



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Field, KM;Fitt, G;Rosenthal, MA;Simes, J;Nowak, AK;Barnes, EH;Sawkins, K;Goh, C;Moffat, BA;Salinas, S;Cher, L;Wheeler, H;Hovey, EJ;Phal, PM

Title:

Comparison between site and central radiological assessments for patients with recurrent glioblastoma on a clinical trial

Date:

2018-10-01

Citation:

Field, K. M., Fitt, G., Rosenthal, M. A., Simes, J., Nowak, A. K., Barnes, E. H., Sawkins, K., Goh, C., Moffat, B. A., Salinas, S., Cher, L., Wheeler, H., Hovey, E. J. & Phal, P. M. (2018). Comparison between site and central radiological assessments for patients with recurrent glioblastoma on a clinical trial. *Asia Pacific Journal of Clinical Oncology*, 14 (5), pp.e359-e365. <https://doi.org/10.1111/ajco.12806>.

Persistent Link:

<https://hdl.handle.net/11343/293836>

Comparison between site and central radiological assessments for patients with recurrent glioblastoma on a clinical trial

Kathryn M Field^{1,2}, Greg Fitt³, Mark A Rosenthal^{1,2}, John Simes⁴, Anna K. Nowak^{5,6}, Elizabeth H Barnes⁴, Kate Sawkins⁴, Christine Goh¹, Bradford A. Moffat², Simon Salinas², Lawrence Cher³, Helen Wheeler⁷, Elizabeth J. Hovey⁸, Primit M Phal¹, CABARET/COGNO investigators

Author affiliations:

1. Royal Melbourne Hospital, Grattan Street, Parkville 3050, Melbourne Victoria, Australia
2. University of Melbourne, Grattan Street Parkville 3052, Victoria, Australia
3. Austin Hospital, 145 Studley Road, Heidelberg, Melbourne, Victoria 3084, Australia
4. National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney, NSW 2006, Australia
5. School of Medicine and Pharmacology, University of Western Australia, Crawley WA 6009 Australia
6. Department of Medical Oncology, Sir Charles Gairdner Hospital, Nedlands, Perth 6009, Western Australia
7. Royal North Shore Hospital, St Leonards, Sydney, NSW 2065, Australia
8. Prince of Wales Hospital, Barker Street, Randwick, Sydney, NSW 2031, Australia

Running title: Site and central radiology review comparison

Corresponding Author: Dr Kathryn Field

Department of Medical Oncology
Victorian Comprehensive Cancer Centre
305 Grattan Street, Parkville Vic 3000
Phone: +61 3 8559 7809
Fax: +61 3 85597739
Email: Kathryn.field@mh.org.au

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ajco.12806](https://doi.org/10.1111/ajco.12806).

This article is protected by copyright. All rights reserved.

Abstract

Aim

Assessment of magnetic resonance imaging (MRI) in glioblastoma can be challenging. For patients with recurrent glioblastoma managed on the CABARET trial, we compared disease status assessed at hospitals and subsequent blinded central expert radiological review.

Methods

MRI results and clinical status at specified time points were used for site and central assessment of disease status. Clinical status was determined by the site. Response Assessment in Neuro-Oncology (RANO) criteria were used for both assessments. Site and central assessments of progression-free survival (PFS) and response rates were compared. Inter-rater variability for central review progression dates was assessed.

Results

Central review resulted in shorter PFS in 45% of 89 evaluable patients ($n=40$). Median PFS was 3.6 (central) versus 3.9 months (site) (hazard ratio 1.5, 95%CI 1.3–1.8, $P<0.001$). Responses were documented more frequently by sites ($n=16$, 18%) than centrally ($n=11$, 12%). Seven of 120 patients continued on trial without site-determined progression for more than 6 months beyond the central review determination of progression. Of scans reviewed by all three central reviewers, 33% were fully concordant for progression date.

Conclusion

While the difference between site and central PFS dates was statistically significant, the 0.3 month median difference is small. The variability within central review is consistent with previous studies, highlighting the challenges in MRI interpretation in this context. A small proportion of patients benefited from treatment well beyond the centrally determined progression date, reinforcing that clinical status together with radiology results are important determinants of whether a therapy is effective for an individual.

Key words

Bevacizumab, carboplatin, clinical trial, glioblastoma, magnetic resonance imaging

Introduction

Glioblastoma is a devastating primary central nervous system tumor, with near-universal mortality.

