Progression-free survival as a surrogate endpoint for overall survival in glioblastoma: a literature-based meta-analysis from 91 trials

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1093/neuonc/not236

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:12152941

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Progression-free survival as a surrogate endpoint for overall survival in glioblastoma: a literature-based meta-analysis from 91 trials

Kelong Han, Melanie Ren, Wolfgang Wick, Lauren Abrey, Asha Das, Jin Jin, and David A. Reardon

Genentech, South San Francisco, California (K.H., M.R., A.D., J.J.); University Medical Center & DKFZ, Heidelberg, Germany (W.W.); F. Hoffmann–La Roche, Basel, Switzerland (L.A.); Dana–Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts (D.A.R.)

Corresponding Author: Kelong Han, PhD, Genentech Inc, 1 DNA Way, South San Francisco, CA 94080 (han.kelong@gene.com).

Background. The aim of this study was to determine correlations between progression-free survival (PFS) and the objective response rate (ORR) with overall survival (OS) in glioblastoma and to evaluate their potential use as surrogates for OS.

Method. Published glioblastoma trials reporting OS and ORR and/or PFS with sufficient detail were included in correlative analyses using weighted linear regression.

Results. Of 274 published unique glioblastoma trials, 91 were included. PFS and OS hazard ratios were strongly correlated; $R^2 = 0.92$ (95% confidence interval [CI], 0.71–0.99). Linear regression determined that a 10% PFS risk reduction would yield an 8.1% ± 0.8% OS risk reduction. $R^2$ between median PFS and median OS was 0.70 (95% CI, 0.59–0.79), with a higher value in trials using Response Assessment in Neuro-Oncology (RANO; $R^2 = 0.96$, $n = 8$) versus Macdonald criteria ($R^2 = 0.70; n = 83$). No significant differences were demonstrated between temozolomide- and bevacizumab-containing regimens ($P = .10$) or between trials using RANO and Macdonald criteria ($P = .49$). The regression line slope between median PFS and OS was significantly higher in newly diagnosed versus recurrent disease (0.58 vs 0.35, $P = .04$). $R^2$ for 6-month PFS with 1-year OS and median OS were 0.60 (95% CI, 0.37–0.77) and 0.64 (95% CI, 0.42–0.77), respectively. Objective response rate and OS were poorly correlated ($R^2 = 0.22$).

Conclusion. In glioblastoma, PFS and OS are strongly correlated, indicating that PFS may be an appropriate surrogate for OS. Compared with OS, PFS offers earlier assessment and higher statistical power at the time of analysis.

Keywords: glioblastoma, meta-analysis, overall survival, progression-free survival, regression, response rate, surrogate endpoint.

Traditionally, the success of new cancer treatments is gauged by their ability to improve overall survival (OS) in large, randomized, phase III trials. However, the use of OS as the primary endpoint is often limited by long trial times and confounding effects of post-protocol events, such as subsequent therapies. It is thus helpful to identify and validate surrogate endpoints to facilitate efficacy evaluation and drug approval. Proposed surrogates for OS include progression-free survival (PFS), time to progression, and objective response rate (ORR). Progression-free survival has many advantages over OS, including earlier assessment of efficacy, greater statistical power at the time of analysis, and lack of influence from post-progression therapies.

The relationship between PFS and OS has been studied in various tumors. Results vary greatly by tumor type, with some reinforcing PFS as a good surrogate endpoint for OS and others indicating weak PFS/OS correlation; it has been shown that PFS may be an appropriate surrogate for OS in colorectal cancer, but may not be a good surrogate in breast cancer.

Glioblastoma is a highly aggressive form of cancer that represents 15.8% of all brain and CNS tumors. Despite decades of research into its treatment, prognosis remains poor, with median OS of only 12–14 months. While the introduction of temozolomide (TMZ), an oral alkylating agent, into first-line standard of care achieved some survival improvement, nearly all patients relapse and treatment options are limited for recurrent patients, with no accepted standard of care. There is therefore an unmet need for effective, novel therapies for glioblastoma. However, with fewer than 20 000 new cases diagnosed in the United States each year, glioblastoma occurs much less frequently than other cancers, and consequently patient accrual is low in glioblastoma studies. Thus, the use of trial endpoints that require prolonged periods of follow-up or mix varying treatments into the primary endpoint are particularly undesirable in glioblastoma studies.

