





Systematic Review

Artificial Intelligence for Non-Invasive Prediction of Molecular Signatures in Spinal Metastases: A Systematic Review

Vivek Sanker^{1,*}, Sai Sanikommu², Alexander Thaller³, Zhikai Li⁴, Philip Heesen⁵, Srinath Hariharan¹, Emil O. R. Nordin¹, Maria Jose Cavagnaro¹, John Ratliff¹ and Atman Desai¹

¹ Department of Neurosurgery, Stanford University, Palo Alto, CA 94305, USA; hsrinath@stanford.edu (S.H.); enordin@stanford.edu (E.O.R.N.); mjcava@stanford.edu (M.J.C.); jratliff@stanford.edu (J.R.); atman@stanford.edu (A.D.)

² Department of Neurosurgery, Miller School of Medicine, University of Miami, Miami, FL 33136, USA; vms136@miami.edu

³ Department of Neurosurgery, Medical University of Graz, 8010 Graz, Styria, Austria; alexander.thaller@medunigraz.at

⁴ Department of Clinical Neurosciences, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK; zl498@cam.ac.uk

⁵ Faculty of Medicine, University of Zurich, 8032 Zurich, Switzerland; heesenphilip@gmail.com

* Correspondence: vsanker@stanford.edu

Abstract

Background: Spinal metastases (SMs) are associated with poor prognosis and significant morbidity. We hypothesize that artificial intelligence (AI) models can enhance the identification and clinical utility of genetic and molecular signatures associated with SMs, improving diagnostic accuracy and enabling personalized treatment strategies. **Methods:** A systematic review of five databases was conducted to identify studies that used AI to predict genetic alterations and SMs outcomes. Accuracy, area under the receiver operating curve (AUC), and sensitivity were used for comparison. Data analysis was performed in R. **Results:** Eleven studies met the inclusion criteria, covering three different primary tumor origins, comprising a total of 2211 patients with an average of 201 ± 90 patients (range: 76–359 patients) per study. EGFR, Ki-67, and HER-2 were studied in ten (90.9%), two (18.1%), and one (9.1%) study, respectively. The weighted average AUC is 0.849 (95% CI: 0.835–0.863) and 0.791 (95% CI: 0.738–0.844) for internal and external validation of the established models, respectively. **Conclusions:** AI, through radiomics and machine learning, shows strong potential in predicting molecular markers in SMs. Our study demonstrates that AI can predict molecular markers in SMs with high accuracy.

Keywords: artificial intelligence; genetic markers; spinal metastasis; molecular signatures



Academic Editor: Kwang Woo Ahn

Received: 19 June 2025

Revised: 10 July 2025

Accepted: 19 July 2025

Published: 23 July 2025

Citation: Sanker, V.; Sanikommu, S.;

Thaller, A.; Li, Z.; Heesen, P.;

Hariharan, S.; Nordin, E.O.R.;

Cavagnaro, M.J.; Ratliff, J.; Desai, A.

Artificial Intelligence for Non-Invasive

Prediction of Molecular Signatures in

Spinal Metastases: A Systematic

Review. *Bioengineering* **2025**, *12*, 791.

[https://doi.org/10.3390/](https://doi.org/10.3390/bioengineering12080791)

[bioengineering12080791](https://doi.org/10.3390/bioengineering12080791)

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license

([https://creativecommons.org/](https://creativecommons.org/licenses/by/4.0/)

[licenses/by/4.0/](https://creativecommons.org/licenses/by/4.0/)).

1. Introduction

The incidence of spinal metastasis is estimated to increase by more than 100,000 cases annually [1,2]. The most common primary sources of spinal metastasis are breast (21%), lung (19%), prostate (7.5%), renal (5%), gastrointestinal (4.5%), and thyroid (2.5%) [3]. Patients may present with pain, spinal instability, pathological fractures, or spinal cord compression with weakness and paralysis [4]. Genetic signatures are the patterns of unique genetic activity or expression seen in certain diseases, including spinal metastasis [5]. These genetic signatures are very helpful in aiding in the diagnosis, prognosis, and treatment of SMs. TP53, PTEN, and RB1, as well as pathways like PI3K, are the most common genetic markers implicated in spinal metastasis, and alterations in these genes and pathways

potentially drive the metastatic process [6]. Understanding these genetic and molecular signatures of spinal metastasis is critical for advancing personalized oncological treatment strategies. Such biomarkers can offer insights into tumor behavior and the likelihood of disease progression, treatment resistance, and possible therapeutic targets. Considering the complexity of spinal metastasis, early and precise identification of genetic and molecular signatures can guide tailored treatment strategies, enhance prognostic accuracy, and potentially improve patient outcomes.