Different patterns of progression and treatment-related effects are commonly seen on imaging, which makes radiological interpretation and formal measurement of disease challenging.^{1,2} Accurate assessment of disease status is critical for clinical trials, in which continuation or cessation of treatment depends on radiological findings in addition to the patient's clinical status. Progression-free survival (PFS), a commonly used endpoint in glioblastoma clinical trials, is calculated using the date at which radiological or clinical progression (or death from any cause without radiologically documented progression) occurs. PFS is important as a surrogate endpoint for overall survival and as an endpoint unaffected by subsequent therapy.^{3,4} Prolonged PFS can indicate clinical benefit and tumor stabilization, even in the absence of objective response.⁵

Clinical trials often include blinded, centralized review of radiological endpoints such as response and PFS, especially in open-label studies, where knowledge of the treatment arm may bias reporting by investigators. The Australian Cancer Network Clinical Practice Guidelines for Neuro-oncology recommend central radiological review owing to reported inter-observer variability in response assessment.^{1,6} However, centralized review does not generally occur in real time, and date discordance between site and central radiological reviews is common.^{7,8} Bias, for example, from awareness of treatment allocation, is not entirely removed with centralized review: if a site determines that progression has occurred and scans cease the central review necessarily censors the patient at this point.⁷ Therefore, the necessity of centralized radiology review in clinical trials has been called into question.

Magnetic resonance imaging (MRI) is the standard method for assessing glioblastoma radiologically.

Over the last decade, with the advent of targeted therapies such as bevacizumab, interpretation of MRI findings in glioblastoma has become more complex. Bevacizumab, a monoclonal antibody targeting

vascular endothelial growth factor (VEGF), may reduce contrast enhancement on T1-weighted MRI.⁹

This may result in an appearance of tumor response (pseudoresponse) on post-treatment imaging

through normalization of vasculature with restoration of blood-brain barrier integrity and a reduction in gadolinium enhancement.¹⁰ Hence, nonenhancing tumor progression has become a key component of tumor assessment, reinforced by recent Response Assessment in Neuro-Oncology (RANO) criteria which include assessment of T2/FLAIR MRI changes¹¹ unlike the previous neuro-oncology assessment method, Macdonald criteria.

In this study, we compared site-based assessment with central radiology review in Part 1 of CABARET, a phase 2 trial in which patients with recurrent glioblastoma were randomized to bevacizumab monotherapy or bevacizumab plus carboplatin.¹² At the time of trial design, RANO criteria had only just been published, and most clinical trials before this had relied on the Macdonald criteria in assessing disease.¹³ The CABARET trial provided a unique opportunity to compare central and site radiology reviews incorporating the new criteria. Our aims were to compare centralized radiology review with the site-determined date of progression and disease response assessment. We hypothesized that, given the specific neuroradiological expertise of the central panel of radiologists, central review may determine disease progression earlier than site reviews.

Materials and methods

Patient eligibility

Eligibility criteria for the CABARET study have been reported.¹² Briefly, 120 consenting adult patients with recurrent glioblastoma from 18 sites across Australia were randomized and received at least one dose of treatment on trial. Inability to undergo contrast-enhanced MRI was an exclusion criterion. Measurable disease was defined according to RANO criteria: at least one site of bidimensionally measurable disease with two perpendicular diameters of at least 10 mm, visible on two or more axial slices that were preferably at most 5 mm thick with 0 mm skip between the slices on the baseline MRI.¹¹ At least 12 weeks had to have passed since completion of radiotherapy to avoid the period in which pseudoprogression is most commonly observed. Patients without

measurable disease (after postsurgical resection for recurrent disease) were eligible to participate, but their best response could only be stable disease, in line with RANO criteria.

Patients were excluded from the central and site radiology analysis described here if they did not have any scans after the baseline scan; if death occurred before radiological progression was determined; or if clinical progression occurred without either site or central radiological progression. If the date of progression was not available for either site or central review for an individual, that patient was not included when the time difference between site and central review was calculated.

MRI assessment timing

MRI was completed at baseline and every two treatment cycles (approximately every 8 weeks). On-study MRI scans were performed, where possible, within 7 days prior to the patient's scheduled treatment.

In follow-up, MRI was conducted every 2 months for patients without objective radiological disease progression according to the site (for example, patients who had withdrawn owing to toxicities). After disease progression, further MRIs were not required.

MRI acquisition

Each site was provided with an acquisition protocol, including details for scan standardization, to facilitate central radiology review (Supplementary materials). Standard imaging of the brain consisted of precontrast and postcontrast enhanced T1-weighted imaging (T1C+), T2, and T2 FLAIR.

Consistency of consecutive MRIs for individual patients was also ensured wherever possible (for example, using the same 1.5T or 3T scanner).

MRI interpretation

RANO criteria were used for disease assessment, with T2/FLAIR changes specifically documented.¹¹

For patients with measurable contrast-enhancing lesions at baseline, the sum of the products of the largest perpendicular diameters for the tumor in the axial plane was calculated and recorded. If there were multiple lesions, all lesions were measured, to a maximum of five. If the largest lesions could not be measured reproducibly, the next largest lesions that could be measured reproducibly were

selected. For index lesions, the products of the diameters of the index lesions (maximum of five lesions) were added together to determine the sum of the products of the diameters (SPD). In cases of progressive disease based on the index-lesions evaluation, the SPD of the nadir was used to calculate the increase in size of the index lesions. In addition, on subsequent imaging, unequivocal new enhancing lesions of any size, not within the original tumor volume, were recorded as new lesions (representing progressive disease).

