Buletin SK 10(1)(2026): 67-77
https://doi.org/10.17576/bsk-2026-10(1)-09 ISSN 2250-1852

Ulasan Artikel/*Review Article*

Machine Learning in Gamma Knife Radiosurgery: Current Applications, Challenges, and Future Perspectives

Pembelajaran Mesin dalam Bedah Radio Gamma Knife: Aplikasi Semasa, Cabaran, dan Perspektif Masa Hadapan

ABDULRAHMAN UMARU[1,2], HANANI ABDUL MANAN[3], RAMESH KUMAR ATHI KUMAR[4*], SITI KHADIJAH HAMSAN[5], NOORAZRUL YAHYA[1]

[1]Diagnostic Imaging and Radiotherapy, CODTIS, Faculty of Health Sciences, National University of Malaysia, Jalan Raja Muda Aziz 50300 Kuala Lumpur, Malaysia.

[2]Department of medical radiography, Ahmadu Bello University, Zaria, Kaduna, Nigeria.

[3]Department of Radiology, Faculty of Medicine, Universiti kebangsaan Malaysia, Jalan yaacob Latif, Bandar Tun Razak, Kuala Lumpur Malaysia

[4]Neurosurgical Unit, Department of Surgery, Hospital Canselor Tuanku Muhriz Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, Cheras Kuala Lumpur, Malaysia.

[5]Department of Radiotherapy and Oncology, Hospital Canselor Tuanku Muhriz Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, Cheras Kuala Lumpur, Malaysia.

*Corresponding author; email: rameshkumar71@hotmail.com

Received date: 5 April 2026
Revised date: 16 June 2026
Accepted date: 25 June 2026

ABSTRACT

Gamma Knife radiosurgery (GKRS) is a well-established treatment modality for intracranial lesions, including brain metastases, benign tumours, and vascular malformations. Despite high local control rates, treatment outcomes remain heterogeneous, and predicting tumor response and radiation-induced toxicity remains a significant clinical challenge. Conventional prognostic models rely mainly on clinical and dosimetric parameters, which may not adequately capture tumour heterogeneity and complex biological interactions. Machine learning (ML) has emerged as a promising approach for analysing high-dimensional clinical and imaging data to improve predictive modeling in radiosurgery. This narrative review examines current applications of ML in predicting radiotherapy outcomes following GKRS, including tumour control, survival, and radiation-induced toxicity. Relevant literature was identified through searches of PubMed, Scopus, and Web of Science using keywords related to GKRS, machine learning, artificial intelligence, and radiomics. Existing studies show that ML models integrating radiomic, clinical, and dosimetric features outperform single-modality approaches in predicting key outcomes such as local tumour control, radionecrosis, and overall survival. Radiomics derived from pretreatment MRI demonstrates potential for identifying imaging biomarkers of treatment response, while multimodal models improve risk stratification. However, current evidence is limited by small, predominantly single-centre retrospective datasets, heterogeneity in imaging and feature extraction methods, and lack of robust external validation, which constrain reproducibility and clinical applicability. Addressing these challenges will require standardised radiomics workflows, adherence to quality frameworks, and multicentre prospective validation to support reliable clinical translation.

Keywords: Gamma Knife radiosurgery, Machine learning, Radiomics, Brain metastases Outcome prediction.

ABSTRAK

Radiosurgeri Gamma Knife (GKRS) merupakan satu kaedah rawatan yang telah terbukti berkesan untuk lesi intrakranial, termasuk metastasis otak, tumor jinak, dan malformasi vaskular. Walaupun kadar kawalan setempat adalah tinggi, hasil rawatan masih menunjukkan variasi, dan peramalan tindak balas tumor serta ketoksikan akibat

Copyright © 2026 Faculty of Health Science UKM. All right reserved

radiasi kekal sebagai cabaran klinikal yang signifikan. Model prognostik konvensional kebanyakannya bergantung kepada parameter klinikal dan dosimetri, yang mungkin tidak mencerminkan sepenuhnya heterogeniti tumor dan interaksi biologi yang kompleks. Pembelajaran mesin (ML) telah muncul sebagai pendekatan yang berpotensi dalam menganalisis data klinikal dan pengimejan berdimensi tinggi bagi meningkatkan pemodelan ramalan dalam radiosurgeri. Ulasan naratif ini menilai aplikasi semasa ML dalam meramalkan hasil radioterapi selepas GKRS, termasuk kawalan tumor, kelangsungan hidup, dan ketoksikan akibat radiasi. Literatur yang berkaitan dikenal pasti melalui carian dalam pangkalan data utama seperti PubMed, Scopus, dan Web of Science menggunakan kata kunci berkaitan GKRS, pembelajaran mesin, kecerdasan buatan, dan radiomik. Kajian sedia ada menunjukkan bahawa model ML yang mengintegrasikan ciri radiomik, klinikal, dan dosimetri mengatasi pendekatan satu sumber dalam meramalkan hasil utama seperti kawalan tumor setempat, radionekrosis, dan kelangsungan hidup keseluruhan. Radiomik yang diperoleh daripada MRI pra-rawatan menunjukkan potensi dalam mengenal pasti biomarker pengimejan bagi tindak balas rawatan, manakala model multimodal meningkatkan stratifikasi risiko. Walau bagaimanapun, bukti semasa terhad oleh saiz sampel yang kecil, kebanyakannya daripada kajian retrospektif satu pusat, variasi dalam protokol pengimejan dan kaedah pengekstrakan ciri, serta kekurangan pengesahan luaran yang kukuh. Kekangan ini menjejaskan kebolehulangan dan kebolehgunaan klinikal. Oleh itu, penyeragaman aliran kerja radiomik, pematuhan kepada kerangka kualiti, serta pengesahan prospektif pelbagai pusat diperlukan bagi menyokong pelaksanaan klinikal yang lebih boleh dipercayai.