Biomarkers such as EGFR, Ki-67, and HER-2 play pivotal roles in spinal metastasis. Epidermal growth factor receptor (EGFR) is a well-studied oncogene with mutations commonly found in non-small cell lung cancer (NSCLC), one of the primary cancers known to cause spinal metastasis [7]. EGFR mutations, especially in exons 19 and 21, are associated with tumor proliferation and responsiveness to tyrosine kinase inhibitors, making EGFR of therapeutic and prognostic relevance in SMs from NSCLC [8]. Ki-67 is linked to cell proliferation and thus serves as a prognostic marker of tumor aggressiveness. Elevated expression, therefore, correlates with rapid growth, poor prognosis, and higher risk of tumor recurrence [9]. Human epidermal growth factor receptor 2 (HER-2) is commonly found in breast cancer, yet its expression has also been identified in spinal metastasis and is associated with increased metastatic potential and targeted treatment responsiveness [10].

Artificial intelligence (AI), including deep learning (DL), machine learning (ML), and various other methods, is being explored to help in the analysis of the vast datasets aiding in research, diagnosis, treatment, and prognosis prediction of various diseases and conditions. Several AI models are being introduced in diagnosing, treating, and prognosing spinal metastasis [11]. AI applications in spinal metastasis utilize the aforementioned techniques as well as radiomics, making use of their distinctive strengths. ML models such as support vector machines (SVMs) and random forests (RFs) are used for classification tasks, e.g., predicting genetic mutations or patient outcomes based on structured clinical or genomic data [12]. SVMs perform well in high-dimensional spaces by identifying optimal hyperplanes for classification [13], while RFs leverage ensemble learning to improve robustness and reduce overfitting [14]. Radiomics refers to the extraction of high-dimensional features from medical imaging, which can be analyzed by ML models for precision diagnosis and treatment [15], or uncover imaging markers associated with molecular alterations. DL models, e.g., artificial neural networks (ANNs) and convolutional neural networks (CNNs), learn hierarchical data representations without the need for much manual input or preprocessing [16]. CNNs are particularly useful in radiomics for automated feature extraction from medical imaging, whereas ANNs can be trained to fully integrate multi-modal inputs [16], improving the prediction of SMs characteristics and outcomes.

In this study, we review and describe the various AI methods used in the literature to identify the genetic signatures associated with SMs.

2. Materials and Methods

2.1. Ethical Review

Ethical review and approval were waived for this study due to it being a systematic review of previously published data and it did not involve human participants or the collection of new data.

2.2. Search Strategy

We searched PubMed, Scopus, Web of Science Advance, Cochrane, and Embase (Ovid) databases to identify relevant studies, using a search query with specific keywords like 'spine metastases', 'artificial intelligence', 'machine learning', 'deep learning', 'neural network', 'radiomics', 'predictive modeling', and 'computer-assisted diagnosis' (Supple-

mentary Table S1). The population under consideration included adult patients with spinal metastases. The inclusion criterion is original studies in the English language reporting the use of AI/DL models in predicting genetic and molecular signatures in spinal metastases.

The exclusion criterion is as follows: Studies unrelated to spinal metastases and those purely investigating primary spinal tumors. Animal studies, reviews, articles in foreign languages, and non-original research articles were also excluded from our analysis to ensure the inclusion of primary research data relevant to our objective. The electronic search ranged from the period’s earliest available date up to 27 January 2025.

2.3. Screening of Studies

Each study’s title and abstract were screened for relevance before proceeding to full-text screening, which was independently assessed by two reviewers (SS and VS). Any discrepancies were addressed through consultation with a third reviewer (SH). The screening of studies adhered to the PRISMA (preferred reporting items for systematic reviews and meta-analyses) guidelines (Figure 1).

