

Evaluating the Prognostic Accuracy of Biomarkers for Glioblastoma Multiforme Using The Cancer Genome Atlas Data

Nan Hu^{1,2,3,4,5}, Haojie Cheng^{2,4}, Kevin Zhang^{2,3,4,5} and Randy Jensen^{2,6}

¹Department of Internal Medicine, Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, UT, USA. ²Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA. ³Study Design and Biostatistics Center, University of Utah, Salt Lake City, UT, USA.

⁴Department of Family and Preventive Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA. ⁵Department of Population Health Sciences, University of Utah School of Medicine, Salt Lake City, UT, USA. ⁶Department of Neurosurgery, Clinical Neurosciences Center, University of Utah School of Medicine, Salt Lake City, UT, USA.

Cancer Informatics
Volume 16: 1–11
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1176935117734844



ABSTRACT:

BACKGROUND: Glioblastoma multiforme (GBM) is the most common and aggressive primary brain tumor. Previous studies on GBM biomarkers focused on the effect of the biomarkers on overall survival (OS). Until now, no study has been published that evaluates the performance of biomarkers for prognosing OS. We examined the performance of microRNAs, gene expressions, gene signatures, and methylation that were previously identified to be prognostic. In addition, we investigated whether using clinical risk factors in combination with biomarkers can improve the prognostic performance.

METHODS: The Cancer Genome Atlas, which provides both biomarkers and OS information, was used in this study. The time-dependent receiver operating characteristic (ROC) curve was used to evaluate the prognostic accuracy.

RESULTS: For prognosis of OS by 2 years from diagnosis, the area under the ROC curve (AUC) of microRNAs, Mir21 and Mir222, was 0.550 and 0.625, respectively. When age was included in the risk prediction score of these biomarkers, the AUC increased to 0.719 and 0.701, respectively. The SAMS1 gene expression attains an AUC of 0.563, and the “8-gene” signature identified by Bao achieves an AUC of 0.613.

CONCLUSIONS: Although some biomarkers are significantly associated with OS, the ability of these biomarkers for prognosing OS events is limited. Incorporating clinical risk factors, such as age, can greatly improve the prognostic performance.

KEYWORDS: Prognostic accuracy, survival analysis, glioblastoma, The Cancer Genome Atlas, microRNA, gene signature, MGMT methylation

RECEIVED: February 23, 2017. **ACCEPTED:** September 2, 2017.

PEER REVIEW: Six peer reviewers contributed to the peer review report. Reviewers' reports totaled 2311 words, excluding any confidential comments to the academic editor.

TYPE: Review

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding was provided in part by the Huntsman Cancer Institute (HCI), Cancer Control and Population Sciences (CCPS) Pilot Grant CCPS HCI-Hu PI 22255. Partial support for this project through the Cancer Biostatistics Shared Resource within the HCI was provided by the HCI Cancer Center Support Grant (CCSG)

P30CA042014 from the National Cancer Institute and the Huntsman Cancer Foundation (HCF). Partial support was also provided by the University of Utah Study Design and Biostatistics Center (SDBC), with funding in part from the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant 5UL1TR001067-02 (formerly 8UL1TR000105 and UL1RR025764).

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Nan Hu, Huntsman Cancer Institute, University of Utah, 2000 Circle of Hope, Salt Lake City, UT 84112, USA. Email: Nan.Hu@hci.utah.edu

Introduction

Glioblastoma multiforme (GBM) is the most common and aggressive primary brain tumor, with an annual incidence rate of 2 to 3 per 10000 in the United States and 3.19 per 10000 worldwide.¹ Glioblastoma multiforme is a World Health Organization (WHO) grade IV astrocytoma of poor prognosis even with the best available multimodal therapies including surgical resection, chemoradiotherapy, and radiosurgery. For patients with GBM, worsening of symptoms is often rapid. Glioblastoma multiforme tumors are infiltrative in nature and often cannot be detected until complete resection is impossible.² Recently, advances have been made in identifying the genetic profile of glioblastomas. Many proteins produced by genes that are implicated in GBM pathogenesis can now be

measured in brain tissues or blood serum.² Although GBM biomarker research is still in its early stages, many promising biomarkers have been identified in recent years. In conjunction with other clinical data (eg, clinical risk factors and neuroimaging modalities such as magnetic resonance imaging), these biomarkers could contribute to the clinical management of GBM by helping classify tumors, assess treatment responses, and make better prognoses. The US National Institutes of Health (NIH)-sponsored *The Cancer Genome Atlas* (TCGA) project has created a comprehensive genetic profile for a variety of cancer types, including brain, lung, and colon. Glioblastoma multiforme was the first cancer to be systematically studied by TCGA. TCGA data set includes not only a vast amount of



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

gene-related biomarkers measurements but also patients' clinical risk factors and survival status. This data set provides us with opportunities not only to detect new prognostic biomarkers but also to assess prognostic performance of the biomarkers.

To evaluate the prognostic performance of new biomarkers, investigators often apply regression models (eg, Poisson regression, logistic regression, and Cox regression) to relate a biomarker with disease or survival events based on parameters such as relative risk, odds ratio, and hazard ratio. These statistical models can test the association between a biomarker and the risk of a disease or survival event. Most previous studies on GBM biomarkers investigated the biomarker effect on patients' overall survival (OS). A few of them were trying to establish prediction models for OS using biomarkers as the predictor variables. Although studies of associations are usually adequate in etiologic research, they do not examine the prognostic ability of biomarkers to correctly predict OS in the future. To assess biomarkers for their classification accuracies, a receiver operating characteristic (ROC) curve is commonly used by plotting pairs of sensitivity and specificity across all possible classification thresholds. Traditional ROC curves evaluate biomarkers with a continuous distribution for classifying dichotomous status, such as diseased versus nondiseased or alive versus dead. The Supplemental Method section within the Supplemental Material provides the definition of sensitivity and specificity, as well as a detailed description on how a traditional ROC curve can be generated based on the biomarker distributions among diseased and nondiseased subjects. Because an ROC curve is a plot across all possible decision thresholds within the range of the biomarker values, it does not require specification of threshold values.³ Area under the ROC curve (AUC) is often used to assess the overall diagnostic performance of biomarkers.⁴

