Schering-Plough (TMZ) comments	PenTAG response
1. **Confirmation of GBM status:** (section 4.8.1.1, page 60) Tumour classification is highly subjective. The large RCT conducted by EORTC and NCIC admitted only patients (n = 573) with histologically confirmed glioblastoma (WHO grade IV) [utilizing local neuro-pathologists, thus reflecting daily clinical practice]	See our comments on p. 60 para 2. Section 4.8.2.3 . There remains the possibility that a small number of patients with chemo-sensitive tumours respond well to TMZ while most patients do not.
2. **MGMT promoter methylation:** The assessment report takes this data of an unplanned retrospective analysis of a subset of patients at face value, rather than considering it as a hypothesis worthy of prospective validation. We recognise the potential importance of this gene expression in the context of optimizing treatment outcomes and are therefore supporting a large RCT to validate these findings.	No comment
3. **Clinical trial population:** The assessment team comments that the clinical trial population for temozolomide is different to the real-world patient population, and therefore questions the generalisability of the findings (section 4.8.1.2, page 62). The approval for TMZ usage in first line GBM has been granted for a population reflecting the population studied in the pivotal trial conducted by the EORTC / NCIC.	Generalisibility refers to the possibility of extrapolating trial results to general clinical populations. We note that older patients are excluded, which is the case. Also that the Athanassiou study only provides data on populations <50 and >50 yrs old which limits our ability to assess the similarity of the studied population to the clinical population (4.8.12, p.62).
4. **Study Blinding and subsequent therapy (performance and/or detection bias):** The assessment team remarks that due to the fact that the EORTC / NCIC trial was not conducted as a double-blinded trial, the assessment of response and progression-free survival might be biased. Conversely, the pivotal study with the BCNU-W was conducted in a double-blinded fashion, and with clearly defined criteria for progression. No difference in progression-free survival could be detected.	True but we cannot know if this would be the case with TMZ. Lack of blinding may be problematic where definitions of progression have subjective elements (as here) and where concomitant treatment and cross over is possible in the trial, as it is here.
Critique of the Schering Plough economic evaluation:	
1. **Costs with disease progression:** The evaluation of the clinical trial data for TMZ shows that the incremental costs between the TMZ + RT and RT-only arms was reduced partly because the latter group received more chemotherapy after progression, and of these, many more received TMZ. The consideration of this treatment pathway in patients with disease progression, with regards to the costs and survival effects of this salvage TMZ treatment in the RT-only arm cannot be ignored.	Please see attached sheet.
2. **Lack of estimation of QALYs:** QALYs were not calculated in the original submission in part due to evidence from [redacted]	This was investigated in sensitivity analysis. See Figure 31 (p. 129) - Threshold analysis shows that TMZ is not cost-effective (at WTP of £30K) even if the utility value in the stable state were 1 (=perfect health).
3. **Survival extrapolation:** For overall survival, the extrapolation distribution was not fitted to the 2-year survival data, but rather to the entire survival curve, thus including patients at risk after 2 years.	Thanks for this clarification which was unclear to us in the original submission. However, this does not alter the fundamental point made in the critique that (a) the industry submission does not fully describe the statistical goodness-of-fit of their extrapolated survival curves to the actual data, and (b) nor does it justify

	the use of different statistical distributions for extrapolating the full cohort and economic subsample.
Comments on the PenTAG economic evaluation:	
A number of errors were identified in the limited time given to review the PenTAG cost-utility model. In light of the structure and difficulty we have found in auditing the model, we have to question the reliability of this tool in evaluating the cost-effectiveness of temozolomide. We would appreciate further review by the NICE appraisal team to assess the internal and external validity of this model.	No comment
Our main concerns regarding the model are as follows:	
1. **Health State Valuation:** We would like to challenge the utility values assigned to 'SMG+RT+TMZ' of 0.8091 and 'SMG+TMZ' of 0.8474. [redacted]	Note that the values quoted are only for the minority of patients who do experience adverse effects due to treatment with TMZ (see table 38 and para 1 p.98) As stated above – The impact of utility values was investigated in sensitivity analysis. See Figure 31 (p. 129) -Threshold analysis shows that TMZ is not cost-effective (at WTP <£30K) even if the utility value in the stable state were 1 (=perfect health). Overall we felt that the utilities obtained were high for a terminal disease.
2. **Time-dependent Risk of Death:** PenTAG assumes that transitions to death are time-dependent as opposed to state-dependent. The group has acknowledged that this assumption is counterintuitive, as during a given cycle, patients with disease progression would be expected to have a higher probability of dying than those with stable disease.	The majority of deaths do occur from the progressive disease state using our model structure (see Figure 26, p. 125). We further investigated this assumption through sensitivity analysis, (Figure 28 p. 127) and it was not found to have a significant impact on the model outputs.
3. **Weibull Distribution Transition Probabilities:** Time dependent transitions probabilities have been estimated using the Weibull distribution and survival data presented by Stupp et al[2]. Whilst the predicted overall survival for RT-only shows a good fit with the observed data, this is not the case for the RT+TMZ treatment arm.	This is shown in Appendix 12 p.218 and 219. In all cases there is good fit. R^2 values are 0.9977 for the TMZ arm and the % error in estimating median survival is just 3.09% (in favour of TMZ). Sensitivity analysis (Figure 29, p. 128) shows that differential survival of 30 weeks is needed for TMZ to become cost effective (at WTP <£30K).
4. **Patients most likely to benefit from treatment:** PenTAG have only considered the cost-effectiveness of TMZ in the overall study population. The NICE scope outlined a remit which included exploring the cost-effectiveness of treatment in those patients most likely to benefit. We would recommend further analyses are conducted (e.g. extent of surgery[3], performance status, MGMT gene silencing[4]).	Given the paucity of data on these groups, we have concentrated on extensive sensitivity analysis. Any subgroup would have to demonstrate considerable survival advantage in order to be considered cost-effective. Figure 29, p. 128 shows that differential survival of 30 weeks is needed for TMZ to become cost effective (at WTP <£30K). We also explored a hypothetical cohort of patients with good prognosis (details in section 5.7.2.3, p.132-133). The ICER for this group was £43K/QALY.
5. **Comment on BMJ/Rawlins paper:** The PenTAG assessment group concludes that treatment with TMZ yields an ICER of £46,000/QALY. GBM is an end-stage cancer. Given the precedent of NICE accepting ICERs >	This is a comment for the consideration of NICE rather than PenTAG. Sensitivity analyses have provided data on a range of WTP thresholds in order to facilitate such decision making.

