UNIVERSITA’ POLITECNICA DELLE MARCHE

FACOLTA’ DI MEDICINA E CHIRURGIA

Tesi di Dottorato di Ricerca in Human Health

XXXVIII Ciclo

TITOLO

Intelligenza artificiale applicata alla radiomica: diagnostica differenziale dei tumori cerebrali intrassiali e risvolti clinici nella loro gestione neurochirurgica

TITLE

Artificial intelligence applied to radiomics: differential diagnosis of intra-axial brain tumors and clinical implications for their neurosurgical management

Tesi di dottorato di:	Tutor:
Dott. Denis Aiudi	Prof. Maurizio Iacoangeli

Anno Accademico 2024-2025

Sommario

1. INTRODUCTION

Glioblastoma (GBM) and primary central nervous system lymphoma (PCNSL) are high-grade brain tumors that frequently overlap in imaging appearance yet diverge profoundly in therapeutic management. For GBM, the contemporary standard pathway is maximal safe surgical resection followed by concurrent chemoradiotherapy and adjuvant temozolomide; by contrast, resection provides no routine therapeutic benefit for PCNSL, where tissue is typically obtained by stereotactic biopsy and first-line treatment is high-dose methotrexate–based chemotherapy, with consolidation strategies individualized to patient fitness and disease response [1–3]. Accurate noninvasive preoperative differentiation therefore carries immediate clinical consequences—avoiding unnecessary craniotomy for PCNSL and, conversely, preventing delays to definitive surgery for GBM [2,4].

Radiomics offers a quantitative framework to address this diagnostic gap by extracting high-dimensional descriptors of tumor intensity, texture, shape, and spatial heterogeneity from standard and advanced MRI sequences. Multiple single- and multiparametric pipelines have demonstrated strong discriminative performance for GBM versus PCNSL, with cross-validated AUCs commonly in the 0.90–0.97 range across combinations of contrast-enhanced T1-weighted, diffusion, FLAIR, and perfusion inputs [5–7]. Notably, carefully engineered feature selection and limited-sequence models (e.g., T1-CE with ADC) can match or approach the accuracy of broader multiparametric inputs, suggesting that robust signatures need not be sequence-prolific to be clinically useful [5]. Deep learning approaches trained on routine CE-T1 images have likewise achieved high accuracy and clinically relevant operating points, reinforcing the feasibility of decision support from widely available acquisitions [7].

At the same time, conventional qualitative markers—homogeneous enhancement, diffusion restriction, and low cerebral blood volume for PCNSL versus ring-like enhancement with central necrosis and elevated perfusion for GBM—remain imperfect in atypical or treated presentations, underscoring the need for quantitative augmentation [4,8]. Current evidence therefore positions radiomics (including deep learning radiomics) as a pragmatic bridge between morphology and molecular pathology, enabling probabilistic triage that can sharpen multidisciplinary decision-making at the point of care. The next translational steps will require harmonized acquisition, feature reproducibility across scanners and centers, interpretable models that expose pathophysiologically plausible predictors, and prospective validation embedded within neuro-oncologic workflows where the clinical endpoints are surgical decision-change, time-to-therapy, and patient outcomes [5–7]. Within such frameworks, radiomics-driven differential diagnosis has the potential to reduce invasive procedures for PCNSL, accelerate operative planning for GBM, and ultimately align patients with the right therapy at the right time [1–4].

In this study, we evaluated the discriminatory capability of a deep learning algorithm trained on gadolinium-enhanced T1-weighted MRI scans (T1w+Gd) to differentiate glioblastomas from PCNSL and to improve both diagnostic and interventional workflows.

1.1 MRI IN NEURORADIOLOGY

1.1.1 Basic principles

Magnetic resonance imaging (MRI) represents one of the most versatile, information-rich, and noninvasive tools in contemporary neuroscience. Its physical basis hinges on fundamental principles of nuclear magnetic resonance coupled with highly engineered strategies for spatial encoding and signal detection. Although MRI has been a clinical mainstay for decades, continual developments in gradient engineering, radiofrequency (RF) design, and reconstruction algorithms have significantly reshaped the landscape in the last 15 years.

At the core of MRI lies the magnetic moment of nuclear spins, primarily protons in biological tissue. When placed in a strong static magnetic field B0, these spins align preferentially along the field direction and precess at the Larmor frequency, $\omega 0 = \gamma B0$, where γ is the gyromagnetic ratio. Excitation with an RF field B1, oscillating at the Larmor frequency, perturbs the equilibrium magnetization and generates a measurable transverse magnetization component. The subsequent return to equilibrium can be described by the Bloch equations, which incorporate longitudinal (T1) and transverse (T2) relaxation. In neuroimaging applications, tissue-specific variations in water content, macromolecular composition, and microstructural anisotropy give rise to contrast mechanisms that encode anatomy and pathology. Recent work has refined the quantification of relaxation and susceptibility properties in gray matter, white matter, and subcortical nuclei, expanding the sensitivity of MRI to microenvironmental features previously considered inaccessible [9].

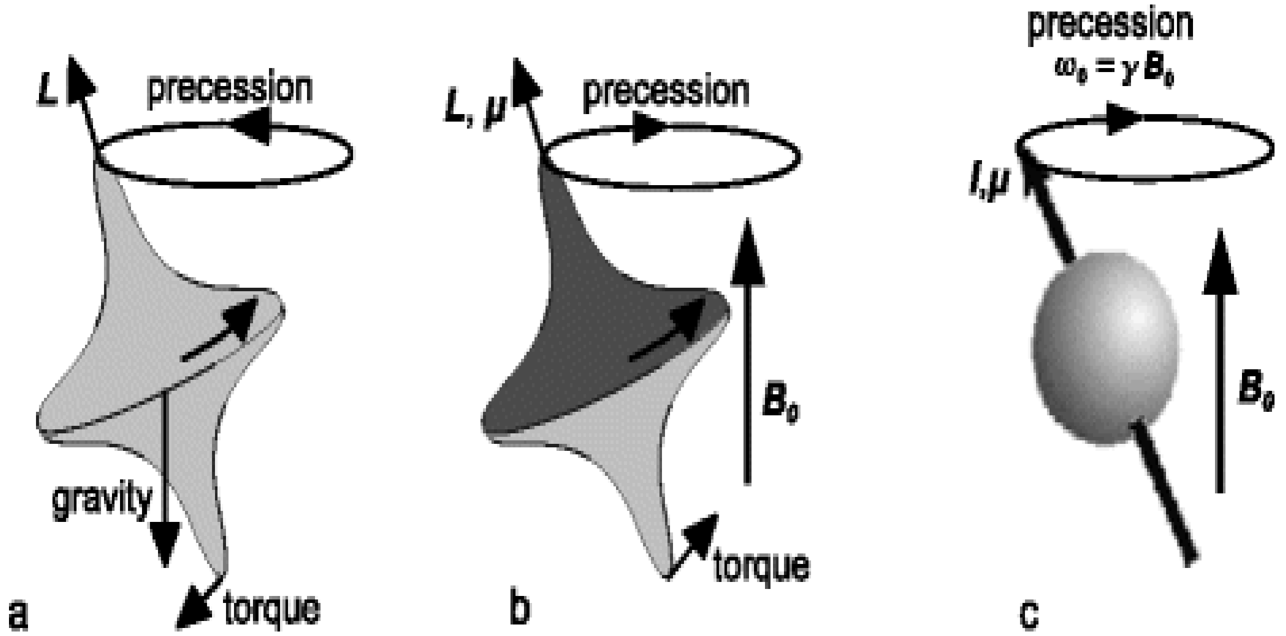


Figure 1.1 Magnetization and Larmor precession

Spatial encoding in MRI relies on magnetic field gradients that induce systematic variations in Larmor frequency across the imaging volume. By manipulating the temporal course of these gradients, MRI systems can sample k-space, the Fourier domain representation of the image. Echo planar imaging (EPI) remains central to functional MRI and diffusion MRI due to its rapid acquisition, although it is susceptible to susceptibility-induced distortions. Advances such as multiband excitation and parallel imaging have dramatically accelerated neuroimaging protocols. In particular, simultaneous multi-slice approaches combined with controlled aliasing have enabled whole-brain fMRI with sub-second temporal resolution, while maintaining image quality compatible with neuroscientific interpretation [10]. These innovations build on rigorous electromagnetic modeling that accounts for coil sensitivity distributions and gradient nonlinearity, further refining quantification and signal stability. Diffusion-weighted MRI (dMRI) introduces additional physics driven by Brownian motion of water molecules. The Stejskal–Tanner sequence uses diffusion gradients to attenuate signals proportionally to molecular displacement, enabling inference on tissue microstructure. In neuroscience, this allows characterization of axonal integrity, myelination, and tract topology. However, contemporary approaches extend beyond the assumptions of Gaussian diffusion. Multi-shell and multi-tensor models had better capture complex fiber geometries, while ultra-high-field MRI has revealed subtle features of diffusion kurtosis and microstructural anisotropy in cortical layers. Innovations in gradient strength, rising to 300 mT/m on advanced human scanners, have dramatically improved sensitivity to restricted diffusion, enabling interrogation of cellular-level features [11]. These advances highlight the symbiosis between fundamental physics and neurobiological interpretation.