Worsening overall performance status and/or neurological decline considered to be attributable to underlying disease, without lack of radiographic evidence of progressive disease and/or an increasing dose of steroids, were incorporated into assessment of progression. Such patients were classified as having clinical progressive disease if they met the RANO criteria for clinical progression. In the case of clinical progressive disease, MRI continued where feasible until the time of documented objective radiological progression.

Site review

The site reviews were also according to RANO criteria; ultimately the site investigator was responsible for documenting and interpreting MRI results although this assessment was facilitated by radiology reporting of the scans at the site. These occurred at the time of clinical review and were used to make decisions about disease response, patient care, and study treatment continuation. It was possible for a patient to continue on Part 1 of the study with site-assessed stable disease or response, even if subsequent central review determined that progression had occurred earlier.

Central review

Central review occurred later and was used to report the endpoints of response and PFS for the primary efficacy results of the trial. Scans were reviewed at the Brain Imaging Laboratory at Royal Melbourne Hospital¹⁴ by a panel of three experienced neuroradiologists (GF, PP, and CG) blinded to treatment allocation. All reviewers had been trained in the RANO criteria; test reviews were conducted on a subset of five patients by all radiologists, where results were compared between reviewers to determine consistency.

Scan data were anonymised and uploaded into a graphical user interface designed and implemented in Matlab (Mathworks MA, USA). Radiologists were trained in the use of this interface, which enabled measurement of enhancing disease and assessment of T2/FLAIR signal abnormality. The system subsequently calculated the SPD and allowed for categorization of the radiological response to the treatment.¹⁵

Each patient was randomly allocated to a primary radiologist (Reviewer 1) by the trial statistician. Radiologists were required to interpret the full series of scans in sequence for each of their allocated patients. Ten percent of patients ($n=12$) were reviewed by all three radiologists, to determine inter-rater reliability. Within the series of scans assigned to each radiologist, 10% of the reviews ($n=12$) were duplicated to determine intrarater reliability. The duplicated scans were allocated a dummy patient ID to keep the reviewer blinded.

For all published analyses of study outcomes and for the comparison of site and central assessments described here, the primary review (Reviewer 1's assessment of disease status) has been used, even if there was discordance between the three reviewers for duplicated reviews. Scans were not re-reviewed if discordance was noted, as this analysis (inter-rater reliability) was conducted after the primary endpoint had been reported and as an exploratory trial objective.

Scan series were reviewed by the radiologists in time sequence, with the exception of a 4-week research-only MRI, which was reviewed last and is not reported here. All scans were reviewed with blinding to both treatment allocation and site assessment of response or progression. Scans were presented to the reviewer one at a time, so the reviewer was not aware how many scans in total the patient had undergone, in order to remain blinded to the site's determination of progression date. The trial statistician combined the radiological assessment with the site-reported steroid dose and clinical status (both of which are components of the RANO criteria) to determine the overall disease response.

Statistics

Given that there was no significant difference between arms for the centrally determined primary endpoint, PFS,¹² we used the entire cohort when comparing the site-determined PFS date with the

central review-determined date. The Kaplan-Meier method was used to describe PFS times.

Proportional-hazards regression was used to calculate the hazard ratio (HR) and 95% confidence

interval (CI) to compare rates of progression determined by the two sources, taking into account the paired nature of the data. Where the progression date differed between site and central assessments, the type of progression determined by central review (T1C+, T2/FLAIR, new lesion, clinical only, or combination) was recorded to determine whether a particular type of radiological change was detected discrepantly. To compare responses, the McNemar-Bowker test of symmetry was performed and a kappa concordance statistic calculated.¹⁶ All analyses used SAS version 9.3 (SAS Institute Inc., Cary, NC, USA).

Results

Between November 2010 and March 2012, 122 patients were enrolled. Two withdrew consent and did not receive any treatment after randomization. The trial closed in December 2014 with two patients remaining on Part 1 of the study, who continued to receive treatment with bevacizumab beyond the date of study closure.

Data for 89 of 120 patients who were randomized and received at least one treatment were available for analysis. Of the remaining patients, nine had no scans after the 4-week (research only) MRI, seven died without documented site or central radiological progression, and 15 had clinical progression only, without site or central radiological progression. Ten of the 89 patients were deemed by the site to have progressive disease, but not yet by central review when they were censored because other anticancer treatment had commenced; these are included when testing for concordance between site and central response and progression but not for time-sensitive analyses.