£30,000/QALY for other treatments of end-stage cancers, £46,000/QALY may be an acceptable ICER in the consideration of TMZ (Rawlins and Culyer, 2004[5]).	

Comments from GDG via NCC for cancer	
Page 1. On the opening page (p1) the authors say that 'existing approaches to chemotherapy have not convincingly demonstrated a survival benefit'. In fact, the evidence from three overviews, and particularly the Stewart overview, does demonstrate that there is a statistically significant benefit to chemotherapy in this situation and has convinced the majority of the establishment in this discipline. In raising doubt on this issue the authors say that 3 later trials did not show benefit. They might not know that these studies, and particularly the largest, MRC, trial have endured heavy methodological criticism. For most of the neuro-oncology community the question is not whether chemotherapy produces a statistically significant effect (it does), rather whether this is clinically worthwhile. It is true in the UK we have felt that the benefit was outweighed by other disadvantages.	This statement is based on existing evidence and is outlined more fully in the section on chemotherapy on p. 16 (3.2.5) I have changed the wording on page 1 to "not conclusively demonstrated" – as stated on p.16 – this meta-analysis included 8/12 trials that were more than 20 years old and only one of the four more recent trials show any survival advantage with chemotherapy.
In the objectives (p1), they suggest that they will investigate adjuvant and concomitant Temozolomide compared to surgery alone. I do not understand how they intend to do that since no comparative study has ever been done and virtually all the surgery alone data derives from a previous era when diagnostic criteria were considerably different. The comparison is of course conventional treatment with surgery + RT vs the same regime plus adjuvant/ concomitant TZ.	This objective is in line with those pre-specified in the scope issued by NICE and our protocol shown in appendix 3 p.155. As no studies of this nature were in the event identified, these have been removed from the summary.
Page 4 The Temozolomide study is criticised for excluding patients with surgical complications and those who died soon after surgery. Since the decision to use Temozolomide and its cost occur after surgery, it is difficult to understand this criticism. The population defining this study and indeed the population who would be eligible for Temozolomide is that population which follows surgery.	This comment simply shows that sicker patients would not have entered the trial.
The Temozolomide study is criticised for including in the analysis 7-8% of patients who were re-categorised at central review as having grade 3 tumours. Much is made of this throughout the document. The authors fail to realise that the diagnosis of malignant glioma is highly subjective. Entering into the study was based on a local diagnosis (as would happen in real life if this agent were licensed and supported). The fact that a central reviewer reclassifies a tumour, does not necessarily mean that this is a 'true' or 'absolute' classification, simply that there is a disagreement with the local pathologist. It gives a consistency to the analysis, since all tumours are reviewed by one panel. Indeed to emphasise this point, the EORTC have recently compared diagnoses on a given panel of tumours made by various senior pathologists who are regularly used in clinical trial central reviews. They found major disagreements amongst these	The generalisibility of the trial data depends on the extent to which local pathologists would diagnose grade III and IV tumours in the same way as the pathologists in the trial. They may identify more or fewer grade III tumours as grade IV. As we don't know this, the trial data may over or underestimate effectiveness in clinical practice. In addition, the population studied in the main trial is not a "confirmed" GBM population (for whom TMZ is licensed). Analysis in the Schering Plough submission suggests a weaker effect in a GBM-only subgroup (TAR p. 68). The risk is that a large number of people, for whom benefit is unlikely, are treated in order that the small number of people who do benefit receive treatment. Identifying patients for whom chemotherapy is beneficial is a research priority in this area.

