PRECISION MEDICINE

original reports abstract

Predictive Model to Guide Brain Magnetic Resonance Imaging Surveillance in Patients With Metastatic Lung Cancer: Impact on Real-World Outcomes

Julie Wu, MD, PhD[1]; Victoria Ding, MS[2]; Sophia Luo, BA[2]; Eunji Choi, PhD[2]; Jessica Hellyer, MD[1]; Nathaniel Myall, MD[1]; Solomon Henry, MS[3]; Douglas Wood, PhD[3]; Henning Stehr, PhD[4]; Hanlee Ji, PhD[1]; Seema Nagpal, MD[5,6,7]; Melanie Hayden Gephart, MD[7]; Heather Wakelee, MD[1,5]; Joel Neal, MD[1,5]; and Summer S. Han, PhD[2,5,7]

PURPOSE Brain metastasis is common in lung cancer, and treatment of brain metastasis can lead to significant morbidity. Although early detection of brain metastasis may improve outcomes, there are no prediction models to identify high-risk patients for brain magnetic resonance imaging (MRI) surveillance. Our goal is to develop a machine learning–based clinicogenomic prediction model to estimate patient-level brain metastasis risk.

METHODS A penalized regression competing risk model was developed using 330 patients diagnosed with lung cancer between January 2014 and June 2019 and followed through June 2021 at Stanford HealthCare. The main outcome was time from the diagnosis of distant metastatic disease to the development of brain metastasis, death, or censoring.

RESULTS Among the 330 patients, 84 (25%) developed brain metastasis over 627 person-years, with a 1-year cumulative brain metastasis incidence of 10.2% (95% CI, 6.8 to 13.6). Features selected for model inclusion were histology, cancer stage, age at diagnosis, primary site, and *RB1* and *ALK* alterations. The prediction model yielded high discrimination (area under the curve 0.75). When the cohort was stratified by risk using a 1-year risk threshold of $>$ 14.2% (85th percentile), the high-risk group had increased 1-year cumulative incidence of brain metastasis versus the low-risk group (30.8% *v* 6.1%, $P < .01$). Of 48 high-risk patients, 24 developed brain metastasis, and of these, 12 patients had brain metastasis detected more than 7 months after last brain MRI. Patients who missed this 7-month window had larger brain metastases (58% *v* 33% largest diameter $>$ 10 mm; odds ratio, 2.80, CI, 0.51 to 13) versus those who had MRIs more frequently.

CONCLUSION The proposed model can identify high-risk patients, who may benefit from more intensive brain MRI surveillance to reduce morbidity of subsequent treatment through early detection.

JCO Precis Oncol 6:e2200220.

INTRODUCTION

Brain metastasis is common among patients with lung cancer, occurring in up to 25%-50% of patients with metastatic lung cancer.[1,2] Although there are multiple management approaches to brain metastasis, they vary in associated morbidity[3,4] and can be influenced by the size and number of metastases.[3-5] For example, surgical removal may be necessary for large brain metastasis. However, given the advent of CNS-penetrant targeted therapy, such as osimertinib,[6] surgery and even radiation may be deferred for patients with limited burden of disease and the appropriate sensitive driver mutations. Thus, earlier detection of brain metastasis may potentially prevent patient morbidity by reducing size of brain metastasis at time of detection, thereby sparing the patient from more invasive methods of brain metastasis management.

Identifying patients at high risk for brain metastasis could enable intervention for early brain metastasis detection. Several previous studies have examined factors such as tumor histology that are linked to a higher incidence of brain metastasis.[7] Tumor genomic factors have also been shown to play a role, with higher incidence of brain metastasis associated with lung cancers with an epidermal growth factor (*EGFR*) or anaplastic lymphoma kinase (*ALK*) driver mutation.[8,9] Although highly informative, these individual factors cannot be directly used to predict the occurrence of brain metastasis in individual patients with lung cancer who might have multiple competing risk factors. Furthermore, most of the previous studies that

ASSOCIATED CONTENT

Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on August 2, 2022 and published at ascopubs.org/journal/po on October 6, 2022: DOI https://doi.org/10.1200/PO.22.00220

CONTEXT

Key Objective

Can we identify patients with lung cancer at high risk of developing brain metastasis to aid brain magnetic resonance imaging surveillance by using patient-level clinical and tumor sequencing data?

Knowledge Generated

Using a study cohort of 330 patients with lung cancer and no brain metastases on initial presentation with metastatic disease, we developed a machine learning–based model to predict a 1-year risk of brain metastases after diagnosis of metastatic disease, which yielded high discrimination and calibration.

Relevance

The proposed model using comprehensive clinical and tumor genomic data can accurately identify high-risk patients for brain metastases, who may benefit from more intensive brain magnetic resonance imaging surveillance.

reported *EGFR* or *ALK* as potential risk factors for brain metastasis did not consider the potential impact of the advent of CNS-penetrant targeted therapies that may reduce risk of brain metastasis in patients with an *EGFR* or *ALK* driver mutation.[10-12] Thus, there is a need to integrate treatment, driver mutation status, and other potential key factors for brain metastasis into a comprehensive model to predict individual risk of brain metastasis for patients without brain metastasis at diagnosis of metastatic lung cancer.

In this study, we integrated clinical, demographic, and genomic factors from a single-institution retrospective data set in a predictive model to identify patients who are at high risk for developing brain metastasis after initial diagnosis with metastatic lung cancer. We considered a broad range of potential predictors of brain metastasis, including demographics (eg, age and smoking history), tumor characteristics (eg, stage and tumor location), treatment history, and tumor sequencing from a broad-based next-generation sequencing panel. We also used the predictive model to identify high-risk patients from our data set and evaluated the relationship between brain magnetic resonance imaging (MRI) screening frequency and outcomes for patients with brain metastasis.

