BJC

Systematic review and meta-analysis of temozolomide in animal models of glioma: was clinical efficacy predicted?

T C Hirst[1], H M Vesterinen[1], E S Sena[1,2], K J Egan[1], M R Macleod*[,3] and I R Whittle[3]

[1]Division of Clinical Neurosciences, University of Edinburgh, FU204, Chancellor's Building, Royal Infimary of Edinburgh, Little France Crescent, Edinburgh, EH16 4SA, UK; [2]Florey Neuroscience and Mental Health Institute, Austin Health, University of Melbourne, Victoria 3084, Australia and [3]Division of Clinical Neurosciences, University of Edinburgh, BramwellDott Building, Western General Hospital, Edinburgh, EH4 2XU, UK

Background: Malignant glioma is an aggressive tumour commonly associated with a dismal outcome despite optimal surgical and radio-chemotherapy. Since 2005 temozolomide has been established as first-line chemotherapy. We investigate the role of in vivo glioma models in predicting clinical efficacy.

Methods: We searched three online databases to systematically identify publications testing temozolomide in animal models of glioma. Median survival and number of animals treated were extracted and quality was assessed using a 12-point scale; random effects meta-analysis was used to estimate efficacy. We analysed the impact of study design and quality and looked for evidence of publication bias.

Results: We identified 60 publications using temozolomide in models of glioma, comprising 2443 animals. Temozolomide prolonged survival by a factor of 1.88 (95% CI 1.74–2.03) and reduced tumour volume by 50.4% (41.8–58.9) compared with untreated controls. Study design characteristics accounted for a significant proportion of between-study heterogeneity, and there was evidence of a significant publication bias.

Conclusion: These data reflect those from clinical trials in that temozolomide improves survival and reduces tumour volume, even after accounting for publication bias. Experimental in vivo glioma studies of temozolomide differ from those of other glioma therapies in their consistent efficacy and successful translation into clinical medicine.

Malignant glioma is very aggressive and notoriously difficult to treat. With an annual incidence of around 5 cases per 100 000 it is relatively uncommon, but still comprises approximately 70% of malignant brain tumours (Wen and Kesari, 2008). Owing to their tendency to infiltrate brain tissue, gliomas are very difficult to treat surgically, and the vast majority recur after apparent 'complete' resection (Wen and Kesari, 2008). One of the first clinical trials (Walker et al, 1978) showed that glioma patients receiving 'optimal conventional care' (no radiotherapy or chemotherapy) had a median survival of 3 months and a 1-year survival of 3%. Various trials between 1970 and the 1990s supported the use of

radiotherapy, but existing chemotherapies were of limited use. Despite showing some promise in animal research (Amarasingh et al, 2009), the nitrosoureas BNCU and CCNU had little clinical efficacy. Indeed, it took a large meta-analysis including >3000 patients (Stewart, 2002) to show a small increase in survival between radiotherapy with adjuvant chemotherapy (46% at 1 year, 20% at 2 years) and radiotherapy alone (40% and 15%, respectively).

Temozolomide is an oral alkylating agent, a prodrug for 3-methyl-(triazen-1-yl)imidazole-4-carboxamide, whose anticancer activity was first described in 1987 (Stevens et al, 1987). In

*Correspondence: Dr MR Macleod; E-mail: Malcolm.macleod@ed.ac.uk

2005, a landmark phase III trial (Stupp *et al*, 2005) described the efficacy the new therapeutic agent temozolomide: when given together with radiotherapy and as adjuvant therapy there were significant increases in median survival (14 *vs* 12 months, $P < 0.001$), 2-year survival and progression-free survival compared with radiotherapy alone. Since then, temozolomide has emerged as the first-line chemotherapeutic agent for the treatment of malignant glioma (van den Bent *et al*, 2006).

Although temozolomide was being used clinically for melanoma in the mid-1990s, supported by substantial experimental data (Bleehen *et al*, 1995), there were no such animal data on temozolomide for brain tumours. The efficacy of temozolomide in glioma was actually first reported in patients somewhat anecdotally (O'Reilly *et al*, 1993), with a more structured study by the same group published in 1996 (Newlands *et al*, 1996). The first animal work in the area was published in 1994 (Plowman *et al*, 1994), and by 1998 and 2000 the first patients were recruited for phase II (van den Bent *et al*, 2003) and III (Stupp *et al*, 2005) trials, respectively. A preliminary literature search of PubMed suggests that by the end of 2000, only seven relevant experimental studies had been published. The basis for the decision to proceed to clinical trial is therefore not clear: possible factors include those experimental data, which did exist; the non-randomised clinical evidence of O'Reilly (O'Reilly *et al*, 1993) and Newlands (Newlands *et al*, 1996); and evidence from phase II studies testing temozolomide with other cancer types (Bleehen *et al*, 1995; Woll *et al*, 1995).

Systematic review and meta-analysis of the efficacy of nitrosoureas (Amarasingh *et al*, 2009) and gene therapy (Conlin, unpublished observations) for experimental glioma models have shown some efficacy, but neither led to substantial improvement in outcomes in human clinical trials (Stewart, 2002; Pulkkanen and Yla-Herttuala, 2005). Identification of differences in the experimental data for nitrosourea, gene therapy and temozolomide studies might therefore provide insight into translational challenges in neuro-oncology.

Here, we use systematic review and meta-analysis of experimental animal research to describe the evidence supporting the application of temozolomide in human glioma. We also compare the evidence available before and after the publication of Stupp's phase III trial (Stupp *et al*, 2005). We were particularly interested in evidence within the experimental temozolomide data that was predictive of its successful translation into clinical use. Secondary aims were (i) to compare temozolomide data from animal experiments before and after 2005 (the publication of the phase III trial); (ii) to seek evidence of the use of measures to avoid bias, and of publication bias; and (iii) to compare the efficacy of temozolomide in animal glioma models with that reported for nitrosoureas (Amarasingh *et al*, 2009) and gene therapy (Conlin, unpublished observations). Our hypotheses were that temozolomide would significantly improve outcome in animal models of glioma; and second that publication and expectation biases would lead to significantly higher estimates of efficacy in studies published after the Stupp trial.

