

The Brain-Lung Immunotherapy Prognostic (BLIP) Score: A Novel Robust Tool for Prognostication in Non-Small Cell Lung Cancer Patients with Brain Metastases

Marcus Skribek^{1,2,*}, Maria-Effrosyni Livanou³, Ioannis Vathiotis³, Viktor Strandman^{1,2}, Axel Thorell², Andreas Koulouris^{1,2,4}, Konstantinos Syrigos³, Simon Ekman^{1,2,#}, Georgios Tsakonas^{1,2,#}

1. Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden
2. Medical Unit Head-Neck, Lung, Skin Cancer, Theme Cancer, Karolinska University Hospital, Stockholm, Sweden
3. Third Department of Internal Medicine, Sotiria Thoracic Diseases Hospital of Athens, National and Kapodistrian University of Athens, Athens, Greece
4. Laboratory of Translational Oncology, Medical School, University of Crete, Heraklion, Greece

Shared senior authors

* Corresponding author: Dr. Marcus Skribek, MD, Department of Oncology-Pathology, Karolinska Institutet, Solnavägen 1, 171 77, Stockholm, Sweden / Medical Unit Head-Neck, Lung, Skin Cancer, Theme Cancer, Karolinska University Hospital, Eugeniavägen 3, 171 76, Stockholm, Sweden. ORCID: 0000-0001-8692-6426. Email: marcus.skribek@ki.se.

Highlights:

- The BLIP score is a new prognostic tool for NSCLC.
- It focuses on patients with brain metastases undergoing immunotherapy.
- The score integrates clinical and molecular factors.
- Internal validation showed strong prognostic power and reliability.
- External validation confirmed effectiveness across diverse patient populations.
- Key factors include histology, actionable mutations, age, and brain metastases count.
- The score stratifies patients into “Good” and “Poor” groups.
- The BLIP score aids in personalized treatment decision-making.

Keywords: NSCLC prognosis, brain metastases, immune checkpoint inhibitors, prognostic score, personalized cancer therapy.

Abstract

Background: Lung cancer remains the leading cause of cancer-related mortality, with brain metastases (BMs) significantly worsening prognosis and quality of life. The advent of immune checkpoint inhibitors (ICIs) has revolutionized the treatment landscape for non-small cell lung cancer (NSCLC). However, precise prognostic tools are essential to optimize clinical decision-making in this context.

Methods:

The Brain-Lung Immunotherapy Prognostic (BLIP) score was developed based on a retrospective cohort of NSCLC patients treated with ICIs at Karolinska University Hospital, Sweden. Prognostic factors were identified using both univariate and multivariate Cox regression analyses. Internal validation was conducted using bootstrap resampling, penalized Cox regression, k-fold cross-validation, and receiver operating characteristics (ROC) analysis. External validation was performed using an independent cohort from Sotiria Thoracic Diseases Hospital of Athens, Greece.

Results:

From a total cohort of 1844 patients screened across both study sites, 152 patients from Karolinska University Hospital and 116 from Sotiria Thoracic Diseases Hospital of Athens, Greece, were included in the final analysis. Key prognostic factors influencing outcomes included histology, actionable mutations, age at BM diagnosis, and the number of BMs. The BLIP score effectively stratified patients into two prognostic groups: “Good” and “Poor”, with a median overall survival (OS) of 15 and 7 months, respectively (hazard ratio [HR]: 0.4; $p < 0.0001$). External validation confirmed these findings, showing a significantly lower risk of death for the “Good” group compared to the “Poor” group (HR: 0.49; $p = 0.0063$). The model’s robust prognostic performance was confirmed with an area under the ROC curve of 0.87, highlighting its accuracy in predicting survival outcomes.

Conclusion:

The BLIP score provides a reliable, validated prognostic tool for NSCLC patients with BMs undergoing ICI therapy. By integrating both molecular and clinical variables, it offers significant improvements over existing models. Prospective validation could further support its use in personalized treatment strategies, improving clinical outcomes and patient management.

1. Introduction

Lung cancer is the leading cause of cancer-related mortality, with non-small cell lung cancer (NSCLC) comprising most cases [1]. Brain metastases (BM) are a frequent and severe complication in NSCLC patients, significantly affecting both prognosis and quality of life. Immune checkpoint inhibitors (ICIs) have revolutionized treatment for advanced lung cancer, improving survival outcomes [1]. However, BMs complicate treatment and prognostic evaluations, necessitating a personalized approach.

Approximately 25% of patients with metastatic malignancies develop BMs [2], with 20–45% of lung cancer patients at risk [3,4]. BMs in NSCLC, particularly in adenocarcinoma and tumors harboring oncogenic driver alterations, are linked to poor prognosis [5]. ICIs, either combined with chemotherapy or as monotherapy, show promise for intracranial responses in NSCLC patients with BMs [6-10]. Nevertheless, management remains challenging due to patient heterogeneity, tumor variation, limited drug penetration into the brain, and resistance to prior therapies [11]. In this context, prognostic scores are crucial for individualized treatment approaches. By integrating clinical parameters, including patient-reported outcomes, biological markers, and disease characteristics, these tools facilitate treatment decisions, clinical trial stratification, and resource allocation.

The Recursive Partitioning Analysis (RPA) is a traditional clinical prognostic tool that categorizes patients based on age, Karnofsky performance status (KPS), primary tumor control, and presence of extracranial metastases [12]. However, it was developed using a relatively small, heterogeneous population of patients who received whole-brain radiotherapy. The Graded Prognostic Assessment (GPA) incorporates age, KPS, extracranial metastases, and the number of metastases, which are validated prognostic factors [13,14]. Over time, the GPA has evolved to include disease-specific factors, enhancing its accuracy, but it does not incorporate molecular biomarkers [13]. Furthermore, neither the RPA nor the GPA takes into account the effects of ICIs. As personalized approaches advance, updates such as disease-specific GPA (DS-GPA) [15], Lung-molGPA [16], and NSCLC GPA [11] have incorporated oncogenic profiles and programmed death-ligand 1 (PD-L1) expression. However, PD-L1 exhibits significant intra- and intertumor heterogeneity, resulting in inconsistent assessments, making it more predictive than prognostic [17-19]. The ALK-Brain Prognostic Index (ALK-BPI), developed for anaplastic lymphoma kinase (ALK)-positive NSCLC patients with BM treated with tyrosine kinase inhibitors, provides targeted insights but is limited to a specific patient population [20]. While several other clinical prognostic scores have been developed and validated [21-26], their utility for NSCLC patients with BMs requires further investigation.

Existing indices fail to capture the complexities introduced by ICIs in patients with BMs. This underscores the pressing need for a specialized prognostic tool tailored to this patient subset. The present study addresses this gap by developing and validating the Brain-Lung Immunotherapy Prognostic (BLIP) score, specifically designed for NSCLC patients with BMs.

2. Methods

2.1. Primary Patient Population and Study Design

This retrospective cohort study included patients diagnosed with NSCLC scheduled to receive ICIs at Karolinska University Hospital (KUH), Stockholm, Sweden, between July 2015 and August 2022. Patients were identified using the Cytodos software (CSAM Health, Norway), which manages the dosing, requisitioning, production, administration, and documentation of anticancer treatments.