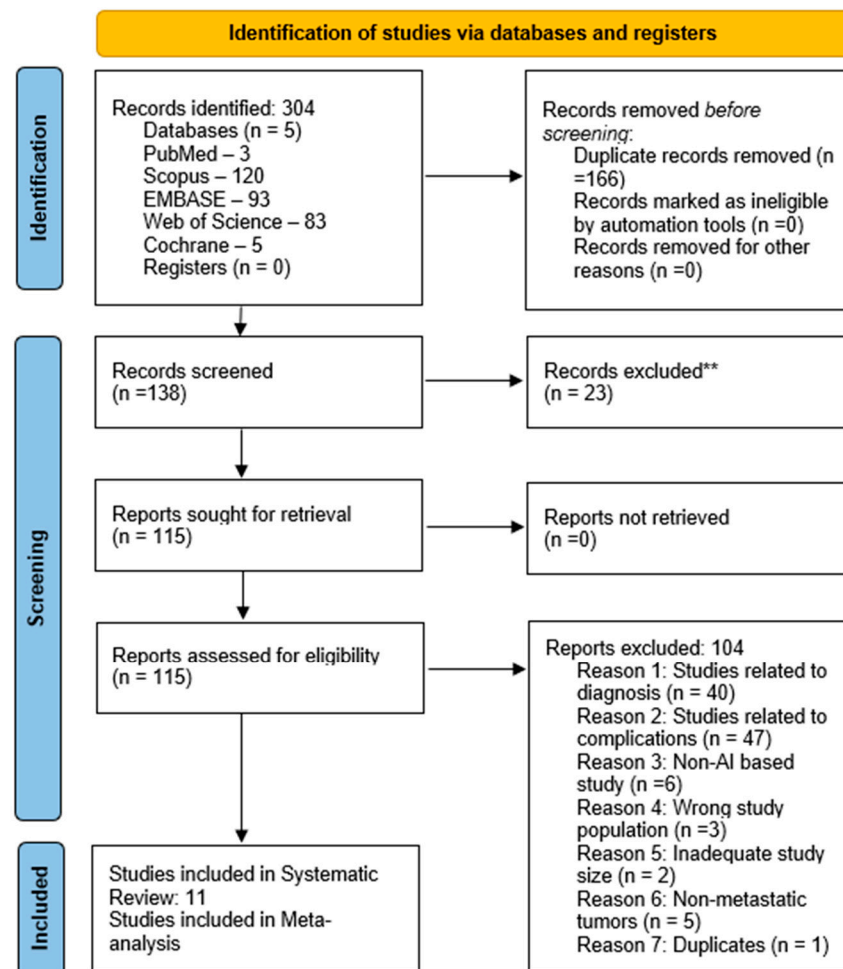


Figure 1. A PRISMA flow diagram is presented to illustrate the screening of studies.

2.4. Data Extraction

Three independent authors (SS, VS, and AT) extracted relevant data from the included studies. The data collected included study design, participant demographics, and the number of participants with respective outcomes and complications. Discrepancies in data extraction were resolved through consensus.

2.5. Data Analysis

Relevant variables were extracted from each of the included articles, such as the primary tumor type, cohort size, prediction model performance matrices: area under the receiver operating characteristic curve (AUC), and type of validation (internal or external validation). The weighted average of the AUC was calculated. All statistical analyses were conducted using Excel, R Statistical Software, version 4.3.1., and Python, version 3.13.3.

2.6. Quality Assessment

The quality assessment was performed using the PROBAST (Prediction model Risk of Bias Assessment Tool) (Supplementary Tables S3 and S4).

PROBAST is designed for assessing the risk of bias and applicability in studies that develop, validate, or update predictive models, which can be analogous to assessing the performance of large language models (LLMs)/natural language processing (NLP) in such settings [17]. It is a structured tool that assesses four domains:

Participants: Evaluating whether the data sources or patient samples used for training and testing are appropriate and representative of the clinical population. **Predictors:** Ensuring that input data or predictors are well defined and appropriately measured. **Outcome:** Ensuring that the outcomes (e.g., model predictions, decisions) are clearly defined and relevant to clinical scenarios. **Analysis:** Evaluating whether the model performance metrics, training/validation processes, and statistical analysis methods are robust and unbiased.

3. Results

This review encompasses eleven studies [18–28] published between 2021 and 2024, including three different primary tumor types (non-small cell lung cancer (NSCLC), breast cancer, and lung adenocarcinoma) with a total of 2211 patients (Table 1). Each study included an average of 201 ± 90 patients, ranging from 76 to 359 patients.

Table 1. An overview of the studies analyzed is presented.

Study	Cohort Size	Mean Age \pm SD	Receptor Analyzed	Primary Tumor Type
Hu et al. (2024) [18]	359	Training cohort = 58.21 ± 9.28 Validation cohort = 59.87 ± 7.23	EGFR	NSCLC
Cheng et al. (2023) [19]	203	Training cohort (EGFR mutant) = 58.12 ± 9.62 (EGFR wild type) = 58.82 ± 10.14 Validation cohort (EGFR mutant) = 58.41 ± 9.04 (EGFR wild type) = 60.35 ± 9.61	EGFR	NSCLC
Niu et al. (2023) [20]	268	NSCLC = 57.88 ± 10.81 Breast cancer = 53.71 ± 9.77	EGFR, Ki-67	NSCLC Breast cancer
Jiang et al. (2023) [21]	265	Training cohort (EGFR-21) = 61.91 ± 10.75 (EGFR 19) = 57.43 ± 9.34 (EGFR wild type) = 59.18 ± 9.99 Validation cohort (EGFR-21) = 61.54 ± 10.85 (EGFR 19) = 58.71 ± 9.72 (EGFR wild type) = 60.13 ± 7.38	EGFR	NSCLC
Cao et al. (2023) [22]	299	Training cohort (EGFR mutant) = 60.19 ± 9.99 (EGFR wildtype) = 59.28 ± 10.35 Validation cohort (internal) (EGFR mutant) = 61.15 ± 11.77 (EGFR wildtype) = 59.60 ± 10.42 (external) (EGFR mutant) = 60.17 ± 7.63 (EGFR wildtype) = 60.06 ± 6.78	EGFR	Lung adenocarcinoma

