POLISH
JOURNAL *of* RADIOLOGY

Review paper

Scoping review of image-based overall survival prediction in glioma using machine learning

Saman Mohammadpour[1,A,B,C,D,E,F,G], Hassan Emami[1,A,B,E,F,G,#], Reza Rabiei[1,A,B,E,#], Arash Roshanpoor[2,A,E], Morteza Tabatabaeefar[3,A,D], Fariborz Faeghi[4,A,E]

[1]Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch, Islamic Azad University, Tehran, Iran
[3]Department of Radiation Oncology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4]Department of Radiology Technology, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[#]These authors contributed equally to this work.

Abstract

Purpose: Accurate prediction of overall survival (OS) in glioma patients is crucial for optimising treatment decisions. Despite advancements in imaging and machine learning, challenges persist due to tumour heterogeneity and confounding factors. This scoping review systematically assesses state-of-the-art image-based OS prediction models for glioma, focusing on tumour characteristics, imaging modalities, preprocessing techniques, and machine learning methods.

Material and methods: This scoping review was conducted following the Joanna Briggs Institute guidelines, comprising five key stages: identifying the research question, searching for relevant literature, selecting studies, charting the data, and collating, summarising, and reporting the results.

Results: The initial search identified 3238 records, of which 70 articles were included in the final analysis. Most studies originated from China, the United States, and India, with datasets averaging approximately 450 cases. To enhance predictive accuracy, various techniques were utilised, including image segmentation, multimodal magnetic resonance imaging (MRI) protocols, and advanced feature extraction methods. Notably, T1-weighted contrast-enhanced MRI and grade-specific glioma analyses improved model performance. Although deep learning models generally outperformed traditional methods, they required large, balanced datasets. Hybrid models showed promising potential; however, their performance was inconsistent due to challenges such as limited image quality and issues with model interpretability.

Conclusions: Increasing sample size alone does not guarantee improved accuracy in glioma prediction models, because data quality and feature selection are critical factors. Incorporating diverse imaging modalities can significantly enhance predictive performance. To ensure greater clinical reliability in decision-making, integrating clinical features with imaging data is essential.

Key words: overall survival, glioma, machine learning, image analysis, image processing, prediction.

Correspondence address:
Dr. Hassan Emami, Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, e-mail: haemami@sbmu.ac.ir
Dr. Reza Rabiei, Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, e-mail: R.Rabiei@sbmu.ac.ir

Authors' contribution:
A Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection

Introduction

Glioma is one of the most prevalent and lethal brain cancers, accounting for approximately 30% of all primary central nervous system (CNS) malignancies [1]. Due to its aggressive nature and poor prognosis, accurate prediction of overall survival (OS) is crucial for guiding early treatment strategies, including surgery, radiotherapy, chemotherapy, and supportive care planning [2].

OS, defined as the time from diagnosis to death or last follow-up, is influenced by various clinical and pathological factors, including tumour grade, location, molecular profile, and treatment response [3,4]. Gliomas are classified by the World Health Organisation (WHO) into low-grade (grades I-II) and high-grade (grades III-IV) tumours, with 5-year survival rates ranging from approximately 80% for low-grade gliomas to less than 5% for high-grade cases [4,5].

Numerous prognostic indicators, such as tumour grade, anatomical location, and intratumoural heterogeneity, can be extracted from imaging data, positioning medical imaging as a critical tool in predicting OS in glioma patients [6]. In particular, advances in magnetic resonance imaging (MRI) have markedly improved the visualisation, segmentation, and quantitative analysis of brain tumours. These developments enable the generation of high-dimensional imaging features that are instrumental in constructing robust prognostic models. Nevertheless, OS prediction frequently depends on manual interpretation by clinicians, which introduces considerable interobserver variability and may compromise predictive accuracy [7,8].

To address the limitations of manual image interpretation, recent research has increasingly turned to machine learning (ML) techniques for extracting quantitative imaging biomarkers and identifying complex patterns associated with glioma progression and OS [9,10]. ML-based image analysis, encompassing radiomics and deep learning, has shown significant potential in improving prediction performance [11]. Key approaches to medical image analysis include geometry-based feature extraction, intensity-based analysis, and machine learning algorithms [12,13]. Among these, ML-based image analysis has garnered substantial attention in recent years, leveraging advanced algorithms to uncover hidden patterns and complex relationships within medical imaging data [12,14].

Despite increasing interest in this field, there is a lack of systematic reviews that comprehensively evaluate the impact of various imaging modalities, tumour segmentation approaches, feature extraction methods, and ML architectures on predictive performance across studies [8].

This scoping review seeks to address this gap by critically evaluating existing image-based glioma OS prediction models. It examines how performance metrics vary across WHO tumour grades, imaging modalities, segmentation strategies, and modelling approaches, providing a structured synthesis of the current landscape and identifying directions for future research.

Material and methods

This scoping review was conducted in accordance with the Joanna Briggs Institute (JBI) guidelines for scoping reviews [15], following five key steps: (1) identifying the research question; (2) searching for relevant literature; (3) selecting studies; (4) charting the data; and (5) collating, summarising, and reporting the results.

Research questions

- Which countries have implemented medical image analysis for glioma overall survival prediction?
- To what extent does sample size influence the predictive performance of image-based models for glioma overall survival prediction?
- How do predictive performance metrics of image-based glioma overall survival models vary across different WHO tumour grades?
- How do predictive performance metrics of image-based glioma overall survival models vary across different imaging modalities?
- How do the predictive performance metrics of image-based glioma overall survival models vary across different tumour segmentation approaches?
- How do the predictive performance metrics of image-based glioma overall survival models vary across different feature extraction methods?
- How do the predictive performance metrics of image-based glioma overall survival models differ among various machine learning architectures?

Toward relevant studies

The study began with an extensive search across the PubMed, Scopus, and Web of Science databases. The search strategy, tailored for each database, used keywords related to glioma, image analysis, and overall survival, as detailed in Supplement 1.

In addition to the systematic database search, the reference lists of relevant articles were manually reviewed, and the 'Related Articles' and 'Similar Research' features on Google Scholar and ResearchGate were used to identify additional studies. All retrieved records were imported into EndNote, where duplicate entries were removed as an initial step. The search period included all available publications up to 12 December 2024, with no restriction on the start date.

Study selection process

In the next phase, 2 reviewers independently screened the remaining articles by assessing their titles and abstracts.

Disagreements were resolved by consulting a third reviewer to reach a final decision. Studies not focused on developing or applying image-based glioma OS prediction models were excluded at this stage. The full texts of the remaining articles were thoroughly evaluated, and studies were excluded if they did not meet the predefined inclusion criteria, which required the following: original research articles published in English with full-text availability, involvement of human subjects, and a specific focus on image-based methods for predicting OS in glioma patients, such as developing or validating predictive models using radiological imaging data. Only studies involving human subjects were included; those based on animal models were excluded during screening.