Inclusion criteria were based on treatment received and diagnosis under ICD-10 code C34. Patients were excluded if they did not have primary NSCLC, had no BMs, did not receive ICIs during BM treatment, or were enrolled in prospective clinical trials.

Retrospective data was gathered from electronic medical records from KUH. The collected variables included sex, age, smoking status, histology, PD-L1 expression, genetic mutations, Eastern Cooperative Oncology Group Performance Status (ECOG PS) at BM diagnosis, number of BMs at BM diagnosis, BM response to prior therapy, location of metastases at lung cancer diagnosis, presence of extracranial metastases at BM diagnosis, BM classification (primary or secondary), symptomatic status of BM at BM diagnosis, details of treatments received, and outcome measures.

The study was approved by the regional ethical review board (approval number: 2020-02636) and adhered to the principles outlined in the Declaration of Helsinki. Due to the retrospective nature and pseudonymization of data, the ethical board waived the need for informed consent. Clinical stages were classified according to the 8th edition of the TNM classification.

2.2. Statistical Analyses and Development of the BLIP Score

Categorical variables were analyzed using descriptive statistics, and statistical significance was set at $p < 0.05$ (two-sided test). Prognostic factors related to intracranial overall survival (OS) were identified using univariate and multivariate Cox proportional hazards regression to estimate hazard ratios (HR) and 95% confidence intervals (95% CI).

All clinical variables were initially assessed using univariate Cox regression. Variables showing statistical or clinical significance were included in the multivariate analysis. The backward stepwise elimination method was used to identify significant prognostic variables, which formed the basis for the BLIP score. Point allocations were based on regression coefficients and adjusted hazard ratios (aHRs).

The Kaplan-Meier method was used to assess the impact of the BLIP score on intracranial OS with binary outcomes, with the log-rank test comparing survival curves. Various cut-offs were tested for statistical significance, selecting the one with the highest prognostic value for the BLIP score. Intracranial OS was defined as the time from BM diagnosis to death, with living patients censored at the last follow-up. Importantly, patients had to receive ICIs while BMs were present to accurately capture the intracranial effect of ICIs, a unique criterion compared to similar studies. To ensure model robustness and assess multicollinearity among predictors, a variance inflation factor (VIF) analysis was conducted.

All statistical analyses were performed using RStudio (version 4.3.2) using the packages: readxl, survival, survminer, dplyr, writexl, car, caret, ggplot2, ROSE, glmnet, boot, pROC, tidyr and rmda (R Foundation for Statistical Computing, Vienna, Austria).

2.3. Internal Validation

The prognostic score was internally validated using multiple statistical methods. Data was split into training (60%) and testing (40%) sets to ensure robustness, providing a reasonable balance given the limited sample size while simulating real-world application. Kaplan-Meier survival curves were generated to visualize survival probabilities over time, with comparisons made

using the log-rank test. Cox proportional hazards regression analyses were conducted to estimate HRs with 95% CI, assessing the model's feasibility.

Model stability was tested using Efron-Gong bootstrap resampling, while penalized Cox regression (using LASSO) addressed overfitting and multicollinearity among predictors. The Partial Likelihood Deviance method was used for lambda (λ) selection in penalized regression, optimizing model parameters to minimize overfitting and maximize performance. Further validation involved 10-fold cross-validation, iteratively training and validating the model on different subsets to ensure performance was not dependent on a single train-test split. Together, these methods offer a comprehensive validation approach, with cross-validation focusing on the generalization aspect and bootstrap on the stability and variability of the model's estimates.

Calibration plots compared predicted probabilities with observed outcomes to assess accuracy. Performance metrics included the concordance index (c-index) for discriminative ability, Brier score to quantify prediction accuracy, and Receiver Operating Characteristics (ROC) curves with area under the curve (AUC) for sensitivity and specificity evaluation. Youden's J statistic was used to determine the optimal threshold by maximizing the sum of sensitivity and specificity. Decision curve analysis evaluated the clinical utility of the prognostic model by assessing the net benefit across different threshold probabilities.

These steps provided a comprehensive internal validation of the prognostic score, ensuring its reliability and robustness in predicting clinical outcomes. Detailed explanations of the analyses are provided in the **Suppl. Data**.

2.4. External Validation

External validation of the BLIP score was performed using an independent cohort from Sotiria Thoracic Diseases Hospital of Athens, Greece. Data was collected retrospectively from medical records between December 2014 and July 2023. The inclusion and exclusion criteria were consistent with those used in the primary cohort at KUH. This approach ensured the robustness and effectiveness of the BLIP score across a different patient population and healthcare system.

Patients were scored using the BLIP score, and survival outcomes were analyzed through Kaplan-Meier estimates, with log-rank tests used to compare survival curves. Cox regression analyses were employed to estimate HRs and 95% CI. All analyses were conducted locally. The study was approved by the regional ethical review board (approval number: 7450).

3. Results

3.1. Patient Characteristics

Out of 914 patients screened at KUH and 930 patients at Sotiria Thoracic Diseases Hospital of Athens, 152 and 116 patients, respectively, were included in the study (**Figure 1**). **Table 1** summarizes the patient characteristics. The primary cohort had a median follow-up of 17.5 months, with 58.6% female participants, whereas the validation cohort had a median follow up of 15.0 months and 77.6% male participants. The validation cohort had a higher proportion of current smokers (73.3% vs. 37.5%). Non-squamous carcinoma was predominant in both cohorts (91.4% in the primary cohort and 79.3% in the validation cohort). PD-L1 expression >50% was observed in 34.2% of the primary cohort and 21.6% of the validation cohort. Actionable mutations were present in 13.8% of the primary cohort and 6.0% of the validation cohort. The median age at BM diagnosis was 68 years in the primary cohort and 64 years in the validation cohort. ECOG PS 0–1 was seen in 83.5% (primary) and 67.2% (validation). Most patients had 1–3 BMs in both the primary (57.9%) and validation (81.9%) cohort. In the primary cohort, BM was diagnosed by CT in 40.1% and MRI in 59.9%, while in the validation cohort, CT was used in 28.4% and MRI in 70.7%. Pembrolizumab was the most used ICI in the primary cohort (72.4%), with nivolumab more common in the validation cohort (42.2%). Progressive disease was the primary reason for ICI discontinuation in both cohorts (75.0% primary, 78.4% validation). See **Suppl. Table 1** for more details.

3.2. Clinical Factors Affecting the BLIP Score

Univariate analysis in the primary cohort identified factors associated with poor survival: squamous histology (HR = 3.17; 95% CI: 1.72–5.84; $p < 0.001$), thoracic metastasis (HR = 1.65; 95% CI: 1.15–2.38; $p = 0.007$), non-actionable mutations (HR = 0.56; 95% CI: 0.32–0.96; $p = 0.034$), age ≥ 65 (HR = 1.77; 95% CI: 1.22–2.56; $p = 0.003$), higher ECOG PS (HR = 2.48; 95% CI: 1.20–5.13; $p = 0.015$), and more than three BMs (HR = 1.52; 95% CI: 1.07–2.16; $p = 0.021$) (**Suppl. Table 2**). In the validation cohort, thoracic metastasis was not prognostic (HR = 0.85; 95% CI: 0.58–1.26; $p = 0.42$).