pathologists. Hence it is clear that the output of central review depends on which pathologist is used. It follows that central pathology review does not give a 'true measure' of the presence of glioblastoma. It merely gives a measure of that pathologists opinion. It may or may not be more valid than the local pathologist. What, hopefully, it does do is give a uniformity to assessment. In real life, patients will be offered Temozolomide on the basis of the local pathologists diagnosis and hence analysis of this trial in these terms gives a more realistic interpretation of the outcome of such treatment and the comparison between treatments.	
Furthermore, much is made of trial results being driven by 'chemo-sensitive tumours' on the assumption that they will influence the outcome favourably for a chemotherapy treated arm. It is equally possible that these chemo-sensitive tumours will influence results in the reverse direction. Whilst this may initially seem paradoxical, the example within this discipline of anaplastic oligodendroglioma is clear and illustrative. This highly chemo-sensitive tumour was thought almost certainly to require adjuvant chemotherapy. When the study was done, no improvement in survival was seen as a result of use of adjuvant chemotherapy in this group of tumours in spite of the chemo-sensitivity. The inclusion of such patients in an adjuvant trial, such as the two described here, may then act to dilute a population that would otherwise show a difference and adversely influence the results of the trial against the extra intervention. The point I make is that *no assumption* can be made that because a tumour is chemo-sensitive it will influence the outcome in a positive direction.	As we state on p. 4 and elsewhere "small numbers of more chemosensitive tumours may have impacted on the findings".
Page 5 The authors admit their model is particularly sensitive to median overall survival benefit. As argued elsewhere in this document, this is not the most appropriate parameter on which to judge the outcome, certainly of the Temozolomide trial where the difference in median survival may be dominated by a resistant population, but a highly beneficial effect might be seen in a sensitive subpopulation which shows up in the later stages of the study, after the time point of medial survival.	In a model assessing treatments for terminal cancer, survival is bound to be a crucial variable. We investigated the effect of different median survival, progression free survival, and survival with "good prognosis" on the model. The results are shown in figure 29, 30 and 35 (p.128-9, 132)
In their discussion (1.61) the authors say the trials reviewed are variable in quality. This does not of course mean that they are necessarily poor quality, they may be variably good! Later they say that 'the impact of specific tumour type needs to be explored further'. They are indicating here separation according to MGMT status. Whilst I certainly agree with this, until it is possible reliably to separate out tumour types which benefit most (and currently it is not), it might be unreasonable to deny a mixed and currently inseparable population access to treatment from which a significant sub-population might benefit, simply because the other population may not.	No comments this is a decision for NICE. However, without better identification of patients likely to benefit from treatment, patients unlikely to benefit may be given it and risk adverse effects.
Page 6 I find statements such as 'evidence for effectiveness of TMZ is	The evidence base is 2 RCTs and 2 case series. This is limited.

limited' of little use. All evidence is limited!	
Page 10 They say that grade I and II tumours are low-grade, slow growing and unlikely to spread. This is simply untrue. Low-grade tumours may infiltrate widely, that is they spread avidly and widely in the brain.	I have clarified this statement in the report to say "grade I and II tumours are unlikely to metastasise."
Page 13 They attach a degree of certainty to the MGMT story that may not be justified. Statements such as MGMT activity will be decreased or absent when the promoter is methylated, offers a degree of certainty that is not yet established from the research. More generally on this issue the authors here are remarkably accepting of the Hegi paper and the potential implications. This study was performed *retrospectively* on a *minority subset* of patients from a *few, selected institutions*, using an assay *which is not validated for clinical use* and which on her own admission is difficult to reproduce. The relationship to MGMT promoter methylation to outcome needs to be validated prospectively before any clinical reliance can be placed on it. (also see remarks under page 5 above).	We have added some more circumspect language to these paragraphs.
Page 13 Again minor errors, high-grade glioma is not associated with tubero sclerosis. Neither are there excess high-grade gliomas in immuno-compromised patients or those with AIDS. Errors like this (which I am sure were not made by their expert advisors) show their naivity when straying from their own fields into clinical areas.	These statements are referenced to a standard text book. However, they can be removed.
Page 16 Statements such as 'the brain and spinal cord are particularly sensitive to radiotherapy' show a rather facile knowledge of the area and are clearly lifted from an undergraduate textbook. They can actually be highly tolerant in the acute situation.	These statements are referenced to a standard text book but can be removed.
Page 21 Whilst the authors criticise heavily the trial work performed in patients with glioma, they are remarkably uncritical of the work of Elizabeth Davis et al with respect to patient views and relatives attitudes. There is no criticism of methods or statistics and no criticism of the environment in which these data were obtained. The conditions in which these patients were managed may not have reflected optimal management conditions nor indeed the generally accepted standard of today.	This information is included as part of the background and as such is not subject to the rigorous quality assessment used with trials included in the systematic review. Methods used in the study are, however, described and results are presented as part of a section looking at quality of life of patients with terminal brain tumours. Qualitative methods are particularly appropriate for investigating patients experiences.
Page 27 Inclusion criteria for the Temozolomide study did not include grade 3 tumours intentionally. Hence the statement under the heading population is erroneous. If grade 3 tumours were entered these were done on the basis of a local pathology report of GBM subsequently altered or a protocol violation.	Inclusion criteria are *pre specified* in line with NHS centre for R&D report no. 4 to prevent bias in the systematic review. Studies meeting these criteria are to be included.
Page 29 External validity Much is made of the generalisability of the data presented here. The presumption is that there is a desire to generalise these findings to all patients with glioblastoma and this may not necessarily be the case. I think no-one is suggesting that the results	Our consideration for generalisibility are clearly outlined on p.29. Presenting sufficient data on patient population to allow the clinician to extrapolate to relevant populations is one such criteria.

