

Long-term trends in glioblastoma survival: implications for historical control groups in clinical trials

Shehryar Sheikh, Tom Radivoyevitch, Jill S. Barnholtz-Sloan, and Michael Vogelbaum

Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, OH, USA (S.S.); Department of Quantitative Health Sciences, Cleveland Clinic Foundation, OH, USA, (T.R.); Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, OH, USA (J.S.B.-S.); Department of Neuro-Oncology, Moffitt Cancer Center, Tampa, FL, USA (M.V.)

Corresponding Author: Michael Vogelbaum, MD, PhD, Moffitt Cancer Center, Department of Neuro-Oncology, 12902 USF Magnolia Dr, Tampa, FL 33602, USA (Michael.Vogelbaum@Moffitt.org).

Abstract

Background. Historical controls continue to be used in early-phase brain tumor trials. We aim to show that historical changes in survival trends for glioblastoma (GBM) call into question the use of noncontemporary controls.

Methods. We analyzed data from 46 106 primary GBM cases from the SEER database (1998-2016). We performed trend analysis on survival outcomes (2-year survival probability, median survival, and hazard ratios) and patient characteristics (age, sex, resection extent, and treatment type).

Results. In 2005-2016 (ie, the post–Stupp protocol era), fitting a parameter independently to each year, there was a demonstrable increase in median survival ($R^2 = 0.81$, $P < .001$) and 2-year survival probability ($R^2 = 0.55$, $P = .006$) for GBM. Trend analysis of the hazard ratio showed a significant time-dependent downward trend ($R^2 = 0.62$, $P = .002$). When controlling, via multivariable Cox regression modeling, for age, sex, resection extent, and treatment type, there was a persistent downward trend in hazard ratios with increases in calendar time, especially in the most recent data.

Conclusion. Contemporary GBM patients face a different overall hazard profile from their historical counterparts, which is evident in changes in measures of patient survival and parametric hazard modeling. Though there was a plateau in these measures before 2005 (pre–Stupp protocol), there is no evidence of a new plateau in recent years even when controlling for known prognostic factors (age, sex, resection extent, and treatment type), suggesting that it may be insufficient to match contemporary patients and noncontemporary controls on the basis of these factors.

Keywords

GBM | glioblastoma | historical controls | survival trends

In recent years, there has been increasing acknowledgment that the use of “historical control” data as a proxy for a true control arm in early-phase efficacy trials is a sub-optimal research design, particularly in neuro-oncology.¹ Though investigators have previously articulated arguments against the use of historical comparisons on the basis of individual early-phase trials that produced misleading results,² no analyses of large databases have been conducted to provide evidence for this notion, and as a result, noncontemporary controls continue to be used.^{3–7} Our work is motivated by the rationale that a large longitudinal

analysis of survival data may provide compelling evidence of changing baseline trends, which could provide strong cautionary evidence against using historical comparisons in early-phase efficacy trials.

To conduct longitudinal trend analyses, we used the latest data from a large prospective national tumor registry. We assessed trends in survival before and after the generalization of the Stupp protocol,⁸ with the rationale that a consistent trend of increasing survival over time (as opposed to a one-time upward shift in survival plateau) would be consistent with improving response to standard-of-care

treatment over time. Furthermore, we assessed whether adjusting the hazard ratio for glioblastoma (GBM) based on known prognostic factors (age at diagnosis, sex, treatment type, and resection extent⁹⁻¹¹) suffices to explain increasing survival, thus commenting on whether matching historical and contemporary patients is a viable strategy in early-phase efficacy studies. Importantly, the aim of our work is not simply to document that there has been an increase in survival metrics after the popularization of the Stupp protocol, which is already a well-documented effect. Rather, we wish to investigate the possibility that there has been a persistent upward trend in GBM survival in recent years that is not completely explained by the popularization of the Stupp protocol and may be due to complex synergistic effects of multiple simultaneous improvements in the care of GBM patients (oncological and nononcological). The existence of this persistent increase in survival would have important implications for the use of historical controls in early-phase efficacy trials.