Multivariate analysis confirmed squamous histology (aHR = 3.77; 95% CI: 1.96–7.24; $p < 0.001$), non-actionable mutations (aHR = 0.52; 95% CI: 0.30–0.89; $p = 0.017$), older age (aHR = 1.57; 95% CI: 1.08–2.30; $p = 0.018$), and more than three BMs (aHR = 1.70; 95% CI: 1.17–2.46; $p = 0.005$) as significant predictors of intracranial OS. VIF values indicated no multicollinearity issues (**Table 2**).

3.3. Creation and Performance of the BLIP Score

Survival outcomes for prognostic factors are shown in **Table 3**. Median intracranial OS was 11.5 months (95% CI: 8–14) for the primary cohort and 12.0 months (95% CI: 10–15) for the validation cohort. Non-squamous histology had a median OS of 12.0 months (95% CI: 9–16) in the primary cohort versus 14.0 months (95% CI: 11–16) in the validation cohort. Squamous histology had a median OS of 5.0 months (95% CI: 2–not reached [NR]) in the primary cohort and 8.0 months (95% CI: 6–NR) in the validation cohort. Patients with actionable mutations had a median OS of 30.0 months (95% CI: 9–NR) in the primary cohort versus 16.0 months (95% CI: 9–NR) in the validation cohort. Points were allocated based on multivariate analysis, creating the BLIP score (**Table 4**). Kaplan-Meier estimates formed two prognostic groups: “Poor” (0–1 points) with a median OS of 7.0 months (95% CI: 4–10) and “Good” (2–4 points) with a median OS of 15 months (95% CI: 11–25). The HR was 0.40, indicating a 60% reduction in death risk for the “Good” group (**Figure 2A**). The difference between groups was highly significant ($p < 0.0001$).

3.4. Internal Validation of the BLIP Score

The dataset was divided into training (60%, N = 92) and testing (40%, N = 60) cohorts (**Suppl. Figure 1**). Both cohorts showed significant survival differences between prognostic groups (training: $p < 0.0001$; testing: $p = 0.003$), demonstrating the BLIP score's prognostic power.

Calibration plots (**Figure 3A**) showed strong agreement between predicted and observed survival probabilities, indicating excellent discrimination. The calibration curve closely followed the ideal line, indicating minimal bias. The Lasso regression model showed robust predictive performance and good calibration. Prognostic group stratification validated the model's clinical utility. The ROSE package addressed class imbalance effectively. Partial likelihood deviance (**Suppl. Figure 2**) remained stable for λ values between -6 and -4.5. The optimal λ minimized deviance while balancing simplicity and accuracy. Predicted survival probabilities at 0, 1, 3, 6, 9, and 12 months for "Good" and "Poor" groups showed clear separation, supporting the model's utility (**Figure 3B**).

Performance metrics included a Brier score of 0.16, indicating accurate predictions, and a c-index of 0.630, showing moderate discrimination (**Suppl. Table 3**). The ROC curve with an AUC of 0.869 indicated excellent discrimination (**Figure 3C**). The optimal threshold (0.496) achieved 82.8% sensitivity and 80.6% specificity, enhancing clinical relevance.

Decision curve analysis (**Figure 3D**) showed the BLIP score's higher net benefit across various thresholds, particularly lower ones. At a 0.2 threshold, the BLIP score's net benefit was ~0.45 compared to 0.20 for treat-all and 0 for treat-none strategies, confirming its utility in reducing overtreatment and improving decisions.

3.5. External Validation of the BLIP Score

In the validation cohort, Kaplan-Meier survival curves (**Figure 2B**) showed distinct outcomes. Patients in the "Poor" prognostic group (0–1 points) had a median intracranial OS of 8 months (95% CI: 6–11), while those in the "Good" prognostic group (2–4 points) had a median OS of 14 months (95% CI: 12–18). The hazard ratio for the "Good" group compared to the "Poor" group was 0.49 (95% CI: 0.29–0.81), significantly reducing the risk of death ($p = 0.0063$).

3.6. Sensitivity Analysis

A sensitivity analysis was conducted to identify the optimal model for evaluation, outlined in **Suppl. Table 3**. The analysis of regularization models revealed that Ridge, Lasso, and Elastic Net regressions each achieved an AUC of 0.8687, indicating their similar ability to rank predictions accurately. However, the Brier scores, which evaluate the accuracy of probabilistic predictions, differed slightly among the models. The Lasso regression model achieved the lowest Brier score of 0.1555, compared to Ridge (0.1574) and Elastic Net (0.1571). This suggests that Lasso regression provides slightly more accurate probability estimates. Despite identical AUC values, the superior calibration of the Lasso model, as evidenced by its lower Brier score, indicates it is the most suitable choice for this analysis.

4. Discussion

The BLIP score offers a novel and robust prognostic tool for NSCLC patients with BM undergoing ICI therapy. Unlike existing models, the BLIP score integrates molecular biomarkers with key clinical factors, capturing the unique dynamics introduced by ICIs. By incorporating variables such as histology, actionable oncogenic drivers, age at BM diagnosis, and the number of BMs, the BLIP score effectively stratifies patients into prognostic groups. Significant differences in intracranial OS were observed, with a median OS of 7 months for the “Poor” group and 15 months for the “Good” group. Internal validation confirmed strong discriminatory power, and calibration plots indicated minimal bias. Furthermore, decision curve analysis underscored the BLIP score’s clinical utility. External validation further solidified its robustness, showing a significant reduction in the risk of death in the “Good” group compared to the “Poor” group.

Existing prognostic tools, such as RPA [12], developed in patients treated with whole-brain radiotherapy, and GPA [13] guide survival predictions for BM but lack molecular data integration, which is now essential for personalized treatment approaches. These models, built on heterogeneous populations, do not fully account for advancements like ICIs, which are pivotal in treating advanced NSCLC with BMs. While DS-GPA [15] incorporates clinical factors, it does not reflect the impact of ICIs, potentially leading to outdated prognostic estimates. More recent tools, such as Lung-molGPA [16] and NSCLC GPA [11], incorporate molecular markers like PD-L1, which significantly influence survival. However, reliance on this specific biomarker can lead to inconsistent assessments due to its variability [17-19]. In our study, PD-L1 was excluded from the multivariate analysis due to its lack of statistical significance, reflecting its uncertain prognostic value compared to its predictive role.

The survival differences between the BLIP score and other tools likely stem from differences in patient selection, treatment modalities, and prognostic factors. Unlike other scores, the BLIP score is tailored to a specific patient population, calculating OS from the date of BM diagnosis while receiving ICIs, making it highly relevant for current clinical practice. This specificity enables a more accurate understanding of BM in the context of modern treatment modalities, including ICIs.