For prognosis of future disease or survival events, however, we often need to consider timing of the disease development. A patient with cancer who is currently alive may die in either the near future or the far future. Therefore, prognosis of a disease or disease-related outcomes, such as OS, progression-free survival, and treatment responses, often needs to deal with a time-varying process. Statistical methods for survival (or time-to-event) analysis can evaluate the prognostic importance of a biomarker through its effect on the disease or survival event. For example, Cox proportional hazards models can estimate the hazard ratio associated with one unit increase in the biomarker value if the biomarker value has a continuous distribution or the hazard ratio comparing the positive and the negative test results when marker values are dichotomized. For prognosis of OS, the mortality status for a given patient will vary across the whole follow-up period. For this reason, the prognostic accuracies (sensitivity, specificity, and AUC) of a biomarker will be time dependent. The traditional ROC analysis will no longer be appropriate under this situation. In addition, the OS time is often subject to censoring, and removing censored subjects from the study can often lead to biased estimates of sensitivity and specificity. To accommodate the problem, the time-dependent ROC method was developed to assess the accuracies of a baseline biomarker

for prognosing of a future event, such as OS. For instance, to assess the performance of a biomarker for prognosing OS by 24 months after GBM diagnosis, we can plot the time-dependent ROC curves and obtain the AUC as well as the sensitivity and specificity at a given threshold. Hu and Zhou^{5,6} summarized recent statistical methods in time-dependent ROC analysis, including works by Etzioni et al,⁷ Heagerty et al,⁸ Heagerty and Zheng,⁹ Cai et al,¹⁰ and Song and Zhou.¹¹

In prognosis studies, patients are monitored until (a) they reach the event of interest (such as disease diagnosis or death), (b) they drop off the study, or (c) the end of the study. The time interval from baseline to event occurrence is called an event time, whereas the time interval from baseline to loss of follow-up or to the end of the study is called a censoring time. For a baseline biomarker Y , Heagerty et al⁸ proposed 2 types of time-dependent sensitivity or true-positive rate (TPR), the cumulative TPR and the incident TPR, as well as the dynamic specificity, or true-negative rate. Based on the 2 kinds of time-dependent TPR, 2 types of time-dependent ROC curves are defined as follows: (a) ROC curve of a biomarker Y at time t based on the cumulative TPR and the dynamic specificity, denoted by $ROC^{C/D}(Y; t)$ and (b) ROC curve of the biomarker Y at time t based on the incident TPR and dynamic specificity, denoted by $ROC^{I/D}(Y; t)$. The definitions of time-dependent TPRs, specificity, and ROC curves (Figure S1 in the Supplemental Methods) as well as interpretation of the time-dependent ROC curves can be found from the Supplemental Method section within the Supplementary Material.

Evaluating prognostic accuracies is an important aspect of studying GBM biomarkers that has not been discussed in the literature. In this study, we use both $ROC^{C/D}(Y; t)$ and $ROC^{I/D}(Y; t)$ to evaluate biomarkers for prognosis of cumulative and incident OS events by or at a series of time points (15, 24, 33, and 42 months since GBM diagnosis). The biomarkers that we evaluated in this study were identified by the previous literature, and the values of these biomarkers for patients with GBM can be obtained directly from TCGA data set.

Materials and Methods

Data collection

The Cancer Genome Atlas Project provides data from more than 500 GBM cases as of 2015. TCGA data set can be easily accessed from its Web page: <https://gdc.cancer.gov>. In this study, we assessed the prognostic performance of the several promising candidate biomarkers and gene signatures reported in the literature.¹²⁻¹⁷ Some previous research using TCGA data set reported the results of genomic and transcriptomic analysis of 206 GBM cases with about 600 genes.¹⁸ The SAMSAN1 gene was identified as an important factor for OS among patients with GBM from another study using TCGA data.¹² In addition, by running a Cox regression model using TCGA data, Bao et al¹³ found that 4 gene signatures ("5-gene," "8-gene," "17-gene," and "61-gene" signatures) were associated with OS.

Table 1. Demographic and clinical characteristics of all patients in the analytical data set and by survival status.

VARIABLE		TOTAL (N=584)	ALIVE (N= 138)	DEAD (N=443)	P VALUE
Sex	Female	225 (38.6)	56 (40.6)	169 (38.1)	.556 ^a
	Male	358 (61.4)	82 (59.4)	274 (61.9)	
Race	Missing	25 (4.3)	5 (3.6)	20 (4.5)	.115 ^a
	Asian	13 (2.2)	7 (5.1)	6 (1.4)	
Ethnicity	Black or African American	40 (6.9)	13 (9.4)	27 (6.1)	
	White	505 (86.6)	113 (81.9)	390 (88)	
	Hispanic or Latino	13 (2.2)	4 (2.9)	9 (2)	
	Not Hispanic or Latino	477 (81.8)	109 (79)	368 (83.1)	
	Not available	93 (16.0)	25 (18.1)	66 (14.9)	.052 ^a
Histologic type	GBM	18 (3.1)	1 (0.7)	17 (3.8)	.159 ^a
	Treated primary GBM	20 (3.4)	2 (1.4)	18 (4.1)	
	Untreated primary (de novo) GBM	545 (93.5)	135 (97.8)	408 (92.1)	
Age (mean ± SD)		57.9 ± 14.4	56.5 ± 14.7	58.3 ± 14.3	.203 ^b
Karnofsky Scores (mean ± SD)		77.4 ± 14.6	78.6 ± 15.1	77 ± 14.4	.338 ^b

Data are reported as count (%) unless otherwise indicated.

^aUsing Fisher's exact test.

^bUsing two-sample Student's t test.

Mir21 and Mir222 were reported to be significantly associated with OS by Rao et al.¹⁴ In addition, MGMT methylation has also been intensively studied in the recent neuro-oncology literature. Thus, we chose to evaluate 2 microRNAs (miRNAs) (Mir21 and Mir222), 1 gene expression (SAMS1), 4 gene signatures (gene signature based on 5, 8, 17, and 61 genes identified by Bao et al),¹³ as well as the MGMT methylation.

We downloaded the miRNA expression microarray data and the messenger RNA (mRNA) expressions (level 3) of GBM samples without specifying any other conditions. The gene expressions of SAMS1 (level 3) were based on Affymetrix microarrays (Human gene U133A). Our analytical data set was obtained by merging TCGA survival and clinical data with the patient-level biomarker data.

Statistical analysis

Patients' characteristics were summarized as mean ± standard deviation for nonskewed continuous variables and as frequency (N) and percentage (%) for categorical variables. The Kaplan-Meier survival function and smoothed hazard rate were plotted. The OS event time is defined as the time interval between GBM diagnosis and mortality and is calculated in terms of months. Overall survival times are censored at either the loss to follow-up or the end of the study (the end of 2015).

