PERGAMON

European Journal of Cancer 36 (2000) 1788–1795

European Journal of Cancer

www.ejconline.com

Health-related quality of life in patients with anaplastic astrocytoma during treatment with temozolomide

D. Osoba [a,*], M. Brada [b], W.K.A. Yung [c], M.D. Prados [d]

[a] *QOL Consulting, 4939 Edenvale Court, West Vancouver, BC, Canada V7W 3H7*
[b] *Institute of Cancer Research and Royal Marsden NHS Trust, Sutton, Surrey, UK*
[c] *Neuro-Oncology, MD Anderson Cancer Centre, Houston, TX, USA*
[d] *University of California, M787 Department of Neurosurgery, 533 Parnassus, Rm. U107, San Francisco, CA 94143-0112, USA*

Received 9 August 1999; received in revised form 7 April 2000; accepted 9 June 2000

Abstract

One of the objectives of this phase II study was to determine whether temozolomide (TMZ) improved the health-related quality of life (HRQL) of patients with recurrent anaplastic astrocytoma (AA). HRQL was assessed at baseline (pretreatment) and every 4 weeks at each treatment cycle using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) (version 2.0) and the Brain Cancer Module (BCM20). Changes from baseline in the scores of seven preselected HRQL domains (role and social functioning, global QL, visual disorder, motor dysfunction, communication deficit and drowsiness) were determined at 6 months as well as prior to, and at the time of, disease progression. The significance of the changes was assessed by calculating statistical significance, effect sizes and the proportions of patients with improvement in their HRQL scores (changes of $\geqslant 10$ points). After 6 months of treatment, patients who were free of progression of disease reported either an improvement or maintenance of all the preselected HRQL domains scores. Patients with disease progression by 6 months usually experienced improvement in HRQL before progression, but there was a sharp decline in most of the preselected domains at progression. We conclude that treatment of recurrent AA with temozolomide is associated with significant HRQL benefits.

Keywords: Anaplastic astrocytoma; Chemotherapy; Health-related quality of life; Tumour response; Temozolomide

1. Introduction

Treatment of primary anaplastic astrocytoma (AA) by surgery and radiation does not result in the cure of a substantial proportion of patients. Furthermore, the treatment of recurrence by chemotherapy is even less successful and prolongation of survival is modest. This is a clinical situation in which the inclusion of health-related quality of life (HRQL) could be very helpful in decision-making because HRQL is substantially adversely affected at relapse. It is similar in some HRQL domains to that of patients with recurrent glioblastoma multiforme (GBM) [1] and advanced metastatic disease in a variety of cancers. Thus, the HRQL results, in addition to the standard outcomes of survival and progression-free intervals, could assist in determining the benefits of treatment. For benefit to be claimed it would be necessary to show that the patients' HRQL scores during treatment or post-treatment follow-up had significantly improved as compared with the pretreatment scores.

The treatment of AA after recurrence is problematic. There is only modest improvement in the length of survival and the toxicity of treatment on the patients' well-being has not been systematically described. In these patients, improvement of HRQL should be used as an outcome measure of equal importance to prolongation of survival. However, there is a paucity of formal studies to determine whether HRQL is improved by treatment of AA at first recurrence.

This report gives the HRQL results of a phase II study of patients with recurrent AA who were treated with temozolomide (TMZ) [2]. HRQL was measured by the European Organization for Research and Treatment

* Corresponding author. Tel.: +1-604-921-793; fax: +1-604-921-7794.
E-mail address: david_osoba@telus.net (D. Osoba).

0959-8049/00/$ - see front matter
PII: S0959-8049(00)00165-9

of Cancer (EORTC) Quality of Life Questionnaire (QLQ-C30(+3)) [3] and the Brain Cancer Module (BCM20) [4]. The primary assessments were intended to determine the nature and magnitude of HRQL changes in seven preselected HRQL domains during the course of treatment and to correlate the HRQL changes with changes in disease status. Our hypothesis was that TMZ treatment after first recurrence would be associated with an improvement in the HRQL scores of one or more of the preselected domains during treatment as compared with pretreatment (baseline) scores.