Kata Kunci: Radiosurgeri Gamma Knife, Pembelajaran mesin, Radiomik, Metastasis otak, Ramalan hasil

INTRODUCTION

Gamma Knife radiosurgery (GKRS) has become an established treatment modality for a wide range of intracranial pathologies, including brain metastases, arteriovenous malformations, vestibular schwannomas, and meningiomas (Alazawy et al. 2024). By delivering highly focused radiation beams with submillimetre precision, GKRS enables effective tumour control while minimizing damage to surrounding normal tissues (Du et al. 2023).

Despite its clinical success, predicting treatment outcomes following radiosurgery remains challenging. Patient outcomes depend on multiple interacting factors, including tumour characteristics, treatment dose, and patient clinical variables (Xia et al. 2025). Conventional statistical models often struggle to capture these complex nonlinear relationships, which has motivated increasing interest in artificial intelligence (AI) and machine learning (ML) approaches (Yahya et al. 2016).

Machine learning techniques have shown promise in analysing large and complex datasets, enabling the identification of patterns that may not be readily detectable using traditional statistical approaches (Kufel et al. 2023). In the context of radiosurgery, ML methods have been applied to predict treatment response, tumour control, survival outcomes, radiation-induced toxicity, tumour recurrence, and treatment planning parameters (Tan et al. 2023). This narrative review summarises current evidence on the application of machine learning techniques in Gamma Knife radiosurgery. It aims to provide a comprehensive overview of current approaches, highlight methodological challenges, and discuss potential clinical applications of machine learning in radiosurgery.

LITERATURE SEARCH AND ARTICLE SELECTION

A structured literature search was conducted using PubMed, Scopus, and Web of Science to identify relevant studies on machine learning applications in Gamma Knife radiosurgery. The search covered articles published up to December 2025. The main search terms included combinations of “Gamma Knife radiosurgery”, “stereotactic radiosurgery”, “machine learning”, “artificial intelligence”, “radiomics”, “deep learning”, “brain metastases”, “radionecrosis”, “treatment response”, “toxicity prediction”, “survival prediction”, and “treatment planning”.

Articles were considered relevant if they discussed the application of machine learning, artificial intelligence, radiomics, or deep learning in Gamma Knife radiosurgery or closely related stereotactic radiosurgery settings. Priority was given to studies addressing outcome prediction, local tumour control, survival, radionecrosis, treatment-related toxicity, treatment planning, model validation, radiomics workflow, and clinical translation. Only articles published in English were considered. Additional relevant publications were identified by manually screening the reference lists of selected articles. Studies not related to radiosurgery, articles with no relevance to machine learning or radiomics, and publications focused solely on conventional radiotherapy without radiosurgical relevance were excluded.

MACHINE LEARNING IN MEDICAL IMAGING AND RADIOTHERAPY

Machine learning has become increasingly integrated into medical imaging and radiotherapy research due to its ability to process high-

dimensional data and extract meaningful patterns. In radiotherapy, ML algorithms have been applied to tasks such as image segmentation, treatment planning optimisation, outcome prediction, and toxicity modelling (Kufel et al. 2023).

Radiomics has emerged as a particularly important application of ML in medical imaging. It involves the extraction of quantitative imaging features from medical images, which can capture tumour shape, texture, and intensity characteristics that may reflect underlying biological processes (Umaru, Abdul Manan, et al. 2025). These features can then be combined with clinical variables to develop predictive models that assist in treatment decision-making.

RADIOMIC FEATURE CATEGORIES AND SOURCES OF VARIABILITY

Radiomic features can generally be grouped into several categories. First-order features describe the distribution of voxel intensity values within the segmented region of interest without considering spatial relationships between voxels. These include measures such as mean intensity, median, standard deviation, skewness, kurtosis, entropy, and energy. Shape-based features describe the three-dimensional geometry of the lesion, including volume, surface area, compactness, sphericity, elongation, and irregularity. Texture features evaluate spatial relationships between neighbouring voxels and are commonly derived from matrices such as the grey-level co-occurrence matrix, grey-level run-length matrix, grey-level size-zone matrix, and neighbouring grey-tone difference matrix. These features may capture intratumour heterogeneity, which is particularly relevant in brain metastases and other intracranial lesions treated with Gamma Knife radiosurgery. Higher-order features are obtained after applying image filters or mathematical transformations, such as wavelet decomposition, Laplacian of Gaussian filtering, or intensity discretisation, to highlight specific spatial patterns that may not be visible on conventional imaging (Larue et al. 2017; Zwanenburg et al. 2016).