MATERIALS AND METHODS

Information sources. We searched PubMed, EMBASE (from 1980) and Medline (from 1950) on the 4 August 2011, using these search terms: Temozolomide AND (glioma OR glioblastoma OR astrocytoma OR ependymoma OR oligodendroglioma OR 'brain tumour'). For EMBASE and Medline the search was limited to animals, and for PubMed we used the Hooijmans (Hooijmans *et al*, 2010) animal search filter.

Study selection. All abstracts and titles were screened by two authors (TCH and KJE) to include papers addressing

temozolomide therapy in *in vivo* xenograft studies. All publications in languages other than English were translated before screening. We included studies that met the following inclusion criteria: (1) temozolomide used as monotherapy; (2) animal model of glioma used; (3) intracranial or subcutaneous implantation of tumour cells; (4) outcome reported as either median survival or tumour volume; and (5) the number of animals per group was stated within the publication.

Data collection. We extracted data relating to experimental design and for each comparison, we recoded median survival or tumour volume in treated and control groups. When data were not clearly or fully described we contacted authors seeking the data; if no reply was received after 1 week these were excluded. Data were entered to the Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies (CAMARADES) data manager application in Microsoft Access 2003.

For experimental design we extracted data for species, strain and sex of animals; glioma cell type, tumour implantation site, tumour implantation method, number or volume of implanted tumour cells and method of tumour volume measurement (where relevant). Glioma models were stratified into groups (Table 1). In addition, we recorded the total dose of temozolomide (in $mg\,kg^{-1}$), delay to treatment (in days) and route of administration. When dose was given in $mg\,m^{-2}$, we converted it to $mg\,kg^{-1}$ using Freirich's (Freireich *et al*, 1966) conversion factor; where animals received continuing treatment courses, we calculated the total dose received by the time of median survival or the last time at which tumour volume was measured. Experimental starting point (i.e., day 0) was defined as the day of tumour implantation. When the delay from implantation to treatment was not given, we contacted authors seeking this information.

For each comparison, we extracted data for the number of animals in each group and either median survival or mean tumour volume, relative tumour volume or percentage change in tumour volume. For measures of tumour volume, we also recorded the s.d. or s.e.m. If any of these measures were missing, the publication was excluded. When outcomes were not quantified in the text, we measured outcome values from Figures using AVPSoft 'Universal Desktop Ruler'.

Table 1. Grouping of glioma models	
Group	**Model**
9L	9L
A172	A172
C6	C6, C6-Fluc
D54	D54, D54-Hif, D54MG
DF-1	DF-1
F98	F98
GBM	GBM6, GBM8, GBM10, GBM12, GBM12 TMZ resistant, GBM14, GBM14 TMZ resistant, GBM22, GBM22 TMZ resistant, GBM26, GBM34, GBM36, GBM39, GBM39 TMZ resistant, GBM43, GBM44, SJGBM2
GL26	GL26
GS	GS2, GS22, GS28
Hs683	Hs683
RG2	RG2
T98	T98
U251	U251, U251SP
U373	U373
U87	U87, U87MG, U87-luc, U87-Fluc, U87MG.delta2-7

Quality assessment. Study quality was evaluated for each publication using a modified 12-point checklist (Amarasingh *et al*, 2009) with one point allocated to each reported item; (1) peer-reviewed publication, (2) sample size calculation, (3) random allocation to groups, (4) blinded assessment of outcome, (5) compliance with animal welfare regulations, (6) statement of potential conflict of interests, (7) consistent volume or number of cells inoculated, (8) consistent site of tumour implantation, (9) reported number of animals in which the xenograft did not grow, (10) number of excluded animals stated, and reasons for exclusion given, (11) explanation of tumour model used, or multiple glioma models used and (12) presentation of evidence that temozolomide acts directly against tumour.

Data analysis. For tumour volume data, efficacy was quantified using the normalised mean difference summary statistic (Sena *et al*, 2010). These effect sizes were then pooled using DerSimonian and Laird random effects meta-analysis (DerSimonian and Laird, 1986) to provide summary estimates of efficacy and to explore differences between groups of experiments.

To summarise median survival data we used the median survival ratio, MSR (treated survival divided by control survival) as an approach consistent with the gold standard hazard ratio method (Michiels *et al*, 2005; Tierney *et al*, 2007). Median survival ratio was log-transformed to give a normal distribution (Simes, 1987), and then log-MSRs were pooled using a modified DerSimonian and Laird random effects model. As no measure of variance was available for MSRs, we weighted studies according to the number of animals used (the number of treated animals plus the number of control animals divided by the number of treatment groups per control group). As measures of variance were not present in the extracted data, we estimated the standard error of summary estimates from the inter-study variance: random effects-weighted s.d. of log-MSR from the random effects – effect size was divided by the square root of the number of comparisons. We used random effects-weighted s.e. because the resulting estimate of variance is broader and therefore more conservative than unweighted s.e. 95% confidence intervals about the random effects mean log-MSR were calculated and the exponential of this log-MSR generated summary MSR data.

For stratified meta-analysis, we used the χ^2 statistic, with n-1 degrees of freedom, to determine the extent to which stratification accounted for the observed heterogeneity. We used Bonferroni correction to account for the number of stratifications – with 12 strata for survival and 13 for volume data (1 extra stratifications for the method of volume measurement) giving critical values for significance of $P = 0.0043$ and $P = 0.0039$, respectively. We used funnel plots, Egger regression (Egger *et al*, 1997) and Trim and Fill analysis (Duval and Tweedie, 2000) to seek evidence of publication bias; for survival data, we used the number of animals as the measure of precision (Peters *et al*, 2006).