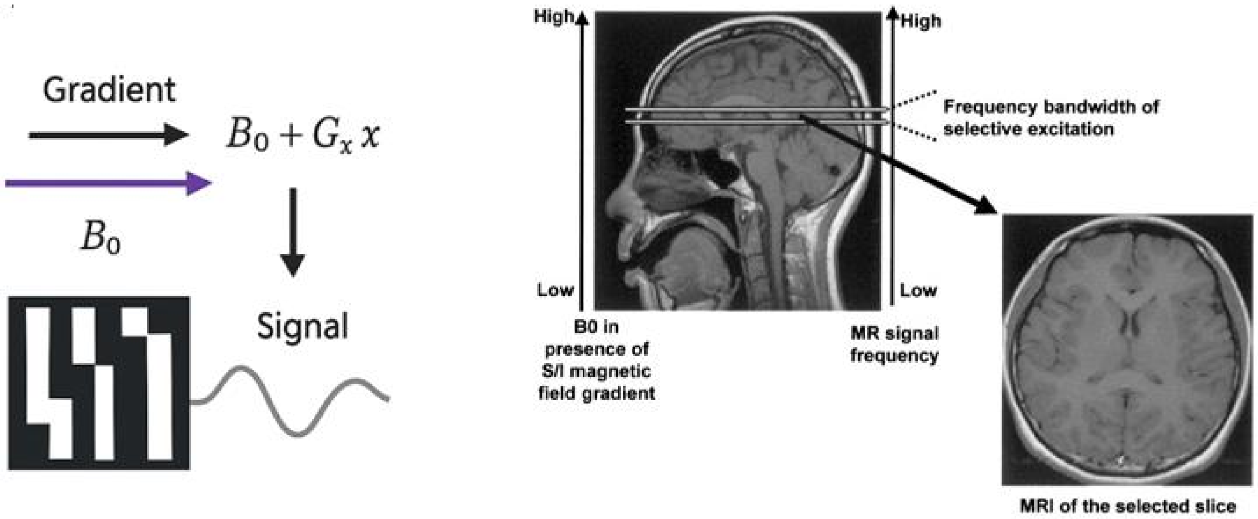


Figure 1.2 MRI image formation

Susceptibility-based contrast mechanisms have also gained prominence. Magnetic susceptibility reflects differences in the diamagnetic or paramagnetic properties of tissue components, including deoxyhemoglobin, iron, myelin, and calcium. Quantitative susceptibility mapping (QSM) reconstructs tissue susceptibility from gradient echo phase information. In the context of neuroscience, QSM has illuminated iron distribution in basal ganglia, cortical myeloarchitecture, and neurodegenerative signatures in Parkinson’s and Alzheimer’s disease. The mathematical inversion underlying QSM is ill-posed, requiring sophisticated regularization strategies and biophysical modeling. Recent studies have employed Bayesian inference and deep learning to improve stability and accuracy, while ultra-high-field acquisitions provide increased signal-to-noise ratio for more reliable susceptibility estimation [12].

High-field and ultra-high-field MRI, typically operating at 7 T and above, introduce additional physical considerations. While SNR increases approximately linearly with field strength, RF field inhomogeneity and susceptibility-induced artifacts also intensify. Novel RF pulse design, including parallel transmission, has been pivotal in correcting B1 inhomogeneity and managing specific absorption rate (SAR). Neuroimaging at these fields yields unprecedented spatial resolution capable of resolving cortical layers and subcortical microstructures. This has expanded the conceptual interface between MRI physics and cortical mesoscale organization, allowing functional laminar imaging and microvascular characterization that were previously impractical in humans [13]. Such developments illustrate how advances in electromagnetic engineering directly fuel neuroscientific discovery.

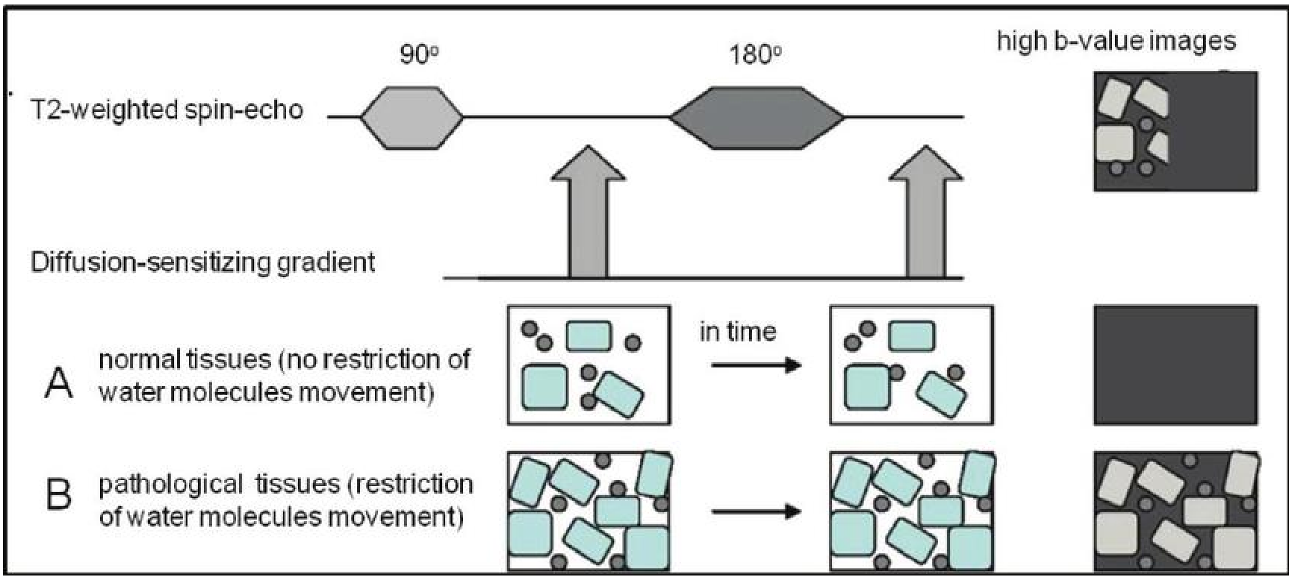


Figure 1.3 Diffusion gradient and diffusion weighting in MRI

MRI's physical underpinnings continue to evolve alongside computational strategies. Reconstruction methods based on compressed sensing, low-rank tensor models, and machine learning reduce noise, accelerate acquisition, and enhance contrast stability. Importantly, these approaches require rigorous validation to avoid hallucinated signal characteristics. Furthermore, MRI hardware innovations, including cryogenic RF coils and adaptive gradient arrays, have begun to redefine the possible contrast-to-noise and spatiotemporal resolutions in neuroimaging. The long-standing interplay between spin physics, signal processing, and neurobiology remains central to the field's rapid progression.

Despite its maturity, MRI remains fundamentally grounded in physical principles that are actively being reexamined and leveraged for new neuroscientific insights. As the field advances, deeper integration of quantum physics, computational modeling, and biological specificity will likely shape the next generation of MRI technologies, extending its reach from macroscopic anatomy to cellular-scale physiology. The past 15 years have demonstrated that MRI physics continues to be a driving force for innovation in neuroscience, reinforcing the modality's uniquely rich capacity for noninvasive brain investigation.

1.1.2 Role of Magnetic Resonance Imaging in the Diagnostic Evaluation of Brain Tumors

Magnetic resonance imaging remains the linchpin of modern neuro-oncology, providing unparalleled contrast resolution, multi-parametric tissue characterization, and advanced quantitative biomarkers. The heterogeneity of intracranial neoplasms requires an imaging framework capable of assessing cellularity, microstructure, vascularity, metabolic composition, and microhemorrhagic burden. Conventional and advanced MRI sequences each contribute distinct and complementary information, collectively enabling lesion detection, differential diagnosis, surgical planning, therapeutic monitoring, and prognostication. Recent work has demonstrated that multi-parametric MRI, when integrated with radiomics and machine-learning analysis, offers diagnostic accuracy approaching that of histopathology in selected applications [14]. Below, the principal MRI sequences used in brain tumor imaging are reviewed, with emphasis on their physical basis, key diagnostic features, and characteristic findings across tumor types.

T1-weighted Imaging

T1-weighted (T1w+Gd) sequences constitute a foundational component of neuro-oncologic imaging. They exploit differences in longitudinal relaxation times to depict anatomical structures with high spatial fidelity. Most intracranial tumors demonstrate iso- to hypointense signal on pre-contrast T1w imaging, reflecting increased water content, reduced lipid density, or necrotic components. Hyperintensity may appear in lesions with proteinaceous or hemorrhagic components. Post-gadolinium T1w imaging remains indispensable for evaluating blood–brain barrier disruption, a hallmark of high-grade gliomas, metastases, and lymphomas. Contrast enhancement patterns can distinguish ring-enhancing necrotic glioblastoma from solidly enhancing metastases or homogeneously enhancing primary CNS lymphoma. Texture descriptors extracted from post-contrast T1w sequences have shown strong prognostic associations in glioma genotyping and outcome prediction [15].

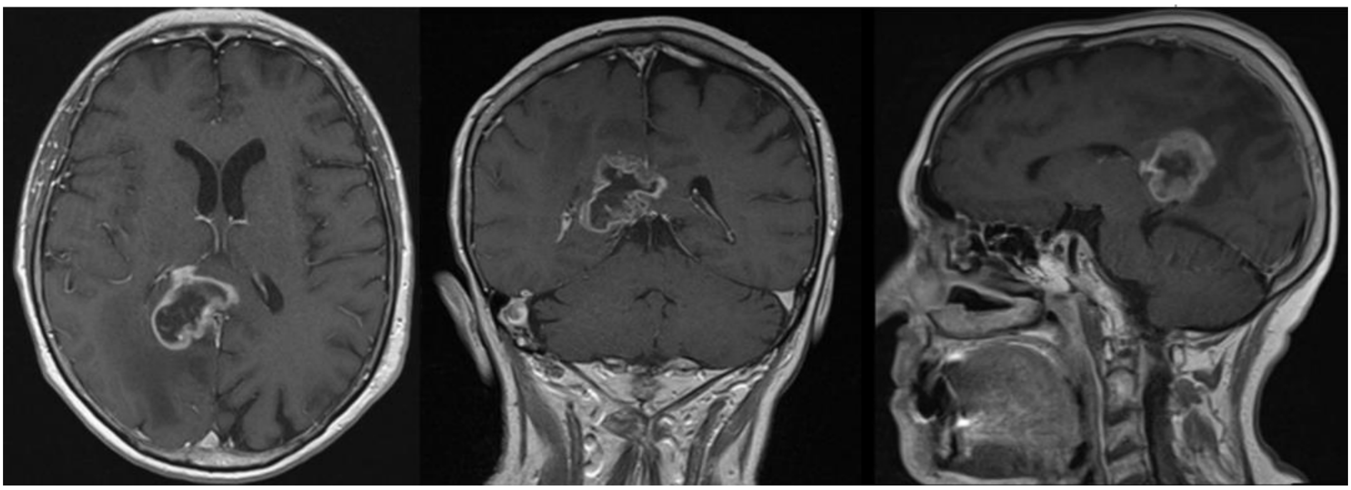


Figure 1.4 Post-gadolinium T1-weighted MRI GBM.

T2-weighted Imaging

T2-weighted (T2w) imaging is highly sensitive to increased free water content and thus delineates edema, cystic change, and tumor infiltration. High-grade gliomas typically exhibit heterogeneous hyperintensity due to necrosis and perifocal edema, while low-grade gliomas demonstrate more uniform T2 hyperintensity reflecting increased water mobility in infiltrated parenchyma. T2w signal characteristics assist in identifying non-enhancing but infiltrative tumors, which may escape detection on T1w post-contrast imaging. In clinical practice, T2w imaging is crucial for assessing lesion burden, identifying satellite foci, and mapping eloquent cortex involvement. Quantitative T2 mapping, increasingly investigated in the last decade, has shown potential for microstructural assessment and treatment response evaluation [16].