2. Patients and methods

The HRQL measurements were carried out as a component of a study of patients who were treated with TMZ for recurrence of AA after failure of primary therapy. The details of the clinical response and survival data have been published elsewhere [2].

2.1. Eligibility criteria

To be eligible for this study, patients had to be at least 18 years of age, with Karnofsky Performance Scores (KPS) $\geqslant 70$. They had to have histologically proven supratentorial astrocytoma at original diagnoses and show unequivocal evidence of tumour recurrence or progression at first relapse by contrast-enhanced magnetic resonance imaging (MRI) or computer tomography (CT) scanning after failed radiation therapy. In addition, they had to be on a stable dose of steroids for at least 10 days before chemotherapy and have an expected life expectancy of at least 12 weeks. Previous chemotherapy was allowed, but had to include a nitrosourea-containing regimen. Reasons for exclusion were significant renal, hepatic or bone marrow impairment, prior treatment with dacarbazine, persisting toxicity from any previous therapy, previous tumours at other sites (excluding basal cell carcinoma), or HIV-positive serology.

2.2. Chemotherapy and toxicity

Briefly, TMZ was given orally to fasting patients at a dose of 150 mg/m^2/day for 5 days (750 mg/m^2 total dose per cycle) if they had prior chemotherapy, and 200 mg/m^2/day for 5 days (1000 mg/m^2 total dose per cycle) if they had not had prior chemotherapy. Treatment cycles were repeated every 28 days. Chemotherapy was to be given for 1 year and could be continued longer in responding patients, if desired.

The criteria for reducing dosage or stopping treatment were grade 3 or higher haematological toxicity or other unacceptable toxicity as determined by the Common Toxicity Criteria Grading, progression of disease or the completion of the prespecified lengths of treatment.

2.3. HRQL measurement

HRQL status was assessed by two self-report questionnaires, the EORTC QLQ-C30 (+3) [3] and the Brain Cancer Module (BCM20) [4]. Both have been shown previously to be reliable and valid instruments in the setting of recurrent high-grade gliomas [5]. The QLQ-C30 (+3) consists of 33 items which form five functioning domains (physical, role, emotional, social, cognitive), a global QL/overall health domain, three symptom domains (fatigue, pain, nausea and vomiting) and six single items (dyspnoea, diarrhoea, constipation, anorexia, insomnia and financial impact). The QLQ-C30 (+3) was an interim version of version 1.0 of the QLQ-C30 [6–9] used to develop version 2.0 [4]. The version 2.0 scoring system was used in this study [10] and the instrument will be referred to as version 2.0 in the remainder of this report.

The BCM20 contains 20 items grouped into four domains (future uncertainty, visual disorder, communication deficit, motor dysfunction) and seven single items (headache, seizure, drowsiness, hair loss, itching, weakness of both legs, difficulty in controlling bladder function) [4]. Some of these symptoms (i.e. hair loss and itching) are related to the toxicity of previous treatment by radiation therapy or side-effects of TMZ, while the remainder are associated with the recurrence of AA.

The HRQL questionnaires (in the appropriate languages) were given to patients for completion prior to the first cycle of chemotherapy (baseline assessment) and subsequently just before each subsequent cycle of chemotherapy for all patients.

The QLQ-C30 was scored according to previously described methods for the QLQ-C30 (version 2.0) [10]. All raw scores were converted to lie in a range between 0 and 100. For the functioning scales and the global QL scale, a higher score indicates better functioning while for the symptom scales/items a higher score indicates more of the symptom or difficulty. The BCM20 was scored in a manner analogous to the QLQ-C30 (version 2.0) [4]: higher scores indicate more difficulty with the symptom.

The changes in scores between baseline and any subsequent scores were calculated by subtracting each patient's baseline score from his/her subsequent score at chosen time points. As patients dropped out of the study, the change in scores for each patient was calculated only for those who were still on study. Virtually all drop-outs were due to disease progression; if a score was not available at the time that progression was declared, then the previous score was used as the score nearest progression. This occurred in 22% of the patients by 6 months.