Table 1. *Cont.*

Study	Cohort Size	Mean Age ± SD	Receptor Analyzed	Primary Tumor Type
Zhang et al. (2024) [23]	110	Training cohort (High Ki-67) = 54.48 ± 9.76 (Low Ki-67) = 54.31 ± 9.06 (HER-2 positive) = 49.73 ± 11.82 (HER-2 negative) = 53.90 ± 8.70 Validation cohort (High Ki-67) = 52.50 ± 9.22 (Low Ki-67) = 49.06 ± 9.51 (HER-2 positive) = 49.01 ± 9.23 (HER-2 negative) = 56.57 ± 10.01	HER-2, Ki-67	Breast cancer
Fan et al. (2022) [24]	183	Training cohort (EGFR mutant) = 58.71 ± 9.34 (EGFR wildtype) = 57.14 ± 11.28 Validation cohort (internal) (EGFR mutant) = 58.14 ± 12.11 (EGFR wildtype) = 59.21 ± 8.95 (external) (EGFR mutant) = 59.06 ± 7.78 (EGFR wildtype) = 60.14 ± 5.61	EGFR	Lung adenocarcinoma
Fan et al. (2022) [25]	192	Training cohort (EGFR mutant) = 58.69 ± 10.31 (EGFR wildtype) = 58.37 ± 9.59 Validation cohort (internal) (EGFR mutant) = 61.5 ± 7.39 (EGFR wildtype) = 56.88 ± 10.10 (external) (EGFR mutant) = 63.06 ± 9.11 (EGFR wildtype) = 60.86 ± 6.64	EGFR	Lung adenocarcinoma
Cao et al. (2022) [26]	76	Training cohort (EGFR-21) = 61.12 ± 11.45 (EGFR 19) = 59.44 ± 8.65 Validation cohort (EGFR-21) = 62.69 ± 11.44 (EGFR 19) = 53.54 ± 10.70	EGFR	Lung adenocarcinoma
Fan et al. (2021) [27]	94	Training cohort (EGFR mutant) = 58.52 ± 9.84 (EGFR wild type) = 61.70 ± 10.75 Validation cohort (EGFR mutant) = 57.26 ± 9.43 (EGFR wild type) = 57.08 ± 10.66	EGFR	Lung adenocarcinoma
Ren et al. (2021) [28]	162	Training cohort (EGFR mutant) = 60.60 ± 10 (EGFR wild type) = 59.10 ± 11.20 Validation cohort (EGFR mutant) = 61.10 ± 9.29 (EGFR wild type) = 60.00 ± 6.64	EGFR	Lung adenocarcinoma

Concerning the status of the receptor analyzed (Table 1), one of the eleven studies reported AUC values for the established radiomics models for the training of the model (Table 2), eleven for internal validation (Table 3), and five studies reported AUC values for external validation (Table 4).

Table 2. The AUC and corresponding 95% confidence interval of studies reporting AUC values for the established radiomics models for the training of the model are presented. 1 = epidermal growth factor receptor.

Study	Output/Prediction	Best Performing Model	AUC	95% CI
Hu et al. (2024) [18]	EGFR ¹ mutation status	Radiomics (RS-SM-Com)	0.929	0.886–0.973

Table 3. The AUC and corresponding 95% confidence interval of studies reporting AUC values for the established radiomics models for internal validation of the model is presented. 1 = epidermal growth factor receptor, 2 = k-fold cross validation, 3 = split sample internal validation, 4 = EGFR¹-tyrosine kinase inhibitor, 5 = non-small cell lung cancer, 6 = breast cancer, 7 = convolutional block attention module, 8 = multi-resolution feature fusion mechanism, 9 = human epidermal growth factor receptor 2.