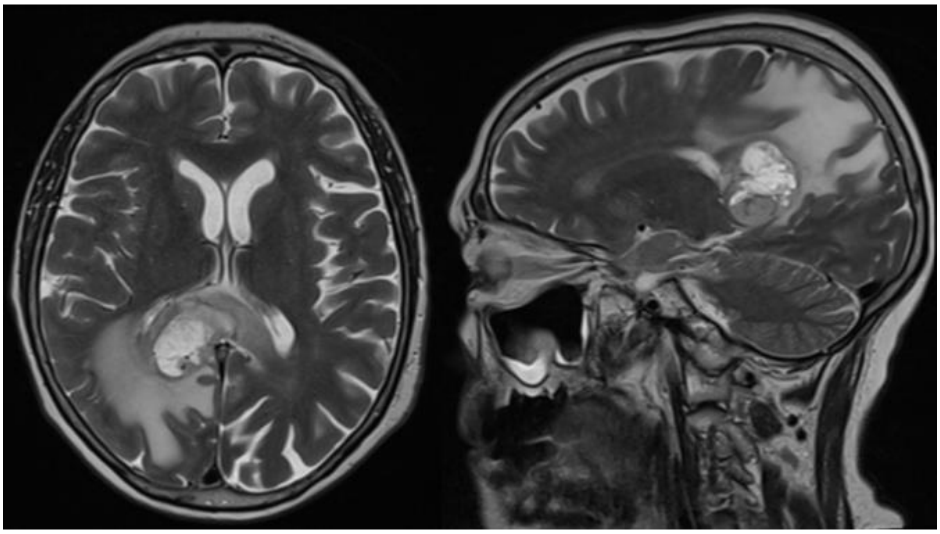


Figure 1.5 T2-weighted MRI GBM.

Fluid-Attenuated Inversion Recovery (FLAIR)

FLAIR sequences suppress cerebrospinal fluid (CSF) signal by applying inversion recovery, enabling enhanced visualization of periventricular and cortical abnormalities. In brain tumor imaging, FLAIR is particularly valuable for distinguishing infiltrative tumor from vasogenic edema. High-grade gliomas often behave as mixed infiltration–edema lesions, producing complex FLAIR hyperintensity patterns. Cortical or subcortical hyperintense lesions on FLAIR are strongly predictive of diffuse glioma infiltration, even in the absence of enhancement. Non-enhancing gliomas, including IDH-mutant lesions, are optimally assessed with FLAIR to determine true disease extent. Recent studies have demonstrated that quantitative FLAIR signal analysis improves performance in differentiating tumor progression from treatment-related changes such as pseudoprogression [17].

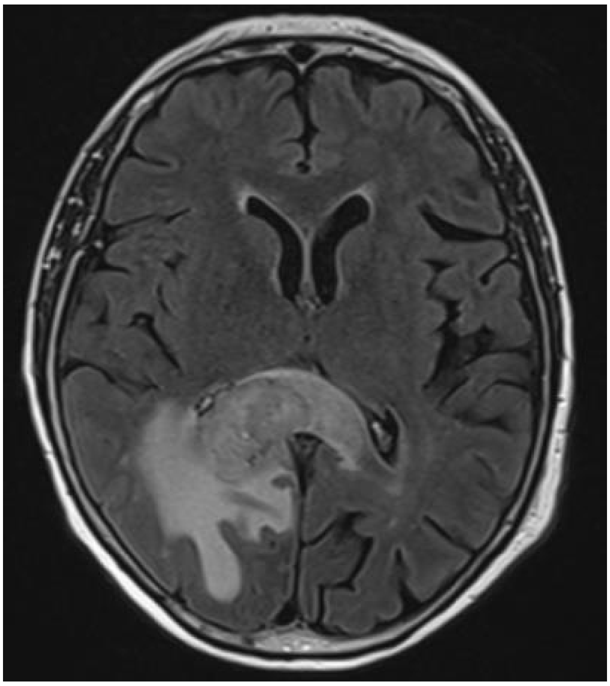


Figure1.6 FLAIR MRI GBM

Diffusion-Weighted Imaging (DWI)

Diffusion-weighted imaging provides critical insight into tissue microstructure by measuring the Brownian motion of water molecules. Tumors with high cellularity restrict water diffusion, generating hyperintense signal on DWI and reduced apparent diffusion coefficient (ADC) values. Primary CNS lymphomas classically show profound diffusion restriction, reflecting densely packed malignant cells. High-grade gliomas often exhibit heterogeneous ADC patterns correlating with proliferative activity. DWI is also essential in the evaluation of abscesses, which exhibit true restricted diffusion due to viscous purulent material. Radiogenomic studies have linked ADC metrics with IDH mutation status, MGMT methylation, and molecular subtypes of gliomas [18].

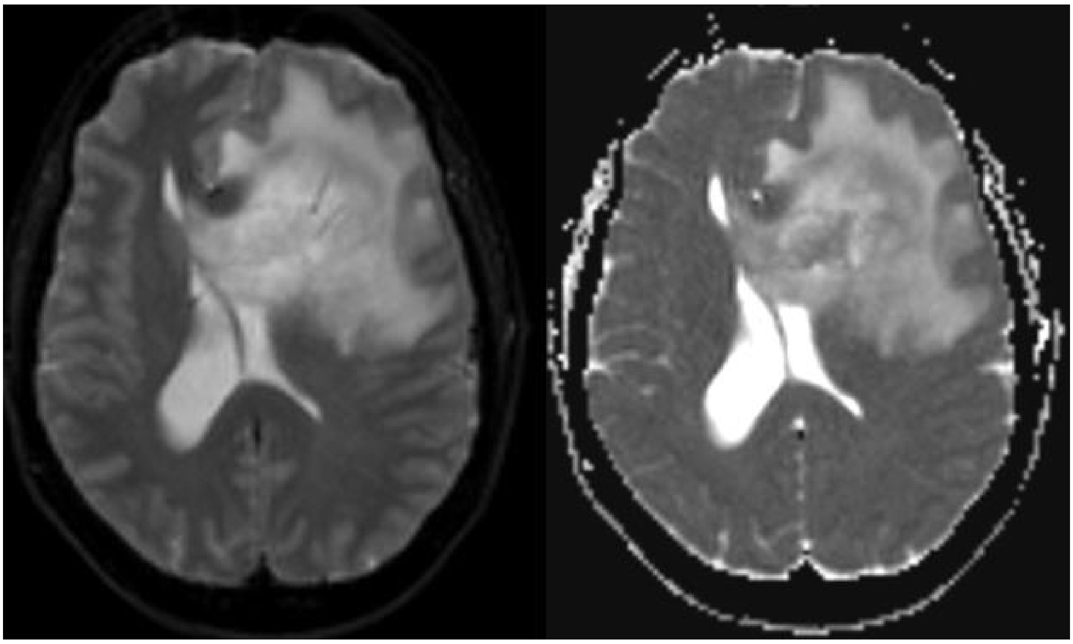


Figure 1.7 diffusion-weighted MRI GBM.

Susceptibility-Weighted Imaging (SWI)

SWI enhances sensitivity to magnetic susceptibility differences arising from paramagnetic blood products, calcifications, and venous structures. In brain tumors, SWI identifies intratumoral hemorrhage, microbleeds, and abnormal vasculature. Hemorrhagic metastases from melanoma, renal cell carcinoma, or choriocarcinoma demonstrate striking SWI signal loss. Glioblastomas often show intralesional susceptibility effects representing microvascular proliferation or hemorrhagic necrosis. The presence and distribution of susceptibility foci have diagnostic value in differentiating oligodendrogliomas (frequent calcifications) from astrocytomas. Advanced SWI techniques, including quantitative susceptibility mapping (QSM), provide additional insights into iron deposition and neovascular architecture [19].

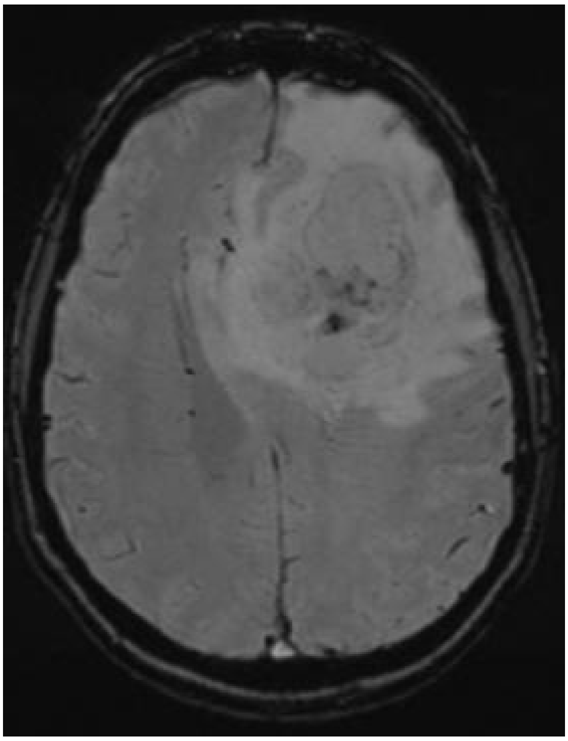


Figure 1.8 SWI MRI GBM.

Perfusion MRI

Perfusion MRI sequences quantify tumor hemodynamics and microvascular proliferation, critical hallmarks of malignancy. Dynamic susceptibility contrast (DSC) perfusion, the most widely utilized method, measures relative cerebral blood volume (rCBV), which strongly correlates with tumor grade. High-grade gliomas exhibit markedly elevated rCBV due to angiogenesis, whereas low-grade gliomas show mild increases or normal perfusion. Dynamic contrast-enhanced (DCE) imaging evaluates vascular permeability parameters such as Ktrans, which reflect blood–brain barrier breakdown. Arterial spin labeling (ASL), a non-contrast technique, provides cerebral blood flow estimates valuable in patients with renal impairment. Perfusion imaging plays a fundamental role in distinguishing pseudoprogression from true progression in glioma patients receiving chemoradiation [20].

