UCSF

Title

Prediction of new brain metastases after radiosurgery: validation and analysis of performance of a multi-institutional nomogram

Permalink

Journal

ISSN

Authors

Ayala-Peacock, Diandra N
Attia, Albert
Braunstein, Steve E
et al.

Publication Date

DOI

Peer reviewed

Prediction of new brain metastases after radiosurgery: validation and analysis of performance of a multi-institutional nomogram

Diandra N. Ayala-Peacock[1,2], **Albert Attia**[2], **Steve E. Braunstein**[3], **Manmeet S. Ahluwalia**[4], **Jaroslaw Hepel**[5], **Caroline Chung**[6,16], **Joseph Contessa**[7], **Emory McTyre**[1], **Ann M. Peiffer**[1], **John T. Lucas Jr**[1], **Scott Isom**[8], **Nicholas M. Pajewski**[8], **Rupesh Kotecha**[4], **Mark J. Stavas**[2], **Brandi R. Page**[9], **Lawrence Kleinberg**[9], **Colette Shen**[9], **Robert B. Taylor**[10], **Nasarachi E. Onyeuku**[1], **Andrew T. Hyde**[10], **Daniel Gorovets**[5], **Samuel T. Chao**[4], **Christopher Corso**[7], **Jimmy Ruiz**[11], **Kounosuke Watabe**[12], **Stephen B. Tatter**[13], **Gelareh Zadeh**[14], **Veronica L. S. Chiang**[7,15], **John B. Fiveash**[10], and **Michael D. Chan**[1]

[1]Department of Radiation Oncology, Wake Forest School of Medicine, Winston-Salem, NC, USA

[2]Department of Radiation Oncology, Vanderbilt University School of Medicine, Nashville, TN, USA

[3]Department of Radiation Oncology, University of California San Francisco, San Francisco, CA, USA

[4]Brain Tumor and Neuro-Oncology Center, Neurological Institute, Cleveland Clinic Foundation, Cleveland, OH, USA

[5]Department of Radiation Oncology, Brown University Alpert Medical School, Providence, RI, USA

[6]Department of Radiation Oncology, Princess Margaret Cancer Center, Toronto, ON, Canada

[7]Department of Therapeutic Radiology/Radiation Oncology, Yale University School of Medicine, New Haven, CT, USA

[8]Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA

[9]Department of Radiation Oncology, Johns Hopkins University, Baltimore, MD, USA

[10]Department of Radiation Oncology, University of Alabama-Birmingham, Birmingham, AL, USA

[11]Department of Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Correspondence to: Diandra N. Ayala-Peacock.

[12]Department of Cancer Biology, Wake Forest School of Medicine, Winston-Salem, NC, USA

[13]Department of Neurosurgery, Wake Forest School of Medicine, Winston-Salem, NC, USA

[14]Division of Neurosurgery, University of Toronto, Toronto, ON, Canada

[15]Department of Neurosurgery, Yale University School of Medicine, New Haven, CT, USA

[16]Department of Radiation Oncology, MD Anderson Cancer Center, Houston, TX, USA

Abstract

Stereotactic radiosurgery (SRS) without whole brain radiotherapy (WBRT) for brain metastases can avoid WBRT toxicities, but with risk of subsequent distant brain failure (DBF). Sole use of number of metastases to triage patients may be an unrefined method. Data on 1354 patients treated with SRS monotherapy from 2000 to 2013 for new brain metastases was collected across eight academic centers. The cohort was divided into training and validation datasets and a prognostic model was developed for time to DBF. We then evaluated the discrimination and calibration of the model within the validation dataset, and confirmed its performance with an independent contemporary cohort. Number of metastases (8, HR 3.53 p = 0.0001), minimum margin dose (HR 1.07 p = 0.0033), and melanoma histology (HR 1.45, p = 0.0187) were associated with DBF. A prognostic index derived from the training dataset exhibited ability to discriminate patients' DBF risk within the validation dataset (c-index = 0.631) and Heller's explained relative risk (HERR) = 0.173 (SE = 0.048). Absolute number of metastases was evaluated for its ability to predict DBF in the derivation and validation datasets, and was inferior to the nomogram. A nomogram high-risk threshold yielding a 2.1-fold increased need for early WBRT was identified. Nomogram values also correlated to number of brain metastases at time of failure (r = 0.38, p < 0.0001). We present a multi-institutionally validated prognostic model and nomogram to predict risk of DBF and guide risk-stratification of patients who are appropriate candidates for radiosurgery versus upfront WBRT.

Keywords

Brain metastases; Distant brain failure; Stereotactic radiosurgery; Multi-Institutional nomogram

Introduction

Salvage therapies for the management of distant brain failure (DBF) following initial stereotactic radiosurgery (SRS) monotherapy, can have a substantial impact on medical cost-effectiveness [1]. SRS is approximately four-fold more expensive than whole brain radiotherapy (WBRT) but has been proven to be more cost-effective in selected patient populations [2, 3]. While the use of upfront WBRT may mitigate the likelihood of DBF [4], it is associated with toxicities such as fatigue [5] and cognitive decline [6]. Consensus guidelines and insurance reimbursement practices for the treatment of brain metastases have traditionally recommended the use of WBRT in patients with greater than four brain metastases, as these patients have a higher rate of DBF [7]. However, using a strict numerical cutoff to dictate which patients can receive SRS may both over-treat and under-

treat certain populations [8], and there is a growing literature that the number of metastases alone is not indicative of outcomes [9].