Actionable and non-actionable alterations were separated to better reflect real-world practice. Although *KRAS* mutations were frequent, targeted therapies were unavailable during the study period. Although thoracic metastases at lung cancer diagnosis were significant in univariate analysis, they were excluded from the model due to unclear clinical relevance—a decision later validated by the external cohort. Additionally, the internal validation cohort’s wide 95% CI likely reflect its smaller size.

Oncogenic drivers are pivotal in determining prognosis and are included in the BLIP score. However, ICIs exhibit limited efficacy in patients with *EGFR*-mutant [27], *ALK*-positive [28] and *BRAF*-mutant [29] NSCLC. Notably, *BRAF*-mutant NSCLC, unlike melanoma, presents unique challenges, emphasizing the need to further investigate the sequencing of tyrosine kinase inhibitors and ICIs in this context [30]. Moreover, *KRAS* mutations complicate prognosis, as patients with these mutations may respond variably depending on the presence or absence of co-mutations (e.g., *KEAP1*, *STK11*, and *TP53*) [31-33]. Smoking further complicates this landscape by influencing genetic co-mutations and other genetic changes [34]. The survival benefit observed in our cohort may reflect real-world variations in the timing of tyrosine kinase inhibitor or ICI use, a practice that is yet to be standardized.

To our knowledge, the BLIP score is the first prognostic tool specifically developed for NSCLC patients with BMs undergoing ICI therapy. It integrates a comprehensive range of clinical and molecular variables, offering a tailored approach that enhances the accuracy of outcome predictions. The use of ECOG PS, rather than KPS, in the BLIP score reflects a variation in practice across different centers and clinical settings, with both scoring systems

still in use depending on tradition and context. This adaptation aligns with contemporary clinical practices where ECOG PS has become more prevalent, particularly in ICI studies.

Both internal and external validations confirmed the BLIP score's reliability, with external validation conducted in a diverse cohort with varying patient characteristics and healthcare systems. This extensive validation, to our knowledge in the largest study of its kind, ensures the score's applicability across different clinical settings, highlighting its value in optimizing personalized treatment strategies for NSCLC patients with BM.

Despite its strengths, this study has several limitations. The retrospective design introduces potential biases and limits causal inference. The relatively small sample size may affect internal validity and increase the risk of Type II errors. Moreover, around 30% of patients in the validation cohort were not tested for oncogenic drivers, potentially influencing the findings. Finally, the limited inclusion of biomarkers may affect the precision of the prognostic model.

In conclusion, the BLIP score represents a significant advancement in the prognostication of NSCLC patients with BM undergoing ICI therapy, addressing critical gaps in existing models. Its rigorous validation, both internally and externally, confirms its reliability across diverse clinical settings. While the retrospective nature and relatively small sample size pose constraints, the BLIP score's ability to accurately stratify patients underscores its potential to guide personalized treatment strategies. Future studies should aim to validate these findings in larger, multicentric (prospective) cohorts and explore the integration of additional biomarkers to further refine this prognostic tool.

Funding

M.S. received funding from the Stockholm Cancer Society (grant number: 009618). G.T. was supported by Region Stockholm (clinical postdoctoral appointment). S.E. received funding from the Swedish Cancer Society (CAN2023/29191057), the Stockholm County Council (987911), King Gustaf V Jubilee Fund (204053), and the Sjöberg Foundation (2022-2024). A.K. was supported and funded by research grants from the European Society for Medical Oncology (ESMO) and the Hellenic Society for Medical Oncology (HeSMO). The funders had no role in the collection, the interpretation of data, or the writing of the manuscript. Any views, opinions, findings, conclusions, or recommendations expressed in this material are those solely of the authors and do not necessarily reflect those of the funders.

Acknowledgements

Gratitude is extended to Dr. Johan Zetterqvist, PhD at the department of Clinical Epidemiology at Karolinska Institutet for his invaluable assistance with the statistical methodology and analysis provided through the CLINICUM program. His expertise and guidance significantly contributed to the success of this research.

Declaration of Interest

The authors report no conflicts of interest.

Data Availability Statement

The data that support the findings in this study are available from the corresponding author, M.S., upon reasonable request.

Author Contributions

Conceptualization: M.S., S.E., G.T.; methodology and data acquisition: M.S., M.L., V.S., A.T., S.E., G.T.; formal analysis: M.S., G.T.; data interpretation: M.S., G.T.; writing—original draft preparation: M.S.; writing—review and editing: M.S., M.L., I.V., V.S., A.T., A.K., K.S., S.E., G.T.; visualization: M.S.; supervision: I.V., K.S., S.E., G.T.; resources and funding acquisition: M.S., S.E., G.T. All authors have read and agreed to the published version of the manuscript.

References

1. Hendriks LE, Kerr KM, Menis J, et al. Non-oncogene-addicted metastatic non-small-cell lung cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2023 Apr;34(4):358-376.
2. Waqar SN, Samson PP, Robinson CG, et al. Non-small-cell Lung Cancer With Brain Metastasis at Presentation. *Clin Lung Cancer*. 2018 Jul;19(4):e373-e379.
3. Le Rhun E, Guckenberger M, Smits M, et al. EANO-ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up of patients with brain metastasis from solid tumours. *Ann Oncol*. 2021 Nov;32(11):1332-1347.
4. Passaro A, Attili I, Morganti S, et al. Clinical features affecting survival in metastatic NSCLC treated with immunotherapy: A critical review of published data. *Cancer Treat Rev*. 2020 Sep;89:102085.
5. Langer CJ, Mehta MP. Current management of brain metastases, with a focus on systemic options. *J Clin Oncol*. 2005 Sep 1;23(25):6207-19.
6. Page S, Milner-Watts C, Perna M, et al. Systemic treatment of brain metastases in non-small cell lung cancer. *Eur J Cancer*. 2020 Jun;132:187-198.
7. Nadal E, Rodríguez-Abreu D, Simó M, et al. Phase II Trial of Atezolizumab Combined With Carboplatin and Pemetrexed for Patients With Advanced Nonsquamous Non-Small-Cell Lung Cancer With Untreated Brain Metastases (Atezo-Brain, GECP17/05). *J Clin Oncol*. 2023 Oct 1;41(28):4478-4485.
8. Paz-Ares L, Ciuleanu TE, Cobo M, et al. First-line nivolumab plus ipilimumab combined with two cycles of chemotherapy in patients with non-small-cell lung cancer (CheckMate 9LA): an international, randomised, open-label, phase 3 trial. *Lancet Oncol*. 2021 Feb;22(2):198-211.
9. Gogishvili M, Melkadze T, Makharadze T, et al. Cemiplimab plus chemotherapy versus chemotherapy alone in non-small cell lung cancer: a randomized, controlled, double-blind phase 3 trial. *Nat Med*. 2022 Nov;28(11):2374-2380.
10. Skribek M, Rounis K, Makrakis D, et al. Outcome of Patients with NSCLC and Brain Metastases Treated with Immune Checkpoint Inhibitors in a 'Real-Life' Setting. *Cancers (Basel)*. 2020 Dec 10;12(12).
11. Sperduto PW, De B, Li J, et al. Graded Prognostic Assessment (GPA) for Patients With Lung Cancer and Brain Metastases: Initial Report of the Small Cell Lung Cancer GPA and Update of the Non-Small Cell Lung Cancer GPA Including the Effect of Programmed Death Ligand 1 and Other Prognostic Factors. *Int J Radiat Oncol Biol Phys*. 2022 Sep 1;114(1):60-74.
12. Gaspar L, Scott C, Rotman M, et al. Recursive partitioning analysis (RPA) of prognostic factors in three Radiation Therapy Oncology Group (RTOG) brain metastases trials. *Int J Radiat Oncol Biol Phys*. 1997 Mar 1;37(4):745-51.
13. Sperduto PW, Berkey B, Gaspar LE, et al. A new prognostic index and comparison to three other indices for patients with brain metastases: an analysis of 1,960 patients in the RTOG database. *Int J Radiat Oncol Biol Phys*. 2008 Feb 1;70(2):510-4.
14. Patchell RA, Tibbs PA, Walsh JW, et al. A randomized trial of surgery in the treatment of single metastases to the brain. *N Engl J Med*. 1990 Feb 22;322(8):494-500.
15. Nieder C, Andratschke NH, Geinitz H, et al. Diagnosis-specific graded prognostic assessment score is valid in patients with brain metastases treated in routine clinical practice in two European countries. *Med Sci Monit*. 2012 Jul;18(7):Cr450-5.
16. Sperduto PW, Yang TJ, Beal K, et al. Estimating Survival in Patients With Lung Cancer and Brain Metastases: An Update of the Graded Prognostic Assessment for Lung Cancer Using Molecular Markers (Lung-molGPA). *JAMA Oncology*. 2017;3(6):827-831.