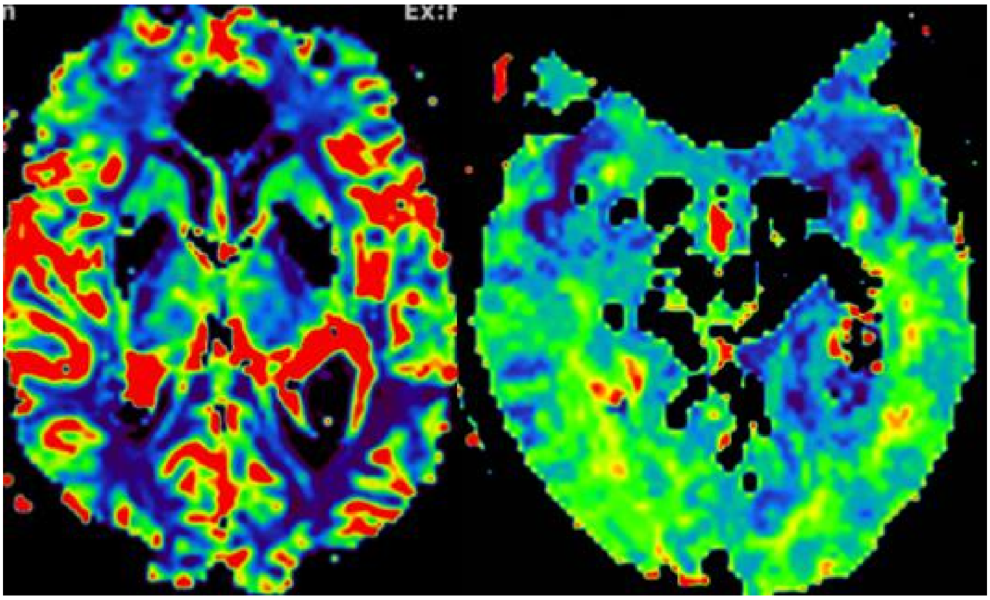


Figure 1.9 Perfusion MRI GBM.

MR Spectroscopy (MRS)

MR spectroscopy provides a metabolic fingerprint of intracranial tumors by quantifying resonances of choline, N-acetylaspartate (NAA), creatine, lactate, lipids, and other metabolites. Elevated choline and reduced NAA ratios are characteristic of neoplastic processes, reflecting membrane turnover and neuronal loss. High-grade gliomas frequently exhibit lipid and lactate peaks representing necrosis and anaerobic metabolism. MRS can differentiate recurrent tumor from treatment-induced necrosis, where the latter typically shows reduced choline and prominent lipid peaks. Recent advances in multi-voxel MRS and ultra-high-field spectroscopy have enhanced spatial resolution and metabolite discrimination, supporting integration of metabolic profiling with radiomic signatures [21].

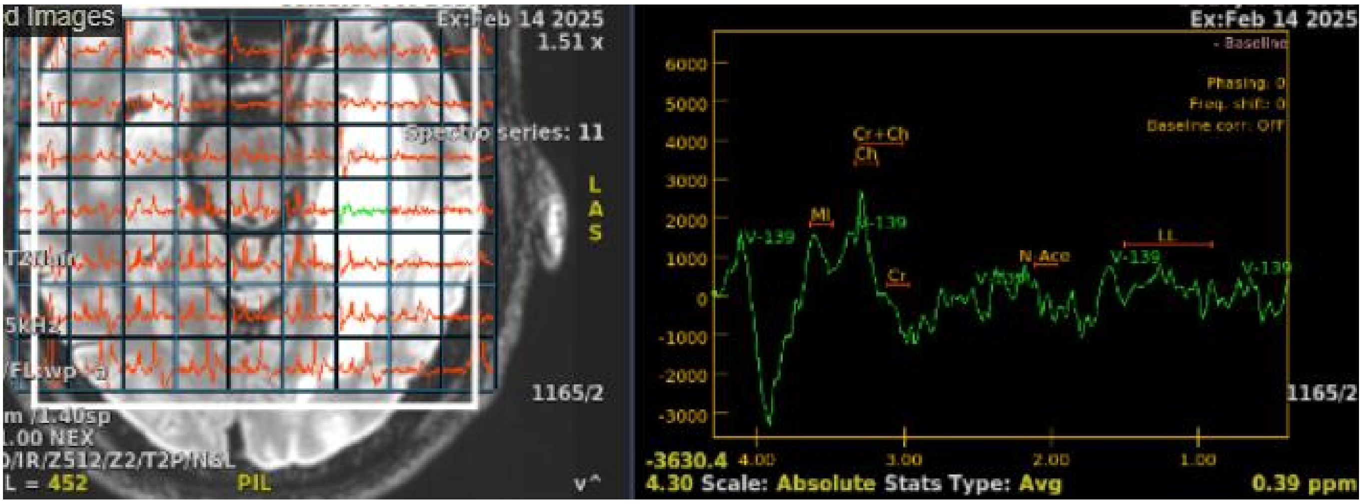


Figure 1.10 Schematic representation of an MR spectroscopy GBM.

Collectively, these MRI sequences form an integrative diagnostic framework that characterizes brain tumors across structural, microstructural, vascular, and metabolic domains. Contemporary studies highlight that multi-parametric MRI, particularly when combined with machine-learning classifiers, yields diagnostic accuracy exceeding 90 percent for several tumor categories and assists in molecular stratification. As neuro-oncology evolves toward precision medicine, the role of MRI continues to expand, with quantitative biomarkers, advanced contrast mechanisms, and artificial intelligence-driven analytics shaping the future of tumor characterization.

1.2 GLIOBLASTOMA

1.2.1 Epidemiology and Risk Factors

Glioblastoma (GBM) represents the most prevalent and biologically aggressive primary malignant brain tumor in adults, accounting for approximately 45 to 55 percent of all malignant gliomas worldwide [22]. Its incidence ranges between 3 and 5 cases per 100,000 individuals per year in North America and Europe, with notable geographic variability suggesting contributions from genetic ancestry and environmental determinants [23]. Incidence increases sharply after the age of 55, peaking in the seventh and eighth decades of life, and shows a persistent male predominance with a male-to-female ratio of approximately 1.6:1 [24]. Pediatric cases are exceedingly rare and demonstrate distinct molecular signatures relative to adult-onset disease [25].

The only well-established environmental risk factor for GBM is exposure to high-dose ionizing radiation, particularly in individuals treated for childhood malignancies or cranial conditions requiring therapeutic irradiation [26]. Most epidemiological studies have failed to demonstrate consistent associations between GBM and lifestyle-related or occupational exposures, including smoking, alcohol consumption, pesticides, or heavy metal exposure [27]. Similarly, although extensive public debate has focused on electromagnetic fields and mobile phone use, large prospective cohort studies have not shown a reproducible increase in GBM risk attributable to these exposures [28].

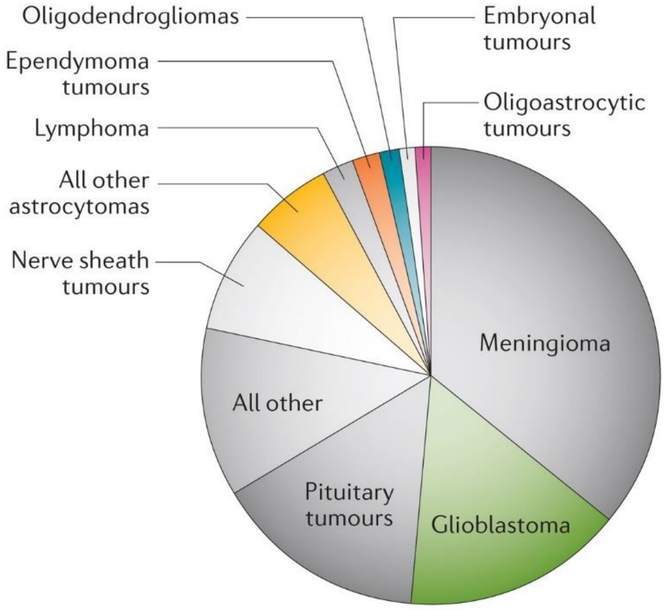


Figure 1.11 Brain tumors epidemiology

Host-related and genetic factors play a more substantial role. Rare familial cancer syndromes, including Li-Fraumeni syndrome, Turcot syndrome, and neurofibromatosis type 1, markedly increase the lifetime risk of gliomagenesis due to inherited mutations affecting DNA repair and tumor suppressor pathways [29]. Genome-wide association studies have identified several susceptibility loci linked to glioma risk, including variants in TERT, RTEL1, EGFR, CCDC26, and PHLDB1, though the absolute increase in risk conferred by each allele is modest [30]. Epigenetic predispositions, including inherited methylation patterns and transcriptional regulatory states, are increasingly recognized as modulators of glioma susceptibility [31].

1.2.2 Pathogenesis

A network of converging molecular derangements affecting genomic stability, cell-cycle regulation, metabolic homeostasis, and microenvironmental interactions drives the pathogenesis of GBM. Canonical genetic events include amplification or mutation of receptor tyrosine kinases (most commonly EGFR), activation of downstream RAS/RAF/MEK and PI3K/AKT/mTOR signaling pathways, and loss of tumor suppressor function through TP53, PTEN, and RB1 alterations [32]. IDH-wildtype GBM, which constitutes the vast majority of adult cases, typically arises de novo and is characterized by hallmark genomic alterations such as TERT promoter mutations and chromosome 7 gain with chromosome 10 loss [33]. In contrast, IDH-mutant glioblastomas evolve through progressive malignant transformation of lower-grade astrocytomas and exhibit a distinct molecular evolution trajectory associated with better prognosis [34].

Epigenetic regulation plays a central role in tumor progression and therapeutic resistance. Global alterations in DNA methylation patterns, dysregulation of chromatin remodeling complexes, and aberrant transcription factor activity enable tumor cells to maintain a stem-like phenotype with enhanced proliferative capacity and resilience to metabolic and genotoxic stress [35]. MGMT promoter methylation, a clinically significant epigenetic biomarker, modulates DNA repair capacity and strongly influences responsiveness to alkylating chemotherapy [36].