Charting, summarising, and reporting

Data extraction was performed by 4 researchers using a form designed in Microsoft Word. The form included fields for the first author and co-authors, year of publication, country of the first author's institution, study type, sample size, imaging modality, image quality, WHO tumour classification, treatment stage, image segmentation method, image data extraction method, feature selection or reduction method, machine learning architectures, features used, performance evaluation metrics, and corresponding results. The first author developed the form, and its content and the data extraction process were explained to the researchers during a meeting to ensure consistency in data collection. Any uncertainties or disagreements were resolved with input from the first and corresponding authors.

Results

An initial literature search retrieved 3238 records from 3 databases: PubMed (369), Scopus (2,024), and Web of Science (845). After removing 438 duplicates using End-Note software, 2703 unique records remained for screening. During the title and abstract screening phase, 2505 records were excluded for not meeting the inclusion criteria. Subsequently, 198 full-text articles were evaluated for eligibility, of which 128 were excluded. Ultimately, 70 studies met all inclusion criteria and were included in the final analysis (Figure 1). These studies, published between 2016 and 2024, are summarised in Supplement 2. The entire selection process followed the PRISMA guidelines to ensure methodological rigor and transparency.

Countries implementing medical imaging analysis for glioma OS prediction: China, the United States, India, South Korea, and Singapore accounted for the highest number of studies, with 24, 17, 9, 4, and 4 studies published, respectively.

Impact of sample size on the predictive performance of glioma OS models: The included studies used varying sample sizes for OS prediction. The mean sample size was

450 ± 205 (mean ± standard deviation). Among these studies, 30 reported accuracy, while 18 provided C-index values for predicting OS in glioma patients. Figure 2 displays the reported accuracy and C-index values along with their respective sample sizes across the studies.

Variation in predictive performance of glioma OS models across WHO tumour grades: A total of 8, 4, 30, and 28 studies focused on predicting OS in high-grade gliomas, low-grade gliomas, grade IV glioblastomas (GBM), and across all glioma grades, respectively (Figure 3).

Variation in predictive performance of glioma OS models across different imaging modalities: The included studies employed various combinations of MRI sequences for survival prediction, such as T1-contrast enhanced (T1-CE), T1-weighted (T1w), T2-weighted (T2w), FLAIR, T1c, diffusion-weighted imaging (DWI), and diffusion tensor imaging (DTI) (Figure 4).

Variation in predictive performance of glioma OS models across different segmentation methods: The reviewed studies employed 3 main segmentation approaches: whole-image analysis (10 studies), manual segmentation (20 studies), and automated segmentation (40 studies). Figure 5 presents the reported accuracy and C-index values stratified by segmentation approach, based on studies that provided these performance metrics.