Statistical methods have been developed to perform time-dependent ROC analysis. Hu and Zhou⁵ reviewed the recent statistical development in this area. The time-dependent ROC

approach developed by Song and Zhou¹¹ was selected (see the "Discussion" section for the reason for this choice). The AUC was used to evaluate the overall prognostic performance of the biomarkers. For Mir21, Mir222, SAMS1, and MGMT methylation, we evaluated the prognostic accuracy for these biomarkers and the predicted scores that combine each of these biomarkers with either age or Karnofsky Performance Score (KPS), the 2 important clinical risk factors reported in literature. To obtain the predicted risk scores, multivariable Cox regression models were generated, and the predicted scores were calculated using the weighted combination of the biomarker value and one of the clinical risk factors (age or KPS). To evaluate the prognostic accuracy of 4 gene signatures, the risk scores of gene signatures for each patient based on 5, 8, 17, and 61 genes were calculated using the linear combination of the gene expression level weighted by the regression coefficients reported in Table 1 (log of hazards ratios) of Bao et al.¹³

The Harrell C statistic¹⁹ was used to estimate overall concordance between the OS time and each of the biomarkers described above. All statistical analyses were performed using statistical package R (www.r-project.org) versions 3.1 to 3.3 (for time-dependent ROC analyses, with code written by N.H.) and Stata (Stata Corp., College Station, TX, USA) version 14 (for all other analyses).

Results

Table 1 reports the demographic features of our study cohort. Our raw data set includes 584 patients with GBM, of which 3

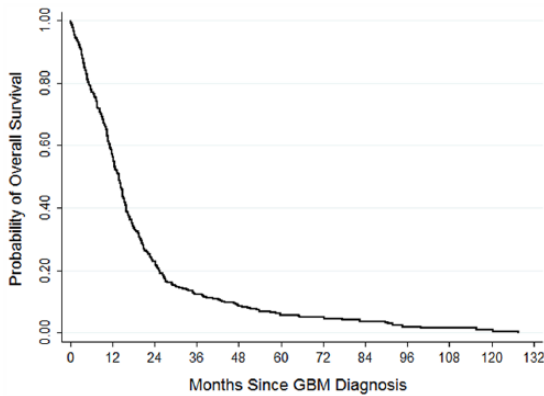


Figure 1. Kaplan-Meier product-limit survival curve of overall survival among all patients with GBM in The Cancer Genome Atlas. The median survival time among these patients is about 14 months. GBM indicates glioblastoma multiforme.

subjects were missing information on OS status and were excluded from all analyses related to survival. The mean age of the 584 patients was 57.9 ± 14.4 years, and their mean KPS was 77.4 ± 14.6 . Our analytical cohort includes 581 patients with observed OS status, of which 138 subjects were alive and 443 patients were dead at the end of the follow-up period. Most of the patients (359 patients, 61.4%) were men. About 87% ($N=505$) of the patients in our analytical cohort were white, 7% ($N=40$) were black, and 2% ($N=13$) were Asian. About 94% ($N=545$) of these patients had untreated primary (de novo) GBM and 3% ($N=20$) had treated primary disease. The median OS time was 14 months (Figure 1). The hazard rate function has 2 local peaks at around 20 and 90 months before it eventually rises toward the end of the study follow-up time (Figure S2 in the Supplemental Results).

MicroRNAs

A total of 565 patients within our analytical cohort had measurements of Mir21 and Mir222. The Harrell C statistic for Mir21 and Mir222 was 54% and 56%, respectively. For both Mir21 and Mir222, when a clinical risk factor (either age or KPS) was included in the predicted risk score (as the classifier), the overall concordance (the Harrell C) rose. The Harrell C statistic of Mir21 increased to 66% and 62% when age and KPS, respectively, were involved. For Mir222, the Harrell C value increased to 66% and 64% when age and KPS were involved, respectively.

Figure 2 shows the time-dependent ROC curves, $ROC^{C/D}(Y; t)$ (panel A) and $ROC^{I/D}(Y; t)$ (panel B), of Mir21 (denoted by Y in ROC curve notations) for prognosing OS events by or at $t=15, 24, 33,$ and 42 months. The ROC curves for predicted risk scores combining Mir21 with clinical risk factors (age or KPS) were also shown in the plot. The overall accuracy of Mir21 for prognosing cumulative OS events by 15 months was 0.539 and increased gradually to 0.566 by 42 months. When age was included in the predicted risk score, AUC for prognosing cumulative OS events increased to 0.688 and 0.756 by 15 and

42 months, respectively. When KPS was included in the prognosis, the overall performance also improved but not as much as the predicted risk score that included age as a clinical risk factor. When prognosis of incident OS events was considered, AUC for Mir21 was around 0.530 at all of the 4 time points. When age was included in the predicted score, the AUC increased to 0.629, 0.611, 0.602, and 0.598 at 15, 24, 33, and 42 months, respectively. When KPS was included in the predicted score, there was also an increase in AUC, although the improvement was much smaller than when age was incorporated. Figure 3 shows the time-dependent AUC of Mir21 based on $ROC^{C/D}(Y; t)$. The AUCs for Mir21 and the 2 predicted risk scores all increased from baseline toward month 60. For example, for AUC of the age-involved predicted risk score increased from 0.58 at baseline to 0.79 by month 60. The time-dependent AUC curve based on $ROC^{I/D}(Y; t)$ can be found from Figure S3 in the Supplemental Results. Figure 4 plots the time-dependent ROC curves of Mir222. For prognosis of cumulative OS events, the AUC of Mir222 was 0.597, 0.625, 0.648, and 0.664 by 15, 24, 33, and 42 months, respectively. However, the AUC of Mir222 for prognosing incident OS events at these time points was 0.574, 0.574, 0.575, and 0.575, respectively. When age or KPS was included in the classifier, AUC for both $ROC^{C/D}(Y; t)$ (panel A) and $ROC^{I/D}(Y; t)$ (panel B) went up, whereas the predicted risk score incorporating age increased more dramatically. Figure 5 plots the AUC of Mir222 based on $ROC^{C/D}(Y; t)$. When using Mir222 expression as the classifier, AUC for prognosing cumulative OS events increased gradually from 0.550 at the beginning to 0.680 by 60 months. When age was included in the predicted risk score, the AUC increased from 0.660 to 0.760 by 60 months. When KPS was involved in the predicted risk score, the AUC ranged from 0.610 to 0.670. It is interesting to note that, after about 30 months, the AUC of Mir222 expressions was greater than that of the predicted score combining Mir222 and KPS. The time-dependent AUC based on $ROC^{I/D}(Y; t)$ can be found from Figure S4 in the Supplemental Results section.

SAMSN1 expression

A total of 524 patients had SAMSN1 gene expressions recorded in TCGA. We calculated overall concordance (the Harrell C) of OS time with SAMSN1 expression as well as OS time with the predicted risk scores involving SAMSN1. Overall concordance between OS and SAMSN1 was only 51%. When KPS was included in the predicted risk score, the Harrell C increased to 61%. When age was involved in the prognosis, the C statistic increased to 66%.