Methods

Dataset

We used the Surveillance, Epidemiology, and End Results (SEER) data released by the National Cancer Institute in April 2019.¹² SEER gathers and publishes data from registries covering approximately 28% of the US population. Exempt IRB approval was obtained from the University Hospitals of Cleveland Institutional Review Board.

Data Selection

All WHO Grade 4 gliomas available in the SEER database were included in a preliminary analysis of survival trends from 1973 through 2016 ($n = 60\,428$). For analyses of more contemporary trends (2005-2016), we focused on a subset of patients for whom all prognostic indicators of interest (age, sex, extent of resection, and treatment type) were available ($n = 46\,106$). To avoid confounding due to non-random incompleteness of data, patients with missing data were not included in the analysis. Only microscopically confirmed primary tumors were included in the analyses.

Surgical resection extent was captured from the SEER database as either radical/total resection, partial resection, or no surgery at the primary site. Treatment type was captured as including radiation (RT) or not, with or without chemotherapy, with either or both potentially unknown.

Data Analysis

Data in SEER text files were converted to R binary files using the R¹³ package SEERaBomb.¹⁴ For trend analyses, survival measures (median survival and 2-year probability of survival) were shown on the y-axis, and year of diagnosis was shown on the x-axis. Trend lines with associated adjusted R^2 and P values were shown to evaluate trends in these measures over time.

Hazard ratios were generated using 2005 as the baseline year (ie, hazard of death for GBM patients in 2005 was set at 1, and any subsequent hazard ratio less than 1 indicates a decrease in the hazard of death, ie, mortality). An adjusted hazard ratio was generated using multivariable Cox regression. Year of diagnosis was treated as an additive Cox factor, and additional covariates included sex, extent of resection, and treatment type; one model was fitted to all the data. Trend analyses for raw and adjusted hazard ratios were conducted as described above.

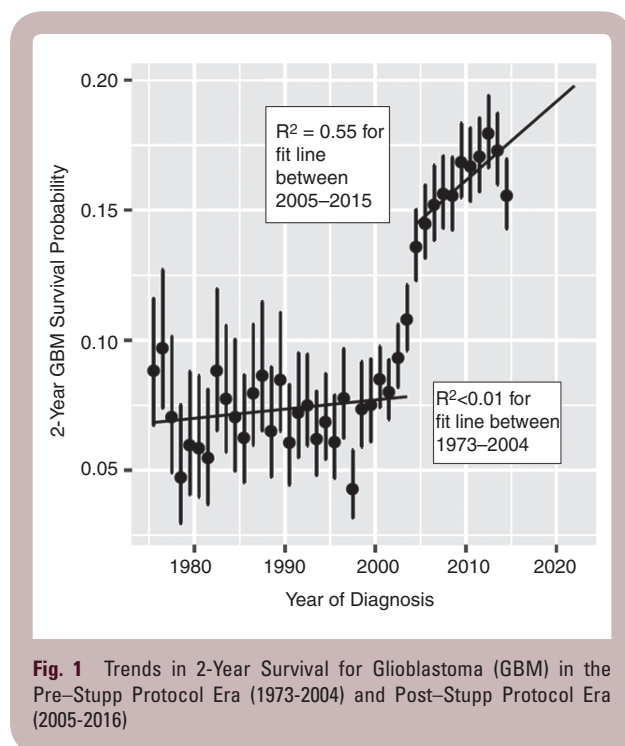
Results

Trends in 2-Year Survival Probability After Glioblastoma Diagnosis

Two-year survival probabilities after GBM diagnoses in 1973-2016 (Fig. 1) showed no clear trend before 2005 ($R^2 < 0.02$ for both). After 2005, however, there was a steady upward trend in 2-year survival probabilities (slope: 0.3%/year, $R^2 = 0.55$, $P < .006$). An analogous trend is seen for median survival after 2005 (0.2 months per year [$R^2 = 0.81$]), representing a modest but significant upward trend ($P < .001$). Of note, the survival probability trend did not plateau after 2005, showing instead a consistent upward trend.