METHODS

Study Population and Selection

Patients diagnosed with lung cancer (any stage) between 2014 and 2019 were identified from the electronic medical record of the Stanford Medical Center. All patients had targeted panel sequencing of their lung cancer using Stanford's Solid Tumor Actionable Mutation Panel (STAMP) as part of their routine clinical care between January 2014 and June 2019, timing of which was determined by the treating physician. Patients were excluded if they did not have distant metastatic disease (either de novo stage IV or recurrence) and biopsy-proven lung cancer. Patients were further excluded if they had synchronous brain metastasis, defined as diagnosis of brain metastasis within 90 days of diagnosis with distant metastatic disease (Fig 1). This study was approved by the Stanford University IRB and received a waiver of informed consent because the study presented minimal risk and could not practicably be conducted without a waiver.

Study Outcome

The primary outcome was the time from date of diagnosis of distant metastatic disease to time of brain metastasis, death, or censoring, whichever occurred first, followed through June 14, 2021. Date of diagnosis of distant metastatic disease was defined by the date of imaging demonstrating distant metastasis (for patients with biopsy-proven lung cancer from the primary tumor) or biopsy-proven metastatic disease, whichever occurred first. Similarly, we defined date of brain metastasis as the date of first brain imaging demonstrating evidence of brain metastasis, as determined from chart review. For imaging findings that were equivocal, chart review was used to determine whether the finding was a brain metastasis on the basis of the documentation from the treating physician. Since

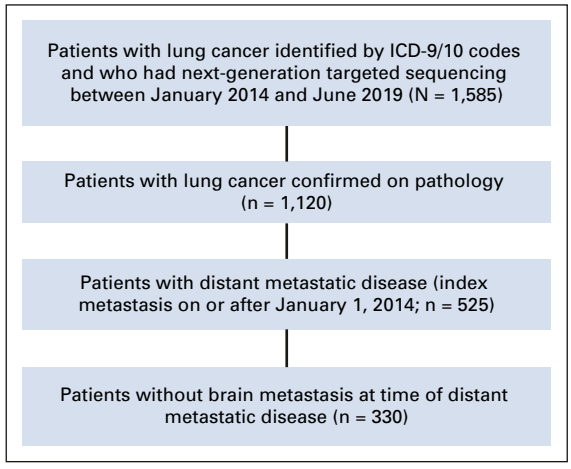

FIG 1. Patient eligibility flow diagram. ICD, International Classification of Diseases.

patients with brain metastasis at the time of their diagnosis of distant metastatic disease were excluded from the study population, all patients with brain metastasis had metachronous brain metastasis, ie, developed more than three months after date of distant metastasis.[13]

Study Variables

Demographic information including sex, race/ethnicity, age at distant metastasis, and smoking status (ever or never) was abstracted from the electronic medical record. Primary tumor characteristics including histology, tumor size, primary site, and initial stage at time of diagnosis were obtained from Stanford registry data. For each patient, all cancer-directed treatment given between initial diagnosis of lung cancer and time of distant metastatic disease (if applicable) was included as their treatment history, including first line of therapy for metastatic disease. First-line metastatic therapy was defined as systemic therapy administered within 90 days of time of distant metastasis. Treatment data were derived from the electronic medical record administration data for intravenous medications (chemotherapy, immunotherapy, and anti–vascular endothelial growth factor). Oral targeted therapy was derived by chart review.

Tumor somatic mutation data were obtained from STAMP,[14] a custom panel that covers 130 genes.[15,16] This gene list includes canonical lung cancer driver oncogenes (eg, *KRAS*, *EGFR*, *BRAF*, *ALK*, *ROS1*, and *RET*) and other frequent alterations (eg, *TP53*, *STK11*, and *RB1*). Genes were included as candidate predictors for feature selection if they were represented at a frequency of $> 3\%$. Individual gene variants (such as *EGFR* p.L858R) were individually represented if present at a frequency of $> 3\%$.

Statistical Analysis

Prediction model development and feature selection. For feature selection in prediction modeling, we applied machine learning approaches on the basis of a set of penalized regression methods (or regularization approaches) for competing risk data, including convex (least absolute shrinkage and selection operator [LASSO] and adaptive LASSO), nonconvex (smoothly clipped absolute deviation), and the minimax concave penalty functions.[17] We considered a total of 45 features for modeling, which included demographics, tumor characteristics, and targeted panel sequencing data. The final features were chosen on the basis of the consensus of the four different penalty functions (LASSO, adaptive LASSO, minimax concave penalty, and smoothly clipped absolute deviation) across 20 imputed data sets (see the Handling Missing Data section), that is, we chose the variables that were selected more than 70% of the time (ie, $\geq$ 14 of 20 imputed data sets) by at least two of four methods to be included in the final model (Data Supplement). *EGFR* and *KRAS* variants were binned. *ALK* driver status was included in the final model because of its established role as a risk factor for brain metastasis,[6,9] including in our cohort,[18] despite being represented in $<$ 3% of the cohort. On the basis of the selected features, we built the final model with complete-case analysis using a cause-specific proportional hazards model using death as competing risks.