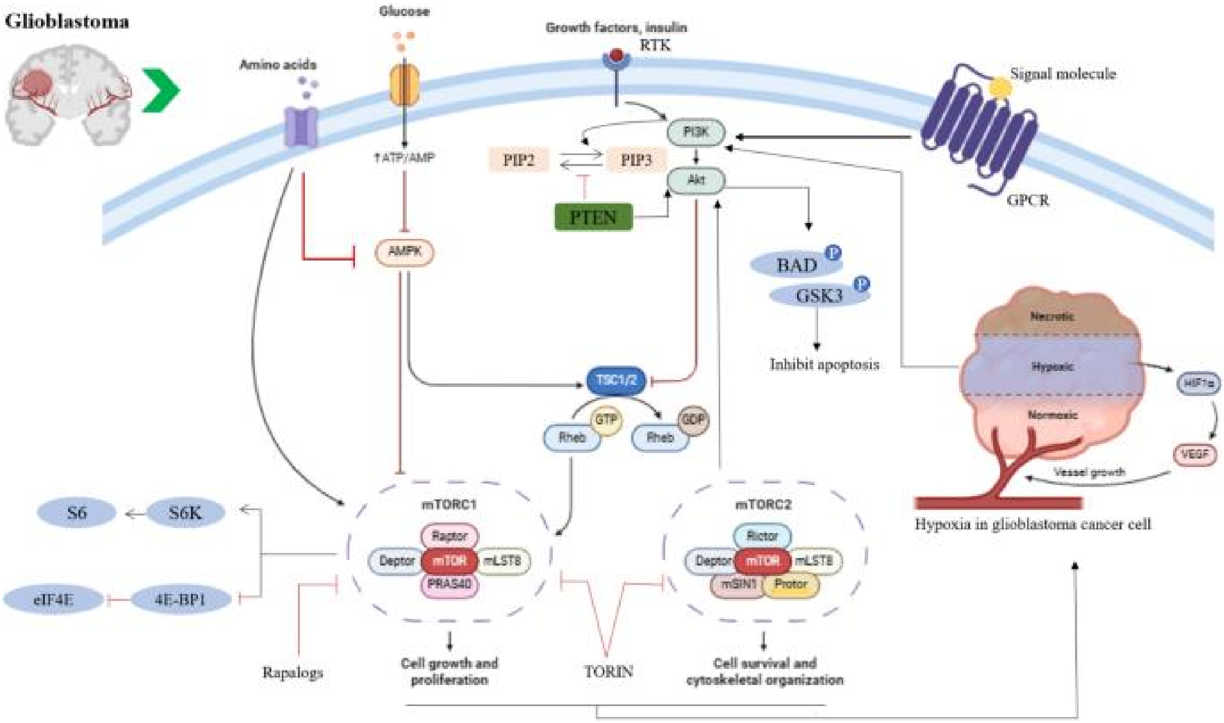


Figure 1.12 GBM Pathogenesis

The GBM microenvironment is characterized by profound immunosuppression, driven by widespread infiltration of tumor-associated macrophages and microglia, secretion of immunomodulatory cytokines, and upregulation of immune checkpoint molecules [37]. Extensive neoangiogenesis, mediated primarily by VEGF overexpression, supports rapid tumor expansion but also generates hypoxic niches that promote metabolic reprogramming toward glycolytic phenotypes and exacerbate treatment resistance [38]. These interlinked biological processes facilitate the tumor's aggressive infiltrative behavior and its notorious capacity for recurrence despite multimodal therapy [39].

1.2.3 Clinical Symptoms

Clinical presentation reflects tumor location, growth kinetics, and the extent of surrounding vasogenic edema. Patients commonly exhibit progressive focal neurological deficits, including motor weakness, sensory disturbances, aphasia, or visual field defects, depending on the affected cortical and subcortical circuitry [40]. Seizures occur in approximately 20 to 30 percent of newly diagnosed patients, particularly when tumors arise in the temporal or frontal lobes [41]. Symptoms of increased intracranial pressure, including severe headaches, nausea, vomiting, and papilledema, often emerge as mass effect and midline shift progress [42]. Cognitive impairment, personality changes, and executive dysfunction are frequently observed and may constitute early manifestations in tumors affecting frontal or limbic networks [43]. Rarely, intratumoral hemorrhage results in hyperacute neurological deterioration necessitating emergent intervention [44].

1.2.4 Diagnosis

Diagnosis integrates clinical evaluation, MRI, and histopathological confirmation.

Magnetic Resonance Imaging

MRI remains the gold standard imaging modality in GBM evaluation. On pre-contrast T1-weighted sequences, GBM typically appears hypointense relative to cortical gray matter due to its high cellularity and infiltrative architecture [45]. Post-contrast T1-weighted imaging demonstrates irregular ring-like or heterogeneous enhancement, reflecting neovascular proliferation and blood-brain barrier disruption [46]. T2-weighted and FLAIR sequences reveal extensive hyperintense peritumoral signal attributable to vasogenic edema and infiltrating tumor cells, frequently extending beyond the enhancing margins [47].

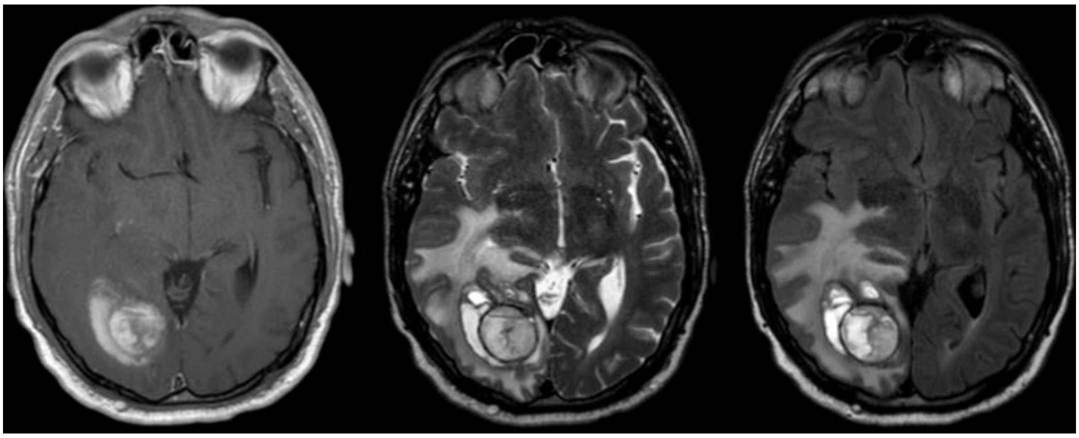


Figure 1.13 T1, T2, FLAIR MRI scans in GBM

Diffusion-weighted imaging (DWI) frequently demonstrates areas of restricted diffusion corresponding to hypercellular regions, whereas apparent diffusion coefficient (ADC) maps show reduced values in aggressive tumor cores [48]. Perfusion-weighted imaging identifies markedly elevated relative cerebral blood volume (rCBV), consistent with microvascular proliferation, and serves as a valuable biomarker for distinguishing GBM from lower-grade gliomas or treatment-related changes such as pseudoprogression [49].

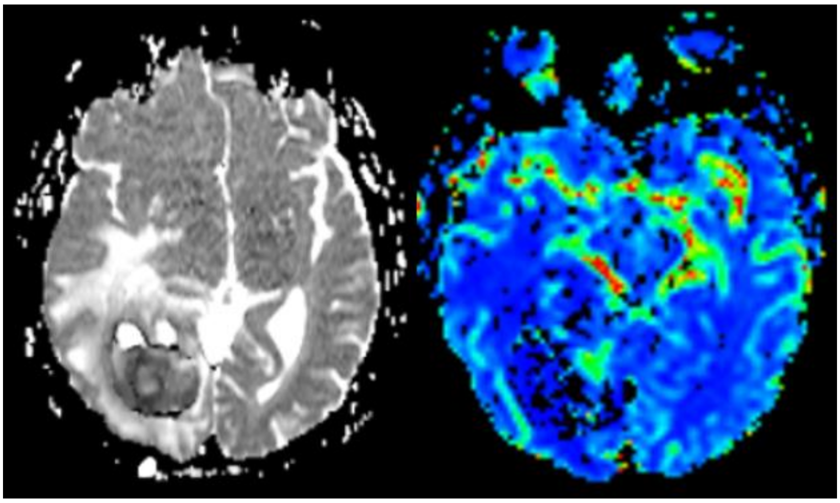


Figure 1.14 MRI DWI and perfusion in GBM

MR spectroscopy provides metabolic characterization, typically demonstrating elevated choline peaks, reduced N-acetylaspartate, and prominent lipid-lactate signals indicative of membrane turnover, neuronal loss, and necrosis [50]. Advanced functional and structural imaging modalities, including diffusion tensor imaging and amino acid PET, offer additional insights but remain adjunctive rather than diagnostic [51].

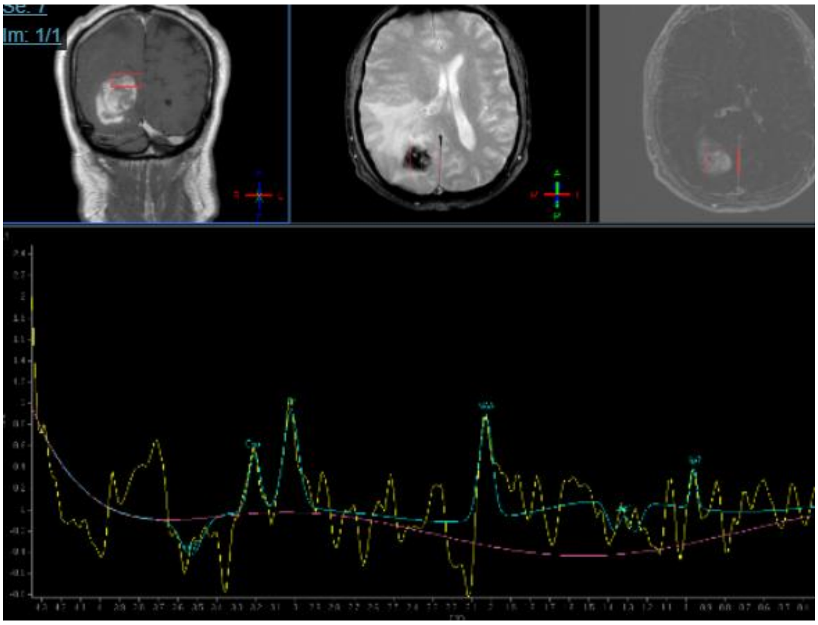


Figure 1.15 MRI Spectroscopy in GBM

Despite its robust diagnostic performance, MRI cannot reliably differentiate GBM from certain metastatic lesions, abscesses, or tumefactive demyelinating processes, thus necessitating histopathological confirmation through stereotactic biopsy or surgical resection [52].

Pathology

Histopathological evaluation remains the definitive diagnostic standard. GBM exhibits hallmark features including dense cellularity, marked nuclear atypia, brisk mitotic activity, microvascular proliferation, and necrosis with pseudopalisading patterns [53]. Immunohistochemical analysis provides additional diagnostic and prognostic information, including GFAP expression patterns, ATRX status, p53 accumulation, and proliferation indices such as Ki-67 [54].

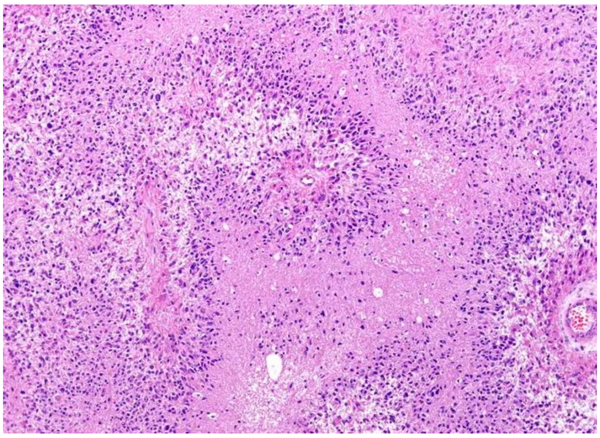


Figure 1.16 Pseudopalisading pattern in GBM

Modern classification systems integrate molecular profiling to distinguish IDH-wildtype from IDH-mutant glioblastomas, identify TERT promoter mutations, evaluate MGMT promoter methylation status, and assess chromosomal alterations such as +7/−10 signatures [55]. These molecular features correlate strongly with clinical behavior, therapeutic responsiveness, and survival outcomes [56].

