and/or child patients with primary or relapsed intracranial AA or GBM; or iii) phase III studies enrolling adult and/or child patients with primary intracranial AA or GBM.

The second step was to create a database. We entered both phase II and phase III trials in our study. If a randomization was done, each arm was entered as one data record. We used one record for each treatment group. If a study used a historical control group, we deleted this group from our database if it contained redundant information or if this was not clear. One of the differences in a typical meta-analysis is that one observation unit is not the individual patient but the study itself.

The first set of columns described the reference i.e. author, year of publication, identifying number, while the second set of columns described the patient population i.e. total number of patients, total number of confirmed histological diagnoses, number of patients with AA, GBM or brainstem glioma, location of the tumors (supratentorial, infratentorial, pons, brainstem, spinal), number of patients with and without recurrence at the start of the study, median Karnowsky index of the patient population, mean and median age of the cohorts, medium size of the tumor, proportion of male patients, previous chemotherapy, previous radiation. A third set of columns described the treatment, number of gross total resections, radiation part of the study, dose and techniques of the radiotherapy, and portion of patients with chemotherapy in this study. We documented some of the cytotoxic agents with cumulative dosages. The last set of columns contained information about the outcome such as number of toxic deaths, time of follow-up, median overall survival (mOS), 1-, 2-, 5-year overall survival, and 1-, 2-, 5-year progression-free survival and response. Overall survival was defined as the time from randomization or diagnosis until death (by any cause). Progression-free survival was defined as the time from randomization or diagnosis until progression or death (by any cause), whichever occurred first. Not all of these data fields could be filled with sufficient information for a meaningful analysis.

The database was analyzed using "SPSS for Windows 9.0" (SPSS Inc., Chicago, USA). First, we used descriptive statistics to summarize and describe the data. This included the analysis, which influencing parameter had a statistically detectable influence on the outcome. Not all outcome parameters were equally frequently reported. We used linear regression and logarithmic regression analysis to describe the connections between different outcome parameters. In those analyses, the data were weighted by the square route of the patient numbers reported in each study. To address the hypothesis, the most frequently reported end-point was picked and the missing data filled in using the weighted multiple linear regression model and the information of other end-point parameters of each particular patient cohort. With this end-point, the study outcomes were compared using the Mann-Whitney Wilcoxon Test. The p-values shown are calculated on the two-tailed level. All, except for those that addressed the hypothesis, were considered observational descriptions.

Results

Patient characteristics. The first searches identified 256 articles that contained treatment strategies for HGG. Thirty-six were found to be ineligible (cell culture and animal experiments). Two-hundred and twenty publications were entered in our database, which contained 337 data records with a total number of 17,213 cases. The mean age was 65 years (±13.8 SD, n=296), the male: female ratio was 1.4:1 (n=228). Sixty-seven % of all studies contained adult patients only, 8% were children, 17% contained both adults and children and 6% of the articles did not contain information about the patients’ age. Two-hundred and twenty-five data records contained patients with exclusively primary disease (excluding
recurrences), whereas 101 records contained exclusively relapse protocols, and 11 records permitted both. Sixty-four % were patients with a GBM (±33.5 SD, n=337) and 27% with AA (±28.8 SD, n=337), whereas 3.4% were patients with brain stem glioma (BSG) (±16.4 SD, n=337). 2.4% of the study populations had some other primary brain tumors e.g. oligodendroglioma (n=337) and in 3.3% the diagnosis was not confirmed histologically (±19.8 SD, n=337). 31.4% of the study population had recurrent disease (±45.7 SD, n=337). The mean of the median Karnofsky performance scores (KPS) was 63.2 (±15.8, n=144).

Treatment characteristics. In 77.1% surgery was a part of the treatment in our database (n=337). If the patient were treated exclusively in first-line, this was 94.2 % (n=225), but in relapse studies it was only 39.6% (n=96). The rate of gross total resection (GTR) in first-line regimens was 22.1% (±22.8, n=156) and 5.8% (±17.5, n=83) on relapse.

Radiotherapy was a part of first-line treatment in 92.0% (n=225), but only in 14.9% (n=101) in relapse studies. In 90.8% of all studies, the radiotherapy was administered in a conventional form (n=337) with a medium dose of 58.9 Gy (±15.7). A seed was implanted in 9.5%.

Chemotherapy was used in primary treatment in 69.3% (n=225), while it was used in 74.3% (n=101) in relapse treatments. Nitrosoureas (e.g. CCNU, BCNU or others) were used in 72.8% as part of the regimens. In studies with recurrent HGG, CCNU was administered in 42.7%, platinum derivatives in 13.3%, temozolamide in 10.7% and procarbazine alone in 8%. The medium dose of nitrosoureas was 205.2 mg/m²/cycle (±291.6SD, n=302) with an average of 4.9 cycles (±2.5 SD). Fourteen out of 17,213 patients were reported to have died from toxic complications of the chemotherapy. Six of those individuals had received 1200 mg BCNU with subsequent bone marrow transplantation or autologous stem cell rescue (n of this study = 27). Three patients expired after the therapy with procarbazine (100 mg/m²/day) and tamoxifen (100 mg/day) for 30 days (n of this study = 53).