A clinically significant change in scores was defined as a change of $\geqslant 10$ (on a scale of 0–100) lasting for at least two HRQL assessments 4 weeks apart This criterion was based on previous studies involving a formal assessment of how much change in QLQ-C30 scores is perceptible to patients (subjectively significant) [11] or is associated with changes in disease status [12]. Other approaches, using other questionnaires in other illnesses also indicate that a change of $\geqslant 10$ should be interpreted as being clinically significant [13–15]. Accordingly, the number of patients in our study who were found to have experienced a subjectively significant improvement were enumerated to obtain the proportions with improvement in HRQL status. The duration of improvement in those with subjectively significant responses was calculated.

Changes in HRQL status were related to changes in tumour and disease status to determine if they were associated. To decrease the possibility of finding statistically or clinically significant associations by chance alone, we decided a priori to limit the number of HRQL domains that would be examined by significance testing. The seven chosen domains were role functioning, social functioning, global QL, visual disorder, motor dysfunction, communication deficit and drowsiness. It was expected that these were likely to be most affected in AA and that if significant improvement was not detectable during treatment in these domains, then changes in the others would be unlikely to be of more clinical importance.

Subsequent to examining the changes in the seven preselected domains, other domains were also examined, using the methodology described above, to obtain a more comprehensive description of the effects of treatment on HRQL However, we did not form any hypothesis about the behaviour of the scores for these other domains.

2.4. Clinical assessment, scanning procedures and tumour status

Complete (CR) (disappearance) and partial tumour (PR) ($\geqslant 50\%$ reduction) responses, stable (SD) ($< 50\%$ reduction and $< 25\%$ increase) and progressive disease (PD) ($\geqslant 25\%$ increase), were determined by enhanced CT or MRI scanning scheduled every 2 months. A clinical response or progression could be declared by neurological examination in the intervals between scheduled scanning. In actual practice, fewer than 10% of clinical responses were assessed only by neurological status.

2.5. Statistical procedures

Standard statistical procedures were carried out using an SAS (SAS Institute, Cary, NC, USA) statistical package. Effect sizes were calculated by subtracting the mean score at baseline from the mean score at a later time and dividing by the standard deviation of the baseline score [16,17]. Effect sizes of 0.2 or more were considered important and have been shown to reflect patients' perception of a change having occurred between two time points [11]. Effect sizes of between 0.2 and 0.49 were considered small, those between 0.5 and 0.8 moderate and those > 0.8 large. P values were calculated using the Student's t-test.

3. Results

The total number of patients enrolled was 162. Of these, 17 (10%) had either no HRQL evaluation at any time or no baseline evaluation. 7 (4%) had a baseline evaluation only, leaving a total of 138 (85%) with both baseline and on-treatment evaluations. The patient characteristics of patients completing HRQL assessments at baseline and at least one additional assessment are shown in Table 1.

Table 1
Characteristics of patients who completed health-related quality of life (HRQL) questionnaires at baseline and at least one subsequent assessment

Age (years) mean (range)	42.5 (19–76)
	n (%)
Gender	
male	78 (57)
female	60 (43)
Race	
white	132 (96)
black	3 (2)
other	3 (2)
KPS	
100	23 (17)
90	43 (31)
80	32 (23)
70	39 (28)

KPS, Karnofsky Performance Status.

Table 2
Baseline health-related quality of life (HRQL) scores for seven preselected domains in patients with recurrent anaplastic astrocytoma (AA)

Domains	*n*	Mean	S.D.
Role functioning	143	66.7	31.9
Social functioning	142	64.0	31.4
Global QL	142	61.4	22.5
Visual disorder	143	14.2	21.4
Motor dysfunction	142	26.4	26.4
Communication deficit	143	24.4	25.5
Drowsiness	143	34.3	29.9

S.D., standard deviation.