1.2.5 Chemotherapeutic Treatment

Temozolomide in combination with radiotherapy remains the backbone of first-line treatment for newly diagnosed GBM, improving median survival and progression-free intervals relative to radiotherapy alone [57]. The therapeutic efficacy of temozolomide is significantly enhanced in patients with MGMT promoter methylation, which impairs tumor cell DNA repair capacity and increases susceptibility to alkylation-induced damage [58]. In the recurrent setting, alkylating agents such as lomustine, procarbazine, and carmustine are utilized, although clinical benefit is typically modest [59].

Targeted therapies, including EGFR inhibitors and antiangiogenic agents such as bevacizumab, have demonstrated radiographic improvements but limited impact on overall survival [60]. Current research focuses on exploiting vulnerabilities in DNA damage response pathways, metabolic dependencies, and immune regulation to develop more effective therapeutic strategies [61].

1.2.6 Radiotherapeutic Treatment

Radiotherapy is integral to standard GBM management, typically administered as fractionated external beam radiation to a total dose of 60 Gy over six weeks [62]. Advanced techniques such as intensity-modulated radiotherapy and image-guided radiotherapy allow precise conformal delivery that spares surrounding eloquent cortex and minimizes neurotoxicity [63]. Hypofractionated regimens have demonstrated comparable efficacy in elderly or frail patients and reduce treatment burden [64].

Biological factors including tumor hypoxia, stem cell niches, and intrinsic DNA repair proficiency contribute to radioresistance and serve as targets for ongoing investigations into radiosensitizers and dose-escalation protocols [65]. Integrating radiotherapy with targeted agents and immunomodulatory approaches remains an active area of translational research [66].

1.2.7 Surgical Treatment

Surgical intervention aims to achieve maximal safe resection, providing cytoreduction, symptomatic improvement, and tissue acquisition for diagnosis and molecular profiling [67]. Extent of resection is strongly correlated with survival, with gross total resection yielding significant prognostic advantage when achievable without compromising neurological function [68].

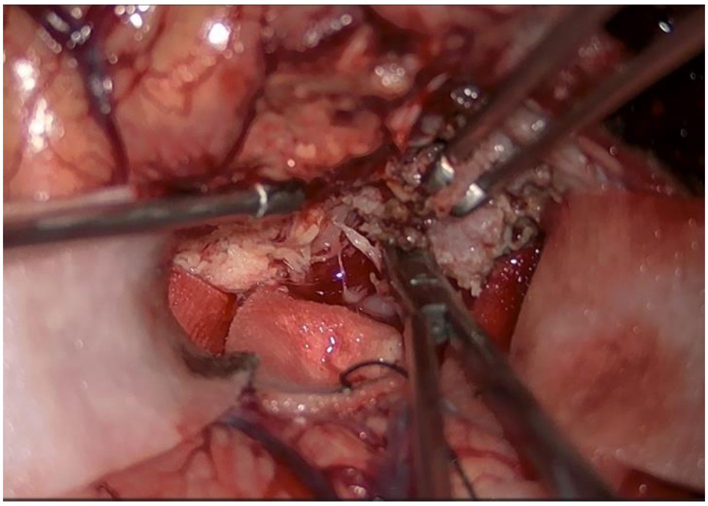


Figure 1.17 Intraoperative image of a GBM during its surgical removal.

Contemporary neurosurgical techniques, including intraoperative MRI, 5-ALA fluorescence-guided resection, functional mapping, and neuronavigation, enhance the precision and safety of tumor removal [69]. Despite these advances, complete resection remains unattainable due to microscopic infiltration extending beyond radiographically visible margins, necessitating integration with adjuvant chemoradiotherapy [70].

1.3 PRIMARY CENTRAL NERVOUS SYSTEM LYMPHOMAS

1.3.1 Epidemiology and Risk Factors

Primary central nervous system lymphoma (PCNSL) is an uncommon extranodal non-Hodgkin lymphoma representing 2 to 4 percent of all primary brain tumors, with an incidence estimated at 0.4 to 0.5 cases per 100,000 persons annually in Western countries [71]. Incidence increased notably during the late twentieth century, largely driven by the HIV epidemic and the expanding population of immunosuppressed individuals, though it has since stabilized among immunocompetent patients [72]. PCNSL shows a modest male predominance, with peak incidence between 55 and 65 years of age [73].

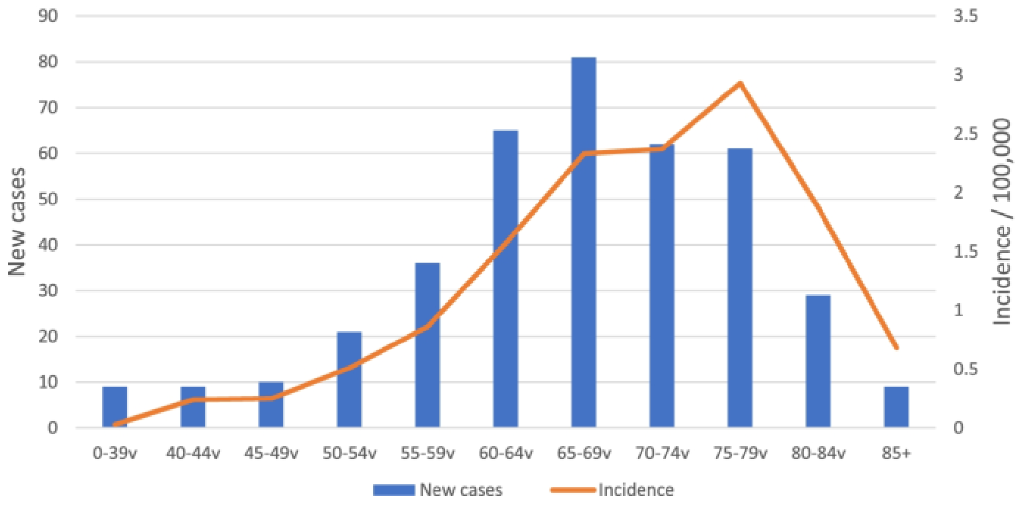


Figure 1.18 Number of cases by age group and age specific incidence rates of PCNSL in 2007-2017

Immunodeficiency remains the most significant risk factor, as demonstrated in HIV-positive patients, post-transplant recipients, and individuals receiving chronic immunosuppressive therapy [74]. Epstein–Barr virus (EBV) plays a central role in lymphomagenesis in immunocompromised hosts, with EBV-driven lymphomas constituting the majority of AIDS-related PCNSL [75]. Conversely, EBV is rarely detected in immunocompetent PCNSL, suggesting divergent pathogenic pathways. Other proposed risk factors include immune senescence, chronic antigenic stimulation, and autoimmune diseases, though their contributions remain uncertain [76].

1.3.2 Pathogenesis

PCNSL most commonly arises as diffuse large B-cell lymphoma (DLBCL), typically of the activated B-cell phenotype. The molecular landscape is dominated by genetic alterations that promote chronic B-cell receptor activation and downstream NF-kB signaling. MYD88 L265P and CD79B mutations are among the most frequent and form a cooperative pathogenic axis enabling constitutive pro-survival signaling [77]. These alterations coexist with dysregulation of chromatin remodeling, antigen presentation, and apoptotic pathways, contributing to immune escape and treatment resistance [78].

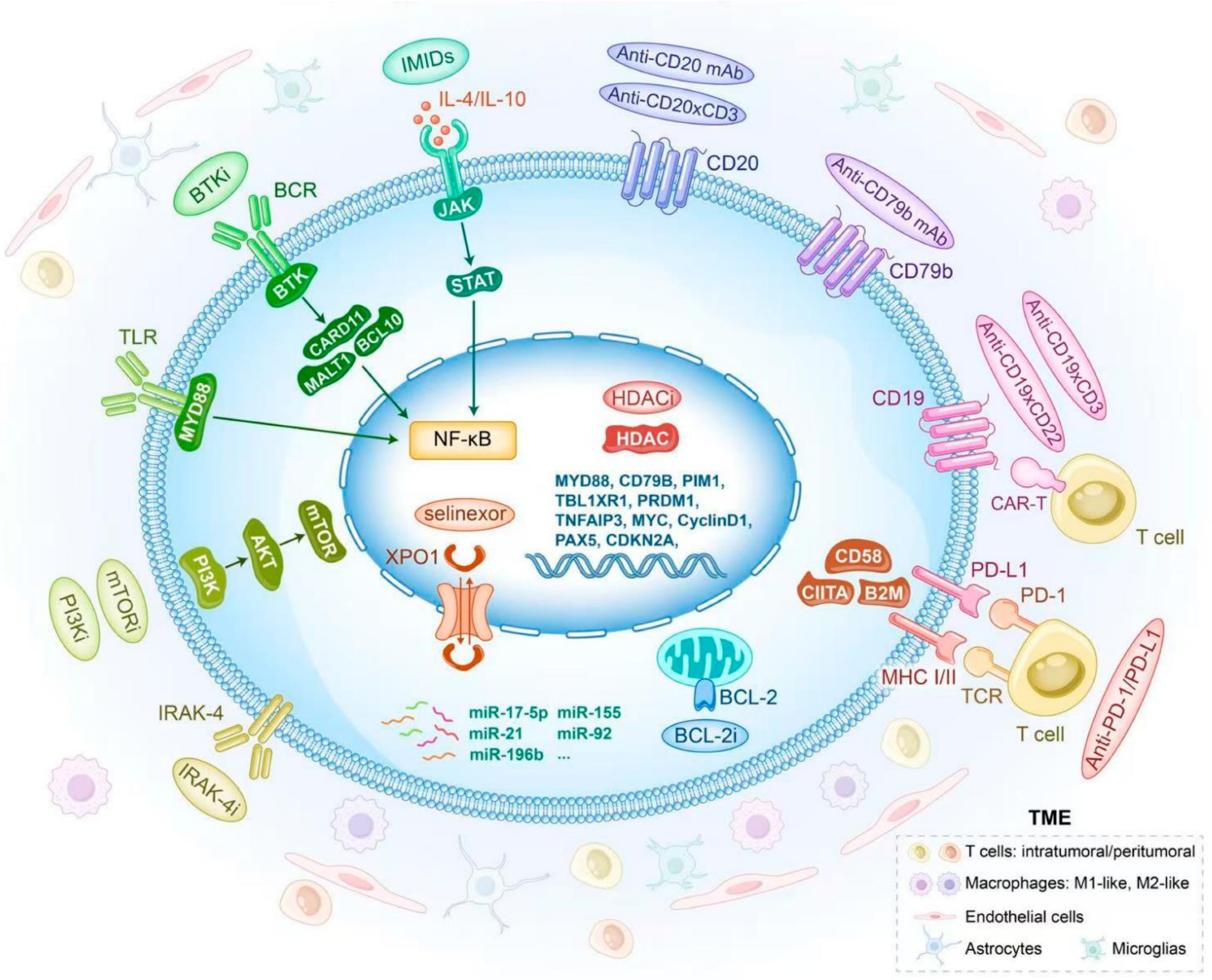


Figure 1.19 PCNSL Pathogenesis and therapeutic targets

The hallmark angiocentric growth pattern results from interactions between tumor cells and vascular basement membranes, facilitated by chemokines, adhesion molecules, and stromal cues unique to the CNS milieu [79]. The immunoprivileged CNS microenvironment, characterized by low MHC expression and reduced T-cell surveillance, further promotes lymphoma survival and expansion [80]. The blood–brain barrier additionally restricts exposure to systemic therapies, shaping both disease behavior and therapeutic outcomes [81].