Trend in Unadjusted Hazard Ratio for Glioblastoma 2005-2016

We used 2005 as the baseline year for comparison. This is reasonable because it coincides with the publication of the Stupp protocol. Furthermore, inspection of the data



showed that the number of patients receiving RT without chemotherapy fell markedly between 2004 and 2005 (suggesting that a time-point earlier than 2005 would be inappropriate). A significant downward trend in hazard ratio was observed (Fig. 2, $R^2 = 0.62$, $P = .002$) (Fig. 2).

Trend in Adjusted Hazard Ratio Over Time

We recognized that patient characteristics, especially regarding key prognostic indicators, may have changed meaningfully over time. Table 1 shows the changes in the makeup of the GBM patient population over time in terms of key prognostic characteristics.

To assess the degree to which changes in these patient characteristics may explain the changes in survival trends, we generated adjusted hazard ratios using multivariable Cox regression for 4 key prognostic factors (age at time of diagnosis, sex, treatment type, and extent of resection). All 4 of these factors were significant covariates in the model ($P < .001$). Thus, the adjusted hazard ratio gives a measure of the hazard that is not explained by variations in these key prognostic factors. Adjusted hazard ratio was graphed over time for trends analysis (Fig. 3). A straight line does not provide a good fit to these data. However, inspection of the data shows clearly that there has been considerable time-dependent variability in the adjusted hazard ratio. This is especially true in the most recent data, for which a downward trend is apparent.

Discussion

Our analyses of GBM survival trends show an increase in “baseline” survival over the last 10 to 15 years. This

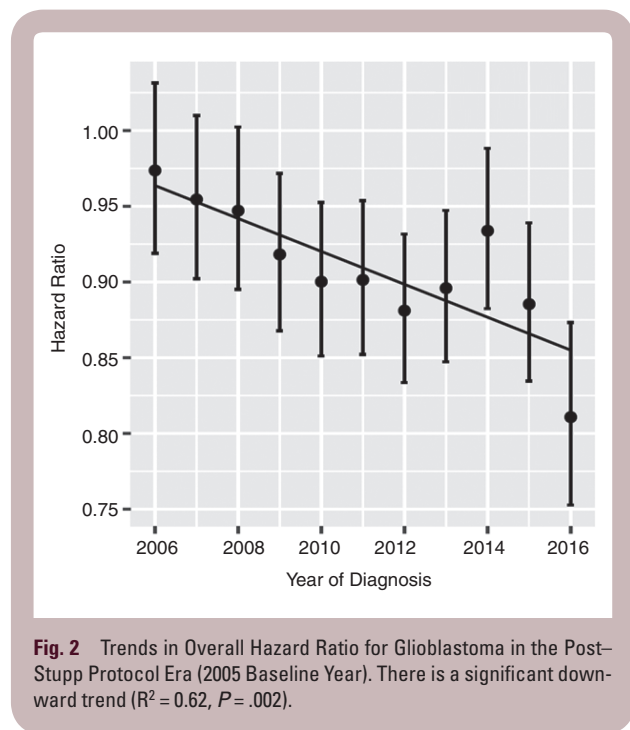


Table 1. Trends in Glioblastoma Patient Characteristics, 1998-2016

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
Patients, n	992	1071	2143	2124	2228	2320	2423	2397	2363	2604	2580	2638	2678	2727	2935	2955	2886	3024	3018	
Percentage receiving radiation and chemotherapy	24	22	24	25	32	37	47	53	58	58	59	60	63	63	63	63	64	64	64	63
Percentage receiving at least partial resection	67	67	69	68	69	72	72	70	72	72	71	72	76	74	77	77	78	77	77	77
Percentage receiving total resection	20	24	23	23	24	34	34	34	35	30	20	20	24	29	33	31	33	34	34	33
Percentage female	47	41	42	43	41	42	42	42	42	42	41	43	43	41	43	43	41	43	41	42
Mean age	64	63	63	63	63	63	63	63	63	63	63	64	63	63	63	63	63	64	64	64