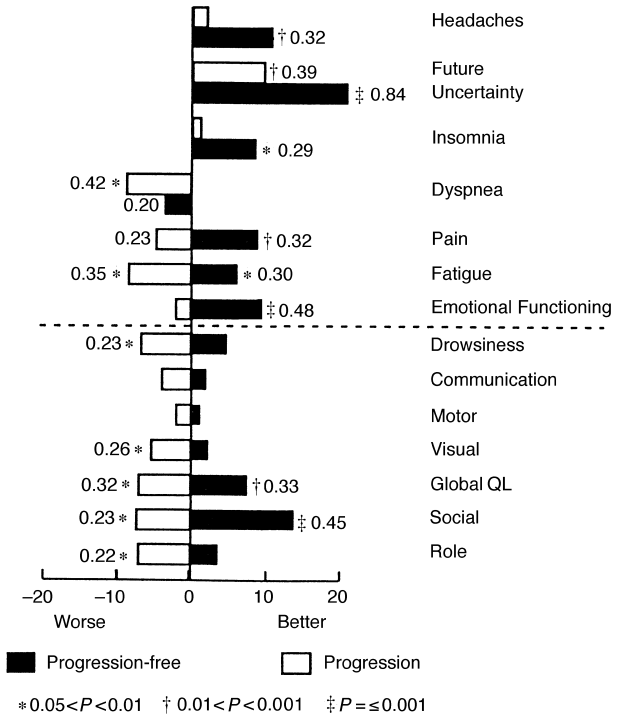

Fig. 1. Association of changes in health-related quality of life (HRQL) scores with disease response status at 6 months. Effect sizes are shown if they were $\geqslant 0.2$. Effect sizes of 0.2–0.49 are small, 0.5–0.8 are moderate and >0.8 are large. *$0.05 < P < 0.01$. †$0.01 < P < 0.001$. ‡$P \leqslant 0.001$.

3.1. Baseline HRQL sores

For the seven preselected domains, the mean baseline scores are shown in Table 2. These indicate considerable difficulty in the functioning domains (role and social), in global QL and in three of the four symptom domains (motor dysfunction, communication deficit and drowsiness).

3.2. Change in HRQL scores with change in disease status

The relationship between changes in HRQL scores and disease status was assessed after 6 months of therapy with temozolomide. At this time, 63 patients (39%) were free of disease progression. Associating changes in HRQL scores with disease status showed maintenance or improvement of scores in the seven preselected domains in patients who did not have disease progression. Social functioning and global QL improved statistically significantly and effect sizes were small (Fig. 1). In contrast, patients with disease progression before and at 6 months reported statistically significant deterioration in five of the seven preselected domains. It is also notable that several of the other scale scores followed a similar pattern, i.e. improvement from, or maintenance of baseline scores in the patients who were free of disease progression as compared with deterioration in the patients with progression. An exception was 'uncertainty about the future' which decreased in the patients with disease progression, but to a lesser magnitude than in patients who were progression free. Another apparent exception may be dyspnoea, which increased slightly in patients who were free of disease progression. The P value was not statistically significant, and the effect size (ES) was small (0.20). The increase in dyspnoea in the patients with disease progression was statistically significant ($P = 0.03$) and the ES was small (0.42).

3.3. HRQL changes before progression of disease

Many patients with recurrent AA will eventually experience disease progression. In this study, HRQL scores at the time of progression (as determined by Gadolinium-enhanced MRI or contract CT scans) were either at baseline values or worse than baseline for the