In the decision analysis for upfront WBRT or upfront SRS, patients who rapidly develop multiple new brain metastases and require early WBRT may not be the best candidates for SRS. In this scenario, there are higher costs without a clinically significant postponement of WBRT-related toxicities. For this reason, we developed a model designed to predict both the risk and severity of DBF within the first several months after upfront SRS without WBRT. Prior attempts have been made to identify patients requiring early WBRT, but these models were generally limited by allowing physician discretion to influence the timing of WBRT [10] and the oversimplification of DBF as a single event rather than a heterogeneous phenomenon that can range from a single new metastasis to diffuse seeding of the brain [11, 12].

In the current study, we present a multi-institutional predictive nomogram for DBF that is both internally and externally validated. This was then validated with a second independent and contemporary dataset. We also compare the predictive ability of the nomogram for DBF against the current guidelines that use a pure numerical cutoff. Finally, we perform an analysis in which we determine nomogram values that predict for the number of metastases at time of DBF, and also define a nomogram value where a patient is at high-risk of being treated with early WBRT.

Materials and methods

Data acquisition

This study was approved by the Institutional Review Board at all participating institutions. Data was reviewed and collected at all eight institutions and initially encompassed patients treated from January 2000 through December 2013. Patients were included in the study if they underwent SRS for first diagnosis of brain metastases in the defined time period, and had one to ten brain metastases present at time of SRS. Electronic medical records were reviewed to determine clinicopathologic characteristics that were previously determined to affect DBF in our prior analysis [11]. Burden of extracranial disease [13] was defined based on prior published works. Radiosurgical doses were generally determined based upon the guidelines published by Shaw et al. for single fraction radiosurgical treatment of brain metastases [14].

At time of SRS planning and delivery, the majority (seven of eight) institutions utilized thin-sliced MRI with Neuroradiology interpretation with half of institutions relying on double/triple contrast to help identify metastases at time of treatment. The majority of patients underwent clinical and MRI follow-up 1–3 months after their SRS procedures and were subsequently followed every 3 months thereafter. DBFs and the number of brain metastases at DBF were determined based on serial imaging evidence of intracranial recurrence. DBFs were identified as any new metastases that developed outside of the previous radiosurgical target volume.

Prognostic model derivation

Across the eight SRS centers, we collected data on 1484 patients treated between 2000 and 2013. Our primary outcome was time to DBF, which is measured from time of initial SRS. We excluded patients missing any covariate used in our previous analysis [11], reducing the sample size to a total of 1354 patients. We then divided the sample into a training (N = 685, four centers) and validation (N = 665, four centers) dataset. We estimated a prognostic index (PI) for time to DBF using the linear predictor ($\mathbf{X\beta}$) from a multivariable Cox proportional hazards (PH) regression model (higher values for the PI indicate a worse rate of DBF). The following variables were included as predictors in the Cox model: age, gender, histology, minimum margin dose delivered to any metastasis during the index SRS session, burden of extracranial disease (none, oligometastatic, and widespread), status of extracranial disease (progressive or stable), and number of lesions. We evaluated the PH assumption for all predictors using hypothesis tests based on scaled Schoenfeld residuals [15].

To evaluate the predictive ability of the PI in the validation dataset, we evaluated the discrimination and calibration of the nomogram as described by Royston and Altman [16] for external validation within the context of censored time-to-event outcomes. We estimated several measures to evaluate model discrimination, including Harrell's c-index [17], Gönen and Heller's K statistic [18], r_D^2 [19], and the explained relative risk measure of Heller [20]. For the c-index and K statistic, a value of 0.5 corresponds to random chance, while r_D^2 and the explained relative risk measure are analogous to conventional r^2 statistics from linear regression, ranging from 0 to 1.

Results are presented in terms of the cumulative distribution function $F(t) = 1 - S(t)$. To evaluate calibration, we mean-centered the PI in both datasets using the PI mean (2.045) estimated in the training dataset. We then divided each dataset into four prognostic groups based on the 16th, 50th, and 84th percentiles of the PI (mean-centered) from the training dataset. We estimated hazard ratios for each prognostic group in both datasets using Cox PH regression. Within the validation dataset, we then graphically compared the observed versus expected rate of DBF using three estimates of the survival function within each prognostic group: a Kaplan–Meier estimate versus two model-based estimates of covariate-adjusted survival. The model-based estimates only differ on whether the baseline cumulative hazard function (the non-parametric component of the Cox model) was estimated from the training dataset versus the validation dataset. Both model-based estimates compute the survival function conditional on the PI at all event times in the validation dataset ($S(t|PI) = S_0(t)e^{PI}$), and then averaging the survival function estimates at each time point within each prognostic group. While Royston and Altman used fractional polynomials to approximate the log of cumulative baseline hazard function [16], we found that fractional polynomials tended to over-fit our data, leading to estimates of the survival function that were not strictly non-increasing (results not shown). Therefore, we used a simple linear approximation to the cumulative baseline hazard function in our calculations (results not shown). A second validation was then pursued with a new contemporary cohort of 193 cases treated from 2011 to 2014 at a ninth academic center. All analyses were performed using the R Statistical Computing Environment (R Core Team [21]).