Study	Output/Prediction	Best Performing Model	AUC		95% CI	
Hu et al. (2024) [18]	EGFR ¹ mutation status	Radiomics (RS-SM-Com)	0.896		0.781–1.000	
Cao et al. (2023) [22]	EGFR ¹ mutation status	Radiomics (RS-Com-EGFR)	0.806 ²	0.745 ³	-	-
	EGFR ¹ Exon 19 mutation	Radiomics (RS-Com-Exon19)	0.872 ²	0.760 ³	-	-
	EGFR ¹ Exon 21 mutation	Radiomics (RS-Com-Exon21)	0.913 ²	0.799 ²	-	-
Cheng et al. (2023) [19]	EGFR ¹ mutation status	Radiomics (RS-Com-EGFR)	0.927 ²	0.812 ³	0.884–0.969 ²	0.709–0.916 ³
	Response to EGFR-TKI ⁴	Radiomics (RS-Com-TKI)	0.880 ²	0.798 ³	0.802–0.958 ²	0.649–0.946 ³
Jiang et al. (2023) [21]	EGFR ¹ mutation status	CM-EfNet (CBAM ⁷ and MFM ⁸ and EfficientNet v2)	0.866 ²	0.851 ³	0.800–0.916 ²	0.750–0.923 ³
	EGFR ¹ mutations in Exons 19 and 21	CM-EfNet (CBAM ⁷ and MFM ⁸ and EfficientNet v2)	0.760 ²	0.711 ³	0.656–0.846 ²	0.552–0.839 ³
Niu et al. (2023) [20]	Differentiating NSCLC ⁵ and BC ⁶ spinal metastasis	Radiomics—logistic regression (Ori-RS)	0.890 ²	0.881 ³	0.843–0.938 ²	0.810–0.953 ³
	EGFR ¹ mutation status	Radiomics—logistic regression (EGFR-RS)	0.793 ²	0.744 ³	0.703–0.833 ²	0.601–0.887 ³
	Ki-67 expression level	Radiomics—logistic regression (Ki-67-RS)	0.798 ²	0.738 ³	0.693–0.902 ²	0.554–0.921 ³
Zhang et al. (2024) [23]	Ki-67 level	Radiomics (RS-Ki-67)	0.812 ²	0.799 ³	0.710–0.914 ²	0.652–0.947 ³
	HER-2 ⁹ mutation status	Radiomics (RS-HER-2)	0.796 ²	0.705 ³	0.686–0.906 ²	0.506–0.904 ³
Cao et al. (2022) [26]	Differentiating Exon 19 and Exon 21 in EGFR ¹ mutation	Nomogram	0.901 ²	0.882 ³	0.783–0.967 ²	0.695–0.974 ³
Fan et al. (2022) [24]	EGFR ¹ mutation status	Radiomics (RS-EGFR)	0.851 ²	0.780 ³	0.774–0.921 ²	0.645–0.916 ³
	EGFR ¹ Exon 19 deletion	Radiomics (RS-19)	0.816 ²	0.789 ³	0.716–0.917 ²	0.636–0.942 ³
	EGFR ¹ Exon 21 mutation	Radiomics (RS-21)	0.814 ²	0.770 ³	0.714–0.914 ²	0.609–0.931 ³

Table 3. *Cont.*

Study	Output/Prediction	Best Performing Model	AUC		95% CI	
Fan et al. (2022) [25]	EGFR ¹ mutation status	Clinical-radiomics nomogram model (nomogram-EGFR)	0.849 ²	0.828 ³	0.776–0.922 ²	0.708–0.949 ³
	T790M mutation status	Clinical-radiomics nomogram model (nomogram-T790M)	0.842 ²	0.823 ³	0.717–0.927 ²	0.633–0.940 ³
Fan et al. (2021) [27]	EGFR ¹ mutation status	Radiomics (multi-regional radiomics signature)	0.879 ²	0.777 ³	0.766–0.947 ²	0.612–0.967 ³
Ren et al. (2021) [28]	EGFR ¹ mutation status	Combined rad score	0.886 ²	0.803 ³	0.826–0.947 ²	0.682–0.924 ³

Table 4. The AUC and corresponding 95% confidence interval of studies reporting AUC values for the established radiomics models for external validation of the model are presented. 1 = epidermal growth factor receptor, 7 = convolutional block attention module, 8 = multi-resolution feature fusion mechanism.

Study	Output/Prediction	Best Performing Model	AUC	95% CI
Hu et al. (2024) [18]	EGFR mutation status	Radiomics (RS-SM-Com)	0.865	0.731–0.998
Cao et al. (2023) [22]	EGFR mutation status	Radiomics (RS-Com-EGFR)	0.738	-
	EGFR Exon 19 mutation	Radiomics (RS-Com-Exon19)	0.825	-
	EGFR Exon 21 mutation	Radiomics (RS-Com-Exon21)	0.811	-
Jiang et al. (2023) [21]	EGFR mutation status	CM-EfNet (CBAM ⁷ and MFM ⁸ and EfficientNet v2)	0.764	0.615–0.914
	EGFR mutations in Exons 19 and 21	CM-EfNet (CBAM ⁷ and MFM ⁸ and EfficientNet v2)	0.687	0.476–0.897
Fan et al. (2022) [24]	EGFR mutation status	Radiomics (RS-EGFR)	0.807	0.595–0.938
	EGFR Exon 19 deletion	Radiomics (RS-19)	0.742	0.478–0.919
	EGFR Exon 21 mutation	Radiomics (RS-21)	0.792	0.530–0.946
Fan et al. (2022) [25]	EGFR ¹ mutation status	Clinical-radiomics nomogram model (nomogram-EGFR)	0.778	0.610–0.946
	T790M mutation status	Clinical-radiomics nomogram model (nomogram-T790M)	0.800	0.548–0.948

The weighted average AUC value among the eleven studies that reported AUC values and corresponding 95% confidence intervals for internal validation of the established models is 0.849 (95% CI: 0.835–0.863). Wherever 95% confidence intervals were not reported, the weighted average of the reported 95% confidence intervals was used.

The weighted average AUC value among the five studies that reported AUC values and corresponding 95% confidence intervals for external validation of the established models is 0.791 (95% CI: 0.738–0.844). Wherever 95% confidence intervals were not reported, the weighted average of the reported 95% confidence intervals was used.