Despite their potential value, radiomic features are sensitive to several sources of variability. These include differences in MRI acquisition parameters, scanner manufacturers, magnetic field strength, slice thickness, voxel size, reconstruction settings, image normalisation, segmentation methods, and feature extraction software. Manual segmentation may introduce inter-observer variability, whereas automatic and semi-automatic segmentation approaches may be affected by algorithm performance and image quality. Preprocessing steps such as resampling, grey-level discretisation, filtering, and normalisation can also substantially influence extracted feature values (Traverso et al. 2018; Umaru, Manan, Kumar, Hamsan & Yahya 2025). Therefore, standardised imaging protocols, reproducible segmentation procedures, harmonised preprocessing pipelines, and transparent reporting of feature extraction methods are essential to improve the reliability and reproducibility of radiomics-based machine learning models in radiosurgery.

In radiosurgery, radiomics and ML approaches have been increasingly explored to improve the prediction of treatment response and complications. By integrating imaging features with clinical and dosimetric data, ML models may provide a more comprehensive understanding of tumour behaviour and treatment outcomes (Umaru, Manan, Kumar, Hamsan, Tan, et al. 2025).

MACHINE LEARNING WORKFLOW IN GAMMA KNIFE RADIOSURGERY

The development of machine learning models in radiosurgery typically follows several key steps, including data acquisition, image preprocessing, feature extraction, model training, and validation (Tulum 2023).

First, clinical and imaging data are retrieved from patients undergoing Gamma Knife radiosurgery. Imaging modalities commonly used include magnetic resonance imaging (MRI), computed tomography (CT), and occasionally functional imaging techniques such as positron emission tomography (Nowakowski et al. 2022). Next, is image preprocessing is performed to ensure data consistency. This may include image normalization, resampling to uniform voxel sizes, noise reduction, and correction for imaging artifacts (Meißner et al. 2023; Umaru, Manan, Kumar, Hamsan & Yahya 2025). Tumour segmentation is then carried out to define the region of interest from which features can be extracted.

Following segmentation, radiomic features are extracted from the tumour region and prepared for model development. The reliability of these features depends strongly on image acquisition parameters, segmentation consistency, preprocessing methods, and the software used for feature extraction. (Tulum 2023).

Once features have been extracted, machine learning algorithms are used to develop predictive models. Common algorithms used in radiosurgery studies include logistic regression, LASSO, support vector machines, random forests, gradient boosting methods, and neural networks (Cao et al. 2025; Liao et al. 2023). To provide a clearer comparison of these modelling approaches, Table 1 summarises commonly used machine learning algorithms in radiomics and Gamma Knife radiosurgery research, including their main characteristics, advantages, limitations, and potential applications.

For effective clinical integration, model validation is a critical step in the ML workflow. Internal validation techniques such as cross-

validation are often used during model development, while external validation using independent datasets is necessary to evaluate generalisability (Lambin et al. 2017) Figure 1 showed the radiomics workflow from image acquisition to model validation.

Although machine learning workflows follow these structured steps, the reliability of predictive models in radiotherapy strongly depends on data quality, feature robustness, and rigorous external validation. Failure to address these factors may lead to model overfitting and limited clinical applicability.

FEATURE SELECTION AND MODEL DEVELOPMENT CONSIDERATIONS

Feature selection is an important step in radiomics-based machine learning because the number of extracted imaging features is often much larger than the number of patients included in radiosurgery studies. Without appropriate feature reduction, models may identify unstable or redundant predictors, leading to overfitting and poor performance on independent datasets. Common feature selection methods include least absolute shrinkage and selection operator (LASSO), minimum redundancy maximum relevance (mRMR), and recursive feature elimination (RFE).

LASSO applies regularisation to shrink less informative feature coefficients toward zero, thereby retaining a smaller set of relevant predictors. The mRMR method selects features that are strongly associated with the outcome while reducing redundancy among selected variables. RFE works by repeatedly training a model, ranking features according to their contribution, and removing the least important features until an optimal subset is obtained (Perniciano et al. 2025; Zhang et al. 2023).

TABLE 1 Comparison of commonly used machine learning algorithms in radiomics and Gamma Knife radiosurgery research

Algorithm	**Main characteristics**	**Advantages**	**Limitations**	**Potential application in GKRS**
Logistic regression	A linear classification model that estimates the probability of an outcome based on selected predictors.	Simple, interpretable, computationally efficient, and useful as a baseline model.	Limited ability to model complex non-linear relationships unless interaction terms or transformations are included.	Prediction of binary outcomes such as tumour control, recurrence, radionecrosis, or survival status.
Ridge and elastic net regression	Regularised regression methods that reduce model complexity by penalising large coefficients. Elastic net combines L1 and L2 penalties.	Useful when predictors are correlated; reduces overfitting and improves model stability.	Requires careful tuning of regularisation parameters and may still underperform when relationships are highly non-linear.	Development of stable radiomics-based prediction models using correlated imaging features.
LASSO regression	A regularised regression method that performs coefficient shrinkage and feature selection.	Useful for high-dimensional radiomics data; reduces overfitting by selecting the most relevant features.	May select unstable features when predictors are highly correlated; performance depends on regularisation strength.	Radiomic feature selection and development of interpretable prediction models.
Support vector machine	A supervised learning algorithm that separates classes using an optimal hyperplane; kernel functions allow non-linear classification.	Effective in small- to medium-sized datasets and high-dimensional feature spaces.	Less interpretable than linear models; requires careful tuning of kernel and hyperparameters.	Classification of treatment response, local failure, or radionecrosis risk.