Variation in predictive performance of glioma OS models across different feature extraction methods: Three distinct feature extraction approaches were utilised across the studies: radiomics analysis (48 studies), deep learning (12 studies), and hybrid methods that combine both radiomics and deep learning (10 studies). Figure 6 displays the reported accuracy and C-index values stratified by feature extraction method, based on studies that reported these performance metrics.
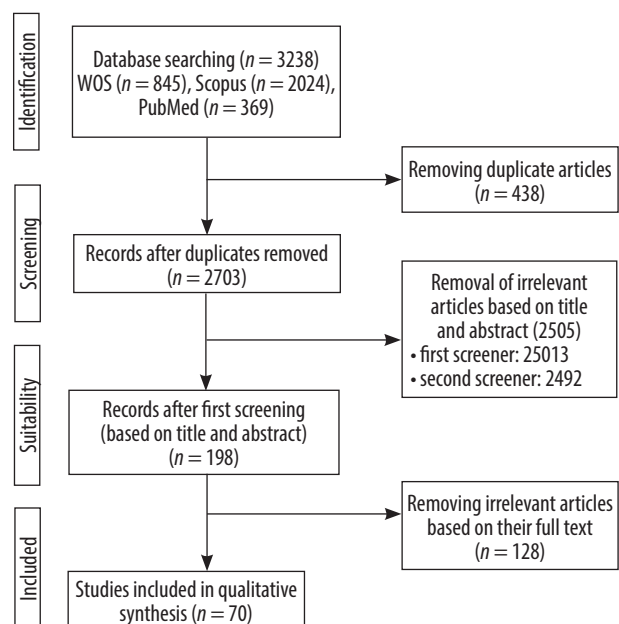
Figure 1. PRISMA flow diagram of study selection process for image-based glioma overall survival prediction review
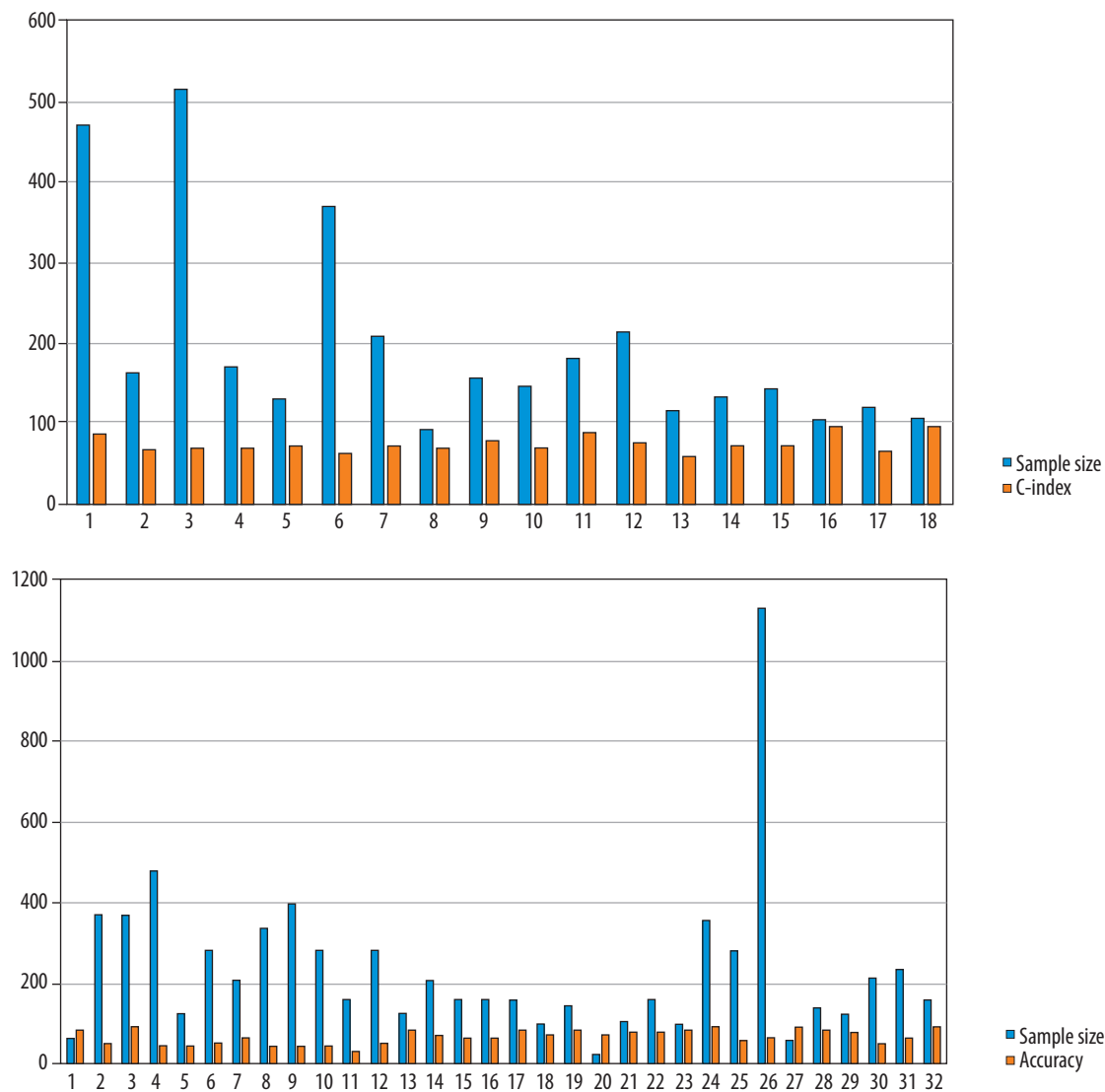
Figure 2. Relationship between sample size and predictive performance (accuracy and C-index) of image-based glioma survival models
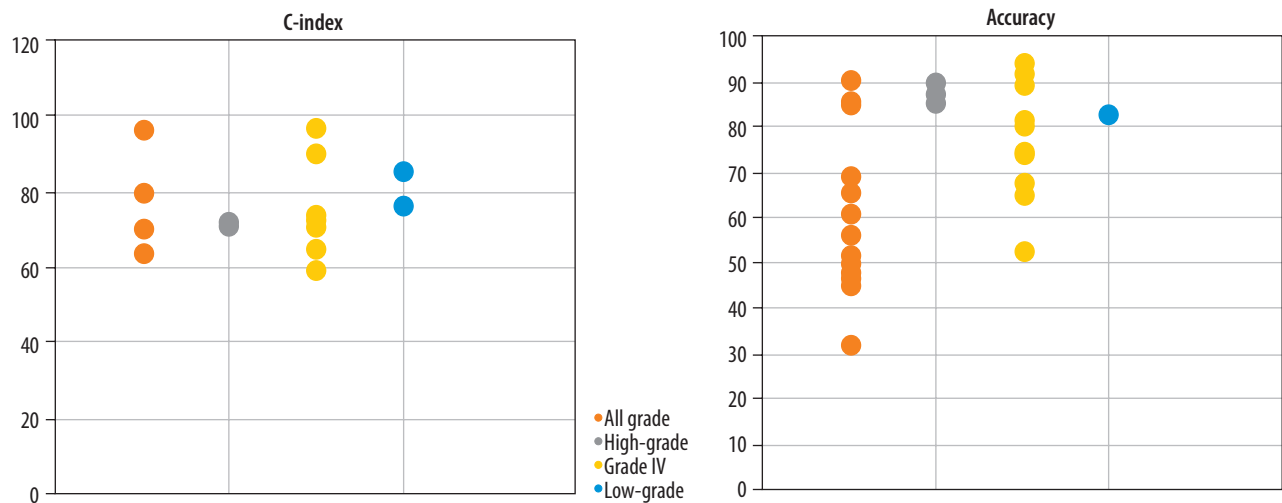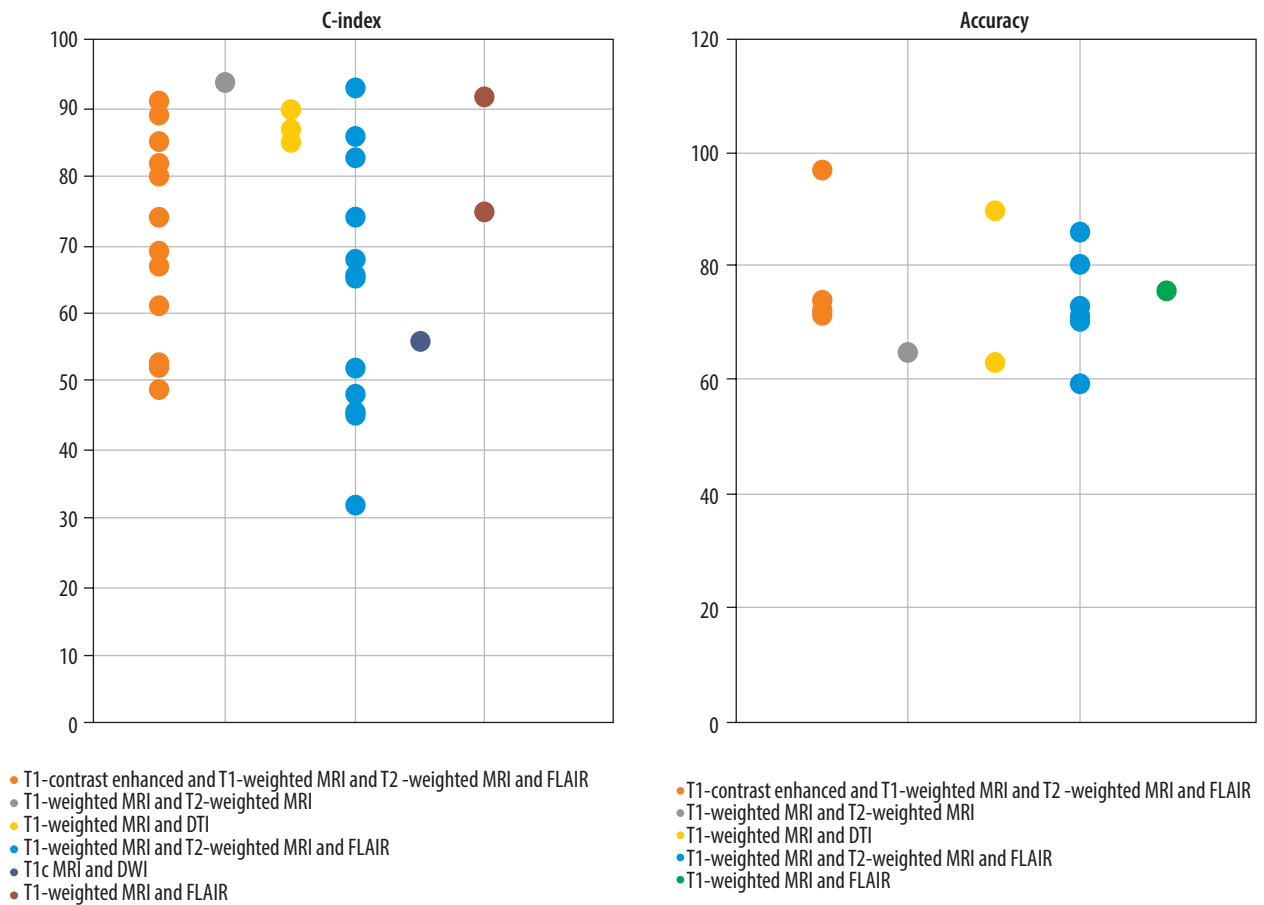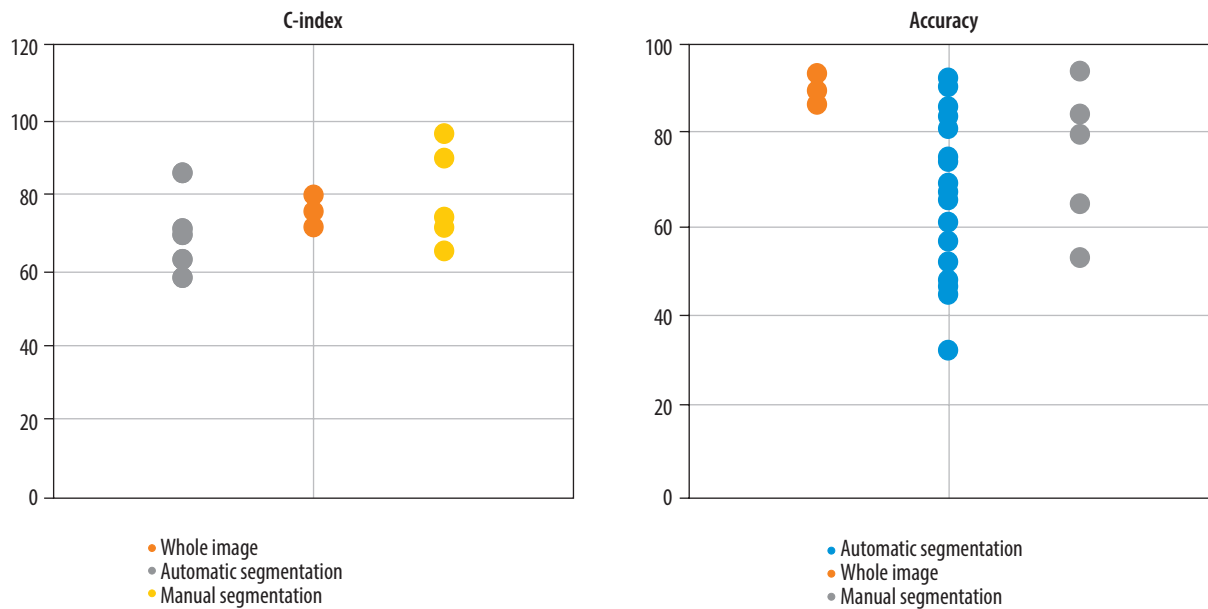