Figure 6 plots the time-dependent AUCs of SAMSN1 based on $ROC^{C/D}(Y; t)$. The ability of SAMSN1 expression to predict cumulative OS events was very limited, with AUC ranging from 0.54 to 0.60. When age was included in the predicted risk score, AUC for prognosing cumulative OS events increased gradually from 0.66 to 0.80 by month 60. When KPS was involved in the prognosis, the AUC was between 0.60 and

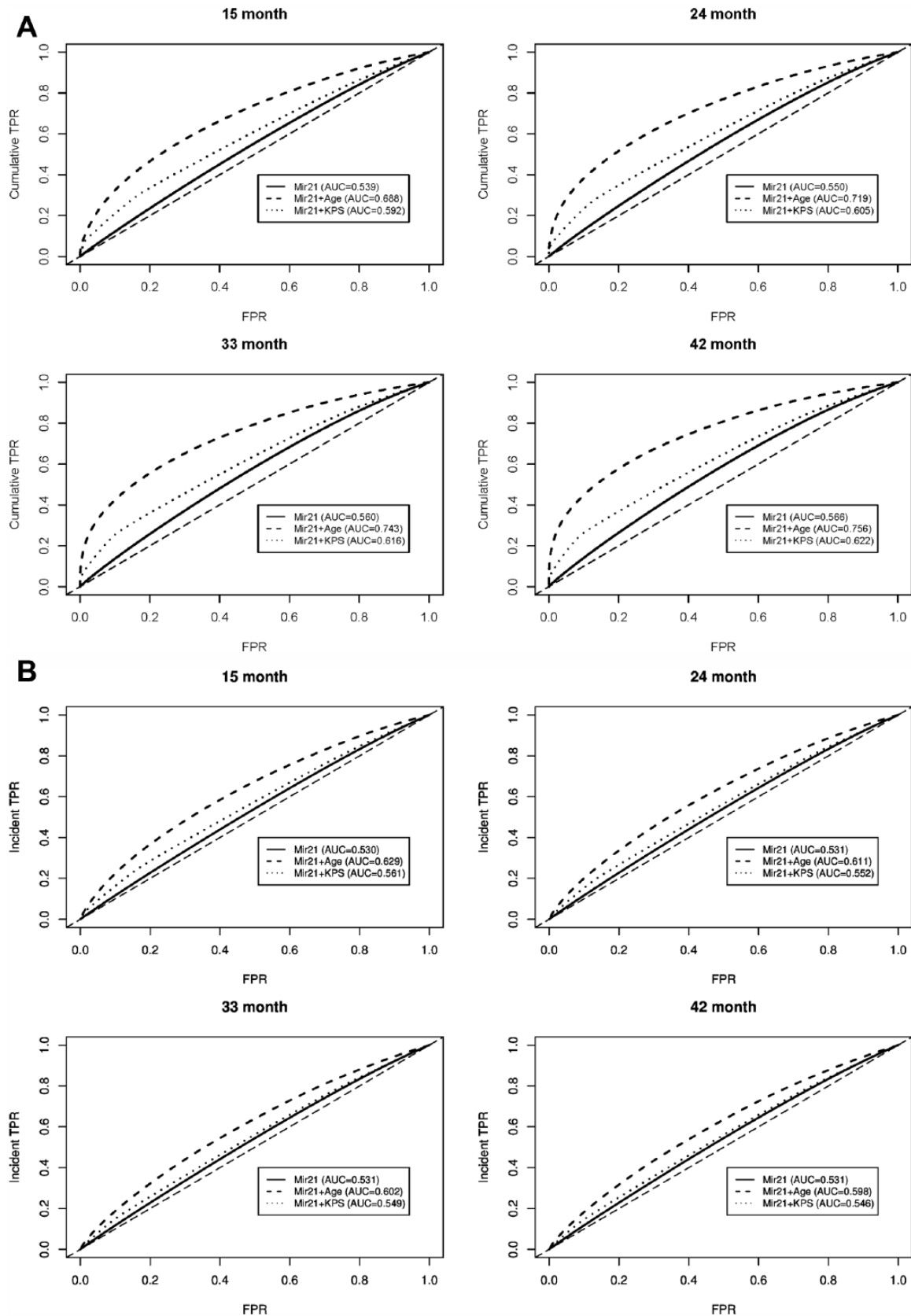


Figure 2. Time-dependent ROC curves of Mir21 based on (A) $ROC^{C/D}(Y; t)$ and (B) $ROC^{I/D}(Y; t)$ by or at 15, 24, 33, and 42 months. AUC indicates area under the ROC curve; KPS, Karnofsky Performance Score; ROC, receiver operating characteristic; TPR, true-positive rate.

0.65. The time-dependent ROC curve of SAMS1N1 expressions is presented in Figure S5 in the Supplemental Results. For prognosing OS by 24 months, the SAMS1N1 expression

can achieve an AUC of 0.563. When age or KPS was included in the predicted risk score, the AUC was 0.716 and 0.615, respectively.

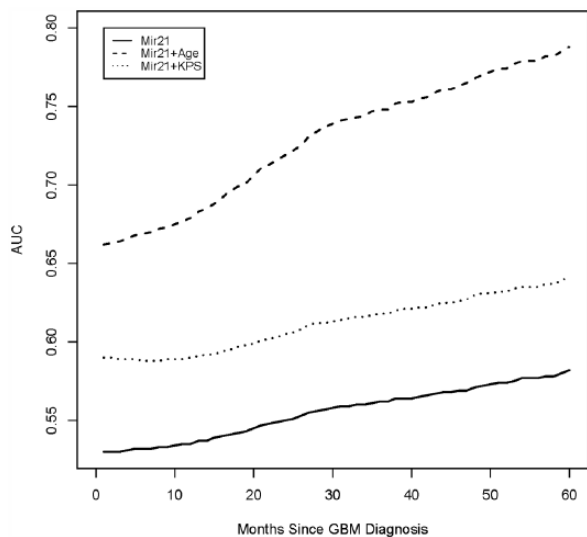


Figure 3. Time-dependent AUC of Mir21 based on $ROC^{C/D}(Y; t)$. AUCs were calculated up until 5 years (60 months) after diagnosis. The Mir21 + age classifier has the best overall accuracy (largest AUC) to prognose cumulative overall survival events over the entire 60-month period, followed by the Mir21 + KPS classifier. AUC indicates area under the ROC curve; GBM, glioblastoma multiforme; KPS, Karnofsky Performance Score; ROC, receiver operating characteristic.

Gene signatures

A total of 540 patients were included in analysis of the 4 gene signatures reported by Bao et al.¹³ All of these 4 gene signatures were statistically significantly associated with OS (log rank test $P < .05$ for all gene signatures). Our analysis showed that among the 4 gene signatures, the 8-gene signature had the best prognostic performance. For prognosis of cumulative OS events by 24 months, the 8-gene signature attained an AUC of 0.613, whereas the AUCs for the 5-gene, 17-gene, and 61-gene signatures were 0.571, 0.594, and 0.594, respectively (Time-dependent ROC curves are presented in Figure S6 in the Supplemental Results). The time-dependent AUC based on $ROC^{C/D}(Y; t)$ is shown in Figure 7. For prognosis of cumulative OS events by 60 months, the 8-gene signature had the largest AUC (0.670) among the 4 gene signatures. In comparison, the AUC for both the 17-gene and 61-gene signatures was 0.650. The 5-gene signature had an AUC of 0.610.