17. Ilie M, Long-Mira E, Bence C, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: a potential issue for anti-PD-L1 therapeutic strategies. *Annals of Oncology*. 2016;27(1):147-153.
18. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481(7381):306-313.
19. Moutafi MK, Tao W, Huang R, et al. Comparison of programmed death-ligand 1 protein expression between primary and metastatic lesions in patients with lung cancer. *J Immunother Cancer*. 2021 Apr;9(4).
20. Zerdes I, Kamali C, Koulouris A, et al. Validation of the ALK-Brain Prognostic Index for patients with ALK-rearranged lung cancer and brain metastases. *ESMO Open*. 2023 Dec;8(6):102069.
21. Prelaj A, Ferrara R, Rebuzzi SE, et al. EPSILoN: A Prognostic Score for Immunotherapy in Advanced Non-Small-Cell Lung Cancer: A Validation Cohort. *Cancers (Basel)*. 2019 Dec 5;11(12).
22. Wang M, Zhang Y, Liu M, et al. Exploration of a novel prognostic model based on nomogram in non-small cell lung cancer patients with distant organ metastasis: implications for immunotherapy. *Translational Lung Cancer Research*. 2023;12(10):2040-2054.
23. Diker O, Olgun P, Balyemez U, et al. Development of a Novel Predictive-Prognostic Scoring Index for Immune Checkpoint Inhibitors in Advanced Non-small Cell Lung Cancer. *Cureus*. 2023 Jan;15(1):e33234.
24. Mezquita L, Auclin E, Ferrara R, et al. Association of the Lung Immune Prognostic Index With Immune Checkpoint Inhibitor Outcomes in Patients With Advanced Non-Small Cell Lung Cancer. *JAMA Oncol*. 2018 Mar 1;4(3):351-357.
25. Zhou Y, Wu B, Li T, et al. Correlation between the Immune Checkpoint Inhibitors Prognostic Index and Outcomes in Nonsmall Cell Lung Cancer: A Multicentre Analysis. *J Oncol*. 2022;2022:7050817.
26. Mountzios G, Samantas E, Senghas K, et al. Association of the advanced lung cancer inflammation index (ALI) with immune checkpoint inhibitor efficacy in patients with advanced non-small-cell lung cancer. *ESMO Open*. 2021 Oct;6(5):100254.
27. Attili I, Passaro A, Corvaja C, et al. Immune checkpoint inhibitors in EGFR-mutant non-small cell lung cancer: A systematic review. *Cancer Treat Rev*. 2023 Sep;119:102602.
28. Leone G, Passiglia F, Bironzo P, et al. Is there any place for immune-checkpoint inhibitors in the treatment algorithm of fusion-driven non-small cell lung cancer?—a literature review. *Translational Lung Cancer Research*. 2020;9(6):2674-2685.
29. Zhang C, Zhang C, Lin J, et al. Patients With BRAF-Mutant NSCLC May Not Benefit From Immune Checkpoint Inhibitors: A Population-Based Study. *JTO Clinical and Research Reports*. 2020 2020/03/01;1(1):100006.
30. Keilholz U, Ascierto PA, Dummer R, et al. ESMO consensus conference recommendations on the management of metastatic melanoma: under the auspices of the ESMO Guidelines Committee. *Annals of Oncology*. 2020;31(11):1435-1448.
31. Li Y, Hu L, Peng X, et al. Resistance to immune checkpoint inhibitors in KRAS-mutant non-small cell lung cancer. *Cancer Drug Resist*. 2022;5(1):129-146.
32. Negrao MV, Araujo HA, Lamberti G, et al. Comutations and KRASG12C Inhibitor Efficacy in Advanced NSCLC. *Cancer Discov*. 2023 Jul 7;13(7):1556-1571.
33. Ricciuti B, Arbour KC, Lin JJ, et al. Diminished Efficacy of Programmed Death-(Ligand)1 Inhibition in STK11- and KEAP1-Mutant Lung Adenocarcinoma Is Affected by KRAS Mutation Status. *J Thorac Oncol*. 2022 Mar;17(3):399-410.

34. Huang Q, Zhang H, Hai J, et al. Impact of PD-L1 expression, driver mutations and clinical characteristics on survival after anti-PD-1/PD-L1 immunotherapy versus chemotherapy in non-small-cell lung cancer: A meta-analysis of randomized trials. *Oncoimmunology*. 2018;7(12):e1396403.