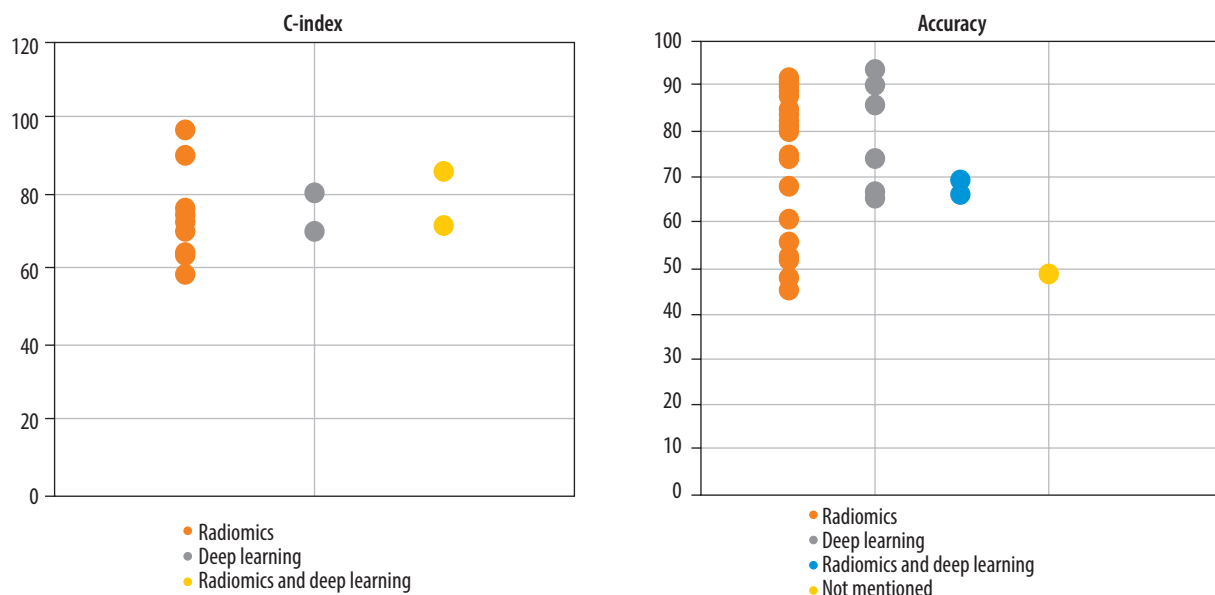**Figure 3.** Predictive performance (accuracy and C-index) of glioma survival models across different WHO tumour grades
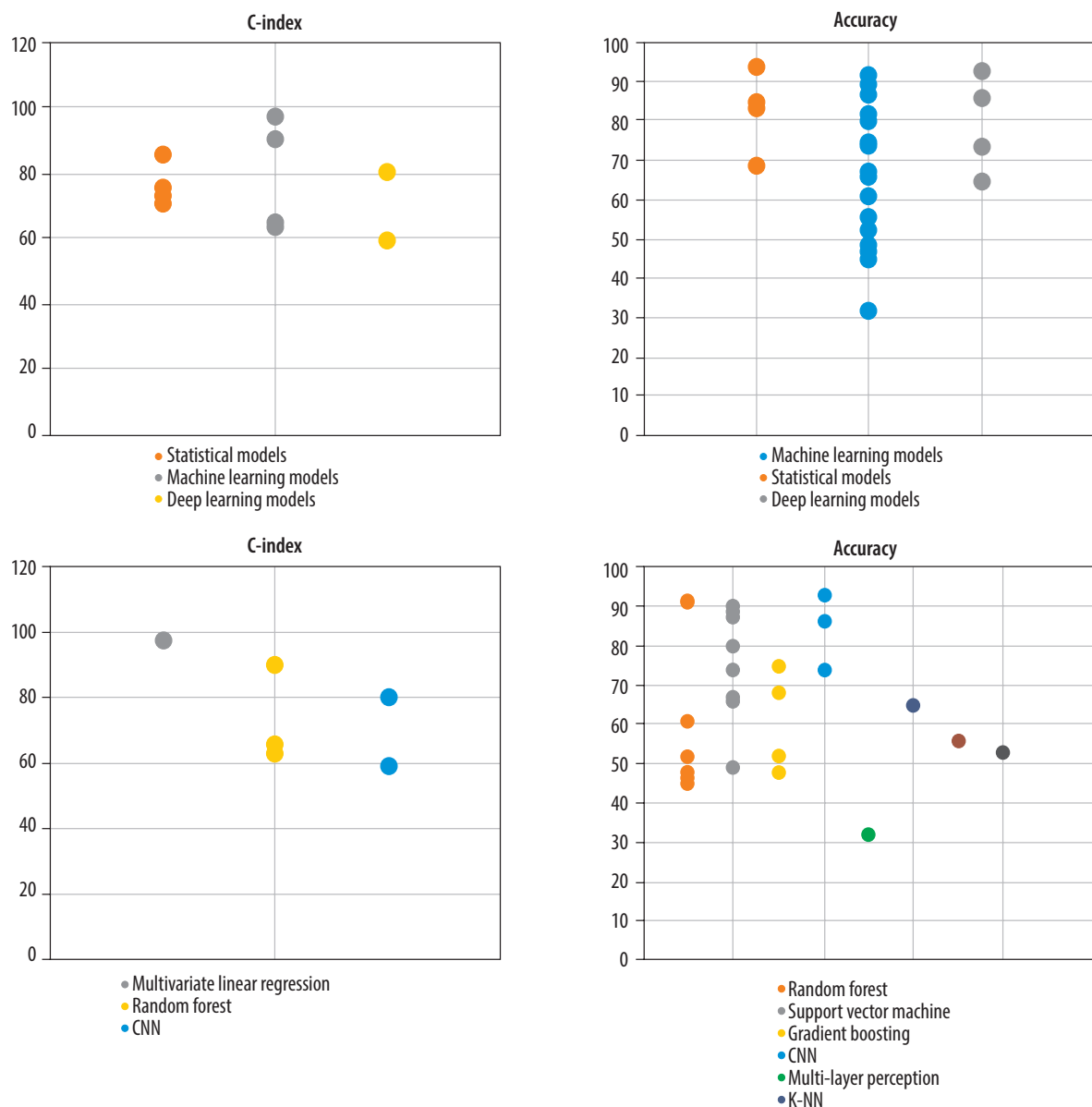
Figure 4. Predictive performance (accuracy and C-index) across various MRI modalities used in glioma survival prediction (T1-CE, T1w, T2w, FLAIR, T1c, DWI, DTI)