RESULTS

Electronic searching identified 298 publications, of which 84 satisfied the inclusion criteria; 24 were excluded at data extraction (Figure 1). Four papers in foreign languages were identified of which one met our inclusion criteria. From the remaining 60 publications, 123 and 26 individual comparisons were identified for median survival and tumour volume outcomes respectively, representing data from 2044 and 399 animals. Temozolomide treatment led to significant improvements in survival (MSR: mean 1.88, 95% CI 1.74–2.03; $\chi^2 = 388$, $df = 122$; $P < 0.0043$) and reductions in tumour volume (50.4% (41.8–58.9); $\chi^2 = 172$, $df = 25$; $P < 0.0039$).
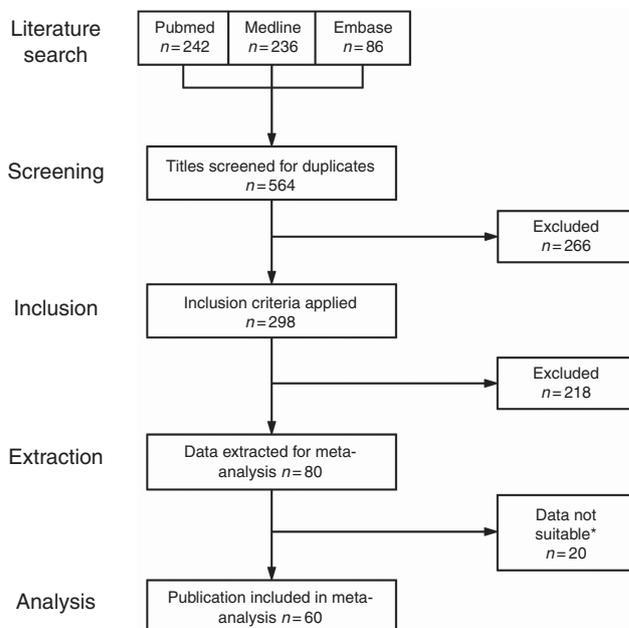
Figure 1. Study selection summary. *We excluded studies at data extraction because they did not present control data ($n = 11$), did not assess the outcomes of median survival or tumour volume ($n = 5$), or the last measured timepoint for tumour volume measurement differed between treated and control groups ($n = 4$).

The median quality score was 6 out of 12 (IQR 5–7, range 3–8). Of the 60 publications, none reported sample size calculations or allocation concealment; 32 reported random group allocation and 5 reported blinding of outcome assessment. All studies used either mice or rats; mice were used more frequently for both survival and volume outcomes (89 *vs* 34; 18 *vs* 8), and 41 different glioma models were tested. Information regarding delay to treatment was not available for seven publications.

Study characteristics. Randomised treatment allocation was associated with an increase in the MSR (2.03 (1.83–2.26), $n = 66$ *vs* 1.71 (1.53–1.90), $n = 57$; $\chi^2 = 13.2$, $df = 1$; $P < 0.0043$) and a reduction in tumour volume (50.1% (40.1–60.0), $n = 19\%$ *vs* 51.1% (33.7–68.4), $n = 7$; $\chi^2 = 18.7$, $df = 1$; $P < 0.0039$). Studies that blinded the assessment of outcome reported a greater reduction in tumour volume (64.4% (9.57–119), $n = 2\%$ *vs* 48.8% (40.3–57.4), $n = 24$; $\chi^2 = 9.70$, $df = 1$; $P < 0.0039$), but no association with MSR was observed. Total quality score had a significant impact on tumour volume ($\chi^2 = 44.5$, $df = 5$; $P < 0.0039$ (Figure 2B) but not survival (Figure 2A)). Method of volume quantification (stratified into excision and histology, MRI, external calliper and unknown) was also associated with differences in reported volume reduction. The single study measuring tumour volume by MRI reported greater temozolomide efficacy compared with those measuring volume by either excision and histology or external caliper (61.0% (50.5–71.6), $n = 1\%$ *vs* 52.6% (38.9–66.3), $n = 9$ and 52.1% (38.4–65.7), $n = 12$, respectively: $\chi^2 = 23.3$, $df = 3$; $P < 0.0039$).