Intra- and inter-rater variability

For intra-rater duplicates, all scans reviewed twice were consistently given the same date of progression by the same radiologist. For inter-rater comparisons, all reviewers agreed that progression had occurred at the same date or had not occurred for four of 12 patients (33%). For another four patients (33%), all agreed that progression had occurred but the date of progression differed for one of

the three reviewers: three with a difference <2 months (that is, one scan period) and one with a difference >6 months. For the other four patients, one ($n=2$ patients) or two ($n=2$ patients) of the three radiologists, but not all three, concluded that progression had occurred.

Progression-free survival

Differences in PFS date are shown in Table 1. Concordance in PFS date between site and central review was achieved for 37% of participants. The PFS date was earlier for 45% by central review and earlier for 18% by the site review. Where a discrepancy occurred, the median difference was 1.84 months, which is the approximate difference between the timing of MRI examinations (every 8 weeks). However, when central review determined earlier progression, the discrepancy was up to 24.6 months.

Where the date was available for both site and centrally determined progression and was discordant ($n=46$), the median PFS was (central versus site): 3.6 months vs 3.9 months, HR 1.52, 95% CI 1.28–1.81, $P < 0.0001$ (Figure 1).

The reasons for progression where centrally determined progression was earlier were heterogeneous and are shown in Table 2. The most common reasons were enhancing disease progression and T2/FLAIR progression. Less common reasons were a combination of T1C+ and T2/FLAIR changes and/or a new lesion.

Of the 33 patients whose central and site dates of progression were concordant, for 23 (70%) the type of radiological findings that determined progression were discordant. Of these, five were determined to have ‘clinical only’ progression by the site, but radiological progression was determined by the central review at the same MRI scan date.

Response

Comparison of the site and central best responses is shown in Table 3. Although we hypothesised that responses would be documented more frequently by site than central review, tests of symmetry showed no evidence that documented responses differed between site and central assessments ($P=0.8$). Overall, for 53 (60%) there was agreement for best response and for 36 (40%) disagreement

(weighted kappa statistic 0.30, indicating fair agreement overall). For six patients (7%), the central review determined partial response and for one patient a complete response where sites documented stable disease as best response; and for twelve patients (13%) the sites determined either a complete or partial response where the central review documented stable ($n=10$) or progressive ($n=2$) disease as the best response.

Patients continuing treatment beyond date of central progression

Seven (6%) of the 120 patients continued to receive study treatment on the basis of the site determination of stability or response for more than 6 months beyond the progression date determined by central review. The median age of continuing patients was 56 years (range 47–65 years), similar to the median age of 55 for all study patients. Two of these patients continued on bevacizumab monotherapy with compassionate access to the drug beyond the date of study closure and data collection (December 2014); at the time of study closure the sites had not determined radiological or clinical progression in these two patients. The central reviewers' progression determination was based on T2/FLAIR change alone for all seven patients. Figure 2 shows examples of patients whose date of progression at central review preceded the date of progression at the site by at least 6 months.

Discussion

In this comparison between site and central radiological review of MRI for patients who received bevacizumab on a clinical trial, we found that central review determined progression earlier than the sites for a substantial proportion (45%) of patients, but on average the sites determined progression at the time of the next trial scan (median 1.9 months). The proportion with discrepancies is similar to that previously reported.⁸ Although the rate of progression was approximately 50% higher when determined by central review (HR=1.52), the difference in median PFS of only 0.3 months is not clinically meaningful. This is also in keeping with available literature indicating that while centralized review may result in discrepancies at the patient level, it often does not substantially alter an overall PFS endpoint at trial level.^{7,17,18}

While intra-rater consistency was absolute for the small number of image sets reviewed twice, there was obvious inconsistency between all three radiologists on the date of progression. This finding is not unexpected: the pivotal BRAIN study of bevacizumab with or without irinotecan used an independent response facility for MRI assessment, requiring two radiologists to read all MRIs, and documented only approximately 50% concordance between reviewers. For discordant results, a third radiologist reviewed scans; further discordance with both earlier reviews was noted in 12%.¹⁹ The lack of consistency is also apparent when comparing the type of radiological progression documented between site and central reviewers for scans with concordant progression dates: most (70%) were different radiological reasons for progression despite the same date of progression being recorded. This exemplifies some of the challenges in MRI reporting, interpretation and application of the RANO criteria.

One of the most thought-provoking aspects of this study is that seven patients were deemed not to have progressed by sites and continued treatment for at least six months beyond the date at which the central review determined progression. All were determined to have progressed on T2/FLAIR changes alone. Clearly, these patients derived a durable survival benefit from therapy. If central review had been used to make decisions about continuing or ceasing therapy, these patients would have completed treatment much earlier. The authors of the RANO criteria themselves indicate that if there is uncertainty regarding progression on imaging, the patient may continue on treatment; if subsequent evaluations indicate progression, then the progression date is backdated to the time point at which the ambiguity occurred.¹¹ Even in the setting of a clinical trial, there is some degree of flexibility with treatment continuation without compromising the final determination of progression date. Based on our experience with CABARET, we concur that this is an appropriate strategy, in particular for T2/FLAIR signal changes alone without other evidence of progression on imaging, such as contrast-enhancing disease or new lesions, and without clinical progression.