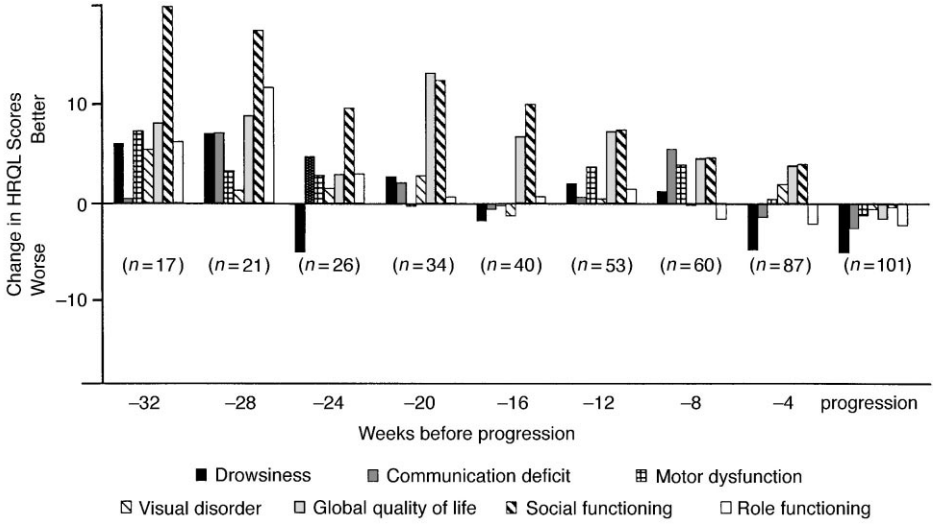

Fig. 2. Mean changes in health-related quality of life (HRQL) scores from baseline to times before disease progression and at progression.

Table 3
Proportions of patients who reported improvement of $\geqslant 10$[a] in preselected health-related quality of life (HRQL) scores, regardless of tumour response, and median duration of response

	n Eligible	*n* Responding (%)	MDR[b] (range)
Role functioning	85[c]	41 (48)	16 (8–56)
Social functioning	92	49 (53)	20 (8–72)
Global QL	114	43 (38)	12 (8–56)
Visual disorder	54	35 (65)	16 (8–64)
Motor dysfunction	91	46 (51)	20 (8–72)
Communication deficit	84	43 (51)	20 (8–64)
Drowsiness	88	38 (43)	12 (8–56)

[a] Patients with maximum (100) functioning scores and minimum (0) baseline symptom scores are excluded because improvement for these patients was not possible.
[b] Median duration of response in weeks.
[c] The proportions of patients with no change or deterioration in HRQL scores is not 100 minus the proportion with improvement shown above since all patients were eligible for deterioration.

seven preselected domains (Fig. 2). To determine the changes in HRQL scores prior to progression, the baseline scores were subtracted from the scores obtained at 4, 8, 12, etc. weeks before progression. These data show initial improvement in scores over baseline during the course of treatment in most of the preselected domains (the main exception being visual disorder) with a gradual decrease in scores as progression neared and worse than baseline scores at progression (Fig. 2).

3.4. Duration of HRQL response and disease status

An HRQL response is deemed to be an improvement of >10 points on a 0–100 scale in any preselected domain lasting for at least 8 weeks. The overall proportions of patients who were eligible for an HRQL response and who achieved it are shown in Table 3. Median duration of HRQL response varied from 12 weeks for global QL and drowsiness to 20 weeks for social functioning, motor dysfunction and communication deficit. The duration of each response associated with disease status is shown in Table 4. Patients were eligible for a response if their baseline HRQL scores were less than 90 in the three functioning scales and greater than 10 in the four symptom scales. The median duration of HRQL response tended to be longest in those with a complete or partial disease response, but responses were almost as long in those with stable disease. The exception was for drowsiness, which was diminished for a longer time in stable disease. The median duration of HRQL response in the very small number of patients with progressive disease who reported HRQL responses was usually shorter than in the other response categories.

Table 4
Mean duration of health-related quality of life (HRQL) responses in seven preselected domains according to disease response

	CR/PR		SD		PD	
	n[a]	MDR[b] (range)	*n*	MDR (range)	*n*	MDR (range)
Role functioning	24	20 (8–56)	16	14 (8–56)	1	40 –
Social functioning	28	26 (8–72)	18	16 (8–56)	3	8 (8–8)
Global QL	23	20 (8–56)	14	12 (8–32)	6	8 (8–12)
Visual disorder	21	16 (8–48)	11	16 (8–64)	3	12 (8–24)
Motor dysfunction	28	26 (8–72)	14	18 (8–44)	4	16 (8–24)
Communication deficit	21	28 (8–44)	18	18 (8–68)	4	10 (8–12)
Drowsiness	23	12 (8–56)	13	12 (8–52)	2	16 (12–20)

CR/PR, complete or partial response; SD, stable disease, PD, progressive disease.
[a] *n* = number of patients meeting the criterion of an improvement of $\geqslant 10$ in a domain score lasting at least 8 weeks.
[b] Median duration of response in weeks.