Model performance and validation. Overall prediction performance was evaluated using the Brier score for competing risk data.[19] Model discrimination was assessed via area under the receiver operating characteristic curve (AUC) for time-to-event data.[20] Model calibration was visualized as observed versus predicted risk and observed 1-year brain metastasis incidence by deciles of estimated risk. For validation of the final model, we performed bootstrap cross-validation using 1,000 resamples that can provide optimism-corrected performance metrics for discrimination, calibration, and predictive accuracy.[21,22]

Risk stratification. We evaluated the risk stratification ability of the prediction model. The study population was stratified as high-risk versus low-risk using the 85th percentile of the estimated risk of 1-year brain metastasis risk (ie, 1-year risk $>$ 14.2%). For each subgroup, we estimated the observed cumulative incidence of brain metastasis using the Aalen-Johansen estimator,[23] and the difference in cumulative incidence across the subgroups was tested using two-sided Gray's test.[24]

Evaluation of clinical outcomes. The goal of our predictive model was to identify patients at high risk of brain metastasis so that they might benefit from an intervention, such as increased frequency of brain MRI surveillance. Therefore, we compared clinical outcomes for high-risk patients (ie, predicted 1-year risk $>$ 14.2%) who developed brain metastases by their brain MRI imaging timing relative to brain metastasis diagnosis, including a brain metastasis size and a rate of surgery for brain metastasis. Brain MRI dates, size of brain metastases, and treatment modality were abstracted from the chart. Brain MRI images were obtained per the local institutional protocol, which includes T1- and T2-weighted MRI with and without gadolinium contrast. We compared brain metastasis size and rate of surgical intervention for patients with increased brain MRI surveillance (last MRI $<$ 7 months before brain metastasis) or less frequent screening (last MRI $>$ 7 months before brain metastasis).

Handling missing data. We assessed the rate of missingness in the included variables. Overall, the missingness rate of demographic and clinical variables was $<$ 5%, except for histology (missing in 17 patients or 5.2% of the cohort) and the following treatment variables: cytotoxic, immunotherapy, anti–vascular endothelial growth factor therapy, and other treatment. We examined the missingness of included variables for correlations with other variables to confirm the missing at random assumption (Data Supplement). Missing data were then imputed with multiple imputation by chained equations[25] with 20 imputations with the

assumption that data were missing at random assumption. These imputed data sets were incorporated into the feature selection algorithm using different regularization methods as described in the previous subsection.

A web-based tool for risk prediction. We implemented the proposed model into a web-based tool, called RAMBO (Risk Assessment for Metastasis to Brain Outcome),[26] which can predict an individual-level probability of developing brain metastasis within one year from the time of distant metastasis diagnosis in patients with lung cancer. This app will be available to public by the time that this work is published.

RESULTS

Of 330 patients with lung cancer with distant metastasis in the study cohort, 84 developed brain metastases over 627 person-years. The median follow-up time was 1.3 (interquartile range 0.65-2.5) years in the overall cohort (Data Supplement). The 1-year cumulative incidence of brain metastasis was 10.2% (95% CI, 6.8 to 13.6; Data Supplement). Patient characteristics in the overall cohort and by study outcome are shown in Table 1. As expected on the basis of the demographics at our institution, a large proportion of Asian patients was observed in the overall cohort (36.4%). Almost half of the overall cohort never smoked (41.6%), and the majority of patients had lung adenocarcinoma (83.4%), with 34.8% and 2.7% of patients having an *EGFR* mutation and an *ALK* driver mutation, respectively.

The features included in the final risk prediction model for brain metastasis are shown in Figure 2. The most important clinical and demographic variables in predicting the development of brain metastasis were histology and stage at diagnosis followed by age at diagnosis, alteration in *RB1* or *ALK*, and primary lung cancer site. Large cell histology was associated with a higher risk of brain metastasis (Data Supplement). A primary lung cancer site near the main bronchus was associated with an increased risk of brain metastasis. One of the most important genomic features in predicting for brain metastasis development included the presence of mutations in *RB1*, which was a key predictor even when accounting for histology (Fig 2 and Data Supplement).

The performance of the proposed model was evaluated using internal validation on the basis of 1,000 resamples through bootstrapping (Fig 3A), which showed good calibration and high discrimination (bootstrapped AUC of 0.75; 95% CI, 0.64 to 0.84). The model was able to accurately predict a 1-year risk of brain metastasis (Brier score 0.08; 95% CI, 0.05 to 0.11). When the study cohort was stratified into high-risk and low-risk groups using a 1-year risk threshold of $>$ 14.2% (85th risk percentile), the high-risk group had a significantly elevated observed incidence of developing brain metastasis versus the low-risk group (30.8% *v* 6.1% for 1-year incidence, $P < .01$; Fig 3B and Data Supplement). The comparison of the clinical and genomic characteristics by high-risk versus low-risk groups is given in Figure 4. Compared with low-risk patients, patients at high risk of brain metastasis were younger and diagnosed at a more advanced stage and their cancer was more likely to have a central primary tumor location and have nonadenocarcinoma histology.

The goal of developing our predictive model was to identify patients at high risk of brain metastasis who might benefit from a tailored intervention, such as increased frequency of brain MRI surveillance. Therefore, we further examined various clinical outcomes of the high-risk group identified by the proposed model stratified by MRI frequency. Of the 48 high-risk patients, 24 patients (50%) developed brain metastasis. Of the 24 patients with brain metastasis in this high-risk group, 12 patients (50%) had their brain metastasis detected more than 7 months after last brain MRI (or date of metastasis, whichever was later), whereas the rest received brain MRIs within the 7-month window. Notably, the patients who missed this 7-month brain MRI surveillance opportunity window showed larger brain metastasis compared with those who had brain MRIs more frequently (58% *v* 33% with the brain metastasis size larger than 10 mm; odds ratio, 2.80; CI, 0.51 to 13; Fig 5). Similarly, those patients who missed this window were more likely to undergo surgery compared with the patients who received brain MRIs within the 7-month window (17% *v* 8%, odds ratio, 2.2; CI, 0.22 to 34).