The results of our study highlight some challenges in using the RANO criteria. For example, in analyzing the images, the first MRI study was used to determine the plane showing the largest tumor

dimensions. This plane and location were used for subsequent measurements of the tumor, so there are limitations related to the slice position and image plane selected initially and the tumor growth

pattern on subsequent imaging, given that analysis was based on the sum of the products of the perpendicular diameters. This was particularly relevant in the nonvolumetric-acquired sequences.

Furthermore, this made it difficult to account for tumor enlargement in different planes on subsequent studies. Another difficulty related to measuring enhancing disease in cystic or necrotic tumors, in terms of confidence in the formal measurements of active tumor. At the time of the trial, RANO criteria were new to both radiologists and clinical investigators; this trial was one of the first to apply and use the criteria, so it is likely a learning curve existed for their application.

An additional challenge is differentiating between disease progression on FLAIR images and treatment-related FLAIR signal change. Typically, in treatment-related leukoencephalopathy the signal change develops gradually in white matter, is relatively ill defined, and has associated volume loss rather than expansion. FLAIR progression related to tumor tends to be more focal and expansile, and may include the overlying cortex. Over multiple scans the difference between the two generally becomes clear, but may be harder to see when comparing only two scans. Figure 2b exemplifies this challenge. Additionally, serial scans may show a clear focal region of expansile FLAIR progression that may nevertheless remain smaller than the larger area of treatment-related signal abnormality, and so could be missed by the measurement criteria. An additional pitfall relates to the slice position in nonvolumetric FLAIR acquisitions, which may not truly reflect the change in FLAIR status.

Should central radiology review be mandated for all clinical trials in glioblastoma? Centralized review does not eliminate inconsistencies in reporting, as exemplified by this trial, and is an additional expense and not in real time, bringing limitations in implementation. One proposed strategy is to use a blinded central audit of a proportion of scans to determine whether there is any evidence of evaluation bias.²⁰ Ultimately, while there may be discrepancies at the patient level, it is at the population level that the efficacy (and effectiveness) of an intervention is determined. If minimal differences at this group level are established, then central review of every scan for every patient ultimately may not be

necessary. On the other hand, if very early evidence of response or progression is sought in clinical trials, then expert neuro-radiology review is ideal. As exemplified by this and other studies, however, even centralized review does not result in uniformity when determining radiological progression, given the complexities of MRI assessment of high grade glioma.

A limitation is that we did not re-analyse any scans where discordance was noted, whether between site and central dates, or between different central reviewers' dates. This would not have changed the primary endpoint of the trial (where central reviewer 1's assessment was always used, and analysis for discordance occurred later); and was beyond the trial scope for this exploratory endpoint.

Several other limitations are worth noting. The sample size and power calculations for the trial were not designed to formally address this research question, which was an exploratory endpoint. Only 89 of the 120-patient cohort were ultimately formally compared, the remainder not comparable for a variety of reasons. Differences in time to progression could only be compared in 79 owing to censoring before central determination of progression. However, all clinical trials in recurrent glioblastoma include patients who progress or die without radiological changes or who are censored for various reasons; therefore we do not believe these findings are likely to differ from those of other studies.

In summary, central radiology review of scans from patients on the CABARET clinical trial resulted in fewer complete or partial responses and a shorter duration of PFS, although ultimately the difference in median PFS of 0.3 months was small. Considerable heterogeneity exists among radiologists when determining the type of radiological progression, even when there is concordance in progression date. In a minority of patients, treatment continued as a result of site reviews for more than 6 months, with durable clinical benefit, reinforcing the conclusion that clinical as well as radiological assessment of therapeutic benefit is relevant in the setting of recurrent glioblastoma.

Acknowledgments

The CABARET trial was conducted under the auspices of the Cooperative Trials Group for Neuro-Oncology (COGNO), coordinated at the National Health and Medical Research Council Clinical Trials Centre, University of Sydney, and supported by Roche Products Pty Limited (Australia). Rhana Pike, from the Clinical Trials Centre, assisted with the manuscript.

CABARET was prospectively registered with the Australian New Zealand Clinical Trials Registry (ANZCTR), ACTRN12610000915055 (anzctr.org.au).