Figure 5. Predictive performance (accuracy and C-index) stratified by tumour segmentation approaches (whole-image, manual, automated)

Figure 6. Predictive performance (accuracy and C-index) based on different feature extraction methods (radiomics, deep learning, hybrid radiomics – deep learning)

Figure 7. Comparison of predictive modelling architectures (statistical models, machine learning, deep learning) in glioma overall survival prediction (accuracy and C-index)

Variation in Predictive Performance of Glioma OS Models Across Different Machine Learning Architectures: Three primary modelling methodologies were applied across the studies to predict survival outcomes in glioma patients: statistical models (18 studies), ML approaches (43 studies), and deep learning techniques (9 studies). Figure 7 illustrates the comparative performance metrics, accuracy, and C-index, stratified by modelling methodology, based on studies that reported both evaluation measures.

Discussion

The first application of medical image analysis for predicting OS in glioma patients was reported in 2016. In that study [16], machine learning techniques were applied to MRI data to predict OS in patients with glioblastoma multiforme (GBM; WHO grade IV glioma). Since this pioneering work, the field has experienced substantial growth, with significant increases in both the number of studies and the diversity of employed methodologies.

Countries implementing medical imaging analysis for glioma OS prediction

China's recent healthcare policies, which emphasise reducing clinical costs, alongside the demonstrated efficiency of medical image analysis in conserving both time and resources, provide strong justification for implementing image-based models in predicting glioma OS. At present, such prognostic evaluations primarily rely on genetic testing, a method that is both costly and time-consuming [17]. According to projections by Grand View Research, China, the United States, and India are expected to become the leading markets for medical imaging and image analysis by 2031 [18]. This anticipated growth is likely to accelerate the clinical integration of these technologies in applications such as glioma survival prediction. Moreover, the large population sizes of these countries [19] may act as a key driver for increased demand for cost-effective and scalable imaging-based solutions. Nevertheless, comprehensive epidemiological studies are still needed to assess glioma prevalence and examine patterns of imaging technology utilisation to better understand this association.

The impact of sample size on the predictive performance of glioma OS models

Sample size is a critical factor in evaluating the accuracy of predictive algorithms, especially for complex models such as radiomics and deep learning applied to survival prediction in glioma patients. Insufficient sample sizes can lead to overfitting in deep learning models, causing them to memorise specific features from the training data without generalising well to new datasets. This overfitting significantly impairs the model's predictive accuracy and real-world applicability. Conversely, the availability of larger sample sizes facilitates the detection of hidden patterns by algorithms, which in turn enhances their predictive accuracy. With increased data availability, deep learning models can uncover more intricate relationships among diverse features, leading to more accurate predictions [20]. However, as illustrated in Figure 2, although sample size influences the performance of machine learning algorithms, studies with larger sample sizes have not consistently reported higher accuracy or C-index values. This indicates that simply increasing the number of samples does not necessarily lead to improved model performance. Such inconsistencies may be attributed to other influential factors, including input image quality, feature extraction methods, data noise levels, and tumour grade [21]. Therefore, in addition to expanding dataset size, greater emphasis should be placed on improving image quality, selecting optimal features, and applying advanced data analysing techniques. Future research should investigate these aspects more thoroughly to further enhance predictive model performance.

Variation in predictive performance of glioma OS models across WHO tumour grades

The WHO classifies gliomas into 4 grades based on histopathological characteristics: Grades I and II are considered low-grade, while Grades III and IV are classified as high-grade tumours [5]. Several studies have examined survival prediction models based on imaging data for glioma patients of different WHO grades. As illustrated in Figure 3, studies that trained a single model on combined data from all glioma grades (i.e. grades I-IV) generally reported lower predictive accuracy. This decline in performance is probably due to the substantial variation in survival outcomes between low-grade and high-grade gliomas, which introduces greater heterogeneity and makes it more difficult for models to identify consistent predictive patterns [22]. Low-grade gliomas typically demonstrate slower growth and progression, whereas high-grade gliomas exhibit aggressive behaviour and are associated with significantly worse prognoses. As a result, combining these 2 biologically and clinically distinct groups within a single predictive model may impair the model's ability to identify consistent and relevant patterns in the data. In such scenarios, the model may become overly influenced by the dominant characteristics of one group, often the high-grade tumours, while underrepresenting critical features of the low-grade group. This imbalance can ultimately reduce the model's generalisability and predictive accuracy across diverse patient populations [23].

Furthermore, the imaging characteristics and tissue architecture of gliomas vary significantly across different tumour grades. For instance, MRI scans of low-grade gliomas often appear less aggressive, displaying minimal variation in signal intensity and exhibiting limited evidence of necrosis or oedema. In contrast, high-grade gliomas typically present with substantial structural abnormalities and pronounced heterogeneity. These distinct phenotypic manifes-

tations pose a challenge for deep learning models, making it difficult to identify consistent, grade-invariant imaging biomarkers. As a result, the accuracy of such models may be compromised when applied across heterogeneous glioma grades. [23,24]. To enhance the accuracy of glioma survival prediction models, researchers may choose to analyse glioma grades separately due to their distinct histopathological and imaging features, or apply advanced techniques like data weighting and transfer learning. These strategies address variability between grades and class imbalance, improving model performance and generalisability. Future research should prioritise multicentre validation of these methods to develop standardised and clinically reliable glioma prediction models.