of concomitant adjuvant Temozolomide should be applied to patients whose characteristics lie way outside the recruitment characteristics for the trial. For example, a 75 year old man with an unresectable glioblastoma and dense hemiparesis would clearly not be a candidate for any treatment, let alone concomitant chemo radiotherapy. Neither would you seek to generalise the gliadel data to inoperable patients, this would be frankly silly! I feel the Peninsula group would have been better spending their time looking at those groups definable within the study who *might benefit*, rather than try to generalise to those groups outwith the study who might not.	
Page 60 The group criticised the Temozolomide trial for its lack of blinding suggesting that this may lead to selective post-trial treatments, which could lead to bias. I would suggest that even if the trial had been blinded, insistent of maintenance of blinding after the trial so that treatment decisions could be made independent of this would be both inappropriate and unethical. Furthermore, the trial reflects what would be done in routine practice. Also since more chemotherapy was given at relapse in the non-experimental arm, this should work to lessen any difference between the groups and gives more credibility to the study rather than less.	We do not suggest unethical practice, merely note that lack of blinding may lower the ability of the trial to avoid systematic biases. We agree that chemotherapy given at recurrence may make it difficult to isolate the impact of TMZ given in newly diagnosed gliomas.
Page 86, Paragraph 2 The logic here is difficult to understand. Patients in the control group do receive **more chemotherapy** and it is **more expensive** and this is what happens in real life. Hence it could be said upfront that treatment with radiotherapy and Temozolomide obviates treatment with chemotherapy at a later stage and **reduces costs**. This is what really happens, it is difficult to understand how a reduction in chemotherapy later can be considered to underestimate the costs of radiotherapy-plus-Temozolomide.	See attached sheet
Throughout the document, great emphasis is given to the value of QALY in estimating the worth or value of a treatment. Whilst this is a concept which might have great credibility amongst health economists, it may not reflect what either clinicians or patients consider as most important. We have then to accept this document from the point of view of health economists, which may not reflect the view of other groups in society.	Cost per QALY is the outcome preferred by NICE in economic evaluations of treatment. We have also provided cost per life year gained for information (see p. 139)

Page 87 An assumption is made that the post progression costs between two arms in the Temozolomide study are equal. This is not reasonable. Since we see that clinicians left to their own devices use less chemotherapy in the Temozolomide arm and hence the post progression costs are reduced in this arm. Furthermore the model takes no account of the fact that a patient living longer in a disease stabilised state, may be able to *contribute* to society, continue employment etc. This is not a fanciful notion. Glioblastomas tend to affect the higher social class patients, many of whom can continue to work in managerial or other capacities for a period following treatment, no account of this is taken in the model. If one took only health costs into account the longer a patient was kept alive the less value it would have in this model and cure would be disastrous!	See attached sheet. The model takes the perspective of the NHS as required by NICE guidelines.
On **page 92**, the group argue for a time independent risk of death model rather than a state dependent risk of death. Their argument is persuasive, but I wonder if it holds true for a dual population such as probably exists for patients with glioblastoma (viz MGMT +/-).	We have explored the impact of this assumption in sensitivity analysis (figure 16, p.115) and there appears to be little impact on the model results.
Page 93 It appears that the model is heavily dependent on median survival time since that is the match that underlies the model. Is this justified when the question being asked concerns two year survival rather than median survival? I note that the fit of the model is weakest in the tail, which is the most interesting part clinically. In paragraph 4, (page 93) there is a statement that they have used data from a review of peri-operative deaths during craniotomy for glioma. Since a decision to use Temozolomide is made after surgery and hence that decision process excludes any patients who have died pre-surgery, what is the justification for this?	The weakness in longer term data is due to the small amount of available evidence on which the model is based. For completeness, both the gliadel and the TMZ model begin with the initial patient treatment – surgery, and radiotherapy. As both arms in the TMZ trial have the same costs and risks associated with this, it is unlikely to affect economic evaluation of TMZ which is based on incremental affects.

Dr Erridge, NHS QIS	
Main concerns:	
the over emphasis on the inclusion of a small number of grade III patients in the EORTC-NCIC trial, this is simply a fact of life when treating this illness. Pathologists have differing opinions on the exact diagnosis. The important issue is that the local pathologist, on whose opinion the management decisions are made, thought the lesion was a GBM.	The problem is that the small number of chemosensitive tumours included in the trial may substantially drive the results. This may lead to a large number of people being treated with no hope of effect to capture a small number of people benefiting to a large degree. We agree that a pragmatic trial design may be a good thing. However, generalisibility is unclear since we do not know that local pathologists will diagnose similar proportions of grade III tumours to grade IV tumours as those in the trial.
the utility calculation grossly over-estimate the impact of this treatment on the patients quality of life. Patients in the trial with grade 3 or 4 toxicity may have only experienced such severe side effects for a short period so it is inappropriate to assume such a low utility value for patients in the RT+TMZ arms of the study. Though the QOL has not yet been published in full (in press) there was not significant difference between the study arms.	Note that the values quoted are only for the minority of patients who do experience adverse effects due to treatment with TMZ (see table 38 and para 1 p.98) As stated above – The impact of utility values was investigated in sensitivity analysis. See Figure 31 (p. 129) -Threshold analysis shows that TMZ is not cost-effective (at WTP <£30K) even if the utility value in the stable state were 1 (=perfect health). Overall we felt that the utilities obtained were high for a terminal disease.
Costs of treatment at relapse are removed from the calculation, which is inappropriate as patients are less likely to receive TMZ again if they have received it in the adjuvant phase. Whereas those who have not received it during this time period are highly likely to do so.	See attached sheet.
Though the 'industry' cost effectiveness model undoubtedly has some problems particularly due to the censuring of the data after two years, it should be remembered that these data are based on actual patients who have received the study medication therefore their data should be given greater weight than a theoretical model. It would be useful for the reviewers to see this original report and for these data to be applied to the PenTAG model.	Our model is also based on survival data from trials of actual patients receiving investigated drugs. (see description of the data used for transitions in section 5.5.1, p. 92)
The total costs have apparently been based on all patients receiving this treatment, whereas in reality less than 50% of patients presenting with a GBM will be suitable for this treatment. The biology of the disease, particularly in the elderly, means that it will not be used out with a clinical trial setting in older and less fit patients	This is appropriate. A cost-utility model examines the incremental costs and benefits for patients suitable for a treatment who do receive it and who do not.
Specific comments	
Section 3	
Guidelines do exist, they were published by the Royal College of Physicians in around 1998.	As we state "evidence based guidelines are few" (3.2.1, p.15) – this is the case. The RCP guidelines (1997) will be referred to, but these are now eight years old.
Radiotherapy – this treatment is generally well tolerated and as evidenced	No comment.