Animal and tumour models. There were no significant differences in temozolomide efficacy between rats and mice for either survival or reduction in tumour volume. Athymic animals were commonly used in survival studies ($n = 50$ publications) and occasionally to study effects on tumour volume ($n = 3$). Other comorbidities used with survival were severe compromised immunodeficient ($n = 1$) and unspecified immunocompromised ($n = 11$). Comorbidity accounted for a significant proportion of between-study heterogeneity for survival ($\chi^2 = 40.1$, $df = 3$; $P < 0.0043$, Figure 3A), with efficacy being significantly greater in
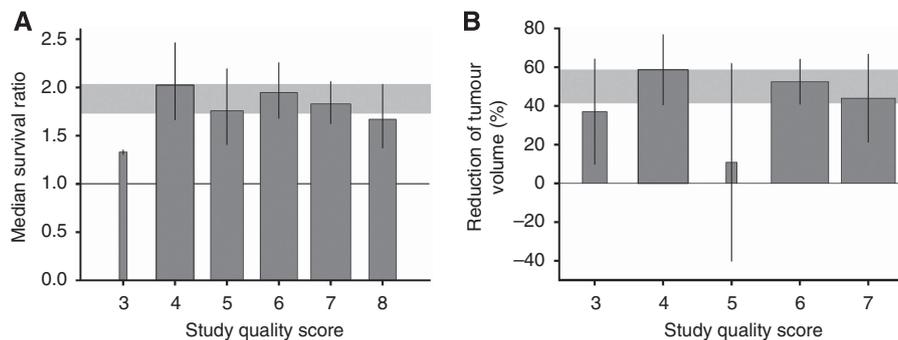
Figure 2. Stratification by aggregated quality score. Quality score accounts for between-study heterogeneity in MSR (**A**, $P < 0.0039$) and tumour volume reduction (**B**, $P < 0.0043$) The grey band represents global 95% confidence intervals; columns represent mean ± 95% CI and column width a measure of number of comparisons within each stratum. The solid line in **A** represents the level of neutral treatment effect.
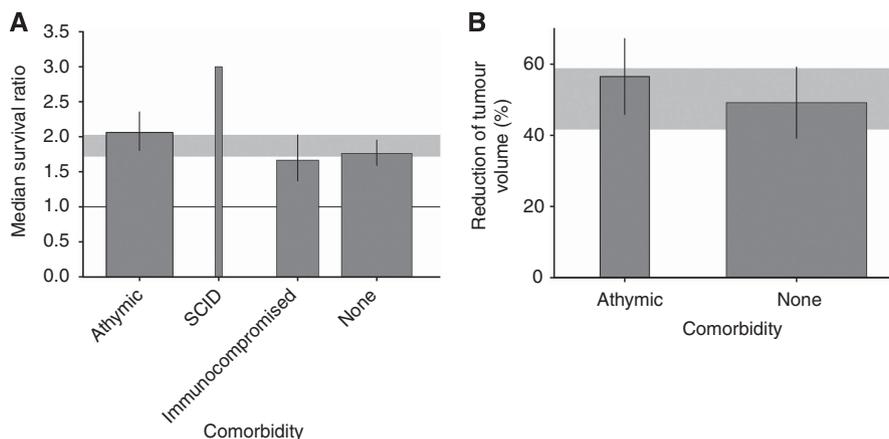
Figure 3. Stratification by comorbidity. Significant heterogeneity was seen between comorbidities in MSR data (**A**, $P < 0.0043$) but not volume data. The grey band represents global 95% confidence intervals; columns represent mean ± 95% CI and column width a measure of number of comparisons within each stratum. The solid line in **A** represents the level of neutral treatment effect.

animals with comorbidities, but not for tumour volume (Figure 3b). Almost all experiments reporting median survival used intracranial glioma models, the remainder using subcutaneous glioma (121 *vs* 2). For tumour volume experiments, 12 comparisons used intracranial and 14 used subcutaneous models. For both survival and volume experiments, the reporting of the amount of tumour cells injected was good – most reporting either a number of cells (124 out of 146 comparisons, ranging from 1000 to 1 000 000 implanted cells, although only 11 comparisons were made with models using fewer than implanted 100 000 cells) or a set volume of implanted cells (14 out of 146, 1 mm^3 in all cases).

Next, we examined the impact of the glioma cell species of origin and the glioma cell type used (Figure 4). Human- and rat-derived cells predominated, with human glioma lines being more commonly used for both outcomes (MSR: $n = 87$, volume: $n = 15$). The species of origin did not account for a significant proportion of the observed heterogeneity for MSR but temozolomide therapy was associated with a larger volume reduction in human-derived glioma cell lines ($\chi^2 = 43.9$, $df = 1$; $P < 0.0039$). In both models, there was significant variation of efficacy between glioma models (MSR: $\chi^2 = 53.6$, $df = 12$; $P < 0.0043$; volume: $\chi^2 = 66.7$, $df = 6$; $P < 0.0039$), with GBM and 9L lines appearing to be most sensitive to temozolomide (for specifications of glioma lines, see Table 1 and Supplementary Figure 2). O-6-methylguanine-DNA methyltransferase (MGMT) status of the glioma models used was investigated and reported in only 11 of 60 publications.

Temozolomide dosing. For temozolomide dosing regimens, we analysed data for total temozolomide dose (Figures 5A and B), delay to treatment (Figures 5C and D) and route of delivery (Figures 5E and F); for treatment duration, number of cycles and other aspects of the dosing regime there was such diversity of approach that we did not consider a stratified analysis likely to be informative. Both outcomes showed significant variability between dose stratified into groups of <50, $51–100$, $101–500$, $501–1000$ and >1000 mg kg^{-1}. The dose of temozolomide explained a significant proportion of the observed heterogeneity, with higher doses generally associated with greater efficacy (MSR: $\chi^2 = 17.2$, $df = 4$; $P < 0.0043$; volume: $\chi^2 = 91.7$, $df = 3$; $P < 0.0039$), although there may have been a decline in efficacy at total doses above 1 g kg^{-1} (Figures 5A and B).

Delay to treatment was stratified into groups of 0, 1–10, 11–20 and >20 days post-inoculation, and this accounted for a significant proportion of between-study heterogeneity for both MSR and volume reduction (MSR: $\chi^2 = 47.0$, $df = 4$; $P < 0.0043$; volume: $\chi^2 = 45.9$, $df = 3$; $P < 0.0039$). For survival, efficacy was greatest the earlier temozolomide treatment was initiated. However, we saw the converse for tumour volume: the longer the delay to treatment the more effective temozolomide was in reducing tumour volume. Those studies in which treatment was initiated >20 days after tumour inoculation were associated with a greater reduction in tumour volume than those where the delay was 1–10 or 11–20 days.