While PFS represents an attractive potential surrogate endpoint, the relationship between PFS and OS has not been extensively...
analyzed in glioblastoma. A recent pooled analysis focused on phase I and single-arm phase II trials, with specific treatments, and evaluated the PFS/OS correlation at the individual level in 5 glioblastoma trials. While good individual-level correlation was demonstrated, the small sample sizes precluded any conclusions regarding correlation at trial level. Our analysis evaluates the validity of PFS and ORR as surrogate endpoints for OS using a meta-analysis of completed published phase II and III glioblastoma trials and consideration of a greater range of variables than previously evaluated.

Materials and Methods

Literature Search and Data Extraction

Completed phase II, III, and IV trials in glioblastoma published between January 1, 1991 and June 4, 2012 were identified through a systematic search on MEDLINE/PubMed and Trialtrove (Citeline) using the following keywords: “oncology and CNS” OR “glioblastoma” OR “GBM” OR “glioblastoma multiforme” AND “survival” OR “PFS” OR “progression free survival” OR “progression-free survival” OR “overall survival” OR “OS” OR “progression”. Relevant sources identified in the bibliographies of reviewed papers were also included. Publications reporting PFS and OS data from unique glioblastoma trials utilizing standard tumor response criteria were included. Abstracts and other nonjournal publications were included if sufficient detail was provided. Duplicate publications of the same trial, pediatric studies, non–English language papers, sources lacking methodology detail, and review/summary papers were excluded. The analysis was performed using the original authors’ and per protocol endpoint definitions. When available, hazard ratios (HRs) for PFS and OS were recorded. Treatment, patient, and clinical endpoint data from each study were included in the database. Endpoints of interest were OS, PFS, and ORR. Due to the small number of glioblastoma trials available, the analysis was not limited to randomized trials, and studies with mixed high-grade glioma subpopulations.

Statistical Methods and Analysis

Weighted linear regression analysis through the origin of the plot was used to evaluate correlation between the following pairs of endpoints: (i) HR in PFS and OS, (ii) median PFS (mPFS) and median OS (mOS), (iii) ORR and mOS, (iv) 6-month PFS and mOS, and (v) 6-month PFS and 1-year OS. Pearsons coefficient was used as a correlation measure between these endpoints. An R2 value of greater than 0.9 is an indicator for a strong correlation, 0.89 to 0.6 for a good to moderate correlation, and below 0.6 for a weak correlation. Points were weighted by the number of patients in the intent-to-treat population. Linear scale or log scale was selected based on the data distribution. Confidence intervals (CIs) for R2 and the weighted fit were calculated using a bootstrapping method (resampling 1000 times) assuming a sample size of 60 patients in additional trials (mean sample size per arm in the included glioblastoma trials).

Differences in the PFS/OS correlations by response criteria that include tumor response and clinical symptomatology (Macdonald and RANO response criteria (P < .49), good correlation between mPFS and mOS was evaluated by treatments, R2 = 0.70 (95% CI, 0.50–0.85) and 0.75 (95% CI, 0.60–0.86) for TMZ-containing and non-TMZ-containing regimens, respectively (Fig. 2C). No significant difference in the slope of the regression line was demonstrated between these 2 treatment types (P = .10). There was no significant difference in the slope of the regression line between mPFS and mOS for BEV-containing and non-BEV-containing treatments (P = .46), with good correlation between mPFS and mOS observed with both (R2 = 0.95 [0.65–0.99] vs 0.70 [0.56–0.80]; Fig. 2D). A significant difference in the slope of the regression line between mPFS and mOS was demonstrated between line settings (newly diagnosed vs recurrent, P = .04; Fig. 2E) and between histology types (glioblastoma only vs mixed histology, P = .02; Fig. 2F). When the correlation between mPFS and mOS was evaluated in trials conducted at different time periods (1991 to present), no significant difference in the slope of the regression line was demonstrated.
Correlation Between Other Endpoints and OS

Objective response rate was poorly correlated with mOS ($R^2 = 0.22$; Fig. 3). For 6-month PFS versus 1-year OS (by study arm), $R^2$ was 0.60 (95% CI, 0.37–0.77), indicating a moderate correlation between the 2 survival rates (Fig. 4A). The correlation between 6-month PFS and mOS yielded an $R^2$ of 0.64 (95% CI, 0.42–0.77; Fig. 4B).

Lead-time Analysis

The lead-time that could be gained by using PFS instead of OS as the endpoint averaged 7.4 months (max 17.6 mo) and 4.2 months (max 8.1 mo) in newly diagnosed and recurrent cases, respectively (Fig. 5). The lead-time increased with increasing mOS: for newly diagnosed cases, it increased from 6–7 months for a mOS of 1 year to ~9–10 months for a mOS of 1.5 years; for recurrent patients, it increased from 3–4 months for a mOS of half a year to ~5–6 months for a mOS of 9 months.