This trial has been reviewed by the appropriate ethics committees and has therefore been performed in accordance with required ethical standards according to the Declaration of Helsinki

All persons gave their informed consent prior to their inclusion in the study

Independent Central Radiological Review Committee P Phal, G Fitt, C Goh

The following study sites participated in the CABARET study and randomized at least 1 patient

(Principal Investigator and site coordinator)

Monash Medical Centre, Victoria – R Freilich, I Arzhintar (17); Royal Melbourne Hospital, Victoria – K Field/M Rosenthal, L Garrett (16); Royal Prince Alfred Hospital, New South Wales – J Simes, A Byrne (13); St Vincent’s Hospital, Victoria – A Dowling, N Ranieri (11); Epworth HealthCare Richmond, Victoria – R Jennens, F Osmond (9); The Queen Elizabeth Hospital, South Australia – WK Patterson, A Phay (8); Calvary Mater Newcastle, New South Wales – F Abell, L Plowman (7); Austin Hospital, Victoria – L Cher, J Flynn (7); Prince of Wales Hospital, New South Wales – E Hovey, H Kilsby (6); Royal North Shore Hospital, New South Wales – H Wheeler, S Kirby-Lewis (6); Royal Adelaide Hospital, South Australia – N Singhal, S Smith and M Whelan (5); Royal Brisbane and Women’s Hospital, Queensland – P Inglis, A Ives (5); Sir Charles Gairdner Hospital, Western Australia – A Nowak, S Lobb (5); Port Macquarie Base Hospital, New South Wales – S Begbie, P Williams (4); Mater Adult Hospital, Queensland – Z Lwin/N Woodward, G Crosbie (1); Royal Hobart Hospital, Tasmania – R Harrup, L Pyszkowski (1); Launceston General Hospital, Tasmania – S Gauden, A Neville (1).

Funding

Roche Products Australia Pty Ltd and National Health and Medical Research Council (1037786).

Author Manuscript

References

1. Vos MJ, Uitdehaag BM, Barkhof F, et al. Interobserver variability in the radiological assessment of response to chemotherapy in glioma. *Neurology*. 2003; 60(5):826-830.
2. Mehta AI, Kanaly CW, Friedman AH, Bigner DD, Sampson JH. Monitoring radiographic brain tumor progression. *Toxins (Basel)*. 2011; 3(3):191-200.
3. Driscoll JJ, Rixe O. Overall survival: still the gold standard: why overall survival remains the definitive end point in cancer clinical trials. *Cancer J*. 2009; 15(5):401-405.
4. Hotte SJ, Bjarnason GA, Heng DY, et al. Progression-free survival as a clinical trial endpoint in advanced renal cell carcinoma. *Curr Oncol*. 2011; 18 Suppl 2:S11-19.
5. Lamborn KR, Yung WK, Chang SM, et al. Progression-free survival: an important end point in evaluating therapy for recurrent high-grade gliomas. *Neuro Oncol*. 2008; 10(2):162-170.
6. Australian Cancer Network. Clinical practice guidelines for the management of adult gliomas: astrocytomas and oligodendrogliomas: Cancer Council Australia/Australian Cancer Network/Clinical Oncological Society of Australia; 2009: http://www.cancer.org.au/content/pdf/HealthProfessionals/ClinicalGuidelines/Clinical_Practice_Guidelines-Adult_Gliomas-AUG09.pdf.
7. Dodd LE, Korn EL, Freidlin B, et al. Blinded independent central review of progression-free survival in phase III clinical trials: important design element or unnecessary expense? *J Clin Oncol*. 2008; 26(22):3791-3796.
8. Ford R, Schwartz L, Dancey J, et al. Lessons learned from independent central review. *Eur J Cancer*. 2009; 45(2):268-274.
9. Thompson EM, Frenkel EP, Neuwelt EA. The paradoxical effect of bevacizumab in the therapy of malignant gliomas. *Neurology*. 2011; 76(1):87-93.
10. Pope WB, Young JR, Ellingson BM. Advances in MRI assessment of gliomas and response to anti-VEGF therapy. *Curr Neurol Neurosci Rep*. 2011; 11(3):336-344.
11. Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol*. 2010; 28(11):1963-1972.
12. Field KM, Simes J, Nowak AK, et al. Randomized phase 2 study of carboplatin and bevacizumab in recurrent glioblastoma. *Neuro Oncol*. 2015; 17(11):1504-1513.
13. Macdonald DR, Cascino TL, Schold SC, Jr., Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. *J Clin Oncol*. 1990; 8(7):1277-1280.
14. Department of Radiology. Brain Imaging Laboratory. 2008(5 Jan 2017). www.melbourne-radiology.org/Brain_Imaging_Lab.html.
15. Salinas S LS, Drummond K, Barras C, Phal P, Desmond P, Moffat B. Validation of the RANO criteria for quantifying therapeutic response of human brain tumors using computer assisted medical diagnosis technology. *Proc. Intl. Soc. Mag. Reson. Med*. 23. 2015.
16. Bowker AH. A test for symmetry in contingency tables. *J Am Stat Assoc*. 1948; 43(244):572-574.
17. Amit O, Bushnell W, Dodd L, Roach N, Sargent D. Blinded independent central review of the progression-free survival endpoint. *Oncologist*. 2010; 15(5):492-495.
18. Zhang J, Chen H, He K, et al. Evaluation of blinded independent central review of tumor progression in oncology clinical trials: A meta-analysis. *Therapeutic Innovation & Regulatory Science*. 2013; 47(2):167-174.
19. Cohen MH, Shen YL, Keegan P, Pazdur R. FDA drug approval summary: bevacizumab (Avastin) as treatment of recurrent glioblastoma multiforme. *Oncologist*. 2009; 14(11):1131-1138.
20. Zhang JJ, Zhang L, Chen H, et al. Assessment of audit methodologies for bias evaluation of tumor progression in oncology clinical trials. *Clin Cancer Res*. 2013; 19(10):2637-2645.