Variation in predictive performance of glioma OS models across different imaging modalities

Notably, all reviewed studies utilised MRI for image analysis, probably due to its superior capability in visualising fine soft tissue details, particularly within the brain parenchyma. By contrast, CT and PET scans are more commonly used for imaging bony structures and detecting trauma-related injuries [25]. Specifically, MRI provides higher-resolution delineation of cerebral soft-tissue architecture, which is crucial for survival analysis. This modality excels at detecting microscopic cellular and histopathological changes, making it an invaluable tool for monitoring disease progression. Because microscopic and histopathological alterations are pivotal in survival prediction studies, MRI's high spatial resolution allows for precise identification and tracking of these changes, thus providing robust data for more accurate prognostic modelling. Conversely, CT and PET have limited capability in resolving such fine details [25,26].

Magnetic resonance imaging is often considered a safer alternative to CT and PET scans because the latter involve exposure to ionising radiation and radioactive tracers, which may pose long-term health risks. In contrast, MRI does not carry these risks and is therefore regarded as a more suitable and lower-risk option for repeated evaluations [27]. Moreover, MRI is more widely available in hospitals and medical centres, making it accessible for both clinical and research purposes. Compared to PET scans, which require specialised equipment and incur higher costs, MRI offers a more practical and cost-effective solution. Given that many studies demand rapid acquisition of high-quality data, MRI is frequently prioritised due to its greater accessibility [27,28].

As illustrated in Figure 4, various MRI modalities, including T1-weighted imaging, T2-weighted imaging, DTI, FLAIR, and T1-contrast enhanced (T1c), are commonly employed for the diagnosis and evaluation of brain tumours and related lesions [29]. T1-weighted Imaging is typically used to assess anatomy and tumours because fats and dense tissues appear bright in these scans. This modality is effective for delineating precise brain struc-

tures but may lack sensitivity for detecting small lesions. Conversely, T2-weighted imaging produces a stronger signal from cerebrospinal fluid, making it well-suited for identifying lesions associated with fluid accumulation, such as inflammation or tumours. However, T2-weighted imaging may not be sufficiently sensitive to detect minor brain injuries [30].

Diffusion tensor imaging, which maps the diffusion of water molecules, is particularly valuable for assessing neural pathways and brain injuries. This advanced technique reconstructs white matter tracts, enabling clinicians to detect disruptions in cerebral connectivity. The FLAIR sequence is another sophisticated imaging approach that suppresses cerebrospinal fluid (CSF) signals, thereby enhancing the visibility of periventricular and cortical lesions adjacent to CSF spaces. This modality is especially effective in detecting small brain lesions that might otherwise be obscured [29,31] tissues such as tumours and inflammatory lesions. This technique provides high-contrast visualisation of highly vascularised tissues, although it requires contrast administration. T1c imaging is particularly effective for tumour detection and characterisation of inflammatory lesions [30]. Studies incorporating T1c MRI alongside T1-weighted, T2-weighted, and FLAIR modalities for survival prediction have reported higher accuracy and C-index values. Since angiogenesis – the growth of new blood vessels – in tumour tissue is a key factor influencing OS in glioma patients, T1c MRI offers valuable insights in this context [32]. This high-contrast modality effectively visualises tissues with substantial blood flow, suggesting its potential to enhance the predictive power of models estimating OS in glioma patients [30].

Variation in predictive performance of glioma OS models across different segmentation methods

For glioma survival prediction, some studies analyse the whole image to capture all features, while others use image segmentation to focus specifically on tumour regions and minimise irrelevant information. This targeted approach improves model accuracy by better identifying tumour characteristics and reducing interference from non-tumour data [33]. Image segmentation in cancer involves dividing medical images into distinct regions to identify and assess tumours or cancerous lesions [34]. This process can be performed manually or automatically. In manual segmentation, clinicians or specialists visually delineate tumour areas. Although this method can be highly accurate, it is time-consuming, requires substantial expertise, and is prone to human error. Automatic segmentation, on the other hand, employs computer algorithms, such as neural networks and deep learning techniques, to automatically detect and segment suspected cancerous regions. This approach is faster and more scalable for large datasets but depends heavily on well-configured algorithms and high-quality training data. Automated segmentation accelerates disease identification and

monitoring, facilitates more precise treatment planning, and is especially valuable for early cancer detection while minimising human error [35].

Among the 18 studies reporting C-index values, those using manual image segmentation demonstrated higher C-index scores. However, contrary to expectations that manual segmentation would also lead to higher accuracy, Figure 5 shows no significant differences in accuracy among studies employing whole-image analysis, manual segmentation, or automated segmentation. The comparable accuracy between manual and automated methods may result from confounding variables affecting accuracy measurements or potential human errors inherent in manual segmentation.

An important prognostic factor in glioma survival prediction is the tumour's anatomical location, with diencephalic tumours associated with significantly poorer OS compared to tumours in other brain regions [36]. Since anatomical location information may be partially lost when analysing segmented images instead of whole images, using full images could potentially compensate for this limitation, leading to predictive accuracy comparable to studies that employ image segmentation techniques.

Variation in predictive performance of glioma OS models across different feature extraction methods

Various methods, including radiomics and deep learning, are employed to predict OS in glioma patients. Radiomics involves extracting predefined features, such as geometric, textural, and structural characteristics, from medical images of brain tissue [37,38]. These features may capture information related to tumour volume, tissue heterogeneity, and the architecture of affected regions, all of which can correlate with glioma prognosis. Typically, such features are analysed using statistical models or machine learning algorithms to predict survival outcomes [37,39].

In contrast, deep learning methods employ convolutional neural networks (CNNs) to automatically extract complex image features that may be imperceptible to human observers. This approach allows for more precise modelling of disease progression and enables more personalised survival predictions. Deep learning is particularly effective at capturing complex relationships between diverse imaging features and survival outcomes in glioma because these models can analyse data across multiple sophisticated levels [37].

The radiomics feature extraction method requires image segmentation and manual feature selection, producing features that are interpretable by clinicians. In contrast, deep learning methods bypass the need for segmentation and manual selection, but the extracted features often remain unintelligible to physicians, functioning essentially as a "black box" [40]. Combining these 2 approaches may enhance model performance in survival prediction [41]. As shown in Figure 5, most studies predominantly use radiomics methods. The study by Afshar et al. highlighted that, unlike other therapeutic areas, radiomics has been

more widely adopted for survival prediction, possibly due to domain-specific knowledge not easily accessible in deep learning approaches [40]. Figure 6 analysis reveals that studies employing deep learning methods are more sensitive to sample size concerning predictive accuracy. Although hybrid methods were expected to outperform single approaches, studies using combined methodologies did not consistently report superior accuracy or C-index values [37,42-45]. Further examination suggested that, while hybrid models often showed better results alongside individual methods, their relatively lower reported performance might be influenced by other confounding factors.