MGMT methylation

A total of 567 subjects from TCGA were included in analysis of MGMT methylation. Figure S7 (in the Supplemental Results) plots the time-dependent ROC curves of MGMT methylation based on both $ROC^{C/D}(Y; t)$ and $ROC^{I/D}(Y; t)$. Similar to other single biomarkers, MGMT methylation had a limited ability to predict both cumulative and incident OS events. For short-term prognosis of cumulative OS events, the AUC was 0.528 and 0.551 by 15 and 24 months, respectively. For long-term prognosis, the AUC increased to 0.562 and 0.567 by 33 and 42 months, respectively. When age was involved in the prognosis, the AUC

for prognosing cumulative OS events increased dramatically to 0.704, 0.729, 0.754, and 0.766 by 15, 24, 33, and 42 months, respectively. When KPS was involved, the AUC at all of the above time points was between AUC of MGMT methylation and that of predicted risk score combining MGMT with age. As for prognosis of incident OS events, the performance of MGMT methylation across the 4 time points varied little. However, when age or KPS was incorporated in the prognosis, the short-term prognostic performance at 15 and 24 months was much better than the long-term performance.

Discussion

Glioblastoma multiforme is the most aggressive and malignant type of glioma. Despite improved medical interventions including radiation, chemotherapy, and surgery, the median survival time of GBM is very low. In our study cohort, the median OS time is only 14 months, which is consistent with the 14.6-month median survival reported by Salzman.²⁰ Detection of new biomarkers or genetic signatures is crucial to improve accuracies for GBM prognosis. In this research, we evaluated the performance of miRNAs, gene expression, gene signatures, and gene methylation for prognosing cumulative and incident OS events using time-dependent ROC curves.

Currently, the role of single miRNAs for prognosis among patients with GBM is poorly understood. MicroRNAs usually function as small noncoding RNAs that potentially increase the instability and may limit translation of target mRNAs.²¹ They are significant genetic regulators and play a very important role in pathogenesis, development, and prognosis of GBM as well as of other cancers.²² It has been proven that miRNAs are expressed differently among patients with GBM. Previous studies have also shown significant prognostic importance of mRNAs on OS.¹² As a single miRNA, Mir21 has been widely involved in the regulation of GBM.²³ Studies have shown that Mir21 is expressed at higher levels in GBM cell lines as compared with normal cells. A high level of Mir21 was found to be significantly associated with shorter survival time.²⁴ The miRNA Mir222 is directly associated with the grades of GBM and is targeted at the important cell cycle regulator, p27/KIP1. By inhibiting Mir222, researchers have shown that p27/KIP1 expression was upregulated and this resulted in a decline of tumor volume.²³ Our study on prognostic accuracy, however, showed that single miRNAs have a very limited ability for an accurate prognosis of cumulative and incident OS events. For instance, the time-dependent AUC of Mir21 for prognosing cumulative OS was always lower than 0.60 within the first 60 months of follow-up.

The neuro-oncology literature reported that the SAMSIN1 gene is overexpressed in glioblastoma tumors when compared with normal brain cells. A high expression of SAMSIN1 indicates an unfavorable outcome of GBM in OS. Yan et al.¹² reported that the median OS time was 11 months among patients with GBM with high SAMSIN1 expression and about 15 months among subjects with low SAMSIN1 expression.

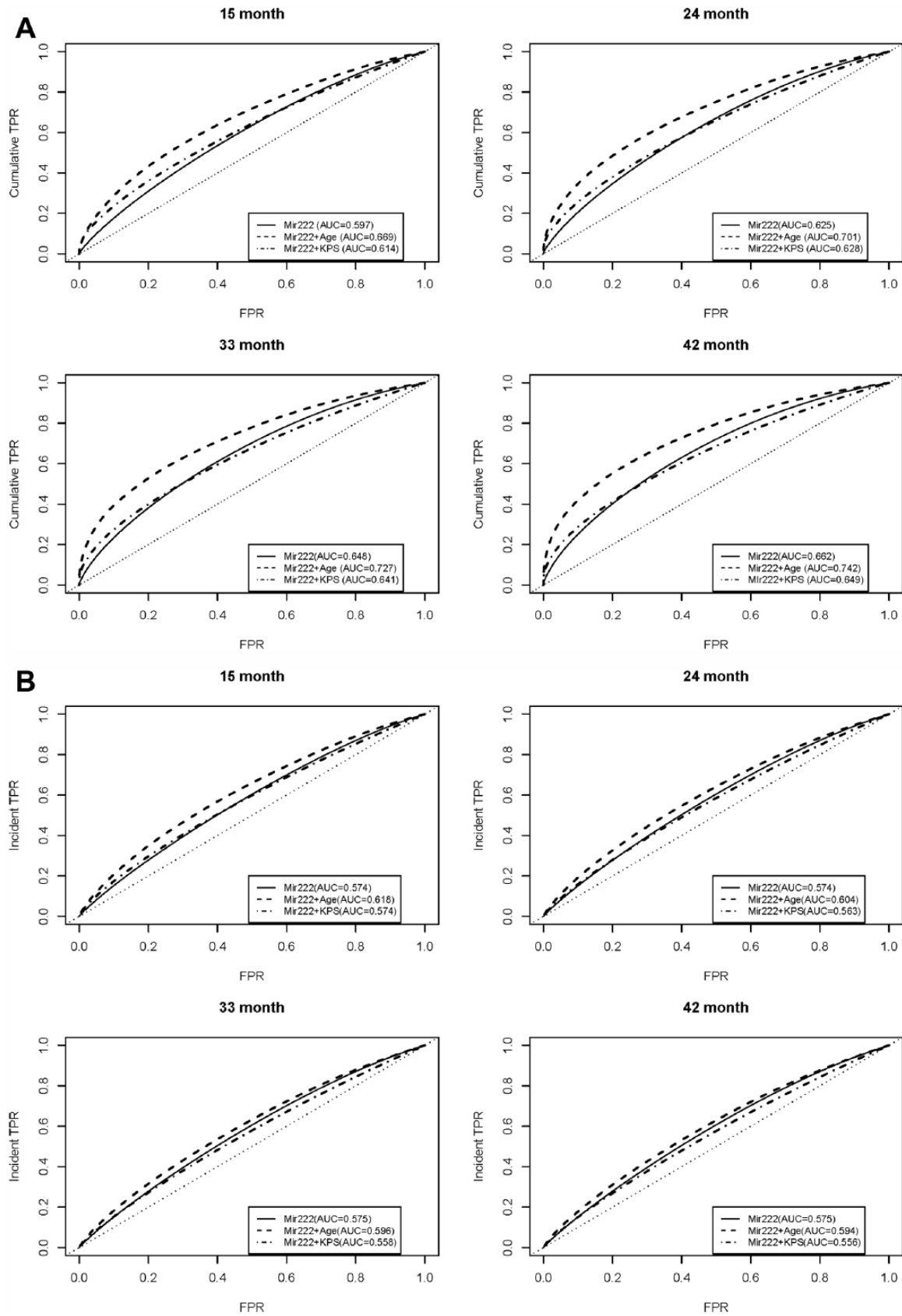


Figure 4. Time-dependent ROC curves of Mir222 based on (A) $ROC^{C/D}(Y; t)$ and (B) $ROC^{I/D}(Y; t)$ by or at 15, 24, 33, and 42 months. AUC indicates area under the ROC curve; KPS, Karnofsky Performance Score; ROC, receiver operating characteristic; TPR, true-positive rate.