k-nearest neighbours	A distance-based algorithm that classifies a case according to the most similar examples in the training data.	Simple and intuitive; can capture local patterns in data.	Sensitive to feature scaling, noise, irrelevant variables, and class imbalance.	Exploratory classification of treatment response or toxicity risk in small datasets.
Decision tree	A tree-based model that splits data into decision rules based on predictor variables.	Easy to interpret and can model non-linear relationships.	Prone to overfitting if not pruned or validated properly.	Identification of clinical, dosimetric, or radiomic decision rules for outcome prediction.
Random forest	An ensemble method that combines multiple decision trees using bootstrap sampling and feature randomness.	Handles non-linear relationships, reduces variance, and provides feature importance estimates.	Can still overfit noisy data and may be less transparent than simpler models.	Prediction of treatment outcomes using combined clinical, dosimetric, and radiomic features.
Gradient boosting methods/XG Boost	Ensemble models that sequentially build decision trees to correct previous prediction errors.	High predictive performance and ability to capture complex feature interactions.	Sensitive to hyperparameter tuning and prone to overfitting if not carefully validated.	Risk stratification, survival prediction, treatment response prediction, and multimodal outcome modelling.
Naïve Bayes	A probabilistic classifier based on Bayes' theorem and an assumption of feature independence.	Computationally efficient and useful for small datasets.	The independence assumption is often unrealistic in radiomics because many features are correlated.	Preliminary classification tasks or baseline comparison in outcome prediction studies.
Artificial neural networks/deep learning	Models composed of interconnected layers that learn complex non-linear patterns from imaging or tabular data.	Suitable for complex image-based learning and automatic feature representation.	Requires larger datasets, has limited interpretability, and may be difficult to validate externally.	Automated image analysis, segmentation, treatment response prediction, and survival modelling.
Cox regression and survival-based models	Models designed to analyse time-to-event outcomes while accounting for censored data.	Appropriate for survival analysis and time-dependent endpoints.	Conventional Cox models assume proportional hazards and may not capture complex non-linear relationships.	Prediction of overall survival, progression-free survival, and time to local failure after GKRS.
Random survival forest	An ensemble survival model that extends random forest methods to censored time-to-event data.	Captures non-linear survival patterns and interactions without requiring proportional hazards assumptions.	Requires adequate sample size and external validation; interpretation may be more complex than Cox regression.	Survival prediction and risk stratification in patients treated with GKRS.

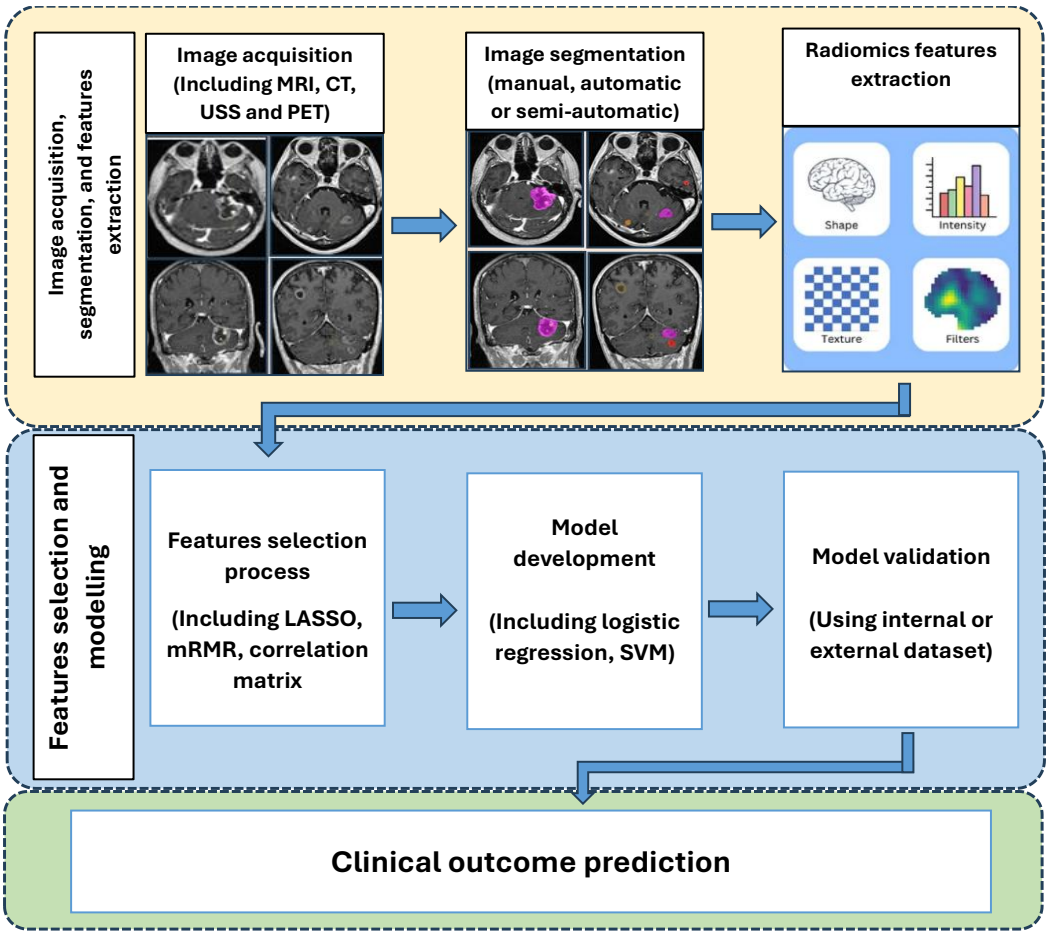

FIGURE 1. Radiomics and machine learning workflow in Gamma Knife radiosurgery, showing image acquisition, image segmentation, radiomic feature extraction, feature selection, machine learning model development, model validation, and clinical outcome prediction. Abbreviations: MRI, magnetic resonance imaging; CT, computed tomography; USS, ultrasonography; PET, positron emission tomography; ROI, region of interest; LASSO, least absolute shrinkage and selection operator; mRMR, minimum redundancy maximum relevance; SVM, support vector machine; NN, neural network.