4. Discussion

HRQL and survival are outcomes of equal importance in recurrent AA. The median survival after recurrence is short and treatment is only modestly effective. Thus, since HRQL measurement encompasses an assessment of functioning ability and toxicity from therapy, the outcomes are useful in determining the overall value of therapy from the patient's perspective.

Our hypothesis was that TMZ treatment for recurrent AA would be associated with an improvement in the preselected HRQL domains during the course of treatment. Thus, on-treatment HRQL scores for each patient were compared with pretreatment (baseline scores) by subtracting baseline scores from on-treatment scores. In this method, each patient acts as his/her own control and must meet preset criteria for improvement over his/her own baseline scores before improvement in HRQL can be declared. This procedure is not subject to the same criticism as is the calculation of mean scores for the patients remaining in the study. In the latter calculation, it would be expected that mean HRQL scores would improve with time, since patients with disease progression drop out of the study early and the remaining patients are more likely to be the ones who have had disease regression. Such patients, who do relatively well, are likely to live longer, have a reduction in symptoms, and have greater opportunity for improvement in HRQL. An exception to these likely outcomes is when on-treatment toxicity is sufficient to counteract any beneficial effects of treatment.

The differences were tested for statistical significance (*P* values) and effect sizes in only seven preselected HRQL domains to avoid chance significance as a result of extensive multiple comparison testing. Other ways to reduce chance significant results are to use a Bonferroni correction, or to set the *P* value cut-off at a more stringent level than 0.05. All these methods are valid, but we chose to use the first one, i.e. to limit the number of

domain scores tested, because we reasoned that the domains we chose, a priori, would be the ones most likely to change and, therefore, of most interest. We did not choose to use a Bonferroni correction because it assumes that the variables being tested are completely independent of each other, whereas some of the QLQ-C30 and BCM20 domains are moderately correlated [4,6–9].

Testing for the null hypothesis, with resultant P values, is commonly practised in cancer clinical trials, but the calculation of effect sizes is uncommon. P values are highly dependent on sample size; the larger the sample, the more likely that a small numerical difference will achieve statistical significance at the $P = 0.05$ level. In contrast, effect sizes are dependent primarily upon the variability (standard deviation) of the sample. It has been estimated that effect sizes of 0.2 are important, but small while those of 0.5 are moderate and 0.8 or greater are large [16,17]. Thus, they are useful when used together with P values, to obtain a better understanding of the meaning of changes in scores between different time points in a relatively small sample of patients.

Using the above approach, the patients being treated with TMZ who remained progression free at 6 months reported a statistically significant improvement in social functioning and global QL domains. The effect sizes for these two preselected domains were small. In addition, apparent improvement was noted in several other domains, but these were not part of our original hypothesis, and require further investigation. In contrast, patients with disease progression reported statistically significant worsening in five of the seven domains, with small effect sizes. It is possible that these changes underestimate of the magnitude of change in this group, since we used the last available score before death or inability to complete the questionnaires rather than imputing a score for this time point. (To provide more detail on the items included in these domains, brief versions of them are shown in Appendix A.)

The value of comparing on-treatment scores with baseline scores is that changes in HRQL scores before disease progression can be examined. In order to do this, patients' scores have to be analysed relative to a fixed time in the disease trajectory that would be the same for all patients. Such a time is the time at which disease progression was declared by Gadolinium-enhanced MRI or enhanced CT scanning. This is preferable to the usual method of analysis in which the starting point is the time of entry into the study. At entry, the patients are at varying degrees of disease recurrence, with some patients having had only sufficient disease growth (volume) to become eligible for the study, while others may have had much more disease volume by the time of study entry. Since the imaging criteria for disease progression were well defined, the population is more homogeneous at this time for the amount of disease than at study entry. To complete the analysis before time of progression, one proceeds 'backwards' to earlier times such as 4 weeks, 8 weeks, etc. before progression.