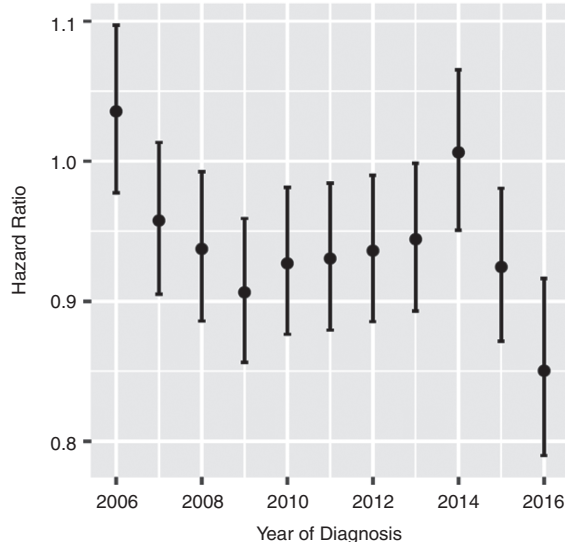


Fig. 3 Trends in Hazard Ratio for Glioblastoma After Adjustment (via Multivariable Cox Regression) for Age, Sex, Treatment Type, and Extent of Resection in the Post-Stupp Protocol Era (2005 Baseline Year) A straight line fit is not appropriate for these data.

finding echoes the reports of previous studies that have also used the SEER database to investigate GBM survival after the popularization of the Stupp protocol. McCarthy et al have reported that overall median survival for patients diagnosed in 2005-2006 was better than those diagnosed in 2001 (15 months and 12 months, respectively).¹⁵ In a similar study, Johnson and O'Neill showed that median survival times had improved when comparing patients diagnosed in 2000-2003 to those diagnosed in 2005-2008.¹⁶ Darefsky et al have reported on the increasing survival disparity between young and old patients in the temozolomide (TMZ) era, noting that 2-year survival for patients diagnosed in 2005-2007 was 39% for patients diagnosed at age 20-44 years and 1% among those diagnosed at age 80 years and older.¹⁷ Though our work confirms the findings of these earlier studies, it is important to note that our main finding is not limited to documenting the improved survival of GBM patients since the popularization of the Stupp protocol (which is already widely accepted). We show that rather than a one-time rise in the survival plateau for GBM patients after the popularization of the Stupp protocol, there is in fact a persistent upward trend. Thus, the Stupp protocol itself does not completely explain the continuing trend of improved survival over time. Our analyses are consistent with the possibility that an overall improvement in the care of GBM patients involving the synergistic effects of improved cancer-related therapies (such as bevacizumab,¹⁸ tumor-treating fields therapy,^{18,19} and immunotherapies²⁰) as well as improvements in noncancer-related patient management has led to improved outcomes.

This finding has important implications for the ongoing debate over the use of historical data for comparative purposes in single-arm, early-phase efficacy studies. We

recognize that many centers conducting early-phase trials attempt to match historical control patients with contemporary trial patients on the basis of matching key patient characteristics, which we have included in our analysis of adjusted hazard ratios. Visual analysis of the trends in adjusted hazard ratios shows clear time-dependent variability, and a downward trend is especially apparent in recent years. Even if patients enrolled in 2016 had been matched with patients as historically proximate as 2014 after adjustment for key prognostic indicators, our data suggest that patients would still not have been adequately matched in terms of their hazard of death.