DISCUSSION

In this study, we developed a machine learning–based model for predicting the risk of brain metastasis among patients with lung cancer with distant metastasis using comprehensive clinical, demographic, and genomic data. We showed that the proposed model can identify patients at high risk of brain metastasis with high discrimination and accuracy. Among the patients with high-risk brain metastases identified through the proposed model, those with more frequent brain MRIs tended to have smaller brain metastases with a reduced rate of surgical intervention compared with those with less frequent brain MRIs. Thus, we demonstrate the potential clinical utility of the model in identifying high-risk patients who may benefit from more intensive brain MRI surveillance and hence reduce morbidity of subsequent brain metastasis treatment.

In our model, we observed that several factors were significant contributors to the development of brain metastasis. In particular, the histology and stage of primary lung cancer were significant drivers of brain metastasis risk. Tumor primary location was important as well, suggesting that a more central location may lead to more frequent dissemination to the brain. Individual genomic factors found to be significant included previously known drivers of brain metastasis, such as mutations in *RB1*. Concurrent *RB1-TP53* is associated with small-cell differentiation and

TABLE 1. Patient Characteristics Stratified by Outcome Status and Overall

Characteristic	Overall (N = 330)	Censored (n = 124)	Brain Metastasis (n = 84)	Death (n = 122)
Demographics				
Age at distant metastasis, years, mean (SD)	67.2 (12.1)	68.2 (12.1)	63.9 (11.0)	68.3 (12.5)
Male, No. (%)	155 (47.0)	60 (48.4)	40 (47.6)	55 (45.1)
Race/ethnicity, No. (%)				
Non-Hispanic/Non-Latino White	150 (45.5)	60 (48.4)	33 (39.3)	57 (46.7)
Asian	120 (36.4)	38 (30.6)	37 (44.0)	45 (36.9)
Others	50 (15.2)	21 (16.9)	10 (11.9)	19 (15.6)
Unknown	10 (3.0)	5 (4.0)	4 (4.8)	1 (0.8)
Ever smoking, No. (%)	192 (58.4)	73 (59.3)	44 (52.4)	75 (61.5)
Primary tumor characteristics				
Primary tumor histology, No. (%)				
Adenocarcinoma	261 (83.4)	101 (87.8)	65 (80.2)	95 (81.2)
Large cell	4 (1.3)	0 (0)	4 (4.9)	0 (0)
Non–small-cell carcinoma, NOS	8 (2.6)	2 (1.7)	5 (6.2)	1 (0.9)
Squamous and transitional cell	26 (8.3)	8 (7.0)	4 (4.9)	14 (12.0)
Other specified carcinomas	14 (4.5)	4 (3.5)	3 (3.7)	7 (6.0)
Primary tumor size, mean (SD)	42.7 (24.9)	40.3 (26.2)	45.9 (23.8)	42.6 (24.6)
Primary tumor stage, No. (%)				
I	33 (10.0)	20 (16.1)	4 (4.8)	9 (7.4)
II	26 (7.9)	11 (8.9)	4 (4.8)	11 (9.0)
III	49 (14.8)	23 (18.5)	15 (17.9)	11 (9.0)
IV	210 (63.6)	68 (54.8)	55 (65.5)	87 (71.3)
Others	12 (3.6)	2 (1.6)	6 (7.1)	4 (3.3)
Primary tumor site, No. (%)				
Lower lobe	96 (29.1)	45 (36.3)	22 (26.2)	29 (23.8)
Main bronchus	13 (3.9)	4 (3.2)	5 (6.0)	4 (3.3)
Middle lobe	19 (5.8)	8 (6.5)	9 (10.7)	2 (1.6)
Overlapping lesion	42 (12.7)	13 (10.5)	8 (9.5)	21 (17.2)
Upper lobe	160 (48.5)	54 (43.5)	40 (47.6)	66 (54.1)
Days from distant metastasis to outcome, mean (SD)	564 (447)	632 (462)	542 (381)	511 (467)
Mutations, No. (%)				
EGFR	115 (34.8)	42 (33.9)	37 (44.0)	36 (29.5)
ALK rearrangement	9 (2.7)	1 (0.8)	6 (7.1)	2 (1.6)
First-line systemic treatment, No. (%)				
Cytotoxic	131 (55.5)	41 (50.6)	37 (54.4)	53 (60.9)
First-generation EGFR TKI	48 (14.5)	13 (10.5)	17 (20.2)	18 (14.8)
Next-generation ALK TKI	6 (1.8)	3 (2.4)	2 (2.4)	1 (0.8)
Immunotherapy	61 (25.8)	27 (33.3)	13 (19.1)	21 (24.1)
Osimertinib	24 (7.3)	10 (8.1)	5 (6.0)	9 (7.4)
Anti-VEGF	14 (5.9)	3 (3.7)	4 (5.9)	7 (8.0)
Crizotinib	9 (2.7)	2 (1.6)	5 (6.0)	2 (1.6)
Other treatment	20 (8.5)	9 (11.1)	4 (5.9)	7 (8.0)

NOTE. Percentages for categorical characteristics are calculated of the number of nonmissing values.

Abbreviations: EGFR, epidermal growth factor; NOS, not otherwise specified; SD, standard deviation; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor.