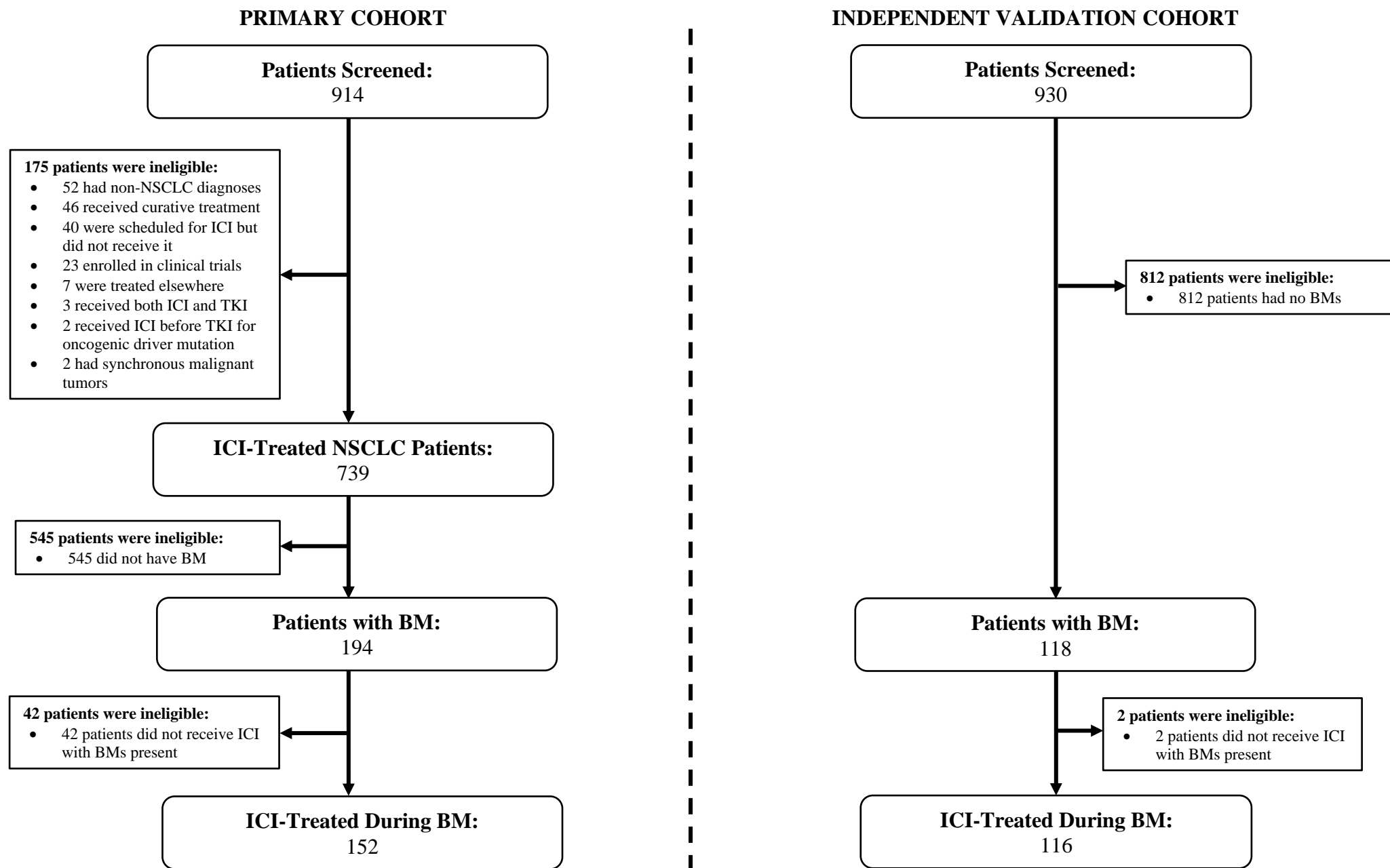


Figure 1: Flowchart of Study Design and Patient Population for NSCLC Patients Treated with ICIs Across Two Cohorts

Abbreviations: NSCLC: non-small cell lung cancer; ICI: immune checkpoint inhibitor; TKI: tyrosine kinase inhibitor; BM: brain metastases.

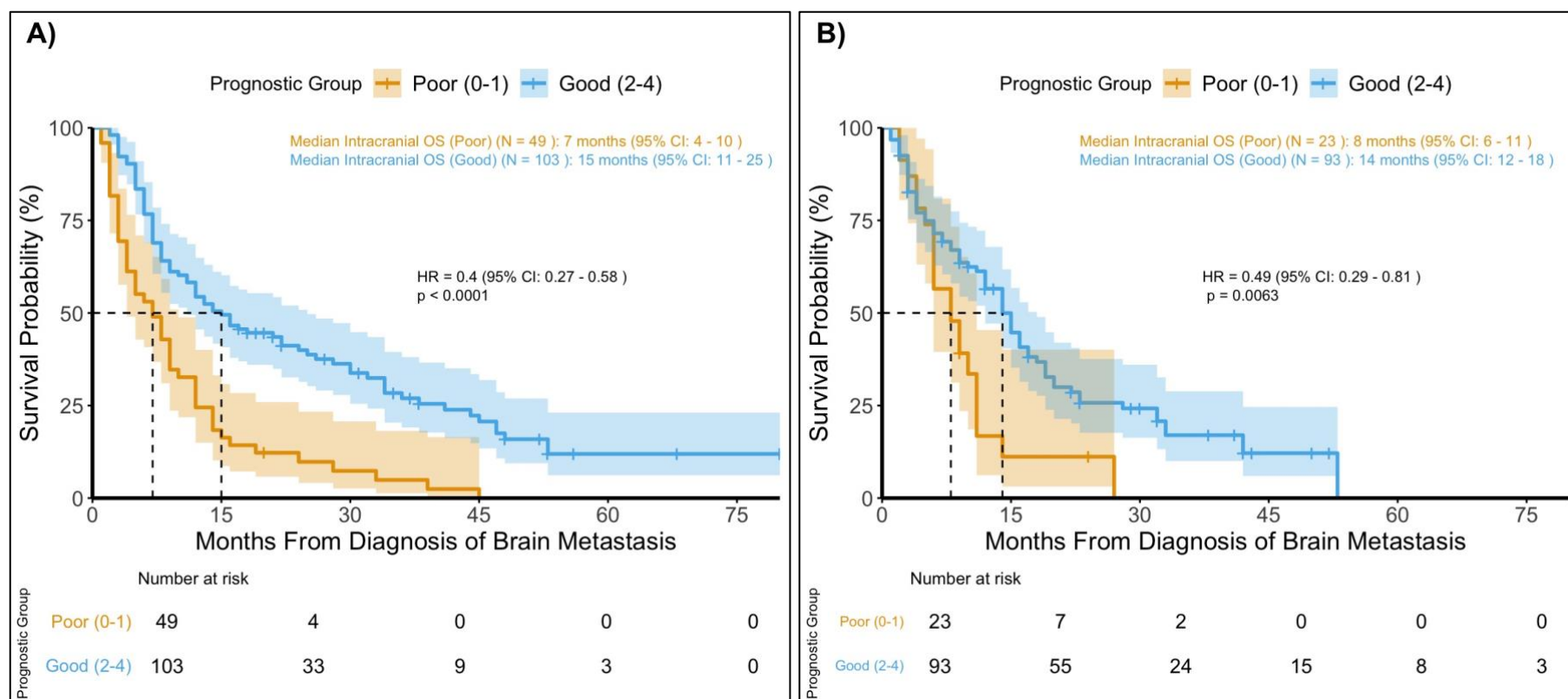


Figure 2: Kaplan-Meier Survival Curves Comparing Intracranial Overall Survival in A) Brain-Lung Immunotherapy Prognostic (BLIP) Groups and in the B) External Validation Cohort

Abbreviations: OS: overall survival; 95% CI: 95% confidence interval; HR: hazard ratio.