Outcome data. The median overall survival (mOS) was the most frequently documented outcome variable in 280 records. Overall survival rates were documented in 253, 194 and 53 cases for 1-year (1 YOS), - 2-year (2 YOS) and 5-year overall survival (5 YOS), respectively. Progression-free survival rates were documented in 42, 30, 18 and 83 cases, for 1-year (1 YPFS), 2-year (2YPFS), 5-year (5 YPFS) and median progression-free survival (MPFS), respectively.

When addressing the hypothesis, we had to decide which end-point should be used. If we had worked with mOS only, we would have lost 57 studies, or 15% of the available data. Therefore, various methods of filling in the missing data were tested. For this purpose, the sources of the information used should be available in as many study reports as possible. Of the 57 missing data for mOS, data were available as shown in Table I. The best available was 1 YOS.

Second, we analyzed the possible relationship between mOS and the other outcome parameters. By using scatter-plot, the relationship could be described in a linear or logarithmic mathematical form (as shown for mOS and 1YOS in Figure 1). One more sophisticated level is the assumption that the survival curve of a cohort of patients is logarithmic:

\[ S(t) = e^{-\lambda t} \]

with \( S(t) \) = survival fraction, and \( \lambda \) = survival coefficient (years⁻¹).

With this assumption, the median overall survival (mOS) and the 1-year overall survival rate are connected as:

\[ mOS = (-\ln 2)/ \ln (S(1)) \]

The accuracy of the assumption (1) can be compared to the linear regression by using those records, which contained both data. The better the assumption, the higher should be the correlation between real mOS and mOS calculated based upon the various assumptions. (3) and (4) show the assumption for 2 YOS and 5 (YOS), respectively.

\[ mOS = -2* \ln 2/ \ln (S(2)) \]

with: \( S(2) = 2 \) year overall survival fraction

\[ mOS = -5 * \ln 2/ \ln (S(5)) \]

with: \( S(5) = 5 \) year overall survival fraction

The next step was to evaluate a possible relationship between the mOS and the progression-free survival data.

\[ mOS = MPFS + TOD \]

with MPFS=median progression-free survival and TOD=time between tumor progression and death at 50% survival rate of the study population.

To calculate the mean of TOD weighted by n of the study population we assumed:

\[ wTOD \text{ mean} = \frac{\sum (TOD * \sqrt{n(study)})}{\sqrt{n(all)}} \]

Using our data this resulted in: \( wTOD \text{ mean} = 5.82 \)
To calculate the predicted mOS from the 1 YPFS (month) we defined:

\[ \text{calmOS (1 YPFS)} = \text{wTOD mean} + \frac{-\ln 2}{\ln (1 \text{ YPFS})/100} \times 12 \]

with wTOD mean = mean of TOD weighted by n, 
1 YPFS = 1 year progression-free survival (%).

Using equivalent assumptions, we also calculated mOS using 2- and 5 YPFS.

In order to decide whether the calculated mOS is more accurate, we subtracted the calculated mOS from the original mOS and determined the mean and standard deviation (SD) from the differences, as shown in Table II. These relationships between the different end-points were used to calculate the estimated median overall survival for those studies, which did not report the median overall survival but reported the other end-points.

The Karnofsky Performance Score was only quantitatively documented in 34 papers. Mostly, the KPS was an inclusion criteria and had to be over 60 for patient registration. A statistically significant relationship between these numbers could not be established in our database. The extent of surgery showed a small positive influence on mOS (mOS = 0.1 x GTR + 14.0, CI 95%, \( p < 0.05 \), weighted by square of patient number, \( n = 156 \)). If radiotherapy (RT) was used, the mOS was significantly better regardless of which way the analysis was performed. The mean mOS without RT was (±11.5 months, \( n = 87 \)) and with radiation the mean was 15.0 months (±2.2 SD, \( n = 199 \), \( p < 0.05 \)).

Finally we addressed the hypothesis by combining the data of those studies which had included nitrosurea treatments (Figure 4). If a nitrosurea-based chemotherapy was used, in first-line treatment, the mean of mOS was 18 months (±3.0 SD, \( n = 105 \)) compared to only 11 months without chemotherapy (±4.2 SD, \( n = 53 \), \( p < 0.001 \)). Both groups had been treated with radiotherapy in similar frequency, and with no significant dose difference. Since we knew, prior to this analysis, that nitrosurea can improve the outcome, this analysis had some features of a self-fulfilling analysis. Therefore, we repeated the analysis after deleting the studies that were the source of that knowledge: after we had deleted all the studies that had clearly shown that nitrosurea had improved the mOS.