4. Discussion

Here, the emerging and significant role of artificial intelligence (AI), particularly radiomics and machine learning (ML), in characterizing genetic and molecular signatures in spinal metastases (SMs) has been highlighted. In the eleven included studies, non-small cell lung cancer (NSCLC) and breast cancer were mainly focused on, with AI models developed that predict biomarkers such as EGFR, HER-2, Ki-67, and T790M. Overall, the models demonstrated promising performance with consistently high AUC values across training and validation cohorts. The ability of AI to analyze complex imaging patterns that may be invisible to clinicians underlines its promise in predicting spinal metastases characteristics. Radiomics—through conversion of medical images into higher-dimensional data—enables quantitative and reproducible analyses that can predict tumor biology non-invasively when combined with ML algorithms. This is invaluable in spinal metastases, where biopsy is often technically challenging and risky due to the anatomical proximity of critical neurovascular structures.

EGFR mutation status emerged as the most commonly predicted biomarker across the included studies, especially in spinal metastases from primary NSCLC. For instance, Hu et al. [18] achieved an AUC of 0.929 in training and 0.865 in external validation using CT and MRI-derived radiomics from both primary and metastatic lesions. Additionally, Cheng et al. [19] demonstrated that combining radiomics features from both primary lung tumors and spinal metastases enhanced prediction of EGFR mutations and response to EGFR-tyrosine kinase inhibitor therapy, with external validation AUCs > 0.80. These findings underline the ability of AI to effectively detect the distinct and clinically relevant molecular profiles that metastatic lesions carry.

Other biomarkers aside from EGFR, such as Ki-67 and HER-2, were also predicted successfully. Zhang et al. [23] used MRI-based features to forecast Ki-67 and HER-2 status in breast cancer spinal metastases, achieving AUCs near or above 0.80 in both training and external validation data sets. The high-performance metrics underscore the robustness and reproducibility of these AI tools, justifying their clinical utility for treatment stratification and prognostication. Importantly, external validation was present in most studies and enhances finding generalizability. Nonetheless, there were performance drop-offs between internal and external cohorts, which is typical in AI research and reflects variability in imaging protocols, scanner types, and patient populations across centers. Such variability proves that multicentric datasets and model standardization are needed to bridge the translational gap from research to real-world clinical application.

The inclusion of deep learning models, such as CM-EfNet by Jiang et al. [21], which integrates advanced modules like the convolutional block attention model (CBAM) and multi-resolution feature fusion, represents a crucial next step in enhancing prediction accuracy and computational efficiency. These models performed superiorly over traditional radiomics with external AUCs reaching up to 0.764 (95%CI: 0.615–0.914), highlighting the evolving sophistication of AI techniques.

Clinically, non-invasive prediction of actionable mutations in SMs may potentially facilitate personalized therapy without necessitating repeated tissue biopsies. This could be particularly relevant in tracking resistance mutations such as T790M in lung adenocarcinoma, as demonstrated in studies by Fan et al. [24,25] and Cao et al. [22], where AI-derived radiomics nomograms showed AUCs above 0.80.

The integration of AI models into routine clinical practice and patient care for SMs remains promising yet faces significant hurdles due to several healthcare and technological constraints. Challenges include a lack of standardized, high-quality datasets, limited interoperability with clinical systems, and concerns over model transparency, interpretability, and generalizability [29,30]. In addition, regulatory approval, general data protection regulations, and clinicians' acceptance and knowledge of AI models pose as roadblocks. Despite this, advances in explainable AI and collaborative data sharing efforts are gradually improving the feasibility of widespread clinical adoption and use [31].

Despite these encouraging results, certain limitations must be acknowledged. First, most studies had relatively small sample sizes (average cohort size: ~200 patients) and resultant overfitting can inflate performance metrics. Second, most data were from Chinese populations, questioning the models' applicability in other ethnic and geographic groups. Third, only a few studies incorporated clinical variables (e.g., smoking history or performance status), which could further enrich model performance and interpretability. Fourth, while PROBAST assessment indicated generally low to moderate risk of bias across domains, the heterogeneity in model architectures, radiomics feature sets, and validation strategies precluded direct comparison of model efficacy. Shared open-access datasets and consensus on performance reporting standards, which have been promoted in other areas of AI in oncology, would likely be beneficial.

5. Conclusions

AI models have shown good potential in predicting key molecular markers in spinal metastases with high accuracy, permitting a preferable non-invasive alternative to biopsies, for both identifying actionable mutations like EGFR and T790M and also in stratifying patients on the basis of biomarkers such as Ki-67 and HER-2 to provide more customized and timely clinical decision-making. However, to further integrate these key tools into daily practice, further large, multi-center datasets reflecting global diversity and differential epidemiology and demographics must be curated, along with standardized radiomics workflows, and multimodal AI models, handling clinical and genomic features, to enhance utility.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/bioengineering12080791/s1>, Supplementary Table S1. Search queries across five databases (PubMed, Scopus, Web of Science Advance, Cochrane and Embase (Ovid)) are shown; Supplementary Table S2. A summary of the studies focusing on genetic and molecular signatures in spinal metastases is presented; Supplementary Table S3. PROBAST Step 1—specifying the systematic review question; Supplementary Table S4. PROBAST Step 2—classifying the type of prediction model evaluation; Supplementary Figure S1. Risk of bias; Supplementary Figure S2. Applicability.