In this group of patients, there was a clear pattern of improvement in most of the preselected domains at times before disease progression, with clear worsening at the time of progression. This might be an expected finding, but the alternative possibilities are either no change from baseline or a worsening during the entire on-treatment period. These alternatives likely would have occurred if there had been no benefit from treatment or if the treatment were sufficiently toxic to counteract and, thereby, camouflage any possible benefit. Since the improvement seen in HRQL scores does not support these latter possibilities, we conclude that there was HRQL benefit for periods of time before disease progression. These benefits were still evident in those who had not had disease progression at 6 months, and would, presumably, disappear at a later time when disease progression eventually occurred.

As mentioned above, when sample sizes are large, small numerical differences may be statistically significant, e.g. a numerical difference of 5 on a 0–100 scale in a sample of 500 patients would give a P value of < 0.001. Therefore, it has been important to determine whether such numerical differences have clinical significance. Previous investigations suggest that patients with a variety of cancers (metastatic breast [11,12], small-cell lung cancer [11] and ovarian cancer (data not shown) are able to detect changes as small as 5 points and that changes of $\geqslant 10$ points are perceptible as moderate differences to patients [11]. A change of approximately 10 also seems to distinguish between different stages and treatment of breast cancer [12]. Furthermore, a similar magnitude of change has been suggested as being the lower limit of perceptible change for patients with asthma [14,15] and has been suggested in a variety of other chronic illnesses [13]. Thus, we reasoned that, in the absence of data derived directly from patients with brain cancer, a similar magnitude of change would be an acceptable cut-off point for the patients in our study.

We set a requirement that the duration of an improvement of $\geqslant 10$ be present for at least 8 weeks, since this included two adjacent cycles of treatment with TMZ and, thus, two consecutive completions of the HRQL questionnaires. We reasoned that HRQL improvement lasting at least 8 weeks was a significant length of time and also that it was less likely to be a chance event than if we had chosen a shorter time, e.g. 4 weeks. Patients who met the two requirements, i.e. an improvement of $\geqslant 10$ lasting a minimum of 8 weeks, were deemed as having an 'HRQL response' in the HRQL domain showing these changes. This criterion is similar to the frequently used assessment of disease status in clinical oncology.

In this analysis, those with baseline HRQL scores of 90 or less in the function domains and those with symptom scores of 10 or more were included. Our hypothesis was that there would be improvement in the seven preselected domains. Therefore, we included those patients who had 'room' for improvement in the analysis of subjective significance. Stability of scores in those with baseline scores of 100 and 10, respectively, was judged not to be sufficient to be able to claim improvement. Thus, patients with stable or deteriorating scores were those who did not meet the criteria for improvement. Data on proportions of patients with stable HRQL or with deterioration may be valid in other studies with different hypotheses.

As might be expected, the largest proportion of patients with subjectively significant HRQL improvement was found in those whose disease was judged to be in complete or partial remission. Such patients also tended to have longer periods of benefit than did patients with stable disease or disease progression. One explanation may be that those with disease regression had a greater opportunity to obtain longer HRQL benefits than did those with stable or progressive disease. However, an equally valid explanation is that improvement in HRQL status is the result of improved disease status and is associated with effective treatment.

It is of interest that a significant proportion of patients with HRQL responses was found in those who were judged to have stable disease, and a few patients with disease progression also reported HRQL responses. The explanation for these findings may lie in the fact that the categories of stable disease include patients with less than 50% disease regression or less than 25% progression. This is a very broad range of disease status and, therefore, less than 50% disease regression (e.g. 30%) may conceivably be accompanied by an improvement in HRQL. Similarly, an increase of disease by 25% may not be sufficient to always produce a deterioration in HRQL, if the progression is in areas of the brain that do not affect functional status or symptomatology.