However, our study showed that when SAMSN1 expression was used on its own as a biomarker for prognosis, its ability is

also limited. For prognosing cumulative OS events by 15 months, the AUC for SAMSN1 was only 0.549.

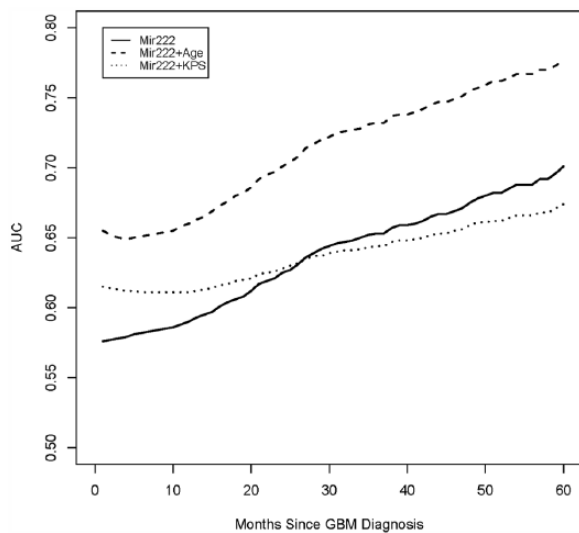


Figure 5. Time-dependent AUC of Mir222 based on $ROC^{C/D}(Y; t)$. AUCs were calculated up until 5 years (60 months) after diagnosis. The Mir222 + age classifier has the best overall accuracy (largest AUC) to prognose cumulative overall survival events over the entire 60-month period. There is a crossover between the time-dependent AUC curve for Mir222 and Mir222 + KPS. This indicates that during the earlier follow-up period ($\sim <30$ months), incorporating KPS in the prognosis will result in better prognostic accuracies than only using the Mir222 expression; however, for prognosis after 30 months, including KPS will not improve the overall prognostic performance. AUC indicates area under the ROC curve; GBM, glioblastoma multiforme; KPS, Karnofsky Performance Score; ROC, receiver operating characteristic.

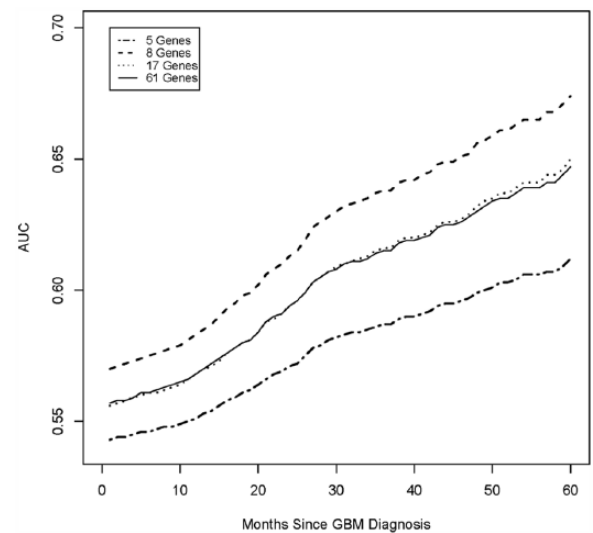


Figure 7. Time-dependent AUC of 4 gene signatures reported by Bao et al¹³ based on $ROC^{C/D}(Y; t)$. The 8-gene signature has the best overall accuracy (largest AUC) to prognose cumulative overall survival events over the entire 60-month period. The performance of the 17-gene and 61-gene signatures is similar as the time-dependent ROC curves for them are close to each other over the 60-month period. AUC indicates area under the ROC curve; GBM, glioblastoma multiforme; KPS, Karnofsky Performance Score; ROC, receiver operating characteristic.

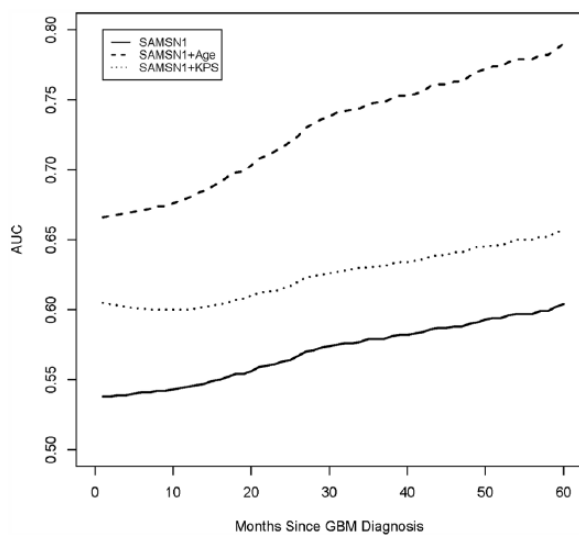


Figure 6. Time-dependent AUC of SAMS1 based on $ROC^{C/D}(Y; t)$. AUCs were calculated up until 5 years (60 months) after diagnosis. The SAMS1 + age classifier has the best overall accuracy (largest AUC) to prognose cumulative overall survival events over the entire 60-month period, followed by the SAMS1 + KPS classifier. AUC indicates area under the ROC curve; GBM, glioblastoma multiforme; KPS, Karnofsky Performance Score; ROC, receiver operating characteristic.

Previous studies reported that certain gene signatures were significantly associated with OS among patients with GBM. Bao et al¹³ identified 4 gene signatures using a risk prediction model. Then, a cutoff value (median risk score for each signature) was

used to divide patients into high-risk and low-risk groups. The authors showed that the median OS among the high-risk group was significantly shorter than that among the low-risk group.¹³ In our study, among the 4 gene signatures, the 8-gene signature performed the best when prognosing both cumulative and incident OS events.