Author Contributions: Conceptualization, V.S.; methodology, V.S. and S.S.; formal analysis, V.S. and A.T.; investigation, V.S., S.S., Z.L. and P.H.; data curation, V.S., S.S. and A.T.; writing—original draft preparation, V.S.; writing—review and editing, V.S., S.S., A.T., Z.L., P.H., S.H., E.O.R.N., M.J.C., J.R. and A.D.; supervision, J.R. and A.D.; project administration, V.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to it being a systematic review of previously published data and did not involve human participants or the collection of new data.

Informed Consent Statement: Not applicable.

Data Availability Statement: All datasets used for this study are presented in the manuscript and its Supplementary Materials.

Acknowledgments: We'd like to acknowledge all the reviewers for their valuable feedback which helped improve the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sciubba, D.M.; Petteys, R.J.; Dekutoski, M.B.; Fisher, C.G.; Fehlings, M.G.; Ondra, S.L.; Rhines, L.D.; Gokaslan, Z.L. Diagnosis and management of metastatic spine disease. A review. *J. Neurosurg. Spine* **2010**, *13*, 94–108. [[CrossRef](#)] [[PubMed](#)]
2. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2020. *CA Cancer J. Clin.* **2020**, *70*, 7–30. [[CrossRef](#)] [[PubMed](#)]
3. Ziu, E.; Viswanathan, V.K.; Mesfin, F.B. Spinal Metastasis. In *StatPearls [Internet]*; StatPearls Publishing: Treasure Island, FL, USA, 2025. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK441950/> (accessed on 14 August 2023).
4. Van den Brande, R.; Cornips, E.M.; Peeters, M.; Ost, P.; Billiet, C.; Van de Kelft, E. Epidemiology of spinal metastases, metastatic epidural spinal cord compression and pathologic vertebral compression fractures in patients with solid tumors: A systematic review. *J. Bone Oncol.* **2022**, *35*, 100446. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
5. Liu, W.; Liang, Y.; Bian, C.; Jiang, L.; Zheng, G.; Dong, J. Gene expression profile analysis of the bone microenvironment in patients with spinal metastases. *Oncol. Lett.* **2018**, *15*, 61–68. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
6. Schroeder, C.; Campilan, B.; Leary, O.P.; Arditi, J.; Michles, M.J.; De La Garza Ramos, R.; Akinduro, O.O.; Gokaslan, Z.L.; Martinez Moreno, M.; Sullivan, P.L.Z. Therapeutic Opportunities for Biomarkers in Metastatic Spine Tumors. *Cancers* **2024**, *16*, 3152. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
7. Li, Y.; Zheng, Z.; Wang, L.; Han, L.; Du, Y.; Zhang, X.; Liu, X.; Xie, J. Association of mutation profiles with metastasis in patients with non-small cell lung cancer. *Front. Oncol.* **2024**, *14*, 1451576. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
8. Mok, T.S.; Wu, Y.L.; Thongprasert, S.; Yang, C.H.; Chu, D.T.; Saijo, N.; Sunpaweravong, P.; Han, B.; Margono, B.; Ichinose, Y.; et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **2009**, *361*, 947–957. [[CrossRef](#)] [[PubMed](#)]
9. Scholzen, T.; Gerdes, J. The Ki-67 protein: From the known and the unknown. *J. Cell Physiol.* **2000**, *182*, 311–322. [[CrossRef](#)] [[PubMed](#)]
10. Slamon, D.J.; Leyland-Jones, B.; Shak, S.; Fuchs, H.; Paton, V.; Bajamonde, A.; Fleming, T.; Eiermann, W.; Wolter, J.; Pegram, M.; et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* **2001**, *344*, 783–792. [[CrossRef](#)] [[PubMed](#)]
11. Ong, W.; Zhu, L.; Zhang, W.; Kuah, T.; Lim, D.S.W.; Low, X.Z.; Thian, Y.L.; Teo, E.C.; Tan, J.H.; Kumar, N.; et al. Application of Artificial Intelligence Methods for Imaging of Spinal Metastasis. *Cancers* **2022**, *14*, 4025. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
12. Kourou, K.; Exarchos, T.P.; Exarchos, K.P.; Karamouzis, M.V.; Fotiadis, D.I. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* **2014**, *13*, 8–17. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
13. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
14. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
15. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [[CrossRef](#)] [[PubMed](#)]
16. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
17. Wolff, R.F.; Moons, K.G.M.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S.; Reitsma, J.B.; Kleijnen, J.; Mallett, S.; PROBAST Group; et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann. Intern. Med.* **2019**, *170*, 51–58. [[CrossRef](#)] [[PubMed](#)]
18. Hu, Y.; Geng, Y.; Wang, H.; Chen, H.; Wang, Z.; Fu, L.; Huang, B.; Jiang, W. Improved Prediction of Epidermal Growth Factor Receptor Status by Combined Radiomics of Primary Nonsmall-Cell Lung Cancer and Distant Metastasis. *J. Comput. Assist. Tomogr.* **2024**, *48*, 780–788. [[CrossRef](#)] [[PubMed](#)]
19. Cheng, Y.; Wang, H.; Yuan, W.; Wang, H.; Zhu, Y.; Chen, H.; Jiang, W. Combined radiomics of primary tumour and bone metastasis improve the prediction of EGFR mutation status and response to EGFR-TKI therapy for NSCLC. *Phys. Med.* **2023**, *116*, 103177. [[CrossRef](#)] [[PubMed](#)]