by the control arm of the EORTC-NCIC trial which demonstrated a 4.9% grade 3 or 4 fatigue but all other grade 3 or 4 toxicities occurred in less than 1% of the population.	
TMZ costs – these apparently assume that all patients will receive the TMZ rather than a more realistic around 50% of patients.	See last sentence on p. 25 – we do quote estimated costs based on half of patients being eligible for TMZ.
Section 4 Systematic review	
Randomisation – the large TMZ study was run by the EORTC an internationally renowned clinical trials organisation with a standard method of telephone randomisation to which all investigators are blinded to the sequence of treatment allocation (computerised).	We base our assessment on methods as reported in the included trial reports.
The inclusion of 7-8% non-GBM patients is irrelevant and a sub-group analysis not required. There are multiple publications demonstrating that there is significant variability in the reporting of brain tumours by even highly specialist neuro-pathologists due to the subtlety of the features required for each diagnosis. Therefore, in any standard population to which this treatment will be applied, there will be a number of patients who may not have a GBM if the pathology were reviewed at another centre. For this reasons most large neuro-oncology centres have a consensus opinion for the final diagnosis.	TMZ is licensed for use in a population with GBM tumours. See our comments above – the generalisability of these results may be limited due to variation in the number of different types of tumour included for TMZ treatment by local pathologist diagnosis.
Bias	
Performance bias - the use of post-progression chemotherapy, the fact that more patients in the RT only arm (72% v 58%) received chemotherapy at progression would have actually reduced the impact of the trial medication.	As we state (p.60, 2nd para. of Performance bias) “unmonitored cross-over may confound evidence about the survival advantage for first line TMZ.” Reviews of the impact of chemotherapy for recurrent tumours have concluded that there is little evidence for many chemotherapy regimens, and very few RCTs. The impact of second line treatment is still uncertain although nitrosoureas and platinum based regimens may have some effect (see for eg Huncharek & Muscat. Anticancer research 1998; 18: 1303-1312)
Attrition bias – it is inevitable that more patients will withdraw from a treatment which lasts six months when compared to one that lasts six weeks. As the primary endpoint is survival such drop-outs are irrelevant.	High levels of attrition may be inevitable in this treatment area - this does not alter the fact that it will lessen the conclusiveness of any study results.
Blinding – it is impossible to blind a study with a myelo-suppressive agent against a placebo as any blood tests taken prior to the next cycle, or if the patient becomes unwell, will immediately unblind the investigators.	Again, blinding may be impossible in such a treatment area, this does not alter the fact that unblinded studies are more susceptible to bias than blinded ones.
Post-operative randomisation– it would not be ethical to randomise such patients pre-operatively. Though this was essential in the BCNU trial, it was not in the TMZ studies and it imperative that any patient offered entry into a clinical trial is in a sufficiently good clinical condition to undergo the study treatment. Only around 50-60% of patients with a pathological diagnosis of GBM (unpublished Scottish audit data) are sufficiently fit to receive such a treatment.	Our point about the randomisation point is two–fold, firstly patients in poorer condition will not be included, and secondly, estimated survival time comes from a later point than if the trial had measured survival from the point of surgery (this is a particular concern in indirect comparison and is one of the reasons, outlined in 4.9 p. 75, why we did not attempt this.)

External validity	
GBM in older patients is a different disease, with predominantly primary GBM with a more aggressive phenotype and hence a shorter survival. To subject such patients to a protracted course of radiotherapy, which would occupy the majority of their life expectancy, would be unethical. It is unlikely that this treatment would be used in the over 70's therefore their exclusion in this trial is appropriate.	Patients over 70 are included in the evidence base – the case series by Lanzetta contains patients aged 25-75. The point of the external validity section is to outline where trial data may not apply to clinical populations, which we have done.
These data cannot be applied to patients with Grade III tumours and indeed a number of follow-up studies by the EORTC, NCRN and other groups are proposed.	As we state (p. 62, last sentence) – we also question whether it can truly be applied to grade IV patients, given that it is not a confirmed grade IV population.
Outcome measures	
The calculation of overall survival and time to progression free survival from randomisation is standard practice in oncology trials.	No comment
Effectiveness	
As stated above, the inclusion of a number of patients felt at central review not to be GBM is irrelevant to everyday clinical practice as this will be inevitable.	See our comments above (p. 9)
The subgroup analysis according to MGMT status – only a proportion of the patients in the whole EORTC-NCIC study had this test performed, particularly it should be noted, none of the French patients (the test failed to work because of the method of tissue preservation). So the opinion of the EORTC Brain Tumour Group and other International experts is that this test cannot be currently be relied upon to select patients for TMZ – a second international study examining two different dose levels of TMZ and prospectively testing the impact of MGMT status is proposed and will open in 2006.	Thank you for this information. The identification of patient subgroups likely to respond well to chemotherapy remains important and we wait the results of the EORTC study with interest.
Toxicity	
The results are reported as per the studies and are within expected and acceptable frequency. The visual disturbance reported in both arms is likely to be due to steroids.	No comment
Comparison of BCNU and TMZ	
I agree that such a comparison would be hazardous and not particularly helpful.	Thank you. No comment.
Cost effectiveness	
The 'industry' cost-effectiveness study on the TMZ study was conducted by a well recognised university department in conjunction with the EORTC BTG. Inevitably the cost-effectiveness data were collected in mainland Europe, as few UK centres recruited to this trial. I agree that by only including the data for the first 24 months after randomisation the costs in the more expensive 'progressive phase' of the study group would have been excluded.	No comment
I am uncertain as to the reasoning behind PenTAG group's concern about	See attached sheet