Several model development challenges must also be considered in Gamma Knife radiosurgery studies. Overfitting is a major concern, particularly when small retrospective datasets are used with high-dimensional radiomic features. To reduce this risk, studies should apply appropriate feature selection, cross-validation, regularisation, and, where possible, independent external validation. Class imbalance is another important issue, especially when the outcome of interest is uncommon, such as radionecrosis or local failure. If not addressed, models may favour the majority class and produce misleadingly high accuracy while performing poorly for clinically important minority outcomes. Approaches such as stratified sampling, class weighting, resampling methods, and reporting sensitivity, specificity, area under the receiver operating characteristic curve, precision-recall metrics, and calibration can provide a more balanced assessment of model performance. Ultimately, model generalisability remains one of the most important requirements for clinical translation, and predictive models should be tested across independent cohorts, different scanners, imaging protocols, and institutional settings before routine clinical use (D'Amico et al. 2019; Staartjes & Schröder 2018).

APPLICATIONS OF MACHINE LEARNING IN GAMMA KNIFE RADIOSURGERY

Machine learning has been applied to several key clinical tasks in Gamma Knife radiosurgery, including outcome prediction, toxicity assessment, and survival modeling.

One of the most widely studied applications is the prediction of local tumour control following radiosurgery. ML models trained on radiomic features extracted from pretreatment MRI scans have demonstrated the ability to identify imaging biomarkers associated with treatment response (Feng et al. 2022; Horvat et al. 2022).

Another important application is the prediction of radiation-induced complications, such as radionecrosis. Early identification of patients at risk of developing radionecrosis may allow clinicians to adjust treatment plans or implement closer follow-up strategies (Hettal et al. 2020; Peng et al. 2018). Machine learning models have also been used to predict overall survival in patients with brain metastases treated with radiosurgery. By integrating imaging features with dosimetric and clinical variables such as age, performance status, and primary tumour type, these models may help stratify patients according to prognosis and guide treatment decisions (D'Amico et al. 2020; Della Seta et al. 2019).

Several studies have demonstrated that combining radiomic features with clinical and dosimetric variables improves predictive performance compared with single-modality models. Multimodal models may capture complementary information, such as tumour morphology, microenvironment heterogeneity, and patient-specific clinical characteristics (Umaru et al. 2026). However, variability in imaging protocols, segmentation strategies, and feature extraction pipelines remains a major source of heterogeneity across studies, which may limit reproducibility and generalisability of radiomics-based prediction models.

MACHINE LEARNING IN TREATMENT PLANNING

Machine learning is also being explored to improve treatment planning in Gamma Knife radiosurgery. Predictive models can assist clinicians in estimating treatment outcomes based on different planning parameters, thereby facilitating more personalised treatment strategies (Munai et al. 2024).

For example, ML models may help identify optimal dose distributions that maximise tumor control while minimising the risk of toxicity. In addition, predictive algorithms can assist in determining appropriate prescription doses based on tumor size, location, and patient-specific factors (Umaru, Manan, Kumar, Hamsan, Tan, et al. 2025). Beyond outcome prediction, machine learning is increasingly being explored for treatment planning optimisation in Gamma Knife radiosurgery. Predictive algorithms can assist clinicians in estimating dose distributions, evaluating plan quality, and identifying optimal treatment parameters before plan delivery. Knowledge-based planning approaches trained on historical treatment data have shown potential for predicting achievable dose volume histograms and improving planning efficiency. In addition, machine learning methods may support automated target segmentation and organ-at-risk delineation, which are critical steps in radiosurgery planning. Although these approaches remain largely investigational, they represent a promising direction for improving planning consistency and workflow efficiency in radiosurgical practice.

CHALLENGES IN MACHINE LEARNING FOR GAMMA KNIFE RADIOSURGERY

Despite the growing interest in machine learning (ML) applications in Gamma Knife radiosurgery, several methodological and practical challenges continue to limit clinical translation.

One of the most frequently reported limitations is the restricted availability of large, high-quality datasets. Many published studies rely on relatively small retrospective cohorts derived from single institutions, which may introduce selection bias and limit the statistical power of predictive models. Small datasets are particularly problematic in radiomics research, where the number of extracted imaging features can substantially exceed the number of patients, increasing the risk of overfitting and reducing model generalisability (Umaru, Abdul Manan, et al. 2025). Multicentre collaborations and data-sharing initiatives are therefore essential to improve the robustness and external validity of ML models. In addition, class imbalance should be explicitly addressed in model development because clinically important outcomes such as radionecrosis, tumour recurrence, or treatment failure may occur less frequently than favourable outcomes, potentially biasing model predictions towards the majority class.