20. Niu, S.; Zhang, H.; Wang, X.; Jiang, W. Radiomics of Spinal Metastases Originating from Primary Nonsmall Cell Lung Cancer or Breast Cancer and Ability to Predict Epidermal Growth Factor Receptor Mutation/Ki-67 Levels. *J. Comput. Assist. Tomogr.* **2023**, *47*, 643–649. [[CrossRef](#)] [[PubMed](#)]
21. Jiang, T.; Sun, X.; Dong, Y.; Guo, W.; Wang, H.; Yue, Z.; Luo, Y.; Jiang, X. Deep learning for preoperative prediction of the EGFR mutation and subtypes based on the MRI image of spinal metastasis from primary NSCLC. *Biomed. Signal Process. Control* **2023**, *79*, 104084. [[CrossRef](#)]
22. Cao, R.; Chen, H.; Wang, H.; Wang, Y.; Cui, E.N.; Jiang, W. Comprehensive analysis of prediction of the EGFR mutation and subtypes based on the spinal metastasis from primary lung adenocarcinoma. *Front. Oncol.* **2023**, *13*, 1154327. [[CrossRef](#)] [[PubMed](#)]
23. Zhang, H.; Niu, S.; Chen, H.; Wang, L.; Wang, X.; Wu, Y.; Shi, J.; Li, Z.; Hu, Y.; Yang, Z.; et al. Radiomics signatures for predicting the Ki-67 level and HER-2 status based on bone metastasis from primary breast cancer. *Front. Cell Dev. Biol.* **2024**, *11*, 1220320. [[CrossRef](#)] [[PubMed](#)]
24. Fan, Y.; Dong, Y.; Sun, X.; Wang, H.; Zhao, P.; Wang, H.; Jiang, X. Development and validation of MRI-based radiomics signatures as new markers for preoperative assessment of EGFR mutation and subtypes from bone metastases. *BMC Cancer* **2022**, *22*, 889. [[CrossRef](#)] [[PubMed](#)]
25. Fan, Y.; Dong, Y.; Wang, H.; Wang, H.; Sun, X.; Wang, X.; Zhao, P.; Luo, Y.; Jiang, X. Development and externally validate MRI-based nomogram to assess EGFR and T790M mutations in patients with metastatic lung adenocarcinoma. *Eur. Radiol.* **2022**, *32*, 6739–6751. [[CrossRef](#)] [[PubMed](#)]
26. Cao, R.; Dong, Y.; Wang, X.; Ren, M.; Wang, X.; Zhao, N.; Yu, T.; Zhang, L.; Luo, Y.; Cui, E.N.; et al. MRI-Based Radiomics Nomogram as a Potential Biomarker to Predict the EGFR Mutations in Exon 19 and 21 Based on Thoracic Spinal Metastases in Lung Adenocarcinoma. *Acad. Radiol.* **2022**, *29*, e9–e17. [[CrossRef](#)] [[PubMed](#)]
27. Fan, Y.; Dong, Y.; Yang, H.; Chen, H.; Yu, Y.; Wang, X.; Wang, X.; Yu, T.; Luo, Y.; Jiang, X. Subregional radiomics analysis for the detection of the EGFR mutation on thoracic spinal metastases from lung cancer. *Phys. Med. Biol.* **2021**, *66*, 215008. [[CrossRef](#)] [[PubMed](#)]
28. Ren, M.; Yang, H.; Lai, Q.; Shi, D.; Liu, G.; Shuang, X.; Su, J.; Xie, L.; Dong, Y.; Jiang, X. MRI-based radiomics analysis for predicting the EGFR mutation based on thoracic spinal metastases in lung adenocarcinoma patients. *Med. Phys.* **2021**, *48*, 5142–5151. [[CrossRef](#)] [[PubMed](#)]
29. Kelly, C.J.; Karthikesalingam, A.; Suleyman, M.; Corrado, G.; King, D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* **2019**, *17*, 195. [[CrossRef](#)] [[PubMed](#)]
30. Topol, E.J. High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* **2019**, *25*, 44–56. [[CrossRef](#)] [[PubMed](#)]
31. Jiang, F.; Jiang, Y.; Zhi, H.; Dong, Y.; Li, H.; Ma, S.; Wang, Y.; Dong, Q.; Shen, H.; Wang, Y. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc. Neurol.* **2017**, *2*, 230–243. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.