1.3.3 Clinical Symptoms

PCNSL typically presents with subacute neurologic decline evolving over weeks. Cognitive impairment, neuropsychiatric symptoms, and behavioral changes are prominent due to preferential involvement of frontal, periventricular, and deep gray matter structures [82]. Focal neurological deficits such as hemiparesis, aphasia, cranial neuropathies, or ataxia may arise depending on lesion topography [83]. Visual symptoms occur in cases of ocular or optic pathway involvement. Seizures are less common than in gliomas, affecting approximately 10 to 20 percent of patients [84]. Increased intracranial pressure may result from mass effect or obstructive hydrocephalus. A well-known clinical feature is transient improvement following corticosteroid administration, attributable to rapid cytolysis and reduction of peritumoral edema [85].

1.3.4 Diagnosis

Definitive diagnosis requires a combination of neuroimaging, histopathological analysis, and, when indicated, CSF evaluation.

Magnetic Resonance Imaging

MRI plays a central diagnostic role. PCNSL lesions are characteristically iso- to hypointense on T1-weighted imaging and demonstrate intense, homogeneous gadolinium enhancement in immunocompetent patients, reflecting dense cellularity and lack of necrosis [86]. T2-weighted and FLAIR sequences typically show iso- to mildly hyperintense lesions with relatively limited edema compared to high-grade gliomas [87].

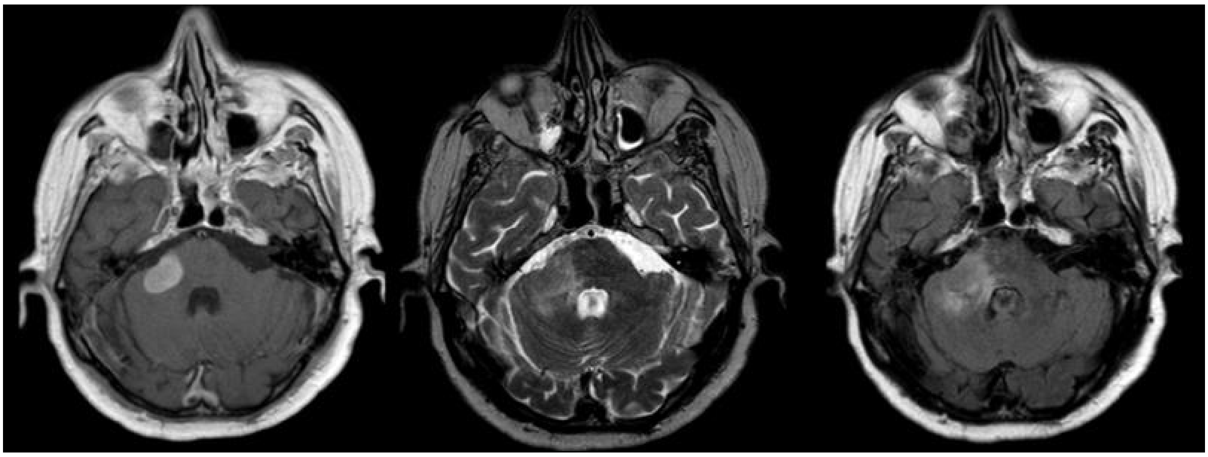


Figure 1.20 T1, T2, FLAIR MRI scans in PCNSL

Diffusion-weighted imaging (DWI) reveals marked restriction, with low apparent diffusion coefficient (ADC) values corresponding to hypercellular tumor regions, a key feature aiding in distinction from glioblastoma [88]. Perfusion MRI generally demonstrates low to moderate relative cerebral blood volume (rCBV), contrasting with the elevated perfusion of hypervascular gliomas [89]. MR spectroscopy often shows elevated lipid-lactate peaks, increased choline, and reduced N-acetylaspartate, reflecting membrane turnover and neuronal loss, though findings may overlap with other malignant lesions [90].

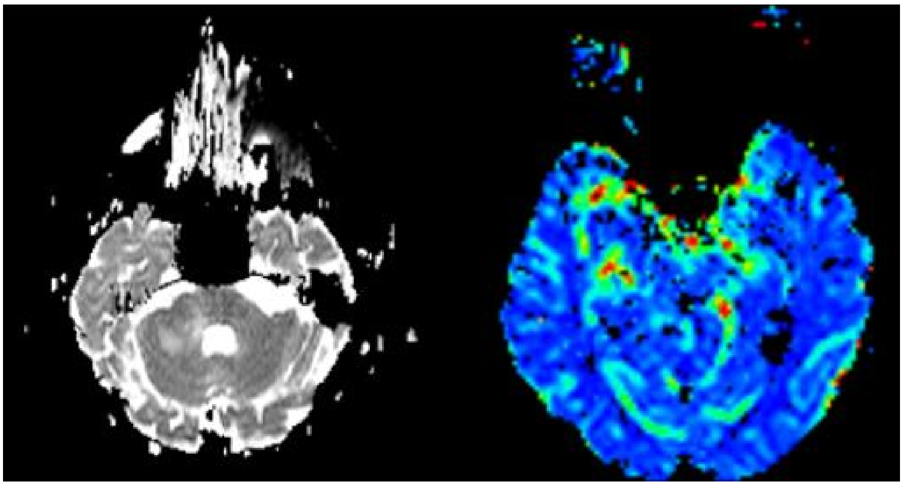


Figure 1.21 MRI DWI and perfusion in PCNSL

Leptomeningeal dissemination may manifest as linear or nodular contrast enhancement and should prompt CSF analysis when safe. Pre-biopsy corticosteroid administration should be minimized, as it can markedly reduce enhancement and alter histopathologic findings, leading to nondiagnostic samples [91].

Pathology

Stereotactic biopsy remains the diagnostic gold standard. PCNSL typically exhibits diffuse sheets of large atypical B cells with prominent nucleoli, abundant mitotic figures, and characteristic angiocentric infiltration [92]. Immunohistochemistry demonstrates expression of CD20, MUM1, BCL6, and BCL2, supporting an activated B-cell phenotype [93]. Molecular studies frequently detect MYD88 and CD79B mutations, supporting their pathogenic significance and, increasingly, their diagnostic value [94]. CSF cytology is supportive but insensitive and is primarily useful in cases of leptomeningeal involvement [95].

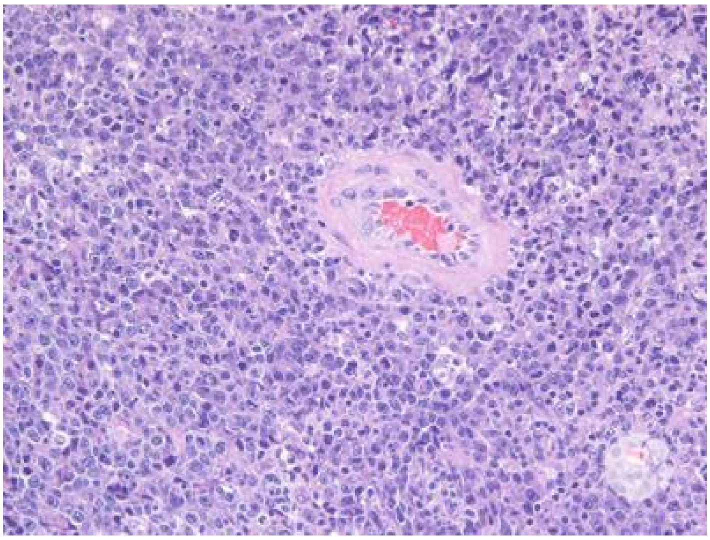


Figure 1.22 Primary CNS diffuse large B-cell lymphoma

1.3.5 Chemotherapeutic Treatment

High-dose methotrexate (HD-MTX) is the backbone of modern PCNSL therapy, achieving superior outcomes compared to historical radiotherapy-alone approaches and forming the basis of current induction regimens [96]. Combination therapy with HD-MTX, cytarabine, rituximab, and thiotepa improves response durability, though optimal induction strategies continue to evolve [97]. Despite limited CNS penetration, rituximab appears to enhance therapeutic efficacy by targeting circulating and perivascular lymphoma cells [98]. Consolidation with high-dose chemotherapy followed by autologous stem cell transplantation has shown promising long-term survival in selected patients, offering an alternative to whole-brain radiotherapy [99]. In relapsed or refractory cases, targeted agents such as ibrutinib and lenalidomide demonstrate activity, though responses are variable and often transient [100].

1.3.6 Radiotherapeutic Treatment

Whole-brain radiotherapy (WBRT), once used as primary therapy, is now primarily reserved for consolidation or salvage treatment due to risk of delayed neurotoxicity, particularly in older adults [101]. Reduced-dose WBRT following effective induction chemotherapy may achieve disease control while minimizing cognitive decline [102]. Advanced modalities including hippocampal-sparing WBRT and intensity-modulated radiotherapy aim to preserve neurocognitive function by limiting dose to memory-critical regions [103]. Radiotherapy alone is generally inadequate for durable remission in immunocompetent patients and is not recommended as monotherapy [104].