When we examined the effect of the route of drug administration. We found that those treated with local temozolomide
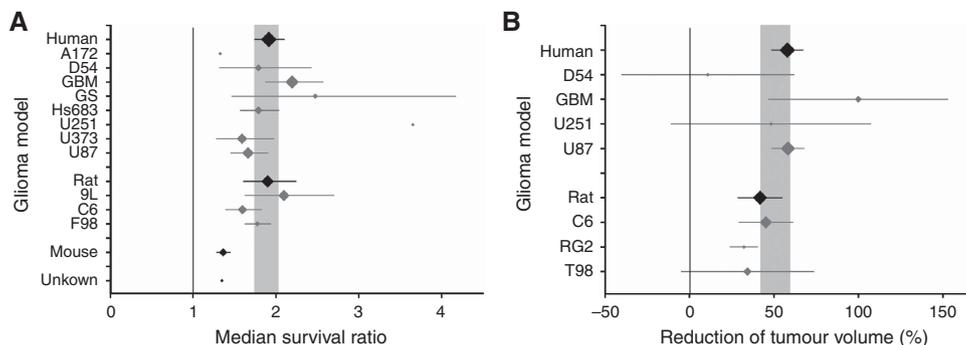
Figure 4. **Stratification by glioma model and the species it is derived from.** Median survival ratio showed no difference between species, where significant heterogeneity was seen between human- and rat-derived tumours in volume data (**A**, $P < 0.0039$). Both outcomes were associated with significant variability between glioma models (**A**, MSR: $P < 0.0043$; **B**, volume: $P < 0.0039$). The grey band represents global 95% confidence intervals; black plots represent species data and grey represent glioma models groups (see table 1). Plots represent mean ± 95% CI and diamond size a measure of number of comparisons within each stratum. The solid line in **A** represents the level of neutral treatment effect.
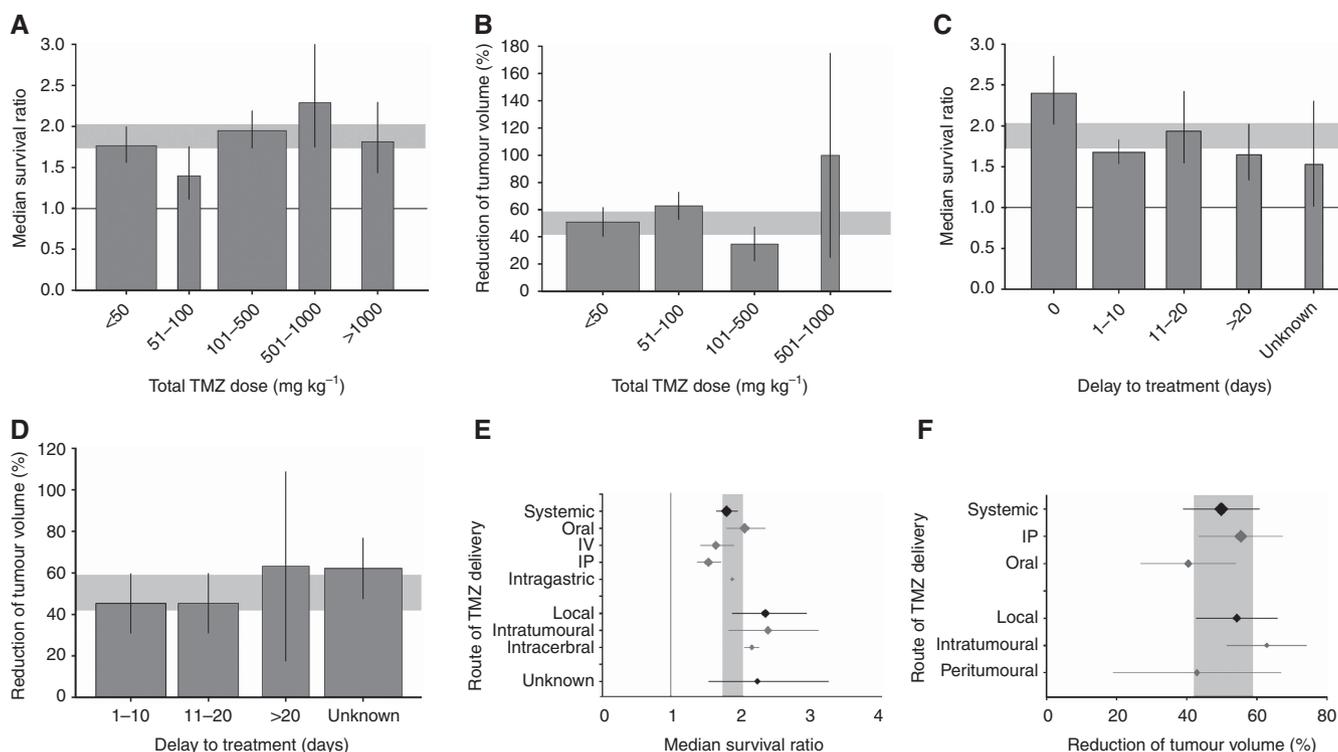
Figure 5. **Stratification by TMZ dosing characteristics.** (**A** and **B**) Significant heterogeneity between total TMZ dose strata was observed in both outcomes (**A**, MSR: $P < 0.0043$; **B**, volume: $P < 0.0039$). (**C** and **D**) Delay to treatment, defined as the time in days between tumour inoculation and first TMZ dose, accounted for significant between-study variability in both outcomes (**C**, MSR: $P < 0.0043$; **D**, volume: $P < 0.0039$). (**E** and **F**) Route of delivery was stratified into local/systemic and then substratified more specifically. Local TMZ therapy was significantly more effective than systemic delivery in MSR data (**E**, $P < 0.0043$) but not with tumour volume (**F**). Different routes of delivery showed significant variability in TMZ efficacy in volume data (**F**, $P < 0.0039$); significant variability within systemic therapies in MSR data (**E**, $P < 0.0043$) but not between local delivery methods. Grey bands represent global 95% confidence intervals; columns (**A**–**D**) and data plots (**E** and **F**) represent mean ± 95% CI and column width (**A**–**D**) and diamond size (**E** and **F**) a measure of number of comparisons within each stratum. Solid lines in **A**, **C**, **E** and **F** represent the level of neutral treatment effect.

survived longer than those treated systemically ($\chi^2 = 17.6$, $df = 2$; $P < 0.0043$), but there was no difference in tumour volume reduction. However, when we looked in more detail at systemic delivery routes for tumour volume reduction we found significant differences ($\chi^2 = 75.7$, $df = 3$; $P < 0.0039$), oral delivery providing the best survival benefit and intraperitoneal delivery having the largest effect on tumour volume.