In retrospect, it is likely that the choice of the seven HRQL domains a priori was not as accurate in this population as we originally thought. The possibility that other domains, such as emotional functioning, fatigue, pain, insomnia and uncertainty about the future, are also important bears further investigation.

We conclude that patients with recurrent AA achieve HRQL responses in at least two important HRQL domains, namely social functioning and global QL, with possible responses in other domains during treatment with TMZ. Further studies, in the form of randomised clinical trials, are required to determine whether TMZ has advantages over other forms of chemotherapy or other therapies for recurrent AA.

Acknowledgements

This study was supported by The Schering-Plough Corporation. We thank Jill Vardy for assistance with assembling the manuscript.

Appendix A

Preselected domains:	Items deal with:
Role functioning	Limitations in doing work or other daily activities, and in pursuing hobbies or other leisure time activities.
Social functioning	Interference with family life and with social activities.
Overall health and global QL	Ratings of overall health and of overall quality of life.
Visual disorder	Double vision, blurred vision and difficulty reading because of vision.
Motor dysfunction	Weakness on one side of the body, trouble with co-ordination and feeling unsteady on feet.
Communication deficit	Finding the right words, difficulty speaking and trouble communicating thoughts.
Drowsiness	Feeling drowsy during the daytime.

References

1. Brada M, Yung WKA, Osoba D. Symptom burden in patients (PTS) with recurrent high-grade gliomas. *Proc Am Soc Clin Oncol* 1999, **18**, 147a (abstract 564).
2. Yung WKA, Prados MD, Yaya-Tur R, *et al*. Multicenter phase II trial of temozolomide in patients with anaplastic astrocytoma or anaplastic oligoastrocytoma at first relapse. *J Clin Oncol* 1999, **17**, 2762–2771.
3. Osoba D, Aaronson NK, Zee B, Sprangers M, te Velde A. Modification of the EORTC QLQ-C30 (version 2.0) based on content validity and reliability testing in large samples of patients with cancer. *Qual Life Res* 1997, **6**, 103–108.
4. Osoba D, Aaronson NK, Muller K, *et al*. The development and psychometric validation of a brain cancer quality-of-life questionnaire for use in combination with general cancer-specific questionnaires. *Qual Life Res* 1996, **5**, 139–150.
5. Osoba D, Aaronson NK, Muller M, *et al*. Effect of neurological dysfunction on health-related quality of life in patients with high-grade glioma. *J Neuro-Oncol* 1997, **31**, 263–278.
6. Aaronson NK, Ahmedzai S, Bergman B, *et al*. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international trials in oncology. *J Natl Cancer Inst* 1993, **85**, 365–376.
7. Osoba D, Zee B, Pater J, Warr D, Kaizer L, Latreille J. Psychometric properties and responsiveness of the EORTC Quality of Life Questionnaire (QLQ-C30) in patients with breast, ovarian and lung cancer. *Qual Life Res* 1994, **3**, 353–364.
8. Ringdal GI, Ringdal K. Testing the EORTC Quality of Life Questionnaire on cancer patients with heterogeneous diagnoses. *Qual Life Res* 1993, **2**, 129–140.
9. Niezgoda HE, Pater JL. A validation study of the domains of the core EORTC Quality of Life Questionnaire. *Qual Life Res* 1993, **2**, 215–319.

10. Fayers P, Aaronson N, Bjordal K, Curran D, Groenvald M. *EORTC QLQ-C30 Scoring Manual* (2nd edn). Brussels, Quality of Life Unit, EORTC Data Centre, 1999.
11. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life-scores. *J Clin Oncol* 1998, **16**, 139–144.
12. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res* 1996, **5**, 555–567.
13. Ware JE, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston, The Health Institute, New England Medical Center, 1994.
14. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clin Trials* 1989, **10**, 407–415.
15. Juniper EF, Guyatt GH, Willan A, *et al*. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994, **47**, 81–87.
16. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989, **27**(Suppl. 3), S178.
17. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. New York, Academic Press, 1997, 8.