Influence of patient populations on outcome data. The more anaplastic astrocytoma (AA per) were entered in a study, the better the mOS was (Figure 2, mOS = 0.53 x AA per + 8.01, CI 95%, \( p < 0.001 \), weighted by square of patient number, \( n = 134 \)). The more glioblastoma (GBM) were in the study, the smaller the mOS was (mOS = -0.13 x GBM per + 24.51, CI 95%, \( p < 0.001 \), weighted by square of patient number, \( n = 213 \)). The older the patient population, the lower the mOS was (Figure 3, mOS = -0.26 x mean age + 27.5, CI 95%, \( p < 0.01 \), weighted by square of study population, \( n = 178 \)).
such as phase III studies addressing this very question), the advantage of these drugs still remained true. If nitrosourea was used the mean of mOS was 18.2 months (±2.2 SD, CI 95%, p<0.001, n=101), while without chemotherapy the mOS was 11.5 months (±1.1 SD, CI 95%, n=51). This analysis confirmed the hypothesis, showing that TASA is a valid possible way to summarize various publications.

Discussion

This analysis summarized 17,213 patients in 220 published articles. To our knowledge, this is the largest analysis of this kind for high-grade glioma yet published. This set of data allowed a novel way to analyze risk factors, confirming some of the known risk factors, but unable to confirm others. Those that could be confirmed included the histological grading (poorer outcome with high-grade tumors, (12) Figure 2) and the patients’ age (better outcome with children, Figure 3 (11)). On the contrary, a lower KPS is a well-accepted strong risk factor for a reduced mOS, but we were not able to confirm this in this database. This might be caused by a lack of variability in the median KPS between different treatment cohorts. Most studies report a median KPS of 80, making a correlation analysis between outcome and KPS difficult.

In analyzing treatment modalities, the dataset showed that radiation therapy improved the outcome of patients with HGG, confirming a generally accepted finding (3-5) and, therefore, confirming the validity of the method. The role of a radical neurosurgical resection of HGG is still somewhat controversial. This analysis confirmed the relevance of resection (Figure 4): if a GTR was done, the time until tumor recurrence was longer and recent studies showed advantages in survival for adult and pediatric patients (15-18). However, high-level evidence to confirm this impression is hard to produce, because a randomization between surgical resection and no surgical resection is not feasible at present. We could only confirm the advantage on a low significance level.

Recent studies and a meta-analysis showed a small, but significant, benefit for patients treated with nitrosourea (7, 9, 10, 14). Our analysis confirmed this (Figure 4). Even when we deleted the phase III studies, which had been specifically designed to address the question and, therefore, had a higher level of evidence, the remaining lower level of evidence data were still sufficient to confirm that nitrosourea leads to a better median overall survival.

Factors influencing the outcome, such as the median age of the study population, differed among studies. For the nitrosourea analysis, this was not important, because there was no difference between the non- and the treated group (e.g. mean age, histology, part of radiation therapy). However, this was only true for this question. Most of the other treatments were
reported less frequently, making the variability between different studies more important. Various mathematical models may be used to control for this variability. Since this is beyond the scope of this publication, they will be reported later.

We conclude that SADA may be a novel method of analyzing the existing large body of literature describing studies with HGG patients. Radiation, GTR and nitrosourea-based chemotherapy could improve the outcome of a study group. We will expand this effort by adding more literature to the database and developing a mathematical model to normalize the outcome, utilizing the information provided by the published description of the patient populations.

Acknowledgements

We thank Ove Peters (Regensburg), Claudia Spix (Mainz), Mark Munsell (Houston) and Dolly Anguilera (Houston) for their helpful discussions, and Barbara Whitehead for her editorial support.

References


Received June 13, 2005
Accepted June 22, 2005

Table II. Differences between mOS and the calculated mOS (mean±SD).

<table>
<thead>
<tr>
<th>mOS cal from:</th>
<th>1 YOS</th>
<th>2 YOS</th>
<th>5 YOS</th>
<th>MPFS</th>
<th>1 PFS</th>
<th>2 PFS</th>
<th>5 PFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>n:</td>
<td>224</td>
<td>166</td>
<td>39</td>
<td>74</td>
<td>30</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>mOS-mOS (cal) ±SD (logarithmic)</td>
<td>–2.05±15.53</td>
<td>1.78±6.43</td>
<td>–6.89±7.26</td>
<td>0.13±7.15</td>
<td>10.48±7.94</td>
<td>9.82±7.35</td>
<td>10.3±7.48</td>
</tr>
<tr>
<td>mOS-mOS (cal) ±SD (linear)</td>
<td>–15.57±10.26</td>
<td>–6.59±8.83</td>
<td>0.79±8.50</td>
<td>–2.33±13.73</td>
<td>10.37±8.56</td>
<td>9.35±8.98</td>
<td>–5.02±6.43</td>
</tr>
</tbody>
</table>