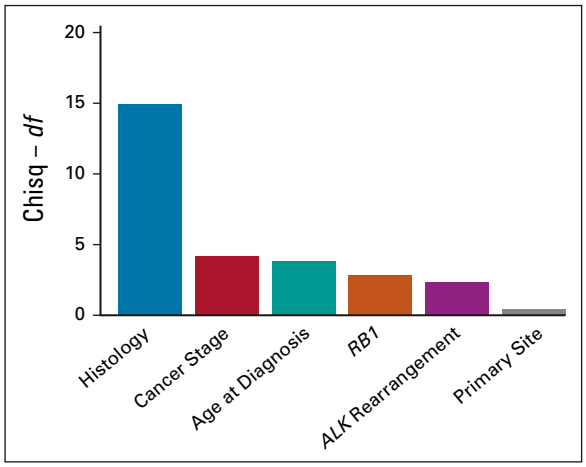

FIG 2. Variable importance of the features included in the proposed risk prediction model for brain metastasis. Importance of variables predicting the risk of brain metastasis. The *y*-axis shows the likelihood ratio test chi-squared statistic (Chisq) subtracted by the degrees of freedom (*df*) for each variable. Variables with the variable importance score (chisq-*df*) greater than zero are listed in descending order on the *x*-axis.

transformation, which carries a significant risk of brain metastasis.[27,28] Interestingly, we did not see a significant contribution from *EGFR*. This may be due to cohort selection, as many *EGFR*-positive patients present with brain metastases, and the advent of CNS-penetrant *EGFR*-directed therapy such as osimertinib that can potentially help reduce the incidence of brain metastasis.[29] Or at a minimum, such therapy may delay brain metastases development such that these patients were not identified as having brain metastases in this analysis. As we only included the first line of metastatic treatment in this model to better approximate the information available to a treating clinician at the time of utilization of this tool, the potential effect from osimertinib in later lines of therapy may be obscured.

Comprehensive prediction modeling, as used in the present study, holds great promise to aid clinical decision making in the management of patients with cancer. As the data from the present study are from a real-world data set, the patients included in this study are representative of those seen in clinic and the predictive model on the basis of these data would therefore be expected to have high external validity. Furthermore, we have built a free access online tool to aid clinicians in identifying patients who may benefit from interventions such as increased brain MRI surveillance. As clinical information is increasingly digitized, the relevant variables in our predictive model could be readily extracted from electronic health records, with the predictive risk score for a patient and surveillance recommendation displayed for the treating clinician. The framework that we have built is readily generalizable to other institutions and data sets.

Strengths of the present study include the comprehensive data using genomic, clinical, and demographic factors and thorough statistical modeling approaches that incorporate competing risks of death, penalized regression methods for feature selection, and validation through bootstrapping. To our knowledge, this is the first comprehensive prediction

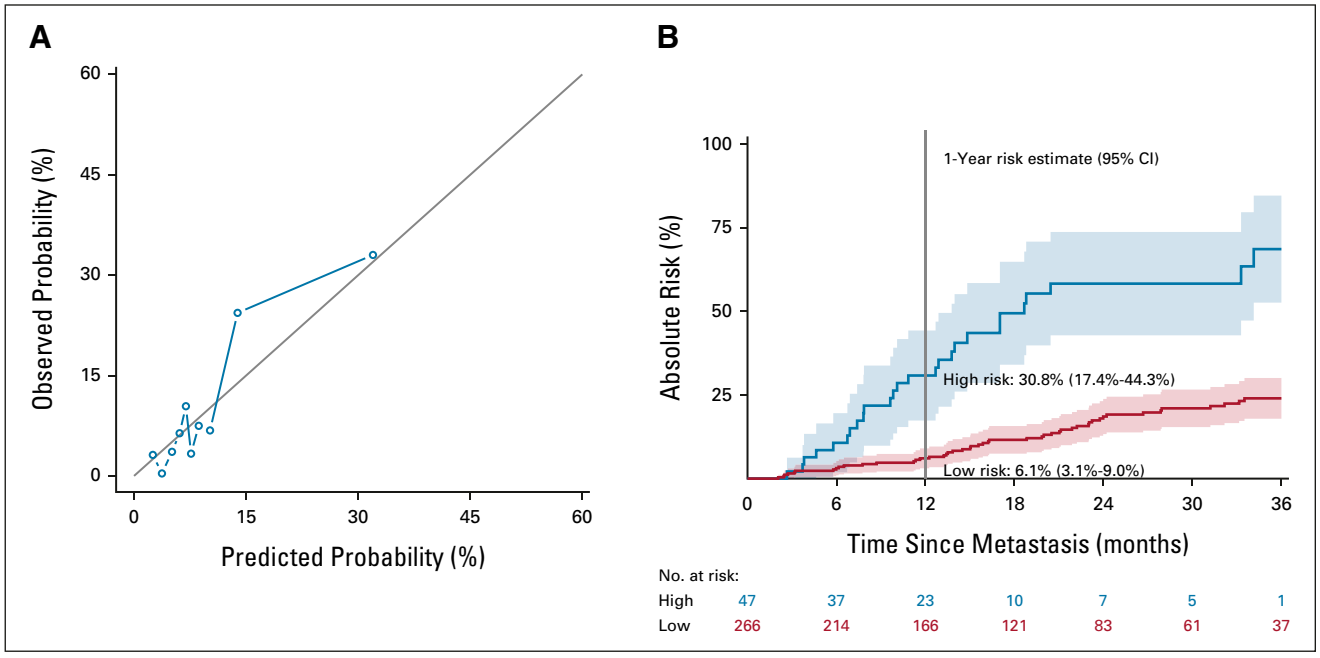

FIG 3. The performance of the prediction model for brain metastasis risk. (a) Calibration plot with discriminative performance (AUC) and prediction accuracy (Brier score) on the basis of an internal validation using 1,000 bootstrapped resamples; (b) risk stratification plot for cumulative incidence by high- versus low-risk groups stratified by a 1-year risk threshold of > 14.2% (ie, 85th percentile of the estimated risk using the proposed model), with 95% CIs and point estimates for 1-year cumulative incidence for each group. AUC, area under the curve.