Results

Prognostic model for distant brain failure

Patient characteristics for the training and both validation datasets are displayed in Table 1. There were several statistically significant differences in characteristics of the populations in the training and validation datasets. Kaplan–Meier estimates of DBF at 3 and 6 months was 19.4 and 36.2% in the training dataset set, and 15.5 and 29.5% in the first validation dataset. For the new cohort (validation 2), KM estimates of DBF at 3 and 6 months were 21.6 and 39.8%, respectively. Cox regression estimates were used in constructing the PI in the training dataset (Supplemental Data). An increasing number of metastases, melanoma histology (HR 1.452, p = 0.0187), and increasing minimal margin dose (HR per one unit increase: 1.066, p = 0.0033) were the major factors associated with an increased rate of DBF. There was not any evidence of substantial deviation from the proportional hazards assumption of the Cox model (global test p value = 0.2968). The PI was used to generate the nomogram in Fig. 1.

Internal validation of the PI

We used bootstrap resampling (1000 replicates) to evaluate the calibration of the model for the PI within the training dataset. We used the bootstrap datasets to generate confidence intervals for the probability of DBF at 3, 6, and 9 months for every possible combination of the covariates used in constructing the PI from the training dataset. Figure 2a–c shows calibration curves delineating the predicted and observed probabilities of DBF. These curves generally indicate concordance between the observed and model-based predictions.

External validation of the PI

Table 2 presents estimates for measures of discrimination for the PI in both the training and validation datasets. In the validation dataset, we estimated Harrell's c-index to be 0.631 (SE = 0.018), with Heller's explained relative risk estimated to be 0.173 (SE = 0.048). These estimates imply a statistically significant ability of the PI to stratify patients according to their risk of DBF. This is further supported by the Harrell's c-index of the contemporary validation set of 0.644 (SE = 0.030). When we divided patients in the validation dataset according to the percentiles of the PI from the training dataset, we similarly see an ability of the PI to stratify patients according to their risk of DBF.

The Cox model fit to the validation dataset solely using the PI as a predictor yielded a coefficient for the PI of 0.985 (SE = 0.140, p = 0.9138) in validation 1 and 0.994 (SE = 0.195, p = 0.976) in validation 2. We then fit an additional Cox model to the validation datasets using all of the predictors included in deriving the PI, conditioned on the PI with its coefficient constrained to be one. A global test of all of the predictors (i.e. all $\beta = 0$) was not statistically significant in validation 1 (p = 0.602), but did indicate a small degree of lack of fit in validation 2 (p = 0.19). Figure 2d represents a graphical depiction of discrimination and calibration for the PI in the original and contemporary validation datasets. The Kaplan–Meier estimates illustrate the ability of the PI to appropriately risk-stratify patients, as the groupings defined by percentiles of the PI are correctly ordered in terms of their risk of DBF, with the exception of the two lowest risk groups being reversed in the second validation.

Comparative performance of nomogram versus current clinical practice

In order to compare the predictive ability of the nomogram to current clinical practices, we considered a simplified prognostic model solely based on the number of lesions (>4 vs. 4) to mimic current clinical practice based on available randomized studies. For that model, Heller's explained relative risk was 0.056 (SE = 0.033) in the derivation dataset, and 0.054 (SE = 0.034) in the validation 1 dataset and 0.080 (SE = 0.063) in validation 2, which are all inferior to the nomogram.

Ability of the nomogram to predict early WBRT and velocity of new brain metastases per year

To determine a threshold value that predicts early use of WBRT, the training dataset was used to search all nomogram values from the 10th to the 90th percentiles. Heller's maximum explained relative risk occurred when dividing the training dataset at the 76th percentile (182 points on the nomogram). The multi-institutional validation dataset similarly showed an increased incidence of WBRT above the 76th percentile. The hazard ratio for requiring WBRT within 3 months of SRS was 3.04 (95% CI 2.18, 4.23) in the training dataset at this threshold, and 2.09 (95% CI 1.23, 3.56) in the validation dataset. Figure 3a, b shows use of WBRT within 3 months for patients above and below the 76th percentile for both the training and validation datasets.

We then analyzed whether the nomogram correlated to the number of new metastases at time of DBF, measured in number of new brain metastases per year (brain metastasis velocity). Figure 3c graphically depicts the correlation between total nomogram points and the brain metastasis velocity (r = 0.38, p < 0.0001).

Discussion

The present study introduces an externally validated multi-institutional nomogram for the prediction of when patients will develop new brain metastases following SRS monotherapy. Such a tool may be helpful in triaging patients to either SRS or WBRT, a decision with significant implications on quality of life and health economics. A recent meta-analysis by Sahgal et al. found that patients younger than 50 years with a single metastasis had a significantly lower risk of DBF than patients with 2–4 metastases [22]. However, the brain metastasis population is heterogeneous, and there is emerging evidence that using the number of metastases as the sole factor to direct clinical decision-making may be an oversimplification of the problem [11, 23, 24]. In our own comparative analysis of the nomogram versus current guidelines, the nomogram was a better predictor of DBF than the number of metastases in the validation datasets, suggesting that there is a population of patients who receive WBRT that may have instead benefited from SRS.