Discussion

This is a systematic evaluation of whether PFS is an appropriate surrogate endpoint for OS in glioblastoma clinical trials. We assembled the largest literature glioblastoma trial database to date, which included almost all published glioblastoma trials (phase II and beyond) since 1991, as well as the latest advances in treatment
The lead-time in newly diagnosed glioblastoma is comparable to using PFS instead of OS as an endpoint in glioblastoma. Notably, we demonstrated that a significant lead-time benefit is achieved. Firstly, PFS offers the opportunity for early assessment. Objective response rate and OS were poorly correlated. Median PFS and OS were also well correlated (R^2 = 0.92, 95% CI: 0.71–0.99). Taken together, these results lend substantial support to the use of PFS as a surrogate endpoint for OS in glioblastoma trials. However, the use of PFS is associated with several limitations that must also be considered. Firstly, it is important to standardize response criteria; a number of standard criteria are being used in glioblastoma trials and have been used in trials building the basis of our knowledge, such as Macdonald, Levin, Response Evaluation Criteria In Solid Tumors (RECIST), and RANO. Secondly, the discrepancy between the time of clinical event (progression or death) and radiologic assessment could be a confounding factor, and therefore the time interval between clinical and radiologic assessments should be minimized and consistent across studies. Thirdly, the association between radiologic progression, clinical benefit, and quality of life remains open for discussion.

Our analysis demonstrated that the percentage risk reduction calculated from the HR of PFS is highly correlated with the percentage risk reduction calculated from the HR of OS in glioblastoma trials, indicating that the treatment effect on PFS can predict the treatment effect on OS in glioblastoma. A great portion (92%) of variability in OS difference can be explained by the PFS difference (R^2 = 0.92). Notably, the 95% CI and prediction interval were relatively narrow. Median PFS and OS were also well correlated (R^2 = 0.70). Taken together, these results lend substantial support to the use of PFS as a surrogate endpoint for OS in glioblastoma trials. While the validity of PFS as a surrogate for OS can be demonstrated in some tumor types, the effect is not consistent. Broglio et al attributed differences to variations in survival post-progression (SPP), where SPP is the time difference between OS and PFS. Overall survival in cancers with long SPPs are more affected by the presence of confounding factors and therefore demonstrate weaker correlations with PFS. Glioblastoma patients have a median SPP of around 7 months, and therefore the PFS versus OS correlation should be fairly strong. Our analysis supports this premise.

Although sample sizes were small in our analysis, every attempt was made to standardize the data, with exclusion of trials including
Fig. 2. (A) Correlation between mPFS and mOS by study arm. (B) Correlation between mPFS and mOS in trials using Macdonald or RANO criteria for response evaluation. All RANO trials contain BEV test regimens, and the 3 BEV-containing trials using Macdonald criteria are indicated in blue. There are 83 arms using Macdonald criteria (red circle or square) and 8 arms (7 unique trials) using RANO criteria (black circle or square). (C) Correlation between mPFS and mOS separated by treatment (TMZ [red] vs non-TMZ [black]). Trials included used the Macdonald/RANO criteria for tumor assessment. (D) Correlation between mPFS and mOS separated by treatment (BEV [red] vs non-BEV [black]). Trials included used the Macdonald/RANO criteria for tumor assessment. (E) Correlation between mPFS and mOS separated by line settings (newly diagnosed vs recurrent). (F) Correlation between mPFS and mOS separated by histology (glioblastoma only vs mixed histology). Abbreviation: PI, prediction interval.
insufficient methodological detail and those not utilizing standardized response assessment and study endpoints.

The effects of TMZ and BEV on the correlation between mPFS and mOS were selected for study because these 2 treatments appeared most often in the literature, and too few trials report other specific treatments. Interestingly, despite differing mechanisms of action, these treatments demonstrated consistent correlations between mPFS and mOS. Although the small sample number in our analysis precludes any definitive conclusions with regard to treatment effect, the results warrant future studies.