Another major challenge is the lack of standardisation in radiomics workflows. Radiomic feature values can be highly sensitive to variations in imaging acquisition parameters, reconstruction settings, segmentation techniques, and preprocessing pipelines (Umaru, Manan, Kumar, Hamsan & Yahya 2025). Even minor differences in MRI protocols or voxel resolution can significantly alter extracted features, thereby affecting model performance. To address these issues, several initiatives have been proposed to improve methodological consistency. The Image Biomarker Standardisation Initiative (IBSI) provides standardised definitions for radiomic features and feature extraction procedures, while the Radiomics Quality Score (RQS) offers a structured framework for evaluating methodological rigor and reproducibility in radiomics studies (Lambin et al. 2017; Zwanenburg et al. 2016). In addition, reporting guidelines such as TRIPOD-AI have been introduced to promote transparency in the development and validation of predictive models (Collins et al. 2024). However, despite the availability of these frameworks, many published studies still employ heterogeneous methodologies,

highlighting the need for greater adherence to standardised protocols.

Closely related to these issues is the challenge of harmonising imaging and radiomic data across institutions. Differences in scanner manufacturers, acquisition protocols, and image reconstruction algorithms can introduce systematic variability in radiomic features. Harmonisation techniques, such as statistical normalisation methods and approaches like ComBat harmonisation, have been proposed to mitigate these effects and improve the reproducibility of multicentre studies. Nevertheless, consistent implementation of such techniques remains limited.

Finally, model interpretability and transparency represent important barriers to clinical adoption. Many machine learning algorithms particularly deep learning architectures operate as complex "black box" systems, making it difficult to understand how individual features contribute to predictions. This lack of interpretability may reduce clinician trust and hinder integration into clinical workflows. Explainable artificial intelligence approaches, including feature importance analysis, SHAP-based interpretation, and visualisation techniques, are increasingly being explored to address this limitation and improve the transparency of ML decisions (Sim et al. 2023; Yoon et al. 2022).

Collectively, addressing these methodological and translational challenges will be essential for enabling the reliable integration of machine learning models into clinical radiosurgery practice.

LIMITATIONS OF THE STUDY

This narrative review has several limitations. First, the evidence synthesised is largely derived from retrospective, single-centre studies with relatively small sample sizes, which may introduce selection bias and limit generalisability. Second, there is considerable methodological heterogeneity across studies, particularly in imaging protocols, segmentation approaches, and radiomic feature extraction pipelines, making comparison and reproducibility challenging.

Although standardisation initiatives such as the Image Biomarker Standardisation Initiative (IBSI), Radiomics Quality Score (RQS), and reporting guidelines like TRIPOD-AI have been developed to improve consistency, their application remains inconsistent, and many studies continue to adopt varying analytical approaches. In addition, most models lack robust external validation, limiting confidence in their performance across different clinical settings.

Furthermore, as a narrative review, this study does not follow a formal systematic review or meta-analysis framework, which may introduce selection bias in the included literature. Finally, the predominance of complex "black-box" machine learning models with limited interpretability remains a broader limitation within the field, potentially hindering clinical adoption.

FUTURE DIRECTIONS

Future research in machine learning for Gamma Knife radiosurgery should prioritize the development of large, high-quality multicentre datasets. Collaborative data sharing will be essential to improve model generalisability and reduce overfitting, particularly in radiomics-based studies where feature dimensionality often exceeds sample size. In parallel, the adoption of standardised imaging protocols and harmonised radiomics pipelines will be critical to enhance reproducibility across institutions.

The integration of multimodal data, including advanced imaging, genomic information, clinical variables, and treatment parameters, represents another key direction. The emerging field of radiogenomics, which links imaging phenotypes with underlying molecular characteristics, may provide deeper insight into tumour biology and radiosensitivity, thereby improving predictive accuracy and enabling more personalised treatment strategies.

Advances in deep learning may further strengthen predictive modelling by automatically capturing complex spatial patterns from imaging data. When applied to large, well-annotated datasets, these approaches may improve prediction of treatment response, radionecrosis risk, and survival outcomes.

For clinical translation, however, machine learning models must demonstrate not only high predictive performance but also interpretability, robustness, and integration into existing workflows. Explainable AI techniques may enhance clinician trust, while prospective validation studies and real-world clinical trials will be essential to determine their true clinical impact.

CONCLUSION

Machine learning is emerging as a powerful tool for advancing precision radiosurgery by enabling improved prediction of tumour control, treatment toxicity, and survival outcomes.

By integrating imaging, radiomic, clinical, and dosimetric data, these models offer the potential to support more personalised treatment strategies for patients undergoing Gamma Knife radiosurgery.

However, significant challenges remain, including limited dataset sizes, variability in radiomics methodologies, and the need for rigorous external validation. Future research should focus on multicentre collaborations, standardised data pipelines, and prospective validation studies to

ensure that machine learning models can be safely and effectively translated into clinical practice.

REFERENCE

Alazawy, N. M. A., Almusawi, M. S., Alabedi, H. H., Faraj, M. K., Ahmad, R. & Albosaabar, M. H. 2024. Gamma knife versus volumetric arc modulated therapy in a linear accelerator in treatment of multiple brain metastasis: Literature review. *Journal of Pioneering Medical Sciences* 13:93-100. 10.61091/jpms202413213

Cao, P., Jia, X., Wang, X., Fan, L., Chen, Z., Zhao, Y., Zhu, J. & Wen, Q. 2025. Deep learning radiomics for the prediction of epidermal growth factor receptor mutation status based on MRI in brain metastasis from lung adenocarcinoma patients. *BMC cancer* 25(1): 443. 10.1186/s12885-025-13823-8.