While DBF is only a component of what drives the clinical decision for WBRT, it is a dominant factor that leads patients to require early WBRT after SRS alone [11, 25]. Nomogram values were predictive of the number of metastases that occur at treatment failure, which is commonly the factor that practitioners use to decide whether to treat further brain metastases with WBRT or further SRS. Other factors including local failure risk, life

expectancy, leptomeningeal involvement, current cognitive function and practitioner bias contribute to this decision. Despite these other factors, DBF is an endpoint that significantly affects the cost of future care as it can lead to multiple salvage treatments given its prevalence and the cost of multiple SRS procedures [1, 2]. The use of salvage WBRT after SRS ranged from 5 to 20% amongst the SRS centers in the current study, suggesting that the standards for use of WBRT vary amongst institutions. However, as nomogram values were predictive of the number of metastases that occurred at treatment failure, patients who are at highest risk of early failure are also those at risk of multifocal failure.

Our analysis identified a high-risk cohort, which demonstrated a 2.1-fold increase in the likelihood of receiving early WBRT in the validation dataset. The identification of patients who will experience rapid DBF is probably the single most clinically useful result of the nomogram. Moreover, the nomogram's calibration is best within the first 3 months after SRS. Those patients who require WBRT within 3 months may not have been appropriately triaged to upfront SRS as rapid DBF had generally occurred. A postponement of WBRT for greater than 3 months may represent a clinically meaningful time period in the survival of a patient with brain metastases since the cognitive toxicities of WBRT are often detectable by 4 months. Given a median life expectancy of 8–11 months for brain metastasis patients, cognition could potentially have been spared for a significant portion of a patient's life expectancy [6]. The divergence of the natural histories of many cancers over time based on responses and tolerances to systemic therapy likely explains why the nomogram is less well-calibrated at later time points when control of systemic disease affects the later DBF rate.

There are several limitations to the current study. The use of multiple retrospective datasets does introduce the possibility of patient selection bias. Variations in provider treatment discretion and follow-up across institutions has likely contributed to the discrepancy between rates of events in the derivation and validation datasets. Moreover, molecular subtyping of the recorded primary malignancies was not available for most histologies in this analysis, but is suspected to contribute to the rate of DBF [26]. A preliminary analysis was performed for the available data suggesting trends in higher rates of DBF for triple negative breast cancer as compared to non-triple negative breast cancers in the pooled training and validation datasets. These differences were not appreciated for BRAF-mutant melanoma and EGFR lung mutations, but we caution its interpretation given the limited numbers present for analysis (data not shown).

Similarly, the contribution of improving systemic therapies continues to evolve with the continued integration of new effective therapeutic agents for extracranial disease. The era and patient population from which the nomogram was derived was witness to the development of new systemic treatments for lung cancer, breast cancer, renal cell carcinoma and melanoma. The likelihood of DBF will continue to change with the advent of newer agents that have the ability to change the likelihood of developing brain metastases. Because of this "moving target" effect, clinical tools such as the current nomogram will need periodic updating to account for the ability to control extracranial disease.

In the current era of precision medicine, it is clear that individualizing care for patients provides an improvement in care. Moreover, the use of nomograms in oncology has

commonly helped to make difficult clinical decisions and properly assign resources [27]. Improvements to the current predictive nomogram in the future will likely come in the form of the integration of a systems biology component. Prospective validation of the nomogram is currently being planned at the multi-institutional level. Particular issues of interest for the prospective validation will be to better determine critical values on the nomogram that may trigger a decision for upfront WBRT and whether nomogram use mitigates the number of patients that are either over- or under-treated with SRS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Hall MD, McGee JL, McGee MC, et al. Cost-effectiveness of stereotactic radiosurgery with and without whole-brain radiotherapy for the treatment of newly diagnosed brain metastases. J Neurosurg. 2014; 121(Suppl 2):84–90. DOI: 10.3171/2014.7.GKS14972

2. Lester SC, Taksler GB, Kuremsky JG, et al. Clinical and economic outcomes of patients with brain metastases based on symptoms: an argument for routine brain screening of those treated with upfront radiosurgery. Cancer. 2014; 120(3):433–441. DOI: 10.1002/cncr.28422 [PubMed: 24452675]

3. Brown PD, Asher AL, Farace E. Adjuvant whole brain radiotherapy: strong emotions decide but rational studies are needed. Int J Radiat Oncol Biol Phys. 2008; 70(5):1305–1309. DOI: 10.1016/j.ijrobp.2007.11.047 [PubMed: 18234426]

4. Aoyama H, Shirato H, Tago M, et al. Stereotactic radiosurgery plus whole-brain radiation therapy vs stereotactic radiosurgery alone for treatment of brain metastases: a randomized controlled trial. JAMA. 2006; 295(21):2483–2491. DOI: 10.1001/jama.295.21.2483 [PubMed: 16757720]

5. Kocher M, Soffietti R, Abacioglu U, et al. Adjuvant whole-brain radiotherapy versus observation after radiosurgery or surgical resection of one to three cerebral metastases: results of the EORTC 22952–26001 study. J Clin Oncol. 2011; 29(2):134–141. DOI: 10.1200/JCO.2010.30.1655 [PubMed: 21041710]

6. Chang EL, Wefel JS, Hess KR, et al. Neurocognition in patients with brain metastases treated with radiosurgery or radiosurgery plus whole-brain irradiation: a randomised controlled trial. Lancet Oncol. 2009; 10(11):1037–1044. DOI: 10.1016/S1470-2045(09)70263-3 [PubMed: 19801201]