The New Approaches to Brain Tumor Therapy (NABTT) CNS consortium published results from single-arm phase 2 trials using the TMZ + RT backbone with the addition of novel agents talampanel,²¹ polyinosinic-polycytidylic acid-poly-L-lysine carboxymethylcellulose,²² and cilengitide.²³ It reported improved outcomes in comparison to the historical European Organisation for Research and Treatment of Cancer (EORTC) cohort. Cilengitide was carried forward to a multicenter, open-label, phase 3 study involving 146 sites in 25 countries and 545 patients. The addition of cilengitide to TMZ + RT did not improve outcomes and the development of the drug in this clinical context has since been abandoned, with Stupp et al noting that the decision to proceed to a phase 3 study was based on previous trials that had used noncontemporary, nonrandomized comparisons.^{24,25} In a similar sequence of events, 3 separate phase 2 trials using historical controls reported encouraging results of the efficacy of rindopepimut (an epidermal growth factor receptor variant III-targeted peptide vaccine)^{4,26,27} leading to a large phase 3 trial enrolling 745 patients in 22 countries, which did not find a survival advantage of adding the new agent to TMZ therapy.²⁸

In considering the provocatively encouraging results of the early NABTT trials, Grossman and colleagues have speculated that survival following treatment with standard TMZ + RT may have improved since 2000 when the EORTC study commenced.²⁹ Our study provides evidence for this notion by showing that there has been an increased survival trend for the overall population of GBM patients in the United States since TMZ + RT became the standard of care. We see several possible reasons for this increase in "baseline" survival over time. Perhaps most important, there has been an increasingly well-founded understanding that TMZ + RT contributes to blood-brain barrier dysfunction, leading to increased peritumoral swelling, mass effect, and contrast enhancement. This "pseudoprogression" is no longer routinely considered grounds for discontinuation of therapy, which has likely contributed to improved outcomes. Furthermore, clinicians have become more adept at managing treatment complications as well as recognizing recurrence. If the trends from our data are extrapolated to future years (which is reasonable given the strength of the trends in the data), we would expect that 2-year survival measures will continue to rise. Consequently, when predicting survival for the control arm at the time of study design, future research studies would be well placed to consider that survival measures are likely to increase over the course of the study even for the control arm. Thus, this finding also has implications for the determination of a minimal clinically significant difference.

The Response Assessment in Neuro-Oncology (RANO) group has previously published guidance on the design of phase 2 trials, in which they recognize the advantages of traditional single-arm designs especially relating to the issue of small sample sizes.¹ We recognize that the prescription for including a control arm in all early-phase efficacy trials is likely to remain aspirational in the coming years. The requirement of a contemporary control arm in all early-phase trials would significantly increase the number of patients required in each trial. Because patients are not eager to be randomly assigned to a control arm, accrual may become more difficult and the total number of new agents investigated may decrease. This would be a particularly difficult situation in this highly fatal disease that has not, in effect, had a major positive drug trial since 2005. As recommended by the RANO group, one way to avoid the requirements of large sample size (associated with direct simultaneous comparison studies) and also minimize the difficulties of interpreting results of single-arm studies is to use randomized, noncomparative phase 2 designs. In these designs, patients are randomly assigned to experimental and standard-of-care arms. The experimental arm is compared with historical controls to assess efficacy, and the standard-of-care arm is compared with historical controls to assess similarity of enrolled patients and historical controls (and thus assess the appropriateness of using comparison with historical controls).

Limitations and Future Directions

A limitation of the SEER database is that patient performance status and general well-being in daily life functions are not recorded. Because the KPS is thought to be an important prognostic factor for GBM patients,¹¹ the absence of these data in the SEER database limits our ability to comment on the role of this variable in explaining the observed survival trends.

The current national registry-based study provides evidence for a global trend of improved GBM survival in recent years, but more work is necessary to assess trends in clinical subpopulations. When noncontemporary controls are thoughtfully employed in modern studies, they are not selected from SEER (a large national registry) but rather from clinical cohorts that have been well characterized on the basis of key factors, including KPS, O⁶-methylguanine-DNA methyltransferase promoter methylation status, isocitrate dehydrogenase 1 status, as well as financial and social determinants (all which are not available in SEER). Our work provides evidence of an increasing trend of survival among the *overall* GBM population that is not adequately explained by known prognostic indicators. The logical next step to investigate the validity of noncontemporary comparative arms would be to conduct analogous survival trend analysis specifically among patients enrolled in trials who have generally passed rigorous homogeneous selection criteria. If there is a persistent upward trend in survival in the trial patients even after adjustment for known prognostic factors, the use of noncontemporary comparison arms would be significantly problematized.