Variation in predictive performance of glioma OS models across different machine learning architectures

Three main categories of models are used for survival prediction: statistical models, machine learning, and deep learning [8]. Traditional statistical approaches, such as Cox proportional hazards models and Kaplan-Meier estimators, remain especially valuable for small or homogeneous datasets due to their simplicity and interpretability. These models offer straightforward interpretations of survival probabilities and hazard ratios without the need for complex computations. However, they have limitations in handling nonlinear relationships and high-dimensional data, and integrating diverse data types like medical images or molecular features [46].

Machine learning models, including algorithms such as Support Vector Machines (SVMs), Random Forests, and Gradient Boosting, have gained attention for their flexibility in analysing heterogeneous data. These approaches can effectively integrate multimodal data (e.g. clinical, radiomics, and molecular features) and capture complex nonlinear patterns. However, their complexity can increase susceptibility to overfitting, reduce interpretability compared to traditional statistical models, and pose challenges in handling incomplete datasets or performing multimodal data fusion [8,46].

Deep learning models, including CNNs and Transformers, have demonstrated remarkable effectiveness in analysing complex, high-dimensional data such as medical images. These models automatically learn intricate patterns without the need for manual feature engineering, and they support seamless integration of multimodal data for accurate prediction. Nevertheless, deep learning models typically require large training datasets, significant computational resources, and are prone to overfitting when trained on small datasets. Moreover, their "black box" nature limits interpretability of the results [8,47].

As illustrated in Figure 7, machine learning methods reported a wide range of accuracy and C-Index values across studies. When examining performance by algorithm, studies utilising SVM and CNN approaches achieved higher accuracy. This finding is supported by the comparative study by Wang et al., which identified SVM and CNN as the top-performing methods, attributing their superiority to

greater resistance to overfitting [48]. Among the few studies reporting C-Index values for machine learning models, no substantial differences were observed across different algorithms.

This study had several limitations, primarily stemming from its reliance on the quality and availability of the published literature. Although comprehensive searches were conducted across reputable databases, it is possible that some relevant but lower-quality, non-English language, or unpublished studies were inadvertently excluded. Furthermore, methodological heterogeneity among the included studies, such as variations in imaging protocols, feature extraction techniques, and evaluation metrics, limited our ability to directly compare and synthesise results in a standardised manner.

Conclusions

The findings of this review highlight that while increasing the sample size can improve the performance of machine learning models for glioma survival prediction, it is not sufficient on its own to ensure high accuracy. Other key factors, such as data quality, signal-to-noise ratio, and optimal feature selection, play critical roles in determining model performance. From a statistical learning perspective, accuracy reflects a model's ability to generalise; models with lower accuracy are less likely to perform well on independent datasets.

Notably, combining different glioma grades (e.g. low-grade and high-grade) into a single predictive model may reduce performance due to substantial biological and imaging heterogeneity between the groups. This increased intra-dataset variance complicates the learning process and undermines generalisability. Moreover, studies demonstrate that the use of diverse MRI modalities, particularly T1-contrast enhanced imaging, can enhance prediction accuracy by capturing tumour vascular characteristics relevant to survival.

Regarding imaging analysis techniques, both full-image and segmented-image approaches have demonstrated comparable effectiveness in certain contexts, suggesting that the inclusion of anatomical context in full-image analysis may offset the benefits of tumour-focused segmentation.

In terms of feature extraction, radiomics remains the most widely used method, offering interpretable results and compatibility with traditional machine learning models. While deep learning shows strong potential in capturing complex, non-linear patterns, its performance is highly sensitive to sample size and training data quality. Although hybrid approaches that combine radiomics and deep learning hold theoretical advantages, current evidence suggests they require further refinement to consistently outperform individual methods.

Ultimately, improving the accuracy and clinical utility of glioma survival prediction models will require an integrative approach. This includes optimising model architectures, enhancing input data quality, and incorporating multimodal data such as clinical, genetic, and imaging features. Future research should also prioritise multicentre validation and standardisation to support broader clinical adoption and improve patient outcomes.

Disclosures

References

1. Luo C, Yang J, Liu Z, Jing D. Predicting the recurrence and overall survival of patients with glioma based on histopathological images using deep learning. Front Neurol 2023; 14: 1100933.

2. Li Q, Du Z, Chen Z, Huang X, Li Q. Multiview deep forest for overall survival prediction in cancer. Comput Math Methods Med 2023; 2023: 7931321.

3. Lobato-Delgado B, Priego-Torres B, Sanchez-Morillo D. Combining molecular, imaging, and clinical data analysis for predicting cancer prognosis. Cancers (Basel) 2022; 14: 3215.

4. Whitfield BT, Huse JT. Classification of adult-type diffuse gliomas: Impact of the World Health Organization 2021 update. Brain Pathol 2022; 32: e13062.

5. Cho HH, Park H. Classification of low-grade and high-grade glioma using multi-modal image radiomics features. Annu Int Conf IEEE Eng Med Biol Soc 2017; 2017: 3081-3084.

6. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv 2018; arXiv 181102629.

7. Agravat RR, Raval MS. A survey and analysis on automated glioma brain tumor segmentation and overall patient survival prediction. Arch Comput Methods Eng 2021; 28: 4117-4152.

8. Poursaeed R, Mohammadzadeh M, Safaei AA. Survival prediction of glioblastoma patients using machine learning and deep learning: a systematic review. BMC Cancer 2024; 24: 1581.

9. Zadeh Shirazi A, Fornaciari E, Bagherian NS, Ebert LM, Koszyca B, Gomez GA. DeepSurvNet: deep survival convolutional network for brain cancer survival rate classification based on histopathological images. Med Biol Eng Comput 2020; 58: 1031-1045.

10. Mohammadpour S, Emami H, Rabiei R, Hosseini A, Moghaddasi H, Faeghi F, et al. Image analysis as tool for predicting colorectal cancer molecular alterations: a scoping review. Mol Imaging Radionucl Ther 2025; 34: 10-25.

11. Lu ZM, Guo SZ. Lossless Information Hiding in Images. Syngress; 2016.

12. Zlochower A, Chow DS, Chang P, Khatri D, Boockvar JA, Filippi CG. Deep learning AI applications in the imaging of glioma. Top Magn Reson Imaging 2020; 29: 115.

13. Sadr H, Nazari M, Yousefzade-Chabok S, Emami H, Rabiei R, Ashraf A. Enhancing brain tumor classification in MRI images: a deep learning-based approach for accurate classification and diagnosis. Image Vis Comput 2025; 159:105555.

14. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017; 19: 221-248.

15. Peters MDJ, Godfrey C, Munn Z, Tricco AC, Khalil H. Scoping Reviews. In: JBI Manual for Evidence Synthesis. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z (eds.). JBI 2024.