including the costs of chemotherapy at progression, as this will inevitably be clinical practice. Currently those patients who have not had chemotherapy at presentation, receive either PCV or TMZ (centre dependant) or enter the BR12 study (comparing the two regimens) at progression, therefore it is important that these costs are included. Their second calculation after the removal of treatment at progression is therefore incorrect as this will not be the clinical picture. If patients are not given TMZ during the early phase of their illness, it is highly likely that it will be given at a later date, thereby reducing difference in costs between the two study arms. In addition, a patient who relapses within a year of adjuvant TMZ is unlikely to be treated again with the same drug as it would be effective. Therefore it is highly relevant to include these costs in the calculation of the costs for management of patients out with a trial setting.	
Other limitations – only data on 224 patients – this was not a commercially sponsored study but was conducted by the EORTC and NCIC. Therefore there were insufficient resources to collect health economic data across the whole population. Also collecting any data, including QOL data, is notoriously difficult in this group of patients, particularly towards the end of their life. A paper on the QOL data in this trial has been written up and will be published soon in Lancet Oncology.	See TAR section 5.3.5.5
PenTAG analysis	
The utility model assumes 18% of patients in the concomitant phase of their illness had nausea, vomiting and infections that might require hospital admission. This is an incorrect assumption. In the trial 0.7% of patients had grade 3 or 4 nausea and 3.1% grade 3 or 4 infection. Only such severe toxicity could necessitate admission to hospital. In addition, a patient maybe graded as having such a level of toxicity when it is present only of a single day. The utility of 0.74 therefore grossly over estimates the impact of this treatment on the patients quality of life. Similar over-estimates have been made of the adjuvant phase.	Note that the values quoted are only for the minority of patients who do experience adverse effects due to treatment with TMZ (see table 38 and para 1 p.98) As stated above – The impact of utility values was investigated in sensitivity analysis. See Figure 31 (p. 129) -Threshold analysis shows that TMZ is not cost-effective (at WTP <£30K) even if the utility value in the stable state were 1 (=perfect health). Overall we felt that the utilities obtained were high for a terminal disease.
Health care costs	
The other 'expert opinion' on potential healthcare costs appear reasonable. However, I am uncertain as to how many cycles of chemotherapy during the adjuvant phase of the treatment were included in the model. It is important to realise that in the trial only 50% of the patients received all six cycles of chemotherapy and careful assessment during this phase is mandatory to ensure progressing patients do not continue to receive this potentially toxic agent and hence significantly reducing health care costs. If the model calculates the proportion of non-progressed patients at each time point and therefore only allows such patients to continue this therapy this been taken into account, but it would be useful to know the median number cycles delivered to the theoretical population.	The PenTAG model does account for the fact that people cease adjuvant therapy, either by their tumour progressing or dying. In the base case analysis, in the TMZ arm of the model 80.8% of the modelled cohort undergoing surgery have survived (and not progressed) in order to start on adjuvant TMZ (cf. 78% in Stupp et al., Table 2). Only 43.7% of those who start TMZ received all 6 cycles of adjuvant TMZ (cf. 47% in Stupp et al., Table 2), with half receiving 5 cycles. As a proportion of all in the treatment (TMZ) arm of the model only 35.3% receive all 6 cycles of adjuvant TMZ.