7. Sawrie SM, Guthrie BL, Spencer SA, et al. Predictors of distant brain recurrence for patients with newly diagnosed brain metastases treated with stereotactic radiosurgery alone. Int J Radiat Oncol Biol Phys. 2008; 70(1):181–186. DOI: 10.1016/j.ijrobp.2007.05.084 [PubMed: 17768015]

8. Sahgal A, Larson D, Knisely J. Stereotactic radiosurgery alone for brain metastases. Lancet Oncol. 2015; 16(3):249–250. DOI: 10.1016/S1470-2045(14)71106-4 [PubMed: 25752547]

9. Yamamoto M, Serizawa T, Shuto T, et al. Stereotactic radiosurgery for patients with multiple brain metastases (JLGK0901): a multi-institutional prospective observational study. Lancet Oncol. 2014; 15(4):387–395. DOI: 10.1016/S1470-2045(14)70061-0 [PubMed: 24621620]

10. Press RH, Prabhu RS, Nickleach DC, et al. Novel risk stratification score for predicting early distant brain failure and salvage whole brain radiotherapy after stereotactic radiosurgery for brain metastases. Cancer. 2015; 121(21):3836–3843. DOI: 10.1002/cncr.29590 [PubMed: 26242475]

11. Ayala-Peacock DN, Peiffer AM, Lucas JT, et al. A nomogram for predicting distant brain failure in patients treated with gamma knife stereotactic radiosurgery without whole brain radiotherapy. Neuro Oncol. 2014; 16(9):1283–1288. DOI: 10.1093/neuonc/nou018 [PubMed: 24558022]

12. Rodrigues G, Warner A, Zindler J, Slotman B, Lagerwaard F. A clinical nomogram and recursive partitioning analysis to determine the risk of regional failure after radiosurgery alone for brain metastases. Radiother Oncol. 2014; 111(1):52–58. DOI: 10.1016/j.radonc.2013.11.015 [PubMed: 24444529]

13. Harris S, Chan MD, Lovato JF, et al. Gamma knife stereotactic radiosurgery as salvage therapy after failure of whole-brain radiotherapy in patients with small-cell lung cancer. Int J Radiat Oncol Biol Phys. 2012; 83(1):e53–e59. DOI: 10.1016/j.ijrobp.2011.11.059 [PubMed: 22342297]

14. Shaw E, Scott C, Souhami L, et al. Single dose radiosurgical treatment of recurrent previously irradiated primary brain tumors and brain metastases: final report of RTOG protocol 90–05. Int J Radiat Oncol Biol Phys. 2000; 47(2):291–298. [PubMed: 10802351]

15. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994; 81(3):515–526.

16. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013; 13:33.doi: 10.1186/1471-2288-13-33 [PubMed: 23496923]

17. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982; 247(18):2543–2546. [PubMed: 7069920]

18. Gönen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. Biometrika. 2005; 92(4):965–970.

19. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Stat Med. 2004; 23(5):723–748. DOI: 10.1002/sim.1621 [PubMed: 14981672]

20. Heller G. A measure of explained risk in the proportional hazards model. Biostatistics. 2012; 13(2):315–325. DOI: 10.1093/biostatistics/kxr047 [PubMed: 22190711]

21. R Core Team. R: a language and environment for statistical computing. [computer program]. R Core Team; Vienna: 2013.

22. Sahgal A, Aoyama H, Kocher M, et al. Phase 3 trials of stereotactic radiosurgery with or without whole-brain radiation therapy for 1 to 4 brain metastases: individual patient data meta-analysis. Int J Radiat Oncol Biol Phys. 2015; 91(4):710–717. DOI: 10.1016/j.ijrobp.2014.10.024 [PubMed: 25752382]

23. Vern-Gross TZ, Lawrence JA, Case LD, et al. Breast cancer subtype affects patterns of failure of brain metastases after treatment with stereotactic radiosurgery. J Neurooncol. 2012; 110(3):381–388. DOI: 10.1007/s11060-012-0976-3 [PubMed: 23001361]

24. Kuremsky JG, Urbanic JJ, Jeff Petty W, et al. Tumor histology predicts patterns of failure and survival in patients with brain metastases from lung cancer treated with gamma knife radiosurgery. Neurosurgery. 2013; 73(4):641–647. DOI: 10.1227/NEU.0000000000000072 [PubMed: 23842552]

25. Gorovets D, Rava P, Ebner DK, et al. Predictors for long-term survival free from whole brain radiation therapy in patients treated with radiosurgery for limited brain metastases. Front Oncol. 2015; 5:110.doi: 10.3389/fonc.2015.0011 [PubMed: 26029666]

26. Johnson AG, Ruiz J, Hughes R, et al. Impact of systemic targeted agents on the clinical outcomes of patients with brain metastases. Oncotarget. 2015; 6(22):18945–18955. DOI: 10.18632/oncotarget.4153 [PubMed: 26087184]