Historically, it has been shown that patients with anaplastic glioma (WHO grade III) have a much better prognosis and survival than patients with glioblastoma. It would be logical to assume that a mixed grade III–IV group would have better survival because the anaplastic glioma patients’ survival would increase the median values for the entire group. Our results confirmed this assumption: the slope of the regression line between mPFS and mOS is significantly higher in trials with mixed grade III–IV glioma compared with glioblastoma only. However, the difference was marginal. Possible explanations for this observation may include the fact that the lower left corner data points, which represent the mixed histology group, also represent recurrent trials (poor PFS and OS) predominantly, and most recurrent glioma patients have progressed from grade III to glioblastoma. Conversely, data in the upper right corner (high PFS, high OS) represent newly diagnosed cases predominantly and support a trend toward better OS for mixed patients (better prognosis) compared with glioblastoma-only patients.

The accrual period in the trials included in this analysis ranged from 1991 to the present. During this period, advances have been made in many aspects of glioblastoma clinical management, such as diagnostics, surgical and imaging technology, treatments, recurrence monitoring, and standard supportive care. However, the correlation between median PFS and OS seems to be consistent across different time periods, which supports the applicability of these results to future trials.

There was only a moderate correlation between 6-month PFS and mOS, which is consistent with the results of Ballman et al, who investigated the relationship between 6-month PFS and 1-year OS in phase II glioblastoma trials. However, it is impossible in our analysis to identify at which time point the PFS rate would be a good predictor for OS because most trials report only 6-month PFS and mPFS and individual patient data are not available to us. In addition, ORR and OS were poorly correlated ($R^2 = 0.22$).

The applicability of our estimate of the linear relationship between the HR of OS and HR of PFS for trials evaluating anti-VEGF...
agents (ie, agents targeting VEGF or VEGFR) may require further validation because none of the trials in the HR correlation analysis contained an anti-VEGF agent, such as BEV (an anti-VEGF antibody) or cediranib (a VEGFR tyrosine kinase inhibitor). VEGF blockade decreases vascular permeability and normalizes vascular perfusion and the blood–brain barrier, often causing decreased contrast because none of the trials in the HR correlation analysis contained an anti-VEGF agent, such as BEV (an anti-VEGF antibody) or cediranib (a VEGFR tyrosine kinase inhibitor). VEGF blockade decreases vascular permeability and normalizes vascular perfusion and the blood–brain barrier, often causing decreased contrast. In conclusion, our meta-analysis of 91 unique glioblastoma trials demonstrated a strong correlation between improvements in PFS and OS. There is also a good correlation between median PFS and OS in glioblastoma trials, regardless of response criteria, treatment, line settings, and histology. However, poor correlation was observed between ORR and OS, indicating that a high ORR may not translate into improved OS. Together these findings indicate that PFS may be an appropriate surrogate for OS in glioblastoma trials. Compared with OS, PFS offers the opportunity for earlier assessment of efficacy and higher statistical power, so establishment of these correlations may facilitate interpretation of interim analyses and future trial design.

In conclusion, our meta-analysis of 91 unique glioblastoma trials demonstrated a strong correlation between improvements in PFS and OS. There is also a good correlation between median PFS and OS in glioblastoma trials, regardless of response criteria, treatment, line settings, and histology. However, poor correlation was observed between ORR and OS, indicating that a high ORR may not translate into improved OS. Together these findings indicate that PFS may be an appropriate surrogate for OS in glioblastoma trials. Compared with OS, PFS offers the opportunity for earlier assessment of efficacy and higher statistical power, so establishment of these correlations may facilitate interpretation of interim analyses and future trial design.

Funding
This analysis was sponsored by Genentech Inc., South San Francisco, CA, USA.

Acknowledgments
We acknowledge Bert Lum for assistance. Support for editorial assistance for this manuscript was provided by F. Hoffmann–La Roche Ltd. These data were previously presented at the American Society of Clinical Pharmacology Therapeutics 2013 annual meeting.

Conflict of interest statement. Kelson Han is an employee of Genentech and holds stock in F. Hoffmann–La Roche. Melanie Ren was an employee of Genentech at the time of writing the manuscript. Wolfgang Wick has acted as a consultant for F. Hoffmann–La Roche, Eli Lilly (uncompensated), and MSD and received research support from Apogenix, Boehringer Ingelheim, Eli Lilly, and MSD. Lauren Abrey is an employee of F. Hoffmann–La Roche and holds stock in F. Hoffmann–La Roche. Asha Das is an employee of Genentech and holds stock in F. Hoffmann–La Roche. Jin Jin is an employee of Genentech and holds stock in F. Hoffmann–La Roche and Eli Lilly. David A. Reardon has received remuneration (other than research funding, honoraria, and consultancy fees) from F. Hoffmann–La Roche/Genentech and Merck & Co.

References