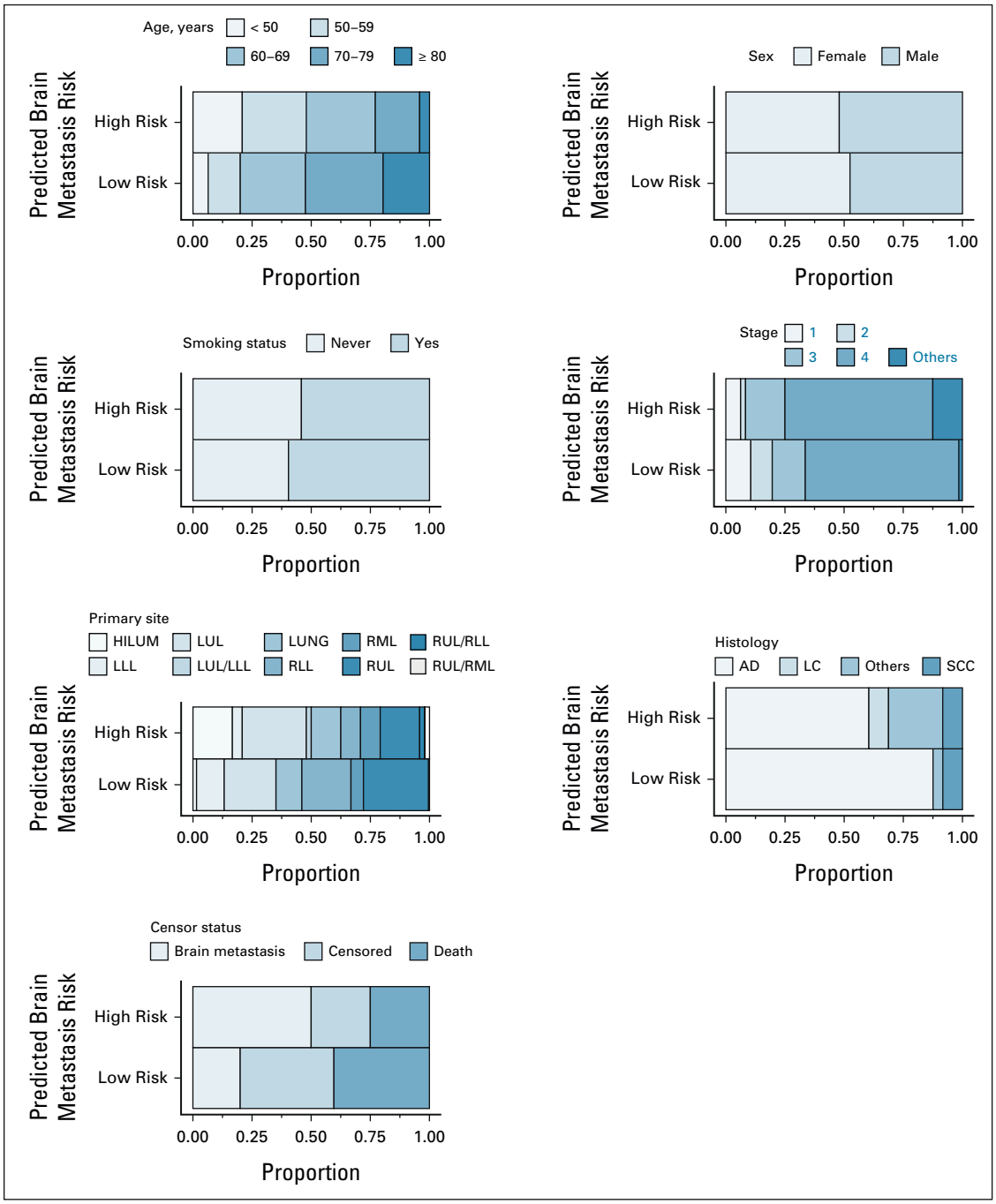

FIG 4. Comparison of patient characteristics in the high-risk versus low-risk groups. High risk determined by a stratification by a 1-year risk of 14.2% (ie, 85th percentile of the estimated risk using the proposed model). AD, *adenocarcinoma*; HILUM, hilum; LC, large cell; LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe; SCC, squamous cell carcinoma.

model to incorporate multiple risk factors into a single predictive score to determine brain metastasis risk for a given patient. Machine learning classifiers, in contrast to traditional regression-based classifiers, can account for multiple high-dimensional features, including comprehensive next-generation sequencing results. In addition, we evaluated the potential clinical utility of the model by comparing outcomes for patients identified to be at high risk for brain metastasis, stratified by brain MRI screening frequency.

Limitations of our study include its retrospective nature and limitation to a single institution. The patients seen in a San Francisco Bay Area academic center may not be representative of patients at other institutions. Brain MRI surveillance patterns may vary by institution and influence timing of brain metastasis detection. Some risk factors or

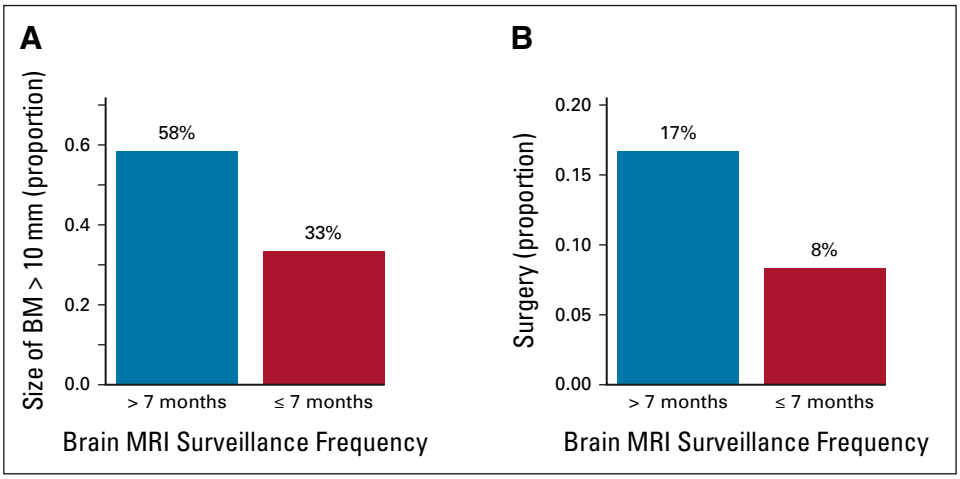

FIG 5. Comparison of clinical outcomes in the high-risk subgroup that developed BN (n = 24): (A) brain metastasis size at diagnosis and (B) surgery at brain metastasis diagnosis. A high-risk subgroup was defined as patients whose predicted 1-year risk was larger than 14.2% (ie, 85th percentile of the estimated risk using the proposed model). Subsets shown are those high-risk patients who developed BN. BM, brain metastasis; MRI, magnetic resonance imaging.