27. Sanghani M, Truong PT, Raad RA, et al. Validation of a web-based predictive nomogram for ipsilateral breast tumor recurrence after breast conserving therapy. J Clin Oncol. 2010; 28(5):718–722. DOI: 10.1200/JCO.2009.22.6662 [PubMed: 20048188]
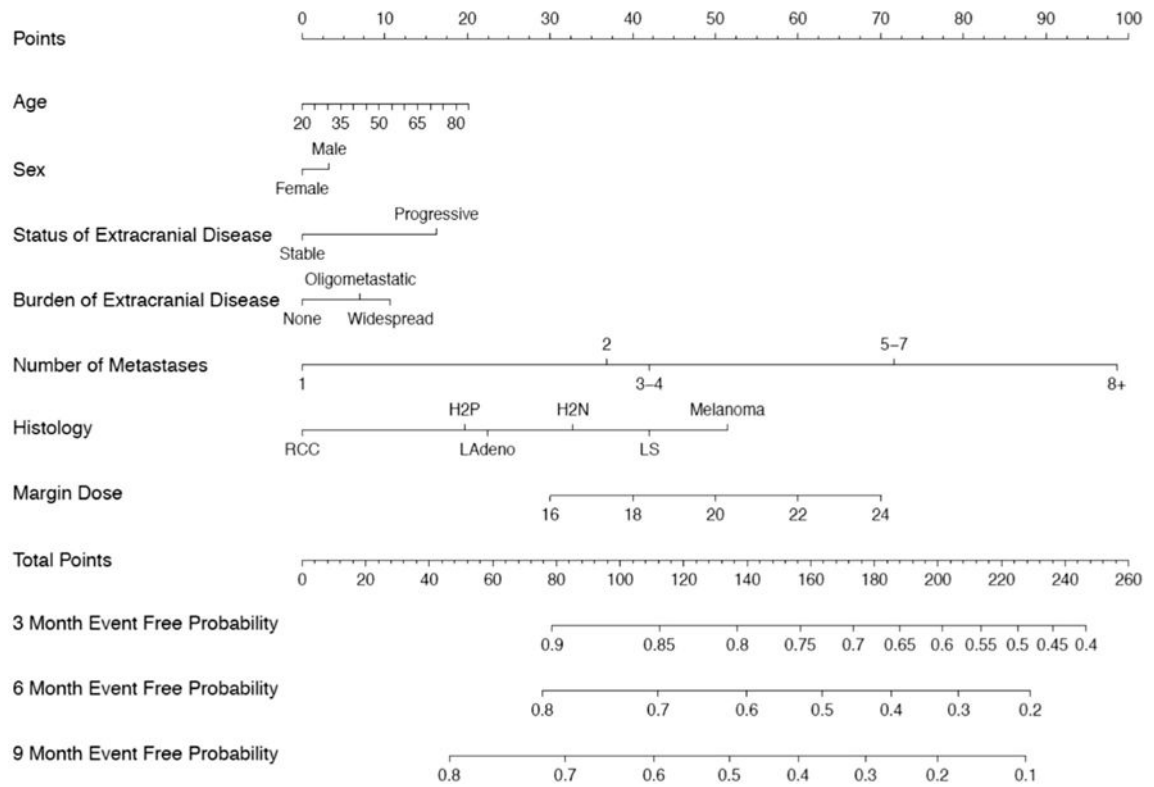
Fig. 1.