Publication bias. For both MSR and tumour volume, there was no difference in reported efficacy before and after the Stupp publication (Figures 6A and B). For both survival and volume data, the intercept of Egger regression was positive (Figures 6C–F) indicating an excess of small imprecise studies overstating efficacy. Trim and Fill analysis supported asymmetry of survival data (Figure 6G) – corrected estimate of efficacy was reduced to 1.56
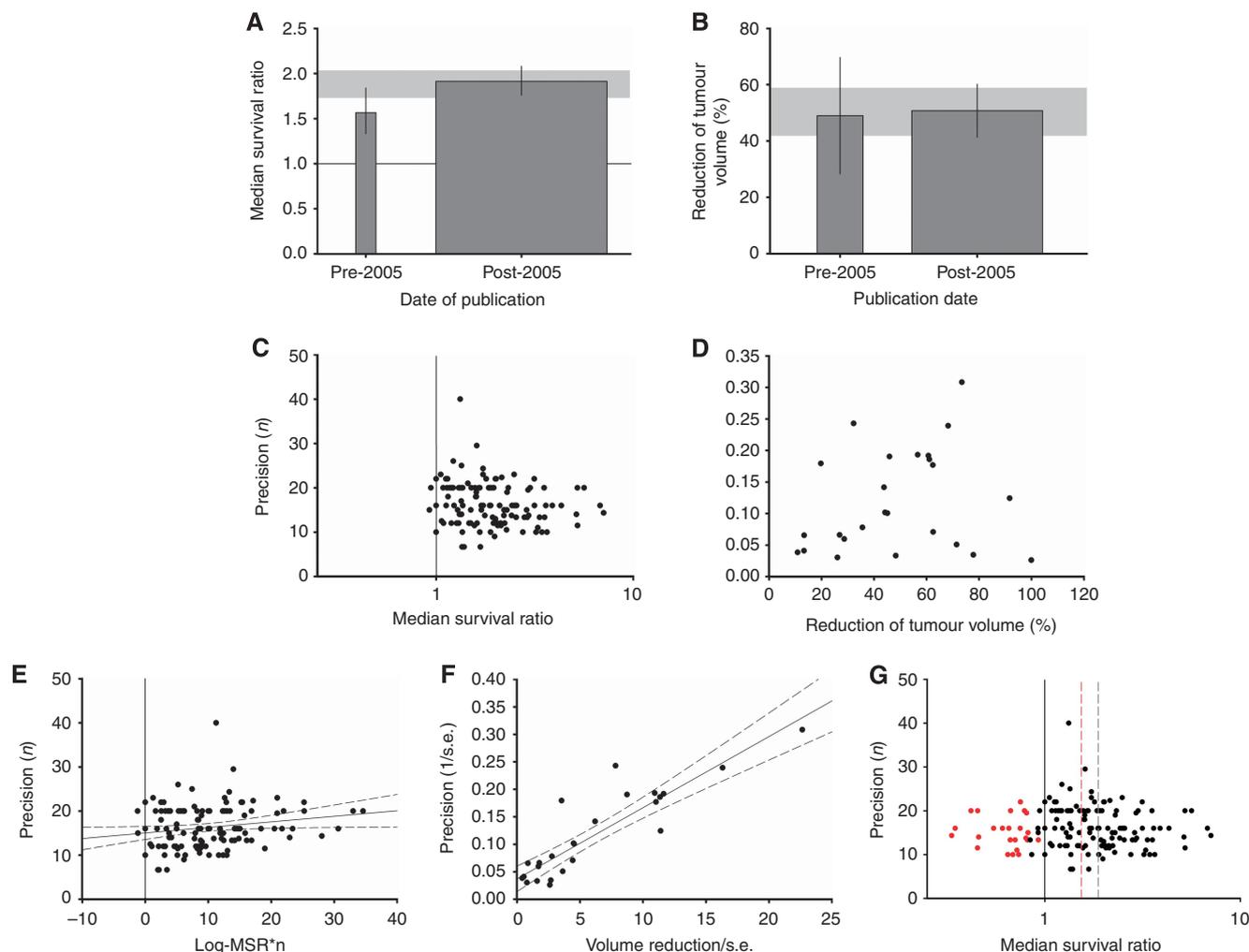
Figure 6. Discrepancies pre-/post-Stupp and publication bias. (A and **B)** There was no difference in TMZ efficacy between pre- and post-Stupp eras in either outcome. The grey band represents global 95% confidence intervals; columns represent mean ± 95% CI and column width a measure of number of comparisons within each stratum. **(C** and **D)** Funnel plots showing effect size (x axis) vs a measure of study precision (y axis). Survival data in the funnel plot appear to be skewed (**D**), imprecise studies generally showing more efficacy than those with larger sample sizes. (**E** and **F**). Egger regression plots, depicting effect size × precision (x axis) vs precision (y axis). Regression revealed positive intercepts for both outcomes (**E**, MSR: $P<0.001$; **F**, volume: $P<0.01$). Dotted lines represent 95% CI of the regression. **(G)** Trim and fill analysis of survival data showed asymmetry of the data set, suggesting a preference towards more efficacious results. Dotted lines represent global estimates of efficacy before (grey) and after (red) Trim and fill analysis. The solid lines in **A, C, E** and **G** represent the level of neutral treatment effect.

(1.42–1.70), compared with an unadjusted estimate of 1.88 (1.74–2.03), after the addition of 24 'missing' studies.