genomic alterations are only present in a small subset (< 3%) of the cohort, such as *RET*, and *ROS1* driver fusions or other clinical variables may not be fully captured. In developing a prediction model, we chose a single starting time point, ie, diagnosis of distant metastatic disease, and used patient information collected at a given time to predict brain metastasis risk. However, risk may be dynamic and dependent on changing treatment. Incorporating changing risk profiles over time and additional prognostic variables, such as performance status, remains an interesting direction for future studies. Regarding the potential application of the model, the clinical utility of brain MRI surveillance in asymptomatic patients must be weighed against the economic cost of additional scans and potential patient anxiety around more frequent scans.

In summary, we developed a predictive model for the risk of brain metastasis in lung cancer using comprehensive clinical and genomic features that can aid clinical decision making. This can help identify patients with lung cancer at high risk of developing brain metastasis, who can benefit from more intensive brain MRI surveillance at earlier stages to reduce morbidity of subsequent brain metastasis treatment.

AFFILIATIONS

[1]Division of Oncology, Department of Medicine, Stanford University School of Medicine, Stanford, CA
[2]Quantitative Sciences Unit, Department of Medicine, Stanford University School of Medicine, Stanford, CA
[3]Department of Biomedical Data Science, Stanford University, Stanford, CA
[4]Department of Pathology, Stanford University, Stanford, CA
[5]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA
[6]Department of Neurology & Neurological Sciences, Stanford University of Medicine, Stanford, CA
[7]Department of Neurosurgery, Stanford University School of Medicine, Stanford, CA

CORRESPONDING AUTHOR

Summer S. Han, PhD, Quantitative Sciences Unit, Stanford University School of Medicine, 3180 Porter Dr, Office 118, Stanford, CA 94304; e-mail: summer.han@stanford.edu.

EQUAL CONTRIBUTION

J.W. and V.D. contributed equally to this work.

SUPPORT

Supported by Division of Cancer Epidemiology and Genetics, National Cancer Institute, Grant No.: U54 CA261717; Grant Recipient: Melanie Hayden Gephart.

AUTHOR CONTRIBUTIONS

Conception and design: Julie Wu, Melanie Hayden Gephart, Summer S. Han
Financial support: Hanlee Ji, Melanie Hayden Gephart, Summer S. Han
Administrative support: Summer S. Han
Provision of study materials or patients: Hanlee Ji, Melanie Hayden Gephart, Heather Wakelee
Collection and assembly of data: Julie Wu, Jessica Hellyer, Nathaniel Myall, Solomon Henry, Douglas Wood, Henning Stehr, Heather Wakelee, Summer S. Han
Data analysis and interpretation: Julie Wu, Victoria Ding, Sophia Luo, Eunji Choi, Nathaniel Myall, Hanlee Ji, Seema Nagpal, Melanie Hayden Gephart, Heather Wakelee, Joel Neal, Summer S. Han
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.
Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Julie Wu
Honoraria: MJH Associates

Nathaniel Myall
Honoraria: Patient Power

Solomon Henry
Stock and Other Owner Interests: Pfizer

Hanlee Ji
Patents, Royalties, Other Intellectual Property: patent held by Sanford University for a sequencing technology
Travel, Accommodations, Expenses: Eisai, Sanofi, Bio-Rad

Seema Nagpal
Consulting or Advisory Role: Novocure, Biocept, Seattle Genetics, Nurix, Pyramid Biosciences, Mirati Therapeutics
Research Funding: Berg Pharma (Inst), Inovio Pharmaceuticals (Inst), Agios (Inst), Pharm Abcine (Inst)

Melanie Hayden Gephart
Stock and Other Owner Interests: SmartLens
Consulting or Advisory Role: Strategy Inc Tamar, Robeaute
Research Funding: Quadriga Biosciences
Patents, Royalties, Other Intellectual Property: SmartLens Patent

Heather Wakelee
Consulting or Advisory Role: AstraZeneca, Mirati Therapeutics, Blueprint Medicines
Research Funding: Genentech/Roche (Inst), Clegene (Inst), AstraZeneca/MedImmune (Inst), Novartics (Inst), Clovis Oncology (Inst), Xcovery (Inst), Bristol Myers Squibb (Inst), ACEA Biosciences (Inst), Arrys Therapeutics (Inst), Merck (Inst), Seattle Genetics (Inst), Helsinn Therapeutics (Inst)
Uncompensated Relationships: Merck
Open Payment Link: Genetech/Roche

Joel Neal
Honoraria: CME Matters, Clinical Care Options, Research to Practice, Medscape, Biomedical Learning Institute, Peerview, Prime Oncology, Projects In Knowledge, Rockpointe CME, MJH Life Sciences
Consulting or Advisory Role: AstraZeneca, Genentecg/Roche, Exelixis, Jounce Therapeutics, Takeda, Lilly, Calithera Biosciences, Amgen, Iovance Biotherapeutics, Blueprint Medicines, Regeneron, Natera, Sanofi/Regeneron, D2G Oncology, Surface Oncology, Turning Point Therapeutics
Research Funding: Genentech/Roche (Inst), Merck (Inst), Novartis (Inst), Boehringer Ingelheim (Inst), Exelixis (Inst), Nektar (Inst), Takeda (Inst), Adaptimmune (Inst), GlaxoSmithKline (Inst), Janssen (Inst), AbbVie (Inst)
Patents, Royalties, Other Intellectual Property: UpToDate—Royalites

No other potential conflicts of interest were reported.