Nomogram for risk of distant brain failure at 3, 6, and 9 months after radiosurgery. *Oligomet* oligometastatic (<5 total metastases), *RCC* renal cell carcinoma, *L Squam* lung squamous cell cancer, *L Adeno* lung adenocarcinoma, *Her2+* her2 neu positive breast cancer, *Her2* (−) her2 neu negative breast cancer
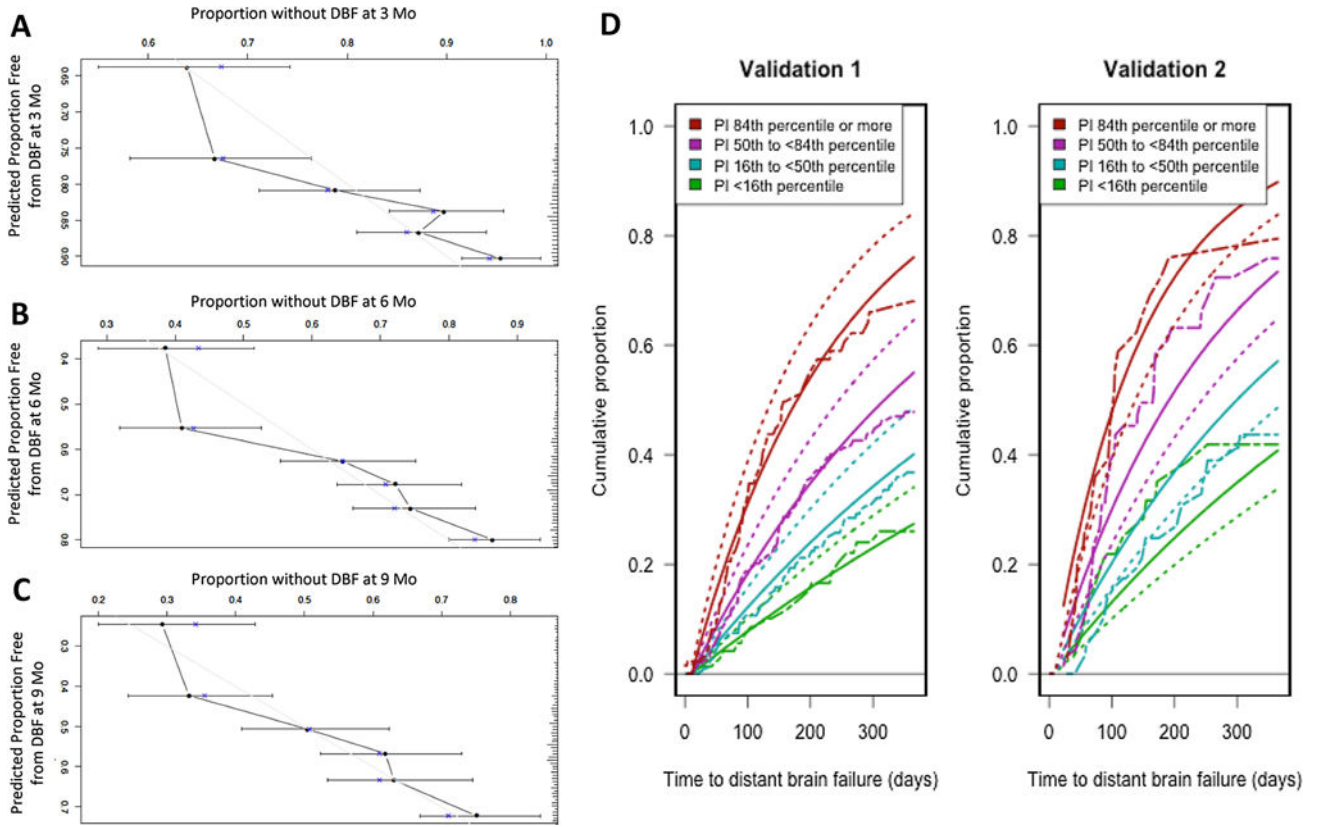
Fig. 2.
Calibration curves for **a** 3, **b** 6, and **c** 9 months. *DBF* distant brain failure, *Mo* months. **d**
Calibration of distant brain failure probabilities in the validation dataset. *X-axis* is in
months. *Y-axis* is distant brain failure probability. *Two dash non-smooth lines* are Kaplan–
Meier estimates of the cumulative distribution function for the PI risk groups in the
validation dataset. *Solid smooth lines* denote predicted, covariate-adjusted distant brain
failure based on approximation to baseline cumulative hazard function estimated within
validation data-set. *Dotted smooth lines* reflect predicted covariate-adjusted distant brain
failure based on approximation to baseline cumulative hazard function estimated within
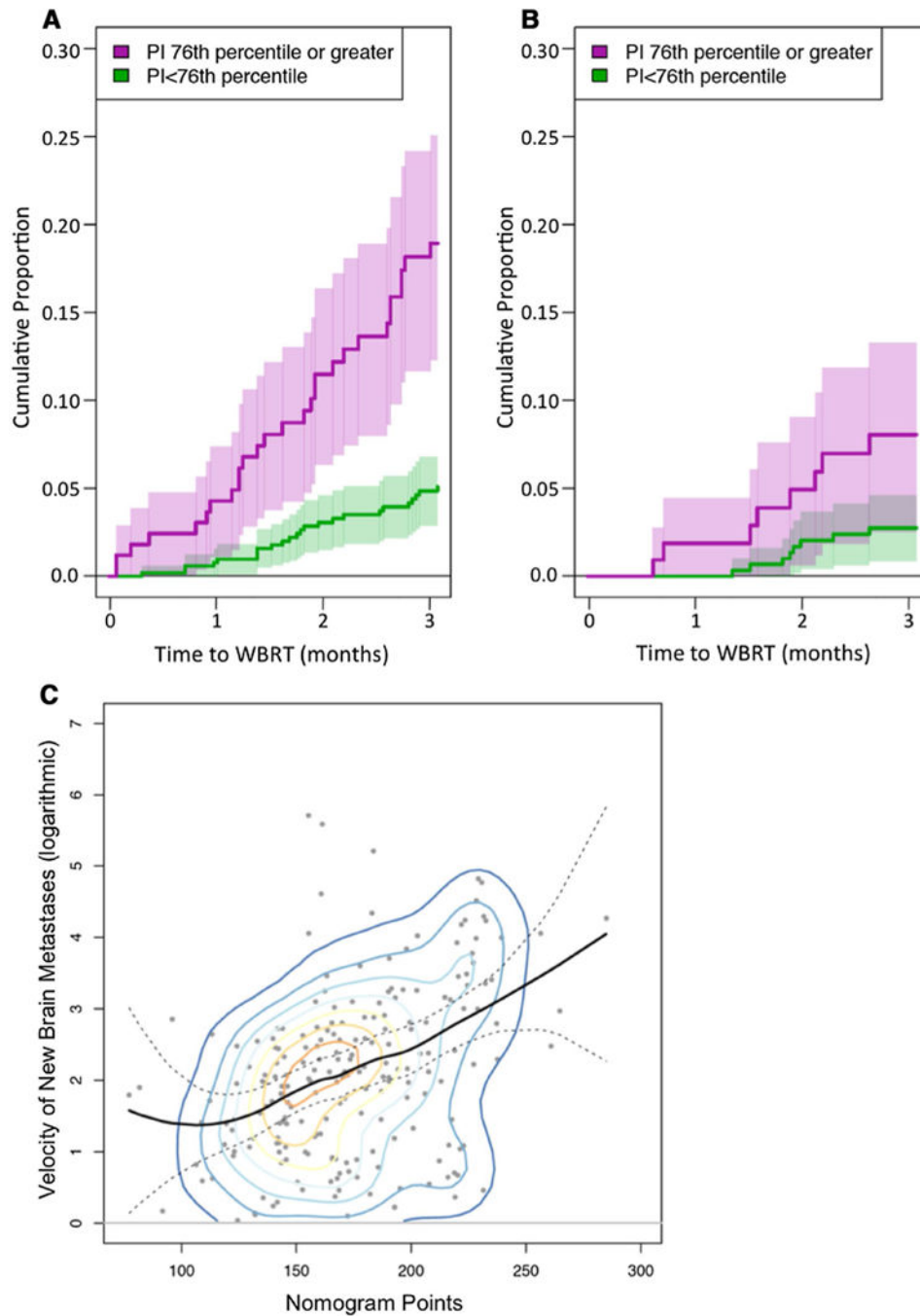training dataset