Our study showed that clinical risk factors (such as age and KPS) could play an important role in prognosing OS events among patients with GBM. Age and KPS are 2 strong known clinical risk factors that can influence OS among patients with GBM and are reported in TCGA data. Previous studies showed statistically significant negative relationships between age and the OS of GBM.^{25,26} That is, older ages are associated with a higher risk of mortality among patients with GBM. Using 40 years as the age cutoff, Walid²⁶ showed that the 5-year survival rate among patients with GBM younger than 40 years of age was 5.67 times the rate for patients at 40 years of age and older. Previous literature also revealed that patients with GBM with KPS less than 80 and more than 40 years of age had the worst survival rate.²⁷ This implies that age and KPS may predict OS of patients with GBM. Our study showed that involving age or KPS as a clinical risk factor can improve the prognostic accuracy of biomarkers and gene signatures. When age was combined with Mir21, Mir222, or SAMS1, the prognostic performance usually improved dramatically. For example, when prognosing cumulative OS events by 24 months, the AUC for SAMS1 expression was only 0.563, but when age was involved, the AUC increased to 0.716. Our time-dependent ROC results for each biomarker, each biomarker in combination with age, and each biomarker in combination

with KPS provide researchers with an easy way to assess how much increment in AUC the clinical risk factors can cause. We also performed time-dependent ROC analysis on biomarkers in combination with both KPS and age. For all of the biomarkers and gene signatures being considered in this study, the prognostic performance is very close to the predicted score with only age included as the risk factor. This indicated that incorporating KPS in the predicted risk score based on age and the biomarker (or gene signature) does not further improve the overall prognostic accuracy. Sex and race are other demographic features recorded in TCGA data. We also investigated the impact of these 2 variables on prognostic accuracies, but none of them further improved the prognostic accuracies for all biomarkers and gene signatures being considered in this study.

Our analytical results showed that time-dependent AUCs based on $ROC^{C/D}(t)$ increased over the follow-up period (Figures 3, 5 to 7), so these AUCs were often greater at later time points than at earlier time points. Theoretically, the number of events is monotone nondecreasing under the cumulative events setting. Hence, there will always be more OS events at the later follow-up stage. When a biomarker is associated with the OS events, there will be a growth in both cumulative event rate conditional on positive test results and the survival rate conditional on negative test results. This, in turn, will elevate both sensitivity and specificity at the same positivity threshold. However, under the setting of incident events, the local estimate of the hazard rate has a large impact on the prognostic accuracies. Hence, even when a biomarker is associated with the OS events, the time-dependent AUC curve will not have a monotone trend over time. Moreover, it may have a drastic local fluctuation (Figures S3 and S4).

Time-dependent ROC analysis provides investigators with a way for prognosing both short-term and long-term survivorship. For prognosis of high-grade brain tumors, such as GBM, 24 months is often used as a threshold to classify the survivorship into short term and long term. The short-term (≤ 24 months) prognostic performance of biomarkers could be quite different from their long-term (> 24 months) prognostic performance. Understanding long-term survivorship in GBM is a critical unmet need in neuro-oncology. Examining the prognostic performance by 33 or 42 months is a way to evaluate potential biomarkers and gene signatures that would predict long-term survivorship.

Time-dependent AUC is a good summary measure to compare the time-dependent prognostic performance among different biomarkers or predicted risk scores. For example, Figure 5 shows an obvious crossover between the time-dependent AUC curve for Mir222 expression and for Mir222 together with KPS. This indicates that at the earlier follow-up period (up until about 30 months), incorporating KPS in the prognosis will result in better prognostic accuracies than using only the Mir222 expression. However, involving KPS cannot improve the overall prognostic performance after 30 months. In addition, the AUC for Mir222 expression increased more dramatically than that for Mir222 expression over the entire 60-month period.

Methodologically, the goal of our study is to demonstrate how to apply the time-dependent ROC analysis to evaluate performance of biomarkers for prognosing cumulative or incident OS events at different time points. Pepe et al²⁸ suggested that 2 basic statistical approaches can be used to evaluate biomarkers. The first method is to model the risk of disease (or disease outcome such as survival) as a function of biomarkers with other predictors using logistic or Cox regression. Using this approach, the value of a biomarker is measured by its effect on OS conditional on other predictors. Pepe et al²⁸ pointed out that this method is adequate in etiologic research but does not address the ability of biomarkers to correctly classify or predict risk for the study population. The second method evaluates biomarkers with measures such as sensitivity, specificity, predictive values, and ROC curves. There is controversy about which approach is more appropriate. Moons and Harrell²⁹ claimed that risk models are much better because ultimately the patients want to know the risk of a disease given their biomarker measurement. Pepe et al³⁰ argued, however, that the public health value of a biomarker lies in the fraction of diseased subjects being detected (ie, sensitivity) and the fraction of nondiseased subjects falsely identified as diseased (ie, $1 - \text{specificity}$). The difficulty lies in the fact that these 2 approaches frequently yield apparently contradictory results. A biomarker that is strongly related to risk may be a poorly performing classifier.³⁰ A biomarker that is a strong risk predictor after controlling for other factors may add little to classification performance. Our study showed that both SAMS1 and MGMT methylation have very limited ability for prognosing OS events, although the literature shows both are highly associated with OS of patients with GBM. For this reason, we believe that it is necessary to investigate the prognostic accuracy for biomarkers in addition to the prognostic importance based on regression coefficients from risk models. This warrants consideration for future studies on prognostic accuracies of GBM biomarkers. In this work, we used the analytical method proposed by Song and Zhou¹¹ because this approach was shown to be the most efficient among several influential time-dependent ROC methods evaluated by Pepe et al.³¹

For assessing concordance of a biomarker with an event time outcome, the Harrell C statistic is often adopted. The Harrell C statistic for right censored time-to-event data is the conditional probability of concordance plus half the conditional proportion of data pairs that are neither concordant nor discordant. The benefit of the time-dependent ROC method over the Harrell C statistic lies in the way the time-dependent ROC method handles censoring. Time-dependent ROC curve methods often use time-to-event models for estimating conditional survival functions, and time-to-event models were developed specifically for dealing with censored time-to-event data. For this reason, time-dependent ROC curves should be more reliable and accurate when evaluating prognostic abilities of biomarkers. In addition, time-dependent ROC curves can estimate sensitivity and specificity at each time point of interest within the study period, whereas the Harrell C can only provide an overall concordance across the entire period.

A couple of alternatives to ROC curves and AUCs have been proposed to evaluate diagnostic tests in a more clinically meaningful way. For example, Halligan et al³² proposed the net benefit as a measure to evaluate radiologic tests. The primary goal of our study is to introduce the time-dependent ROC analysis in neuro-oncology and apply it to prognostic studies of GBM biomarkers. Our analysis focused on the early stage of biomarker studies (ie, biomarker discovery and validation). At the stage of biomarker discovery and validation, ROC analyses are intensively used to evaluate the performance of the biomarkers; hence, the prognostic accuracy should be established before clinical meaningful measures (such as cost) can be considered. Although Halligan et al³² claimed that the “net benefit” is more meaningful in evaluating the diagnostic performance of radiological tests, they agreed that ROC and AUC are most useful in the early stages of diagnostic test assessment, especially for tests not requiring subjective interpretation.