Collins, G. S., Moons, K. G., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B. & Van Smeden, M. 2024. TRIPOD+ AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *bmj* 385. 078378. https://doi.org/10.1136/bmj-2023-078378.

D'amico, N. C., Merone, M., Sicilia, R., Cordelli, E., D'antoni, F., Zanetti, I. B., Valbusa, G., Grossi, E., Beltramo, G. & Fazzini, D. 2019. Tackling imbalance radiomics in acoustic neuroma. *International Journal of Data Mining and Bioinformatics* 22(4): 365-388. https://doi.org/10.1504/ijdmb.2019.101396.

D'amico, N. C., Sicilia, R., Cordelli, E., Tronchin, L., Greco, C., Fiore, M., Carnevale, A., Iannello, G., Ramella, S. & Soda, P. 2020. Radiomics-based prediction of overall survival in lung cancer using different volumes-of-interest. *Applied Sciences (Switzerland)* 10(18): 6425. https://doi.org/10.3390/app10186425.

Della Seta, M., Collettini, F., Chapiro, J., Angelidis, A., Engeling, F., Hamm, B. & Kaul, D. 2019. A 3D quantitative imaging biomarker in pre-treatment MRI predicts overall survival after stereotactic radiation therapy of patients with a singular brain metastasis. *Acta Radiologica* 60(11):1496-1503. 10.1177/0284185119831692

Du, P., Liu, X., Xiang, R., Lv, K., Chen, H. Y., Liu, W. F., Cao, A. H., Chen, L., Wang, X. F., Yu, T. G., Ding, J., Li, W. C., Li, J., Li, Y. X., Yu, Z. K., Zhu, L., Liu, J. & Geng, D. Y. 2023. Development and validation of a radiomics-based prediction pipeline for the response to stereotactic radiosurgery therapy in brain metastases. *European Radiology* 33(12): 8925-8935. 10.1007/s00330-023-09930-4.

Feng, A., Yuan, P., Huang, T., Li, L. & Lyu, J. 2022. Distinguishing Tumor Recurrence From Radiation Necrosis in Treated Glioblastoma Using Multiparametric MRI. *Academic Radiology* 29(9): 1320-1331. 10.1016/j.acra.2021.11.008.

Hettal, L., Stefani, A., Salleron, J., Courrech, F., Behm-Ansmant, I., Constans, J. M., Gauchotte, G. & Vogin, G. 2020. Radiomics Method for the Differential Diagnosis of Radionecrosis Versus Progression after Fractionated Stereotactic Body Radiotherapy for Brain Oligometastasis. *Radiation Research* 193(5): 471-480. 10.1667/RR15517.1.

Horvat, N., Veeraraghavan, H., Nahas, C. S. R., Bates, D. D. B., Ferreira, F. R., Zheng, J., Capanu, M., Fuqua, J. L., Fernandes, M. C., Sosa, R. E., Jayaprakasam, V. S., Cerri, G. G., Nahas, S. C. & Petkovska, I. 2022. Combined artificial intelligence and radiologist model for predicting rectal cancer treatment response from magnetic resonance imaging: an external validation study. *Abdominal Radiology* 47(8): 2770-2782. 10.1007/s00261-022-03572-8.

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M. & Lis, A. 2023. What is machine learning, artificial neural networks and deep learning?—Examples of practical applications in medicine. *Diagnostics* 13(15): 2582. https://doi.org/10.3390/diagnostics13152582

Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., De Jong, E. E., Van Timmeren, J., Sanduleanu, S., Larue, R. T., Even, A. J. & Jochems, A. 2017. Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology* 14(12): 749-762. 10.1038/nrclinonc.2017.141

Larue, R. T., Defraene, G., De Ruysscher, D., Lambin, P. & Van Elmpt, W. 2017. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *The British Journal of Radiology* 90(1070): 20160665. 10.1259/bjr.20160665.

Liao, C.-Y., Lee, C.-C., Yang, H.-C., Chen, C.-J., Chung, W.-Y., Wu, H.-M., Guo, W.-Y., Liu, R.-S. & Lu, C.-F. 2023. Predicting survival after radiosurgery in patients with lung cancer brain metastases using deep learning of radiomics and EGFR status. *Physical and Engineering Sciences in Medicine* 46(2): 585-596.10.1007/s13246-023-01234-7.

Meißner, A.-K., Gutsche, R., Galldiks, N., Kocher, M., Jünger, S. T., Eich, M.-L., Nogova, L., Araceli, T., Schmidt, N. O. & Ruge, M. I.

2023. Radiomics for the non-invasive prediction of PD-L1 expression in patients with brain metastases secondary to non-small cell lung cancer. *Journal of Neuro-Oncology* 163(3): 597-605. 10.1007/s11060-023-04367-7.

Munai, E., Zeng, S., Yuan, Z., Yang, D., Jiang, Y., Wang, Q., Wu, Y., Zhang, Y. & Tao, D. 2024. Machine learning-based prediction model for brain metastasis in patients with extensive-stage small cell lung cancer. *Scientific reports* 14(1): 28790. s41598-024-80425-y.