Conclusion	
I am concerned that the health care costs collected from the actual trial have not been used in the models. I am uncertain as to the reasons behind this. Was the EORTC BTG approached for these data directly so they could be incorporated in the PenTAG model? If not, such an approach should be made before any final conclusions about the cost utility of this regimen are made.	We were cautious about using costs based on resource consumption details taken primarily from Canada, Germany and the Netherlands which may differ considerable from the UK (see 5.3.5.3, from p. 83 for other limitations). Thank you for the offer of access to such data.
BCNU wafers	
I have less concern about the analysis for the cost effectiveness of BCNU wafers as the data, particularly regarding the impact on survival, are much weaker.	No comment
General comments	
The imbalance in pathological type in the studies was unavoidable as the pre-operative diagnosis would have been a 'best guess' from the radiological appearances. Frozen section, on-table pathology, cannot provide a detailed diagnosis and can only identify whether or not the lesion is a high grade glioma. In addition, it was suggested that a separate analysis should be conducted examining the 1p19q of the anaplastic oligodendroglioma (AO). It should be noted that the chemosensitivity of patients with AO does not correlate as well for gene loss as it does for grade II oligodendroglioma.	Imbalance in tumour type may lead to biased results (see p.35, "selection bias")
As with the TMZ study the exclusion of patients over 65 is reasonable as patients over this age are infrequently fit enough to undergo a tumour resection and hence have wafers inserted.	See comments above – we report elements that may make the results not applicable to groups of patients in clinical practice.
For reasons stated above the survival analysis should be performed on the whole group, not just the GBM cases. However, any survival advantage identified by these studies is small and non significant by 12 months. Though it should be noted that even though the potential concerns with the non-protocol analysis, the FDA did feel there was sufficient evidence to grant a licence for the use of BCNU wafers in newly diagnosed patients	No comment
As the intervention appears to have minimal impact on overall and progression free survival it is unlikely to be a cost-effective intervention.	No comment
Obviously it is important in any cost effectiveness analysis not only to look at the costs of the intervention but the overall health care costs. This should have therefore been included in the 'industry' submission	As stated section 5.3.4.3 from p. 81.
The PenTAG model assumes a utility of 0.73 for patients in the study arm due to presence of seizures and blurred vision, however seizures were actually more frequent in the control arm (13.3 v 9.2%) and blurred vision a consequence of steroids. This is therefore an over-estimate of the impact of BCNU wafers on the patients quality of life.	The impact of utility values was investigated in sensitivity analysis. See Figure 19 (p. 117) -Threshold analysis shows that gliadel is not cost-effective (at WTP <£30K) even if the utility value in the stable state were 1 (=perfect health). Overall we felt that the utilities obtained were high for a terminal disease.
Conclusions	
With a very marginal impact on outcome, the conclusion that BCNU wafers	Thank you, no comment.

are not cost effective seems reasonable.	

Link submission	PenTAG response
Executive Summary	
Overall Survival: PenTAG intimate that carmustine implants have not been proven to confer a significant survival advantage. However an increase in median survival of 2.3 months compared to placebo was shown in the ITT group by a pre-specified, stratified by country, log-rank analysis (p=0.03). Even in the unstratified analysis, the p value for median survival was 0.079, close to the arbitrarily accepted p value of 0.05 for statistical significance. The absolute clinical benefit for carmustine implants, a 29% reduction in the risk of death, is independent of stratification. Additionally long term survival data for carmustine implant compared to placebo demonstrates a statistically significant (p=0.01) 5-fold increase in 3-year survival (9.2% vs. 1.7%) in favour of carmustine implants.	We have used a fitted Weibull curve in the economic model – this means that the apparent difference between survival in the two arms of the Westphal trial is treated as real, rather than due to chance. As Link state on p. 4 para 1 – this result is constant regardless of the which subsequent analysis is used – stratified or unstratified. However, we do note that according to FDA transcripts the stratified analysis was NOT specified in the protocol. (http://www.fda.gov/ohrms/dockets/ac/01/transcripts/3815t2.rtf). We believe that apparent differences are driven by a small number of long term survivors with chemosensitive tumours (see p. 43-44) of which there is an imbalance between the treatment and control arms in favour of gliadel. Long term survival data are based on very small numbers (11 patients in total alive at 2 years) and subject to tail effects as described on p. 44 of our TAR. No significant survival advantage was seen for patients with grade IV tumours (see table 10, p.48).
Progression Free Survival (PFS): PenTAG use the results for PFS as determined by radiological imaging as an indicator of symptom free survival. For the reasons outlined in our original submission and in this document, the use of radiological imaging for PFS is both inappropriate and inaccurate for patients with glioma who have undergone surgery and/or radiotherapy. We therefore submit that an alternative measure, specifically neuroperformance decline, is a more appropriate measure of symptom free survival in this group of patients and should be used.	We agree that accurately defining the threshold between stable and progressive disease may be difficult (p.148) In the Westphal trial, on which our cost-effectiveness study is based, PFS was defined by neuroimaging in 70% of cases (p.47). No significant difference in median PFS were found between the two arms. The trial also assessed time to decline across 11 neuroperformance indicators. When these are analysed using the stratified analysis, there appear to be differences in decline however, unstratified analysis shows no difference in time to decline in 10 of the 11 measures. (see p. 191). Figure 18 (p.117) shows that PFS advantage would need to be at least 20 weeks for gliadel to be cost-effective (at WTP threshold of £30K).
Price of carmustine implants: this has been incorrectly quoted as £687.50 per implant and should be corrected to the current cost of £650.38 per implant.	The quoted price was correct at the time of writing, and came from the BNF no 49. The lower price is in the BNF 50. As sensitivity analysis on p. 119 (Figure 21) shows – cost of gliadel would have to fall by 40% for gliadel to be cost-effective (at WTP £30K/QALY).
Therapeutic gap: patients must recover from neurosurgery before starting a course of radiotherapy and this time will vary. There is therefore a therapeutic gap after surgical resection and prior to radiotherapy, when there is nothing to halt tumour regrowth. Local delivery of chemotherapy with carmustine implants at this time bridges this therapeutic gap providing active cytotoxic treatment.	We discuss this gap in the discussion section of the report, p.149 but note that either chemo regimen may be used here.
Economic analysis: The modelling carried out by PenTAG, while sound in structure, is based on a number of assumptions which are simply incorrect	(I assume they are referring to median survival – as stated on p.12 of Link's full submission)