Table 1: Progression-free survival date differences between site and central assessments

Progression-free survival date	<i>n</i> (%)†	Median difference (mo)‡	Lower/upper quartile (mo)‡	Minimum/maximum (mo)‡
Site first	16 (18)	1.8	1.6–2.1	0.8–3.7
Concordant	33 (37)	n/a	n/a	n/a
Central first	40 (45)	1.9	1.6–4.2	0.7–24.6

† Total *n*=89 includes:

n=75 where date available for both site and central progression

n=4 where date available for central progression, but patient had not progressed on site review at the time of study closure

n=10 where date available for site progression, but patient was subsequently censored at time of commencing other anticancer therapy before any central progression was documented

‡ Calculated for patients with date available for both site and central progression at time of study closure (December 2014)

Table 2: Type of centrally determined progression when it was earlier than site-determined progression (*n*=40)

Type of progression	Number	%
T1C+ only†	14	35
T2/FLAIR only	11	28
T1C+ and new lesion	6	15
T1C+ and T2/FLAIR	5	13
New lesion	4	10

†C+ = contrast enhancing

Table 3: Assessed best response ($n=89$)†

Central review assessment	Site review assessment			
	Complete response	Partial response	Stable disease	Progressive disease
Complete response	0‡	0	1	0
Partial response	0	4‡	6	0
Stable disease	2	8	37‡	7
Progressive disease	1	1	10	12‡

† $n=89$ where data available for both site and central best response

‡ Concordant

Response assessed as per RANO criteria, inclusive of T1C+, T2/FLAIR, clinical status and steroid dose

Figure legends:

Figure 1

Kaplan-Meier curve comparing site review (dotted line) and central review (solid line) for progression-free survival