16. Macyszyn L, Akbari H, Pisapia JM, Da X, Attiah M, Pigrish V, et al. Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques. Neuro Oncol 2016; 18: 417-425.

17. Yan J, Lin HH, Zhao D, Hu Y, Shao R. China's new policy for healthcare cost-control based on global budget: a survey of 110 clinicians in hospitals. BMC Health Serv Res 2019; 19: 84.

18. Medical Imaging Market. Size, Share & Trends Analysis Report By Technology (X-ray Devices, Computed Tomography, Ultrasound, Nuclear Imaging). By Application, By End-use, By Region, And Segment Forecasts, 2024-2030. San Francisco, USA 2025. Available from: https://www.grandviewresearch.com/industry-analysis/medical-imaging-systems-market/toc.

19. Woldometer. Countries in the world by population (2025). Available from: https://www.worldometers.info/world-population/population-by-country.

20. Zhong J, Liu X, Lu J, Yang J, Zhang G, Mao S, et al. Overlooked and underpowered: a meta-research addressing sample size in radiomics prediction models for binary outcomes. Eur Radiol 2025; 35: 1146-1156.

21. Sohn CK, Bisdas S. Diagnostic accuracy of machine learning-based radiomics in grading gliomas: systematic review and meta-analysis. Contrast Media Mol Imaging 2020; 2020: 2127062.

22. Claus EB, Walsh KM, Wiencke JK, Molinaro AM, Wiemels JL, Schildkraut JM, et al. Survival and low-grade glioma: the emergence of genetic information. Neurosurg Focus 2015; 38: E6.

23. Haydar N, Alyousef K, Alanan U, Issa R, Baddour F, Al-Shehabi Z, et al. Role of Magnetic resonance imaging (MRI) in grading gliomas comparable with pathology: a cross-sectional study from Syria. Ann Med Surg (Lond) 2022; 82: 104679.

24. Horská A, Barker PB. Imaging of brain tumors: MR spectroscopy and metabolic imaging. Neuroimaging Clin N Am 2010; 20: 293-310.

25. Hooper GW, Ansari S, Johnson JM, Ginat DT. Advances in the radiological evaluation of and theranostics for glioblastoma. Cancers (Basel) 2023; 15: 4162.

26. Pope WB, Brandal G. Conventional and advanced magnetic resonance imaging in patients with high-grade glioma. Q J Nucl Med Mol Imaging 2018; 62: 239-253.

27. Al Murdef HMH, Al Jumhur ASH, Hamim AMHB, Al Rashah MAS, Alzanati MMH, Mahdi N, et al. Comparison of CT and MRI for brain imaging. J Int Crisis Risk Commun Research 2024; 7: 337-351.

28. Ibrahim R, Samian SD, Mazli M, Amrizal M, Aljunid SM. Cost of magnetic resonance imaging (MRI) and computed tomography (CT) scan in UKMMC. BMC Health Serv Res 2012; 12 (Suppl. 1): P11.

29. Shukla G, Alexander GS, Bakas S, Nikam R, Talekar K, Palmer JD, et al. Advanced magnetic resonance imaging in glioblastoma: a review. Chin Clin Oncol 2017; 6: 40.

30. Zhou Z, Huang D, Bao J, Chen Q, Liu G, Chen Z, et al. A synergistically enhanced T(1)-T(2) dual-modal contrast agent. Adv Mater 2012; 24: 6223-6228.

31. Minja FJ, Bhimani A, Wydra E. New techniques in MRI – fluid-attenuated inversion recovery (FLAIR) imaging, diffusion tensor imaging (DTI), and MRI-guided Laser-induced thermotherapy (LITT) for brain lesions. Int Anesthesiol Clin 2016; 54: 94-108.

32. Ahir BK, Engelhard HH, Lakka SS. Tumor development and angiogenesis in adult brain tumor: glioblastoma. Mol Neurobiol 2020; 57: 2461-2478.

33. Sun L, Zhang S, Chen H, Luo L. Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning. Front Neurosci 2019; 13: 810.

34. Ramesh K, Kumar GK, Swapna K, Datta D, Rajest SS. A review of medical image segmentation algorithms. EAI Endorsed Trans Pervasive Health Technol 2021; 7: e6.

35. Cardenas CE, Yang J, Anderson BM, Court LE, Brock KB. Advances in auto-segmentation. Semin Radiat Oncol 2019; 29: 185-197.

36. Liu TT, Achrol AS, Mitchell LA, Du WA, Loya JJ, Rodriguez SA, et al. Computational identification of tumor anatomic location associated with survival in 2 large cohorts of human primary glioblastomas. AJNR Am J Neuroradiol 2016; 37: 621-628.

37. Xu C, Peng Y, Zhu W, Chen Z, Li J, Tan W, et al. An automated approach for predicting glioma grade and survival of LGG patients using CNN and radiomics. Front Oncol 2022; 12: 969907.

38. Chato L, Latifi S. Machine learning and radiomic features to predict overall survival time for glioblastoma patients. J Pers Med 2021; 11: 1336.

39. Kalaiselvi T, Selvi SK. Investigation of image processing techniques in MRI based medical image analysis methods and validation metrics for brain tumor. Curr Med Imaging 2018; 14: 489-505.

40. Afshar P, Mohammadi A, Plataniotis KN, Oikonomou A, Benali H. From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities. IEEE Signal Process Mag 2019; 36: 132-160.

41. Fu L, Ma J, Ren Y, Han YS, Zhao J (eds.). Automatic detection of lung nodules: false positive reduction using convolution neural networks and handcrafted features. Medical Imaging 2017: Computer-Aided Diagnosis; 2017: SPIE.

42. Khan MKH, Guo W, Liu J, Dong F, Li Z, Patterson TA, et al. Machine learning and deep learning for brain tumor MRI image segmentation. Exp Biol Med (Maywood) 2023; 248: 1974-1992.

43. Salçin K. Detection and classification of brain tumours from MRI images using faster R-CNN. Tehnički glasnik 2019; 13: 337-342.

44. Fu J, Singhrao K, Zhong X, Gao Y, Qi SX, Yang Y, et al. An automatic deep learning-based workflow for glioblastoma survival prediction using preoperative multimodal MR images: a feasibility study. Adv Radiat Oncol 2021; 6: 100746.

45. Wang Y, Shao Q, Luo S, Fu R. Development of a nomograph integrating radiomics and deep features based on MRI to predict the prognosis of high grade Gliomas. Math Biosci Eng 2021; 18: 8084-8095.

46. Zhao R, Krauze AV. Survival prediction in gliomas: current state and novel approaches. In: Gliomas. Debinski W (ed.). Brisbane (AU), Exon Publications 2021.

47. Wiegrebe S, Kopper P, Sonabend R, Bischl B, Bender A. Deep learning for survival analysis: a review. Artif Intell Rev 2024; 57: 65.

48. Wang P, Fan E, Wang P. Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. Pattern Recognit Lett 2021; 141: 61-67.