Conclusions

Since the introduction of the Stupp protocol in 2005, there has been a trend of increasing survival among GBM patients, and a new steady plateau of survival has not been established. This finding is consistent with the possibility that patients are responding increasingly well to the standard-of-care over time. Early-phase GBM trials often use historical controls matched with contemporary patients based on prognostic indicators. Our analyses suggest that this strategy may not be adequate for matching patients because there are differences in the residual hazard of death even after adjusting for prognostic indicators.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgment

A portion of the work in this manuscript was presented June 2, 2018, as a poster at the annual scientific conference of the American Society of Clinical Oncology in Chicago, Illinois.

Conflict of interest statement. Shehryar Sheikh, Tom Radivoyevitch, and Jill S. Barnholtz-Sloan do not disclose any financial and/or personal relationships with other people or organizations that could inappropriately influence/bias this work. Michael Vogelbaum declares that, at the time of submission, his institution has received funding from the following sources to support other research projects unrelated to the current work: Tocagen, DNAtrix, Medicenna, and Infuseon Therapeutics. He also has indirect ownership interest in Infuseon Therapeutics, and has received honoraria from Tocagen, Blue Earth Diagnostics, and Celgene.

References

1. Galanis E, Wu W, Cloughesy T, et al. Phase 2 trial design in neuro-oncology revisited: a report from the RANO group. *Lancet Oncol*. 2012;13(5):e196–e204.
2. Grossman SA, Schreck KC, Ballman K, Alexander B. Point/counterpoint: randomized versus single-arm phase II clinical trials for patients with newly diagnosed glioblastoma. *Neuro Oncol*. 2017;19(4):469–474.
3. Desjardins A, Gromeier M, Herndon JE II, et al. Recurrent glioblastoma treated with recombinant poliovirus. *N Engl J Med*. 2018;379(2):150–161.