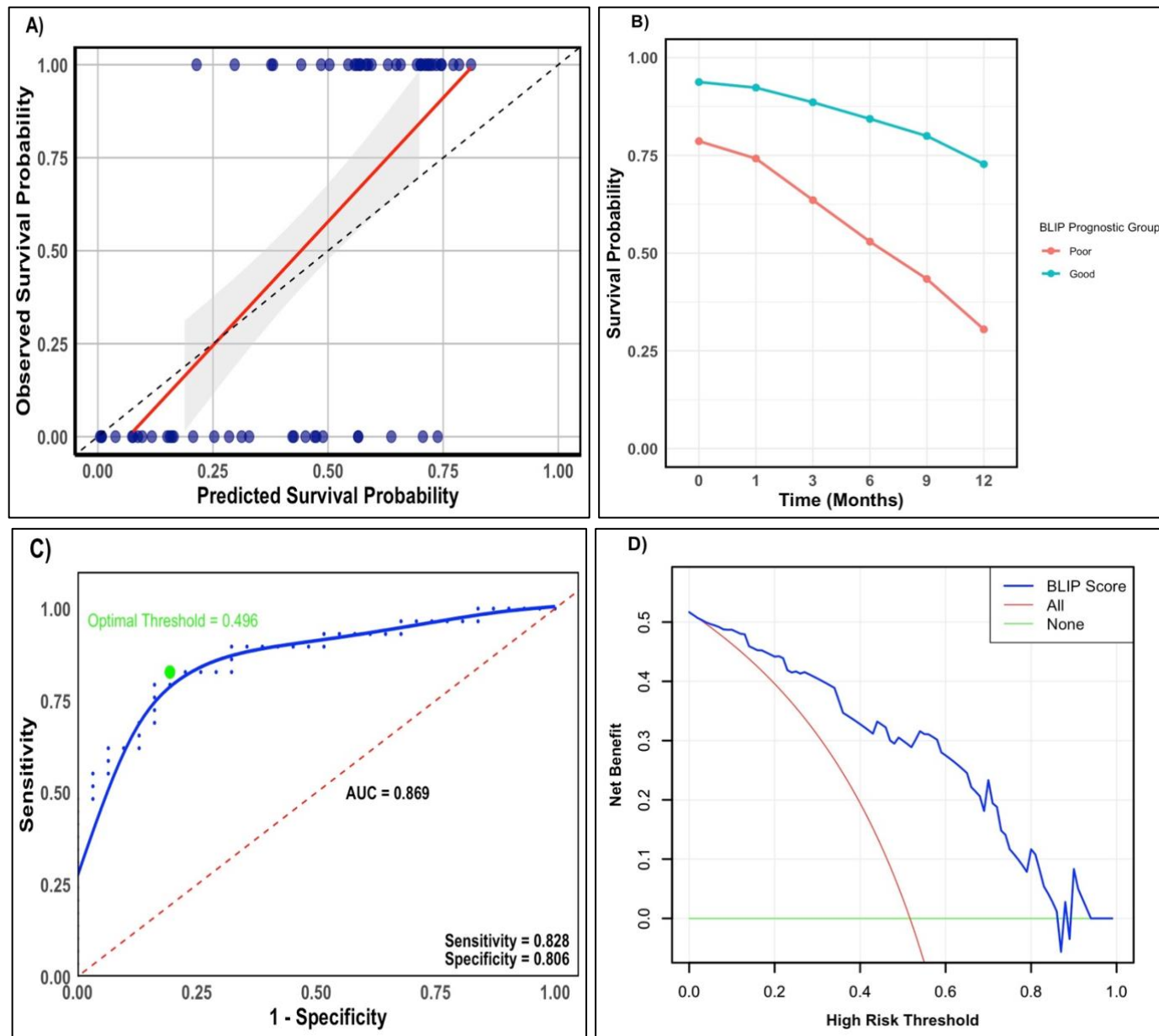


Figure 3: Internal Validation of the Brain-Lung Immunotherapy Prognostic (BLIP) Score: A) Calibration Plot for 12-Month Survival Probabilities, B) Predicted Survival Probabilities, C) ROC Curve, and D) Decision Curve Analysis

Abbreviations: AUC: area under the curve; BLIP: Brain-Lung Immunotherapy Prognostic Score; ROC: receiver operating characteristics.

Variable	Primary Cohort (N = 152)	Validation Cohort (N = 116)
Sex		
Male	63 (41.4%)	90 (77.6%)
Female	89 (58.6%)	26 (22.4%)
Smoking status		
Non-smoker	20 (13.2%)	1 (0.9%)
Former smoker	75 (49.3%)	20 (17.2%)
Current smoker	57 (37.5%)	85 (73.3%)
Missing data	0 (0.0%)	10 (8.6%)
Histology		
Squamous cell carcinoma	13 (8.6%)	24 (20.7%)
Non-squamous cell carcinoma	139 (91.4%)	92 (79.3%)
PD-L1 expression		
Low (<1%)	40 (26.3%)	37 (31.9%)
Intermediate (1–49%)	41 (27.0%)	21 (18.1%)
High (≥50%)	52 (34.2%)	25 (21.6%)
Not tested	19 (12.5%)	33 (28.4%)
Genetic mutation status*		
No mutations	61 (40.1%)	60 (51.7%)
Actionable mutations	21 (18.1%)	7 (6.0%)
<i>EGFR</i>	10 (6.6%)	3 (2.6%)
<i>BRAF</i> ^{V600E}	4 (2.6%)	2 (1.7%)
<i>MET</i> (exon 14 skipping mutation)	4 (2.6%)	0 (0.0%)
<i>ALK</i>	3 (2.0%)	2 (1.7%)
Not tested	6 (3.9%)	35 (30.2%)
Non-actionable mutations	91 (59.9%)	14 (12.1%)
Primary BM	93 (61.2%)	59 (50.9%)
Extracranial metastasis at BM diagnosis		
Absent	11 (7.2%)	NA
Present	141 (92.8%)	NA
Age at BM diagnosis (years)		
Median (IQR; range)	68 (14; 25–84)	64 (14; 33–82)

ECOG PS at BM diagnosis		
0	64 (42.1%)	18 (15.5%)
1	63 (41.4%)	60 (51.7%)
2	17 (11.2%)	16 (13.8%)
3	6 (3.9%)	7 (6.0%)
4	2 (1.3%)	3 (2.6%)
Missing data	0 (0.0%)	12 (10.3%)
Number of BM		
1	49 (32.2%)	57 (49.1%)
2–3	39 (25.7%)	38 (32.8%)
4–5	24 (15.8%)	12 (10.3%)
>5	36 (23.7%)	9 (7.8%)
Leptomeningeal involvement	4 (2.6%)	0 (0.0%)
Size of largest BM (mm)		
Median (IQR; range)	15 (14.5; 1–60)	12 (15.5; 3–47)
Largest BM ≥3 cm		
Yes	23 (15.1%)	14 (12.1%)
Missing data	5 (3.3%)	1 (0.9%)
Neurological symptoms associated with BM		
	95 (62.5%)	39 (33.6%)
Diagnosis of BM		
CT	61 (40.1%)	33 (28.4%)
MRI	91 (59.9%)	82 (70.7%)
Clinical benefit from previous therapy for BM		
No previous treatments	61 (40.1%)	53 (45.7%)
No clinical benefit (PD)	32 (21.1%)	14 (12.1%)
Clinical benefit	23 (15.1%)	18 (15.5%)
SD	6 (3.9%)	15 (12.9%)
PR	14 (9.2%)	2 (1.7%)
CR	3 (2.0%)	1 (0.9%)
NA	36 (23.7%)	31 (26.7%)
Age at ICI Initiation		
Median (IQR; range)	68 (14; 28–84)	64.5 (14.3; 33–83)