DISCUSSION

In this systematic review and meta-analysis of 60 publications involving 2377 animals, we found that overall temozolomide therapy almost doubled survival and halved tumour volume. Furthermore, observed efficacy seemed to be influenced by randomisation, blinding, total quality score, animal comorbidity, glioma model (including species), total dose, delay to treatment and route of temozolomide delivery.

We identified 17 papers published before 2005 describing the efficacy of temozolomide. Although this is a reasonable basis on which to proceed to clinical trials it is a cause of some concern that many of these early publications, which were of most interest to us, did not present control data or measures of variance, and indeed

nine did not present sufficient data to allow them to be included in this meta-analysis.

Overall study quality was modest; the median number of study quality checklist items scored was 6 (of a possible 12) and no study scored higher than 8. Studies that did not report simple measures to reduce bias such as randomisation and blinding of outcome assessment gave significantly higher estimates of efficacy, confirming an important role in experimental design in glioma similar to that established in other neurological disease models (Frantzias et al, 2011; van der Worp et al, 2007).

Previous meta-analyses in the animal modelling of glioma have been conducted using median survival as the primary outcome, but there is as yet no consensus about the most appropriate analysis of such data. We used MSRs as an approximation of the gold standard hazard ratio approach (Michiels et al, 2005; Tierney et al, 2007). By estimating standard error from the size of each study, we were able to combine data using a random effects approach. As weights and the heterogeneity statistics were therefore calculated using study size to estimate standard error it is possible that this

might have confounded the relative weighting given to individual studies. However, given the high heterogeneity observed, in practice this approach has trended towards a simple average, with studies being given roughly equal weight. As MSR analysis does not provide a measure of within-study variance, we calculated s.e. and therefore 95% confidence intervals using inter-study variance to provide the most reliable measure of overall variance from the available data. A major limitation of our analysis is, because it is univariate, it does not provide insight into how different variables interact with each other. An example of this is the dose regimen and the interaction between different aspects of this; however, there were insufficient data to allow multivariate analysis to help disentangle the impact of multiple factors.

Differences between different experimental glioma models have previously been discussed (Whittle *et al*, 1998; Amarasingh *et al*, 2009). There is a large amount of variation in neuropathology, natural history and drug sensitivity (Whittle *et al*, 1998; Candolfi *et al*, 2007; Zhang *et al*, 2010) and this underpins the importance of glioma model selection during experimental design. Of the 60 studies, 18 justified their choice of glioma model, or used multiple models for comparison. Our data show a significant variation in temozolomide sensitivity between tumour models. Previous evidence suggests that human-derived tumours are more sensitive to chemotherapy than those originating in rodents (Amarasingh *et al*, 2009): temozolomide also appeared to be more effective at reducing tumour volume in human-derived tumours, however, there was no difference in survival data and there may be an artefact of between-model heterogeneity. Nevertheless, it remains clear that tumours show a great degree of variability in their sensitivity to treatment. This may reflect either the degree of differentiation within the tumour or temozolomide resistance through a variety of mechanisms (Zhang *et al*, 2010), including MGMT expression. Therefore, genomic analysis of specific tumour models and analysis of drug sensitivities and resistances may give insights into the future development of tailored chemotherapy.

Temozolomide is clinically effective in malignant glioma (Stupp *et al*, 2005; van den Bent *et al*, 2006) and this meta-analysis shows temozolomide also improves survival and reduces tumour volume in experimental glioma; whereas a substantial proportion of this evidence comes after the publication of the pivotal clinical trial this remains an example of concordance between experimental and clinical findings (Perel *et al*, 2007; Sena *et al*, 2010). There was no significant change in reported efficacy for temozolomide after the publication of the phase III temozolomide randomised control trial, but there was evidence for a significant publication bias.

The nitrosoureas previously used in glioma such as carmustine (BCNU) and lomustine (CCNU) had limited efficacy (Stewart, 2002). In a meta-analysis of nitrosoureas in experimental glioma efficacy was less than reported here for temozolomide with greater variability, and survival was not significantly increased (Amarasingh *et al*, 2009). Similarly, gene therapy has been studied in glioma models, but has not translated successfully into clinical use. Pulkkanen and Yla-Herttuala (2005) concluded that clinical efficacy of gene therapy in several small RCTs was variable, and overall efficacy remains questionable. Data from a meta-analysis of gene therapy in experimental glioma models (Conlin, under review) came to a similar conclusion; gene therapy does prolong survival in experimental glioma, but there was substantial variability, that is to say inconsistency of effect. This contrasts with the findings reported here for temozolomide, and suggests that consistency of effect across a range of circumstances may be important for successful translation.

In conclusion, temozolomide is effective in experimental glioma models, although this is qualified by concerns of the internal validity (randomisation, blinding) and of publication bias. Temozolomide appears to be more consistent in its efficacy than either nitrosoureas or gene therapy, and this observation may help guide future translational research in neuro-oncology.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Amarasingh S, Macleod MR, Whittle IR (2009) What is the translational efficacy of chemotherapeutic drug research in neuro-oncology? A systematic review and meta-analysis of the efficacy of BCNU and CCNU in animal models of glioma. *J Neurooncol* **91**(2): 117–125.

Bleehen NM, Newlands ES, Lee SM, Thatcher N, Selby P, Calvert AH, Rustin GJ, Brampton M, Stevens MF (1995) Cancer Research Campaign phase II trial of temozolomide in metastatic melanoma. *J Clin Oncol* **13**(4): 910–913.

Candolfi M, Curtin JF, Nichols WS, Muhammad AG, King GD, Pluhar GE, McNiel EA, Ohlfest JR, Freese AB, Moore PF, Lerner J, Lowenstein PR, Castro MG (2007) Intracranial glioblastoma models in preclinical neuro-oncology: neuropathological characterization and tumor progression. *J Neurooncol* **85**(2): 133–148.