REFERENCES

1. Barnholtz-Sloan JS, Sloan AE, Davis FG, et al: Incidence proportions of brain metastases in patients diagnosed (1973 to 2001) in the Metropolitan Detroit Cancer Surveillance System. J Clin Oncol 22:2865-2872, 2004
2. Villano JL, Durbin EB, Normandeau C, et al: Incidence of brain metastasis at initial presentation of lung cancer. Neuro Oncol 17:122-128, 2015
3. Scoccianti S, Ricardi U: Treatment of brain metastases: Review of phase III randomized controlled trials. Radiother Oncol 102:168-179, 2012
4. Mantovani C, Gastino A, Cerrato M, et al: Modern radiation therapy for the management of brain metastases from non-small cell lung cancer: Current approaches and future directions. Front Oncol 11:772789, 2021
5. Fenske DC, Price GL, Hess LM, et al: Systematic review of brain metastases in patients with non–small-cell lung cancer in the United States, European Union, and Japan. Clin Lung Cancer 18:607-614, 2017
6. Thomas NJ, Myall NJ, Sun F, et al: Brain metastases in EGFR- and ALK-positive non-small cell lung cancer: Outcomes of CNS penetrant tyrosine kinase inhibitors (TKIs) alone versus in combination with radiation. J Thorac Oncol 17:116-129, 2022
7. Waqar SN, Samson PP, Robinson CG, et al: Non-small cell lung cancer with brain metastasis at presentation. Clin Lung Cancer 19:e373-e379, 2018
8. Shin DY, Na II, Kim CH, et al: EGFR mutation and brain metastasis in pulmonary adenocarcinomas. J Thorac Oncol 9:195-199, 2014
9. Johung KL, Yeh N, Desai NB, et al: Extended survival and prognostic factors for patients with ALK-rearranged non-small-cell lung cancer and brain metastasis. J Clin Oncol 34:123-129, 2016
10. Mok T, Camidge DR, Gadgeel SM, et al: Updated overall survival and final progression-free survival data for patients with treatment-naive advanced ALK-positive non-small-cell lung cancer in the ALEX study. Ann Oncol 31:1056-1064, 2020
11. Soria JC, Ohe Y, Vansteenkiste J, et al: Osimertinib in untreated EGFR-mutated advanced non–small-cell lung cancer. N Engl J Med 378:113-125, 2018
12. Ramalingam SS, Vansteenkiste J, Planchard D, et al: Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC. N Engl J Med 382:41-50, 2020
13. Ho VKY, Gijtenbeek JMM, Brandsma D, et al: Survival of breast cancer patients with synchronous or metachronous central nervous system metastases. Eur J Cancer 51:2508-2516, 2015
14. Molecular Genetic Pathology: https://stanfordlab.com/content/stanfordlab/en/molecular-pathology/molecular-genetic-pathology.html/
15. Yang SR, Lin CY, Stehr H, et al: Comprehensive genomic profiling of malignant effusions in patients with metastatic lung adenocarcinoma. J Mol Diagn 20:184-194, 2018
16. Nadauld LD, Ford JM, Pritchard D, Brown T: Strategies for clinical implementation: Precision oncology at three distinct institutions. Health Aff (Millwood) 37:751-756, 2018
17. Fu Z, Parikh CR, Zhou B: Penalized variable selection in competing risks regression. Lifetime Data Anal 23:353-376, 2017
18. Wu J, Ding V, Luo S, et al. The impact of molecular driver mutations onbrain metastasis risk depends on timing of brain metastasis relative to diagnosis. Presented at Targeted Therapies of Lung Cancer Meeting, Santa Monica, CA, February 22-26, 2022
19. Gerds TA, Andersen PK, Kattan MW: Calibration plots for risk prediction models in the presence of competing risks. Stat Med 33:3191-3203, 2014
20. Choi E, Luo SJ, Aredo JV, et al: The survival impact of second primary lung cancer in patients with lung cancer. J Natl Cancer Inst 114:618-625, 2022
21. Harrell FE, Lee KL, Califf RM, et al: Regression modelling strategies for improved prognostic prediction. Stat Med 3:143-152, 1984
22. Harrell FE, Lee KL, Mark DB: Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15:361-387, 1996
23. Putter H, Spitoni C: Non-parametric estimation of transition probabilities in non-Markov multi-state models: The landmark Aalen-Johansen estimator. Stat Methods Med Res 27:2081-2092, 2018

24. Gray RJ: A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 16:1141-1154, 1988
25. White IR, Royston P, Wood AM: Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 30:377-399, 2011
26. RAMBO: https://hanlab.shinyapps.io/RAMBO/
27. George J, Lim JS, Jang SJ, et al: Comprehensive genomic profiles of small cell lung cancer. Nature 524:47-53, 2015
28. Goncalves PH, Peterson S, Vigneau FD, et al: Risk of brain metastases in patients with non-metastatic lung cancer: Analysis of the Metropolitan Detroit Surveillance, Epidemiology, and End Results (SEER) data. Cancer 122:1921-1927, 2016
29. Reungwetwattana T, Nakagawa K, Cho BC, et al: CNS response to osimertinib versus standard epidermal growth factor receptor tyrosine kinase inhibitors in patients with untreated EGFR-mutated advanced non-small-cell lung cancer. J Clin Oncol 36:3290-3297, 2018