Study limitations

One limitation of this study is that our study data set from TCGA lacks treatment information for the patients with GBM. This limits the extrapolation of our findings to all patients, especially those that do not undergo standard-of-care therapy. In addition, this prevented us from performing a subanalysis by treatment to evaluate whether biomarkers perform differently across different treatments. In addition, some known important glioblastoma predictive factors such as extent of resection and postsurgical complications are not included in TCGA data set. Lack of information on these risk factors could limit our findings in the prognostic ability of the biomarkers and gene signatures.

Conclusions

Although some biomarkers and gene signatures are statistically significantly associated with OS of patients with GBM, their ability to accurately prognose cumulative or incident OS events is very limited. Predicted risk scores that combine biomarkers with clinical risk factors, such as age, can greatly improve the prognostic performance of these biomarkers.

Acknowledgements

The authors thank Ms Mingming Zhang, RN, MS, BSN, for assistance with editing tables and figures. They thank all reviewers for their insightful comments and suggestions.

Author Contributions

NH conceived and designed the experiments. NH and HC analyzed the data and wrote the first draft of the manuscript. NH, HC, and KZ contributed to the writing of the manuscript. NH, HC, KZ, and RJ agree with manuscript results and conclusions and jointly developed the structure and arguments for the paper. NH and RJ made critical revisions. All authors reviewed and approved the final manuscript.

Disclosures and Ethics

All authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication and have permission from rights holders to reproduce any copyrighted material.

REFERENCES

- Thakkar JP, Dolecek TA, Horbinski C, et al. Epidemiologic and molecular prognostic review of glioblastoma. *Cancer Epidemiol Biomarkers Prev*. 2014;23:1985–1996.
- Farias-Eisner G, Bank AM, Hwang BY, et al. Glioblastoma biomarkers from bench to bedside: advances and challenges. *Br J Neurosurg*. 2012;26:189–194.
- Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*. 1993;39:561–577.
- Zhou X-H, McClish DK, Obuchowski NA. *Statistical Methods in Diagnostic Medicine*. Vol 569. Hoboken, NJ: John Wiley & Sons; 2002.
- Hu N, Zhou X. A review of time-dependent ROC curve for evaluating the prognosis capacity of biomarkers and semiparametric regression methods. Paper presented at: Proceeding of Joint Statistical Meeting; July 31-August 5, 2010; Vancouver, BC, Canada.
- Hu N, Zhou X-H. ROC analysis in business decision-makings. In: Wang J, ed. *Encyclopedia of Business Analytics and Optimization*. Hershey, PA: IGI Global; 2014:2104–2114.
- Etzioni R, Pepe M, Longton G, Hu C, Goodman G. Incorporating the time dimension in receiver operating characteristic curves: a case study of prostate cancer. *Med Decis Making*. 1999;19:242–251.
- Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56:337–344.
- Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61:92–105.
- Cai T, Pepe MS, Zheng Y, Lumley T, Jenny NS. The sensitivity and specificity of markers for event times. *Biostatistics*. 2006;7:182–197.
- Song X, Zhou X-H. A semiparametric approach for the covariate specific ROC curve with survival outcome. *Stat Sin*. 2008;18:947–965.
- Yan Y, Zhang L, Xu T, et al. SAMS1 is highly expressed and associated with a poor survival in glioblastoma multiforme. *PLoS ONE*. 2013;8:e81905.
- Bao ZS, Zhang CB, Wang HJ, et al. Whole-genome mRNA expression profiling identifies functional and prognostic signatures in patients with mesenchymal glioblastoma multiforme. *CNS Neurosci Ther*. 2013;19:714–720.
- Rao SA, Santosh V, Somasundaram K. Genome-wide expression profiling identifies deregulated miRNAs in malignant astrocytoma. *Mod Pathol*. 2010;23:1404–1417.
- Chen W, Yu Q, Chen B, Lu X, Li Q. The prognostic value of a seven-microRNA classifier as a novel biomarker for the prediction and detection of recurrence in glioma patients. *Oncotarget*. 2016;7:53392–53413.
- Zhang C, Han L, Zhang A, et al. Global changes of mRNA expression reveals an increased activity of the interferon-induced signal transducer and activator of transcription (STAT) pathway by repression of miR-221/222 in glioblastoma U251 cells. *Int J Oncol*. 2010;36:1503–1512.
- Malzkorn B, Wolter M, Liesenberg F, et al. Identification and functional characterization of microRNAs involved in the malignant progression of gliomas. *Brain Pathol*. 2010;20:539–550.
- Brennan CW, Verhaak RG, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell*. 2013;155:462–477.
- Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.
- Salzman M. Glioblastoma and malignant astrocytoma. In: Kaye A, Laws E Jr, eds. *Brain Tumors: An Encyclopedic Approach*. Edinburgh: Churchill Livingstone; 1995:449–477.
- Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol*. 2014;15:509–524.
- Niyazi M, Pitea A, Mittelbronn M, et al. A 4-miRNA signature predicts the therapeutic outcome of glioblastoma. *Oncotarget*. 2016;7:45764–45775.
- Karsy M, Arslan E, Moy F. Current progress on understanding microRNAs in glioblastoma multiforme. *Genes Cancer*. 2012;3:3–15.
- Gabrieli G, Wurdinger T, Kesari S, et al. MicroRNA 21 promotes glioma invasion by targeting matrix metalloproteinase regulators. *Mol Cell Biol*. 2008;28:5369–5380.
- Burger PC, Green SB. Patient age, histologic features, and length of survival in patients with glioblastoma multiforme. *Cancer*. 1987;59:1617–1625.

26. Walid MS. Prognostic factors for long-term survival after glioblastoma. *Perm J*. 2008;12:45–48.
27. Lamborn KR, Chang SM, Prados MD. Prognostic factors for survival of patients with glioblastoma: recursive partitioning analysis. *Neuro Oncol*. 2004;6:227–235.
28. Pepe MS, Feng Z, Huang Y, et al. Integrating the predictiveness of a marker with its performance as a classifier. *Am J Epidemiol*. 2008;167:362–368.
29. Moons KG, Harrell FE. Sensitivity and specificity should be de-emphasized in diagnostic accuracy studies. *Acad Radiol*. 2003;10:670–672.
30. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004;159:882–890.
31. Pepe MS, Zheng Y, Jin Y, et al. Evaluating the ROC performance of markers for future events. *Lifetime Data Anal*. 2008;4: 86–113.
32. Halligan S, Altman DG, Mallett S. Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach. *Eur Radiol*. 2015;25: 932–939.