Nowakowski, A., Lahijanian, Z., Panet-Raymond, V., Siegel, P. M., Petrecca, K., Maleki, F. & Dankner, M. 2022. Radiomics as an emerging tool in the management of brain metastases. *Neuro-Oncology Advances* 4(1): 141. 10.1093/noajnl/vdac141.

Peng, L. K., Parekh, V., Huang, P., Lin, D. D., Sheikh, K., Baker, B., Kirschbaum, T., Silvestri, F., Son, J., Robinson, A., Huang, E., Ames, H., Grimm, J., Chen, L. D., Shen, C., Soike, M., Mctyre, E., Redmond, K., Lim, M., Lee, J., Jacobs, M. A. & Kleinberg, L. 2018. Distinguishing True Progression From Radionecrosis After Stereotactic Radiation Therapy for Brain Metastases With Machine Learning and Radiomics. *International Journal of Radiation Oncology Biology Physics* 102(4): 1236-1243. 10.1016/j.ijrobp.2018.05.041

Perniciano, A., Loddo, A., Di Ruberto, C. & Pes, B. 2025. Insights into radiomics: impact of feature selection and classification. *Multimedia Tools and Applications* 84(26): 31695-31721. 10.1007/s11042-024-20388-4.

Sim, J. Z. T., Fong, Q. W., Huang, W. & Tan, C. H. 2023. Machine learning in medicine: what clinicians should know. *Singapore medical journal* 64(2): 91-97. 10.11622/smedj.2021054.

Staartjes, V. E. & Schröder, M. L. 2018. Class imbalance in machine learning for neurosurgical outcome prediction: are our models valid? *Journal of Neurosurgery: Spine* 29(5): 611-612. 10.3171/2018.5.SPINE18543.

Tan, D., Mohamad Salleh, S. A., Manan, H. A. & Yahya, N. 2023. Delta-radiomics-based models for toxicity prediction in radiotherapy: A systematic review and meta-analysis. *Journal of Medical Imaging and Radiation Oncology* 67(5): 564-579. 10.1111/1754-9485.13546.

Traverso, A., Wee, L., Dekker, A. & Gillies, R. 2018. Repeatability and reproducibility of radiomic features: a systematic review. *International Journal of Radiation Oncology* Biology* Physics* 102(4): 1143-1158. 10.1016/j.ijrobp.2018.05.053.

Tulum, G. 2023. Novel radiomic features versus deep learning: differentiating brain metastases from pathological lung cancer types in small datasets. *The British Journal of Radiology* 96(1146): 20220841. 10.1259/bjr.20220841

Umaru, A., Abdul Manan, H., Kumar Athi Kumar, R., Hamsan, S. K. & Yahya, N. 2025. Utilizing radiomics as predictive factor in brain metastasis treated with stereotactic radiosurgery: systematic review and radiomic quality assessment. *iRADIOLOGY* 3(2): 132-143. https://doi.org/10.1002/ird3.70007

Umaru, A., Abdul Manan, H., Tan, D., Kumar, R. a. K., Hamsan, S. K. & Yahya, N. 2026. Enhancing treatment response prediction in brain metastases: A multi-modal radiomics approach using MRI and CBCT. *Heliyon* 12(11): e45076. 10.1016/j.heliyon.2026.e45076.

Umaru, A., Manan, H., Kumar, R., Hamsan, S. & Yahya, N. 2025. The impact of intensity discretisation and filtration on the performance of the radiomic and machine learning models in brain metastasis patients treated with gamma knife radiosurgery. *Radiography* 31(5): 103010. 10.1016/j.radi.2025.103010.

Umaru, A., Manan, H. A., Kumar, R. K. A., Hamsan, S. K., Tan, D. & Yahya, N. 2026. Development and external validation of an MRI radiomics-based machine learning model to predict tumour recurrence in brain metastases treated with Gamma Knife radiosurgery. Radiotherapy and Oncology 216(111362.https://doi.org/10.1016/j.radonc.2025.111362

Xia, X., Wu, W., Tan, Q. & Gou, Q. 2025. Interpretable Machine Learning Models for Differentiating Glioblastoma From Solitary Brain Metastasis Using Radiomics. *Academic Radiology* (32):9 5388-5400. https://doi.org/10.1016/j.acra.2025.05.016.

Yahya, N., Ebert, M. A., Bulsara, M., House, M. J., Kennedy, A., Joseph, D. J. & Denham, J. W. 2016. Statistical-learning strategies generate only modestly performing predictive models for urinary symptoms following external beam radiotherapy of the prostate: A comparison of conventional and machine-learning methods. *Medical Physics* 43(5): 2040-2052. 10.1118/1.4944738.

Yoon, C. H., Torrance, R. & Scheinerman, N. 2022. Machine learning in medicine: should the pursuit of enhanced interpretability be abandoned? *Journal of Medical Ethics* 48(9): 581-585. 10.1136/medethics-2020-107102.

Zhang, W., Guo, Y. & Jin, Q. 2023. Radiomics and its feature selection: a review. *Symmetry* 15(10): 1834. https://doi.org/10.3390/sym15101834.

Zwanenburg, A., Vallières, M., Abdalah, M.A., Aerts, H.J., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R.J., Boellaard, R. and Bogowicz, M., 2020. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology, 295(2), pp.328-338. https://doi.org/10.1148/radiol.2020191145