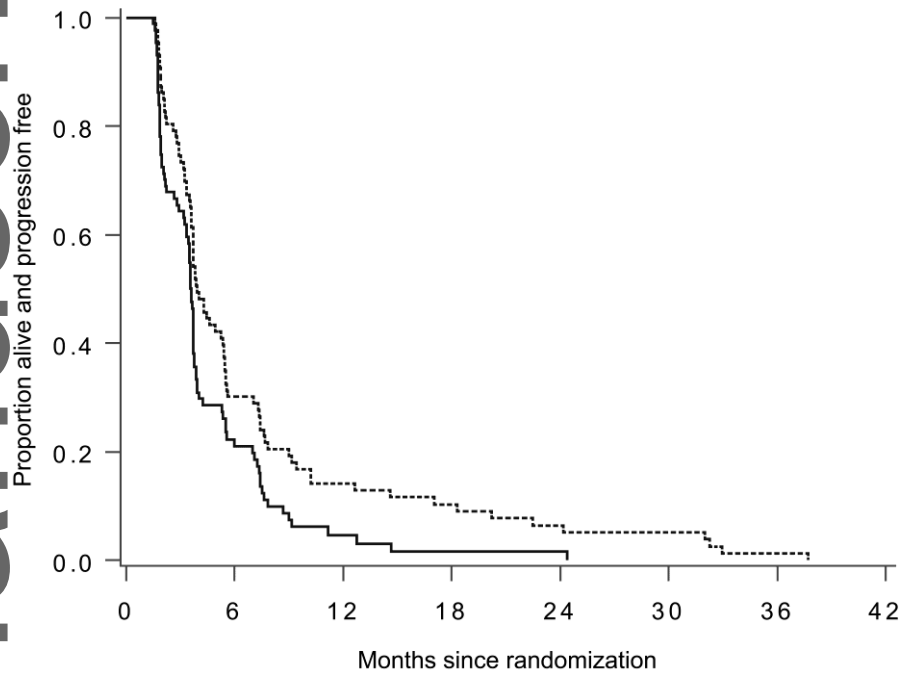


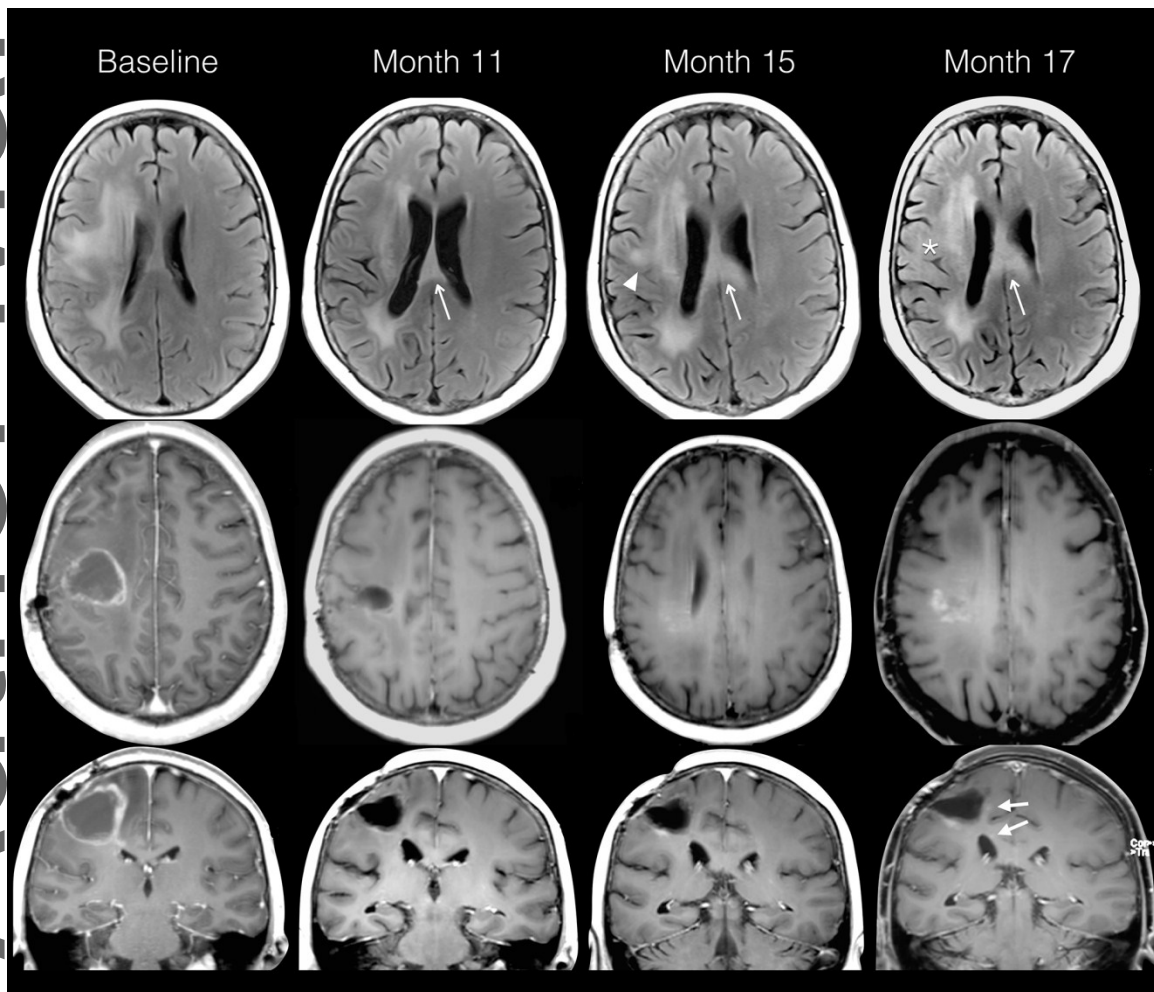
Figure 2

T1C+ (contrast enhancing) and T2/FLAIR MRI images for two patients whose centrally determined date of progression was >6 months before the site-determined date.

2a) At month 11, axial FLAIR, and axial and coronal contrast-enhanced T1-weighted sequences show almost complete response to treatment of the enhancing disease present on baseline imaging.

Between months 11 and 15, increasing expansile FLAIR hyperintensity in the corpus callosum (open arrows) and subcortical white matter (arrowhead) was deemed FLAIR progression at central review.

By month 17, when disease progression was called by the site, FLAIR hyperintense expansion within the corpus callosum had also extended to the right frontal cortex (*), and there was also increased enhancement inferior to the cavity plus new punctate foci of enhancement in the corpus callosum (closed arrows).



2b) Increased FLAIR hyperintensity about the resection cavity at month 2 (arrowheads) was called as progression by central review. However, this regressed on subsequent scans. FLAIR signal change throughout bilateral supratentorial white matter (*) slowly increased over subsequent months, but due to its ill-defined nature and lack of positive mass effect this was felt to represent treatment effect rather than progression. There was no recurrence of enhancing disease throughout the study period.