Regimen		
Monotherapy	81 (53.3%)	72 (62.1%)
Combination therapy	71 (46.7%)	44 (37.9%)
Name of ICI		
Pembrolizumab	110 (72.4%)	47 (40.5%)
Nivolumab	28 (18.4%)	49 (42.2%)
Atezolizumab	13 (8.6%)	3 (2.6%)
Ipilimumab/Nivolumab	1 (0.7%)	17 (14.7%)
Reason for therapy discontinuation		
Disease progression	114 (75.0%)	91 (78.4%)
Toxicity	25 (16.4%)	6 (5.2%)
Non-cancer related death	4 (2.6%)	2 (1.7%)
Patient decision	3 (2.0%)	1 (0.9%)
Treatment duration over 2 years	2 (1.3%)	4 (3.4%)
Planned therapy break	1 (0.7%)	0 (0.0%)
Other	3 (2.0%)	6 (5.2%)
Death	129 (84.9%)	90 (77.6%)
Median follow-up time	17.5 months	15.0 months

Table 1: Patient Characteristics

Abbreviations: PD-L1: programmed death-ligand 1; EGFR: epidermal growth factor receptor; BRAF: B-Raf proto-oncogene; MET: MET proto-oncogene; ALK: anaplastic lymphoma kinase; IQR: interquartile range; ECOG PS: eastern cooperative oncology group performance status; BM: brain metastasis; NA: not available; PD: progressive disease; SD: stable disease; PR: partial response; CR: complete response; ICI: immune checkpoint inhibitor.

* = Note that some patients had multiple mutations

Variables	<i>p</i> -value	aHR (95% CI)	Coefficient	VIF
Histology (squamous versus non-squamous)	<0.001 ***	3.765 (1.958–7.240)	1.326	1.138
Presence of actionable oncogenic driver	0.017 *	0.516 (0.299–0.889)	-0.663	1.006
Age at BM diagnosis (≥65 versus <65)	0.018 *	1.574 (1.080–2.295)	0.454	1.034
Number of BM (>3 versus 1–3)	0.005 **	1.698 (1.171–2.461)	0.529	1.115

Table 2: Multivariate Cox Proportional Hazards Regression Analysis of Variables Associated with Overall Survival Using Backward Stepwise Elimination

Abbreviations: aHR: adjusted hazard ratio; 95% CI: 95% confidence interval; VIF: variance inflation factor; BM: brain metastasis.

* = $p < 0.05$

** = $p < 0.01$

*** = $p < 0.001$

Prognostic factor	Primary cohort (N = 152) Median intracranial OS	Validation cohort (N = 116) Median intracranial OS
Whole patient population	11.5 months (95% CI: 8–14)	12.0 months (95% CI: 10–15)
Histology		
Non-squamous cell carcinoma	12.0 months (95% CI: 9–16)	14.0 months (95% CI: 11–16)
Squamous cell carcinoma	5.0 months (95% CI: 2–NR)	8.0 months (95% CI: 6–NR)
Presence of actionable oncogenic driver		
No or unknown	10.0 months (95% CI: 8–14)	12.0 months (95% CI: 9–15)
Yes	30.0 months (95% CI: 9–NR)	16.0 months (95% CI: 9–NR)
Age at BM diagnosis		
<65	16.0 months (95% CI: 9–36)	14.0 months (95% CI: 10–19)
≥65	9.0 months (95% CI: 7–12)	11.0 months (95% CI: 8–15)
Number of BM		
1–3	13.5 months (95% CI: 9–24)	14.0 months (95% CI: 10–16)
>3	9.0 months (95% CI: 7–12)	9.0 months (95% CI: 6–NR)

Table 3: Survival Outcomes Associated with Prognostic Factors

Abbreviations: OS: overall survival; 95% CI: 95% confidence interval; NR: not reached; BM: brain metastasis.

Prognostic factor	Point allocation	Patient score
Histology		
Non-squamous	1	
Squamous	0	
Presence of actionable oncogenic driver (<i>EGFR</i>, <i>BRAF</i>^{V600E}, <i>MET</i> exon 14 skipping, <i>ALK</i>)		
Yes	1	
No or Unknown	0	
Age at BM Diagnosis		
<65	1	
≥65	0	
Number of BM		
1–3	1	
>3	0	
Total score	Prognostic group	Median intracranial OS
0–1	Poor	7.0 months (95% CI: 4–10)
2–4	Good	15 months (95% CI: 11–25)

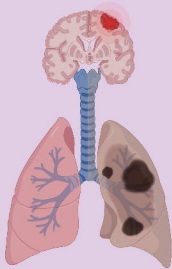
Table 4: Calculation Worksheet for the Brain-Lung Immunotherapy Prognostic (BLIP) Score

Abbreviations: OS: overall survival; 95% CI: 95% confidence interval; EGFR: epidermal growth factor receptor; BRAF: B-Raf proto-oncogene; MET: MET proto-oncogene; ALK: anaplastic lymphoma kinase; BM: brain metastasis.

The Brain-Lung Immunotherapy Prognostic (BLIP) Score

KEY POINTS

- New prognostic tool for NSCLC patients with brain metastases
- Integrates clinical and molecular factors
- Stratifies patients into "Good" and "Poor" prognostic groups



METHOD

Primary Cohort (N = 152)

Prognostic factor	Point allocation
Histology	
Non-squamous	1
Squamous	0
Presence of actionable oncogenic driver (EGFR, BRAF ^{V600E} , MET exon 14 skipping, ALK)	
Yes	1
No or Unknown	0
Age at BM Diagnosis	
<85	1
≥85	0
Number of BM	
1-3	1
>3	0

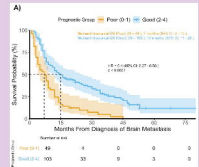
Internal Validation

Independent Validation Cohort (N = 116)

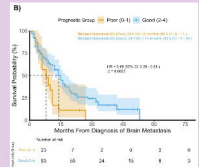


FINDINGS

Survival outcomes in the primary cohort



Survival outcomes in the independent validation cohort



CONCLUSION

- The BLIP score improves personalized treatment decisions for NSCLC patients with brain metastases.
- Validated internally and externally with significant predictive accuracy.

REFERENCE: Skribek et al., 2024