Fig. 3.
Probability of early whole brain radiotherapy for **a** training dataset and **b** initial validation dataset. The 76th percentile (182 points on the nomogram) represents the threshold at which there is maximum explained relative risk of WBRT within 3 months of radiosurgery. **c** Total nomogram points versus brain metastasis velocity (metastases per year on natural log logarithmic scale) for patients who experienced DBF in validation dataset. The *solid back line* represents a locally weighted polynomial regression fit (95% confidence interval between *dashed lines*). *Overlaid contour lines* (*color*) represent the relative densities of the

plotted points. A statistically significant correlation was shown between total nomogram points and brain metastasis velocity ($r = 0.38$, $p < 0.0001$)

Table 1

Characteristics of training and validation datasets

	Training N = 685	Validation 1 N = 669	Validation 2 N = 193	p Value
Age (years), mean ± SD	60.1 ± 12.3	61.4 ± 12.4	61.1 ± 12.5	0.15
Female sex, No. (%)	321 (46.9)	350 (52.3)	94 (48.7)	0.13
Race, No. (%)				0.001
White	596 (87.0)	586 (87.6)	162 (83.9)	
African-American	53 (7.7)	33 (4.9)	25 (13.0)	
Other	36 (5.3)	50 (7.5)	6 (3.1)	
Histology, No. (%)				0.019
Her2 negative breast cancer	66 (9.6)	46 (6.9)	21 (10.9)	
Her2 positive breast cancer	48 (7.0)	67 (10.0)	10 (5.2)	
Lung adenocarcinoma	262 (38.2)	215 (32.1)	77 (39.9)	
Lung squamous cell carcinoma	40 (5.8)	37 (5.5)	12 (6.2)	
Melanoma	180 (26.3)	185 (27.7)	47 (24.4)	
Renal cell carcinoma	89 (13.0)	119 (17.8)	26 (13.5)	
Minimal margin dose, median [IQR]	20.0 [18.0, 21.0]	19.0 [17.5, 22.3]	19.5 [17.0, 22.0]	0.38
Burden of extracranial disease, No. (%)				<0.001
None	131 (19.1)	64 (9.6)	40 (20.7)	
Oligometastatic	276 (40.3)	165 (24.7)	84 (43.5)	
Widespread	278 (40.6)	440 (65.8)	69 (35.8)	
Progressive status of extracranial disease, No. (%)	282 (41.2)	438 (65.5)	85 (44.0)	<0.001
Number of metastases, No. (%)				0.307
One	327 (47.7)	324 (48.4)	91 (47.2)	
Two	158 (23.1)	146 (21.8)	36 (18.7)	
Three or Four	138 (20.1)	136 (20.3)	37 (19.2)	
Five to Seven	44 (6.4)	37 (5.5)	20 (10.4)	
Eight or more	18 (2.6)	26 (3.9)	9 (4.7)	
Year of SRS				<0.001
2000–2005	159 (23.2)	186 (27.8)	0 (0.0)	
2006–2010	403 (58.8)	257 (38.4)	0 (0.0)	
2011–2014	123 (18.0)	226 (33.8)	193 (100.0)	

SD Standard Deviation, *IQR* interquartile range, *SRS* stereotactic radiosurgery without whole brain radiotherapy

Table 2

Discrimination measures and hazard ratios for prognostic groups based on percentiles of the prognostic index (PI) evaluated in the training and validation datasets

Measure	Training Estimate (SE)	Validation 1 Estimate (SE)	Validation 2 Estimate (SE)
Harrell c-index	0.659 (0.018)	0.631 (0.018)	0.644 (0.030)
Gönen and Keller K	0.628 (0.003)	0.629 (0.003)	0.635 (0.006)
Explained variation $\left(r_D^2\right)$	0.133 (0.029)	0.124 (0.033)	0.140 (0.055)
Heller's estimated explained relative risk	0.178 (0.043)	0.173 (0.048)	0.196 (0.080)
	Training HR (95% CI)	Validation 1 HR (95% CI)	Validation 2 HR (95% CI)
Prognostic index category (<16th percentile)	Referant	Referant	Referant
16th to <50th percentile	1.48 (0.97–2.25)	1.49 (0.86–2.58)	0.93 (0.47–1.83)
50th to <84th percentile	2.91 (1.96–4.32)	2.27 (1.33–3.85)	2.32 (1.24–4.35)
84th percentile or greater	4.44 (2.86–6.89)	4.17 (2.40–7.24)	3.07 (1.56–6.02)

Standard errors (SE) for the c-index, K, and r_D^2 were estimated from 200 bootstrap samples. Hazard ratios (HR) are relative to the reference group of PI <16th percentile *CI* confidence interval