4. Schuster J, Lai RK, Recht LD, et al. A phase II, multicenter trial of rindopepimut (CDX-110) in newly diagnosed glioblastoma: the ACT III study. *J Neurooncol.* 2015;17(6):854–861.
5. Reardon DA, Desjardins A, Peters KB, et al. Phase II study of carboplatin, irinotecan, and bevacizumab for bevacizumab naïve, recurrent glioblastoma. *J Neurooncol.* 2012;107(1):155–164.
6. Khasraw M, Lee A, McCowatt S, et al. Cilengitide with metronomic temozolomide, procarbazine, and standard radiotherapy in patients with glioblastoma and unmethylated *MGMT* gene promoter in ExCentric, an open-label phase II trial. *J Neurooncol.* 2016;128(1):163–171.
7. Lassman AB, Pugh SL, Gilbert MR, et al. Phase 2 trial of dasatinib in target-selected patients with recurrent glioblastoma (RTOG 0627). *J Neurooncol.* 2015;17(7):992–998.
8. Stupp R, Mason WP, van den Bent MJ, et al; European Organisation for Research and Treatment of Cancer Brain Tumor and Radiotherapy Groups; National Cancer Institute of Canada Clinical Trials Group. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987–996.
9. Tian M, Ma W, Chen Y, et al. Impact of gender on the survival of patients with glioblastoma. *Biosci Rep.* 2018;38(6):BSR20180752.
10. Lamborn KR, Chang SM, Prados MD. Prognostic factors for survival of patients with glioblastoma: recursive partitioning analysis. *J Neurooncol.* 2004;6(3):227–235.
11. Lacroix M, Abi-Said D, Fourney DR, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. *J Neurosurg.* 2001;95(2):190–198.
12. Surveillance, Epidemiology, and End Results (SEER) Program. Research Data (1975–2016), National Cancer Institute, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission. www.seer.cancer.gov. Accessed April 17, 2019.
13. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2013. <http://www.R-project.org/>. Accessed April 4, 2019.
14. Radivoyevitch T, Sachs RK, Gale RP, et al. Defining AML and MDS second cancer risk dynamics after diagnoses of first cancers treated or not with radiation. *Leukemia.* 2016;30(2):285–294.
15. Koshy M, Villano JL, Dolecek TA, et al. Improved survival trends for glioblastoma using the SEER 17 population-based registries. *J Neurooncol.* 2012;107(1):207–212.
16. Johnson DR, O'Neill BP. Glioblastoma survival in the United States before and during the temozolomide era. *J Neurooncol.* 2012;107(2):359–364.
17. Darefsky AS, King JT Jr, Dubrow R. Adult glioblastoma multiforme survival in the temozolomide era: a population-based analysis of Surveillance, Epidemiology, and End Results registries. *Cancer.* 2012;118(8):2163–2172.
18. Johnson DR, Leeper HE, Uhm JH. Glioblastoma survival in the United States improved after Food and Drug Administration approval of bevacizumab: a population-based analysis. *Cancer.* 2013;119(19):3489–3495.
19. Toms SA, Kim CY, Nicholas G, Ram Z. Increased compliance with tumor treating fields therapy is prognostic for improved survival in the treatment of glioblastoma: a subgroup analysis of the EF-14 phase III trial. *J Neurooncol.* 2019;141(2):467–473.
20. Jackson CM, Lim M, Drake CG. Immunotherapy for brain cancer: recent progress and future promise. *Clin Cancer Res.* 2014;20(14):3651–3659.
21. Grossman SA, Ye X, Chamberlain M, et al. Talampanel with standard radiation and temozolomide in patients with newly diagnosed glioblastoma: a multicenter phase II trial. *J Clin Oncol.* 2009;27(25):4155–4161. doi:10.1200/JCO.2008.21.6895.
22. Rosenfeld MR, Chamberlain MC, Grossman SA, et al. A multi-institution phase II study of poly-ICLC and radiotherapy with concurrent and adjuvant temozolomide in adults with newly diagnosed glioblastoma. *J Neurooncol.* 2010;12(10):1071–1077.
23. Nabors L, Mikkelsen T, Batchelor T, et al. NABTT 0306: a randomized phase II trial of EMD 121974 in conjunction with concomitant and adjuvant temozolomide with radiation therapy in patients with newly diagnosed glioblastoma multiforme (GBM). *J Clin Oncol.* 2009;27(suppl 15):2001.
24. Stupp R, Hegi ME, Gorlia T, et al; European Organisation for Research and Treatment of Cancer (EORTC); Canadian Brain Tumor Consortium; CENTRIC study team. Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated *MGMT* promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol.* 2014;15(10):1100–1108.
25. Stupp R, Picard M, Weller M. Does cilengitide deserve another chance?—Authors' reply. *Lancet Oncol.* 2014;15(13):e585–e586.
26. Sampson JH, Heimberger AB, Archer GE, et al. Immunologic escape after prolonged progression-free survival with epidermal growth factor receptor variant III peptide vaccination in patients with newly diagnosed glioblastoma. *J Clin Oncol.* 2010;28(31):4722–4729.
27. Sampson JH, Aldape KD, Archer GE, et al. Greater chemotherapy-induced lymphopenia enhances tumor-specific immune responses that eliminate EGFRvIII-expressing tumor cells in patients with glioblastoma. *J Neurooncol.* 2011; 13(3):324–333.
28. Weller M, Butowski N, Tran DD, et al; ACT IV trial investigators. Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. *Lancet Oncol.* 2017;18(10):1373–1385.
29. Grossman SA, Ye X, Piantadosi S, et al; NABTT CNS Consortium. Survival of patients with newly diagnosed glioblastoma treated with radiation and temozolomide in research studies in the United States. *Clin Cancer Res.* 2010;16(8):2443–2449.