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* **7**(3): 177–188.

Duval S, Tweedie R (2000) Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**(2): 455–463.

Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**(7109): 629–634.

Frantzias J, Sena ES, Macleod MR, Al-Shahi Salman R (2011) Treatment of intracerebral hemorrhage in animal models: meta-analysis. *Ann Neurol* **69**(2): 389–399.

Freireich EJ, Gehan EA, Rall DP, Schmidt LH, Skipper HE (1966) Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. *Cancer Chemotherapy Reports* **4**(50): 219–244.

Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M (2010) Enhancing search efficiency by means of a search filter for finding all studies on animal experimentation in PubMed. *Lab Anim* **44**(3): 170–175.

Michiels S, Piedbois P, Burdett S, Syz N, Stewart L, Pignon JP (2005) Meta-analysis when only the median survival times are known: a comparison with individual patient data results. *Int J Technol Assess Health Care* **21**(1): 119–125.

Newlands ES, O'Reilly SM, Glaser MG, Bower M, Evans H, Brock C, Brampton MH, Colquhoun I, Lewis P, Rice-Edwards JM, Illingworth RD, Richards PG (1996) The Charing Cross Hospital experience with temozolomide in patients with gliomas. *Eur J Cancer* **32**: 2236–2241.

O'Reilly SM, Newlands ES, Glaser MG, Brampton M, Rice-Edwards JM, Illingworth RD, Richards PG, Kennard C, Colquhoun IR, Lewis P *et al.* (1993) Temozolomide: a new oral cytotoxic chemotherapeutic agent with promising activity against primary brain tumours. *Eur J Cancer* **29**: 940–942.

Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, Macleod M, Mignini LE, Jayaram P, Khan KS (2007) Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ* **334**(7586): 197.

Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L (2006) Comparison of two methods to detect publication bias in meta-analysis. *JAMA* **295**(6): 676–680.

Plowman J, Waud WR, Koutsoukos AD, Rubinstein LV, Moore TD, Grever MR (1994) Preclinical antitumor activity of temozolomide in mice: efficacy against human brain tumor xenografts and synergism with 1,3-bis(2-chloroethyl)-1-nitrosourea. *Cancer Res* **54**(14): 3793–3799.

Pulkkanen KJ, Yla-Herttuala S (2005) Gene therapy for malignant glioma: current clinical status. *Mol Ther* **12**(4): 585–598.

Sena ES, Briscoe CL, Howells DW, Donnan GA, Sandercock PA, Macleod MR (2010) Factors affecting the apparent efficacy and safety of tissue plasminogen activator in thrombotic occlusion models of stroke: systematic review and meta-analysis. *J Cereb Blood Flow Metab* **30**(12): 1905–1913.

Simes RJ (1987) Confronting publication bias: a cohort design for meta-analysis. *Stat Med* **6**(1): 11–29.

Stevens MF, Hickman JA, Langdon SP, Chubb D, Vickers L, Stone R, Baig G, Goddard C, Gibson NW, Slack JA *et al.* (1987) Antitumor activity and pharmacokinetics in mice of 8-carbamoyl-3-methyl-imidazo[5,1-d]-1,2,3,5-tetrazin-4(3 H)-one (CCRG 81045; M & B 39831), a novel drug with potential as an alternative to dacarbazine. *Cancer Res* **47**(22): 5846–5852.

Stewart LA (2002) Chemotherapy in adult high-grade glioma: a systematic review and meta-analysis of individual patient data from 12 randomised trials. *Lancet* **359**(9311): 1011–1018.

Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, Mirimanoff RO (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* **352**(10): 987–996.

Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR (2007) Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* **8**: 16.

van den Bent MJ, Hegi ME, Stupp R (2006) Recent developments in the use of chemotherapy in brain tumours. *Eur J Cancer* **42**(5): 582–588.

van den Bent MJ, Taphoorn MJ, Brandes AA, Menten J, Stupp R, Frenay M, Chinot O, Kros JM, van der Rijt CC, Vecht Ch J, Allgeier A, Gorlia T (2003) Phase II study of first-line chemotherapy with temozolomide in recurrent oligodendroglial tumors: the European Organization for Research and Treatment of Cancer Brain Tumor Group Study 26971. *J Clin Oncol* **21**(13): 2525–2528.

van der Worp HB, Sena ES, Donnan GA, Howells DW, Macleod MR (2007) Hypothermia in animal models of acute ischaemic stroke: a systematic review and meta-analysis. *Brain* **130**(Pt 12): 3063–3074.

Walker MD, Alexander Jr. E, Hunt WE, MacCarty CS, Mahaley Jr MS, Mealey Jr J, Norrell HA, Owens G, Ransohoff J, Wilson CB, Gehan EA, Strike TA (1978) Evaluation of BCNU and/or radiotherapy in the treatment of anaplastic gliomas. A cooperative clinical trial. *J Neurosurg* **49**(3): 333–343.

Wen PY, Kesari S (2008) Malignant gliomas in adults. *N Engl J Med* **359**(5): 492–507.

Whittle IR, Macarthur DC, Malcolm GP, Li M, Washington K, Ironside JW (1998) Can experimental models of rodent implantation glioma be improved? A study of pure and mixed glioma cell line tumours. *J Neurooncol* **36**(3): 231–242.

Woll PJ, Crowther D, Johnson PW, Soukop M, Harper PG, Harris M, Brampton MH, Newlands ES (1995) Phase II trial of temozolomide in low-grade non-Hodgkin's lymphoma. *Br J Cancer* **72**(1): 183–184.

Zhang J, Stevens MF, Laughton CA, Madhusudan S, Bradshaw TD (2010) Acquired resistance to temozolomide in glioma cell lines: molecular mechanisms and potential translational applications. *Oncology* **78**(2): 103–114.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)