or contentious and which are very damaging to the estimated cost effectiveness of carmustine implant. In particular the estimated mean survival using a Weibull curve approximation underestimates the mean survival observed in the key trial, symptom free survival has been estimated inappropriately and the costing methodology adopted contradicts the views of NICE's own methodology committee. Correction of these errors and the pricing error halves the estimated cost/QALY shown in the PenTAG report.	Assessment of the goodness-of-fit of the Weibull curve is outlined in Appendix 12, and show that it is good, with an R^2 of 0.9986 for gliadel and a % error of just 5.75 (7.3 vs 10 weeks). Sensitivity analysis shows that gliadel is unlikely to be cost-effective at usual levels of WTP ever if the survival advantage were about 25 weeks. (Figure 17, p.116) We agree that there are difficulties in all methods of assessing PFS. Progression was defined by neuroimaging in 70% of cases in the Westphal trial so this is the measure we used in the model. No difference was seen, with both arms reporting 5.9 months median PFS (see table 9 p.47). Progression based on 11 indicators of sym[tom deterioration shows significant differences between arms using a stratified analysis but not the unstratified analysis. (see p.191) Sensitivity analysis shows that PFS advantage would need to be about 20 weeks for gliadel to be cost-effective at WTP of £30K (Figure 18, p.117). Costing methodology used by PenTAG incorporated costs related to high grade glioma and its treatment. This is recommended in existing guidelines for economic evaluations in health care (for e.g. Drummond et al, 2005, Gold et al 1996) We were unable to find any contradictory advice in NICE guidance. Much more controversy surrounds the inclusion of costs for health care aspects unrelated to the investigated condition and its treatment and we have not done this in our analysis.
We present additional cost effectiveness analyses in Section 5 of this document and ask that the Appraisal Committee consider them carefully in conjunction with their invited experts and not accept the PenTAG cost-effectiveness assumptions without serious consideration.	Given the inherent uncertainty in model parameters it is generally possible to contrive scenarios where an ICER estimate falls below a given threshold. The two changes to model parameters which cause the most substantive effect on the ICER value are an optimistic assumption of the effect of Gliadel in prolonging PFS (7.4 as opposed to the PenTAG estimate of 1.3 weeks), and a reliance on the original Link Pharmaceuticals estimates for the additional treatment costs of Gliadel versus placebo rather than the PenTAG estimates. We feel that these scenario analyses represent a series of connected assumptions in favour of Gliadel.
Other Issues:	
External Validity	Our comments on generalisibility stand – the trials exclude older patients and, as acknowledged by Link on p.15, thus it may be difficult to extrapolate finings to this population.
Blinding of Westphal study	As stated, the blinding of the trial is imperfect.
Imbalance of Grade III vs. Grade IV tumours	Our comment on the imbalance of tumour types between arms stands, although we acknowledge that definitive definitions are difficult. We simply point out that a central analyst suggested the imbalance may have been even higher than that reported. No further data is available.
Effects of placebo implants	The impact of placebo wafers on AEs remains unknown – the comparison

	undertaken buy Link cannot be definite and faces the usual problems of indirect comparison. In addition, AEs with placebo wafers are based on a very small number of patients.
CSF leaks	We have assumed that the number of CSF leaks is the same in both arms of the model (see table 39, p. 99). Therefore costs will apply equally to both arms.

Helen Neil – Royal Pharmaceutical formulary	**PenTAG response**
Section 3.4.1 (p23) the cost for an individual Gliadel wafer according to BNF 50 is £650.38	As stated above, BNF 49 was current at the time of writing.
Section 3.4.2 (p24) Temodar is incorrect; it should read *Temodal.* Also consider rephrasing the information on the licensing of temozolomide in children e.g. "TMZ is licensed for use in children 3 years and older" or "TMZ is unlicensed in children under 3 years"	Thank you - we have corrected this typo. Reworded to read “The TMZ licence excludes children under 3 years old.”

Geoff Saunders – British oncology pharmacy Association	**PenTAG response**
Clinical Effectiveness	
Concerns raised about inducing chemo resistance if these drugs are used early in the course of the disease appear to be based on comments from review articles, is there a strong evidence base for this opinion?	(I assume this relates to page 17) this relates to nitrosurea regimens and is taken from a recent review of the literature. (Brandes, 2003)
Product licence for temozolamide states that it is indicated for the treatment of newly diagnosed glioblastoma multiforme concomitantly with radiotherapy and subsequently as monotherapy treatment. As there is considerable interpersonal variation between the differentiation of grade III and grade IV tumours then inclusion of a small number of grade III tumors within the temozolamide study group is valid as this reflects reality.	No comment
Cost Effectiveness	
Utility values used are based on panel's perceptions of health states rather than patients own perceptions.	This is appropriate where a “societal” value of health is required.
There is a perverse incentive not to offer **any** treatment if cost of maintaining patients in progression-free state is dominant	Uncertain of the point here.
The sensitivity analyses for both assessments demonstrate that the models are particularly sensitive to most parameters, i.e. high levels of uncertainty	As we state on p.144.
Is it possible to predict whether measurement of MGMT expression would lead to a more cost effective use of temozolamide?	Currently there is no standard method of measuring MGMT activity so it is not possible to estimate costings.
Agree -results from economic model should be treated with extreme caution given the uncertainty in the model and about key inputs.	No comment