1.3.7 Surgical Treatment

Surgical resection plays a limited role in PCNSL due to its diffuse, infiltrative biology and angiocentric growth pattern. Gross total resection does not improve survival and carries unnecessary risk; thus, biopsy alone is recommended for diagnosis in nearly all cases [105]. Surgery may be indicated only in exceptional circumstances, such as life-threatening mass effect or hemorrhage [106]. Advances in stereotactic neurosurgical techniques have improved safety and precision, allowing diagnostic sampling even from deep or eloquent brain regions [107].

1.4 ARTIFICIAL INTELLIGENCE (AI)

1.4.1 Machine Learning

An Overview of Artificial Intelligence: History and Definitions

The notion of a thinking machine capable of autonomous learning and adaptation, comparable to human cognition, has fascinated humanity since ancient Greece. For centuries this concept remained largely utopian (and, for some, dystopian), until the past century, with the most rapid advances occurring over the last decade across both scientific and non-scientific domains.

A formal definition of artificial intelligence (AI) emerged only in 1950, when the British mathematician Alan Turing proposed criteria for machine intelligence in his seminal paper *Computing Machinery and Intelligence* [90]. He introduced what is now known as the “Turing Test” (or “imitation game”), a conceptual framework for determining whether a machine’s behavior can be indistinguishable from that of a human. Turing's ideas reflected a scientific environment increasingly interested in the possibility of artificial cognition, inspiring subsequent generations of researchers.

The academic birth of AI is typically traced to the 1956 Dartmouth Conference, attended by leading scientists including Claude Shannon, Marvin Minsky, and John McCarthy [91]. They hypothesized that computers could be programmed to reason and solve problems, asserting that if the mechanisms underlying learning were described with sufficient precision, a machine could in principle replicate them. This formed the conceptual foundation for what they termed "artificial intelligence."

Despite growing enthusiasm, the first practically applicable AI systems only emerged in the 1980s, including rule-based platforms for industrial logistics and expert systems capable of competing with human specialists. One of the most widely recognized milestones occurred in 1997, when IBM's Deep Blue defeated world chess champion Garry Kasparov, marking a pivotal moment in public awareness of AI's potential [92].

Today, AI permeates a wide range of domains, including finance, economics, large-scale digital platforms (Amazon, Netflix, Google), language processing, home automation, artistic generation, and—crucially—medicine, where it supports diagnostic workflows, prognostic modeling, and treatment planning [93]. Two main factors have driven the exponential growth of AI over the last decade: (1) the massive expansion of computational capacity, unimaginable in Turing's era; and (2) the social transformation brought by the Internet, where enormous quantities of shared data have become essential for training and refining machine-learning algorithms.

AI refers broadly to computational systems designed to automate intellectual tasks normally performed by humans. It is not synonymous with machine learning (ML) or deep learning (DL), which are specialized methodological subsets within the broader AI domain (Figure 7.1).

Machine learning focuses on creating systems that learn autonomously and improve performance based on input data. This evolution of AI stems from the need to emulate the human ability to acquire knowledge subjectively and intuitively. Human cognition requires vast amounts of world knowledge, much of which is individually acquired and difficult to formalize explicitly for computational use.

Deep learning, itself a subset of machine learning, relies on multilayered hierarchical learning from input data. A computer interprets datasets using conceptual hierarchies: complex representations emerge from simpler ones through successive transformations. If this architecture is visualized, the layers form a vertically deep structure, hence the term "deep learning."

DL arose from the need to manage the high dimensionality of modern datasets, and its learning mechanisms are partly inspired by the mammalian brain, where layered neuronal circuits adaptively process sensory information. However, DL methodologies also suffer from the “black-box” problem: although an input yields an output, the internal decision-making process—specifically which features drive the result—remains largely opaque. In contrast, simpler linear algorithms offer lower predictive power but far greater interpretability.

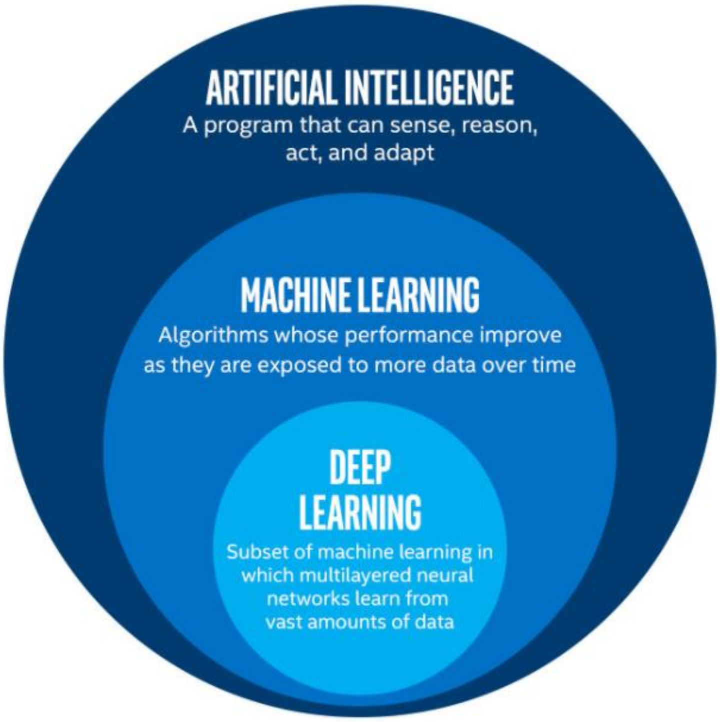


Figure 1.23 Subsets of AI

Algorithms

Because artificial intelligence does not rely—at least at present—on human-like instinct, intuition, or volitional decision-making, it must depend on alternative cognitive paradigms grounded in algorithms: structured logical–mathematical sequences that guide the system during its learning processes. The principal goal of algorithmic computation is problem-solving, spanning tasks as simple as sorting lists of numbers or names to challenges as complex as human genome mapping or prognostic outcome prediction [94].

All algorithms are built upon fundamental conceptual components that are subsequently assembled and manipulated according to the model’s structure:

- Input: the set of data (features) forming the dataset or training set on which the AI system is trained.
- Weights (W): contextual parameters allowing the model to classify or prioritize information hierarchically.
- Output: the final objective or prediction produced by the AI system.

Machine-learning and deep-learning algorithms can be broadly categorized as follows:

1. Supervised learning

The most widely adopted paradigm, supervised learning involves the developer instructing the algorithm on how to achieve a predefined output based on labeled data. Each input example includes its correct classification, and the system must learn to reproduce it. Logistic regression, Random Forest, Naïve Bayes classifiers, neural networks, and deep-learning architectures belong to this category [95].

2. Unsupervised learning

In this more autonomous framework, the system identifies patterns within unlabeled data to generate novel classification schemes without predefined reference labels. Representative examples include DBSCAN, decision-tree clustering approaches, and Bayesian belief networks [96].

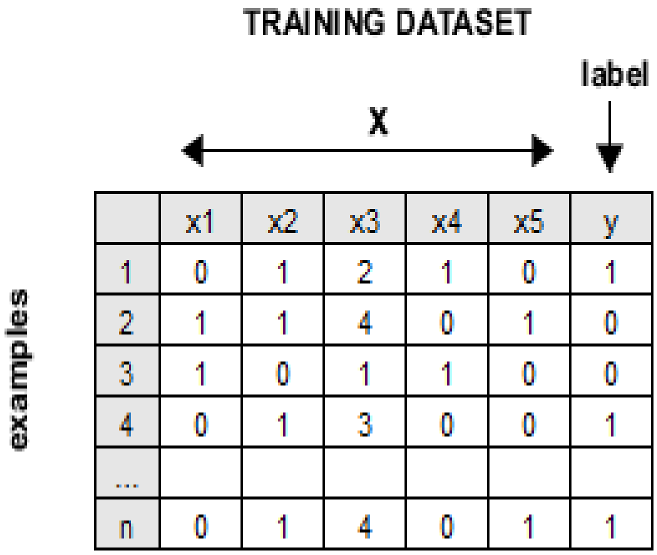
3. Reinforcement learning

Reinforcement learning relies on interpreting feedback as “reward” or “penalty.” An objective is defined, and decisions that bring the system closer to that objective are positively reinforced. Unlike supervised and unsupervised learning, reinforcement learning does not rely on a structured training set. Examples include Q-learning, Deep Q-Networks, and genetic algorithms [97].

Additional categories include semi-supervised learning, self-learning architectures, and feature-learning systems, although these are less commonly used in clinical research. The following sections detail the algorithms directly applied in the present work.

Logistic Regression

Logistic regression is a supervised machine-learning algorithm that uses the logistic (sigmoid) function to convert real-valued inputs into a continuous probability ranging from 0 to 1. Training requires a dataset composed of n examples, each characterized by m features x and a label y indicating the correct classification (Figure 1.24).



	x1	x2	x3	x4	x5	y
1	0	1	2	1	0	1
2	1	1	4	0	1	0
3	1	0	1	1	0	0
4	0	1	3	0	0	1
...						
n	0	1	4	0	1	1

Figure 1.24 Construction of the Dataset

The algorithm determines a weight vector *W* to associate with each attribute vector X_m, maximizing correct classifications (or equivalently minimizing misclassifications). Formally:

$$\mathbf{z = W \cdot X = w1x1 + \ldots + w_m x_m}$$

This linear combination represents the model's internal response for each training example. In logistic regression, *z* becomes the argument of the logistic function, which rescales it to the interval [0, 1]. This value serves as the activation function of neural-network nodes **f(z)** (Figure 1.25).

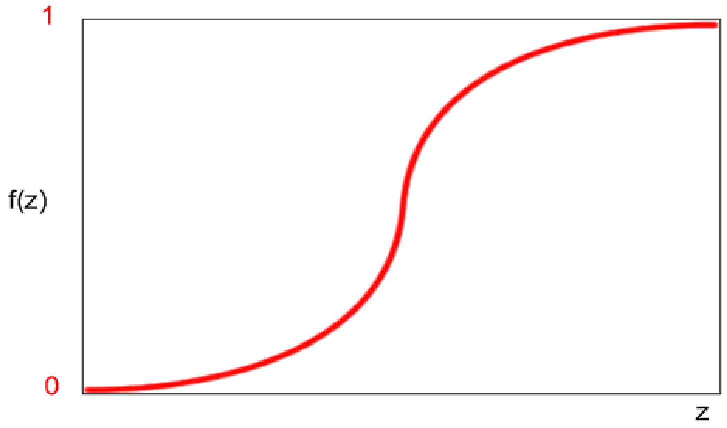


Figure 1.25: Logistic activation function in neural-network nodes

A threshold is then applied to determine whether the node is "activated." To identify the optimal distribution of weights *W*, the algorithm maximizes the likelihood function:

$$\max L(W) = \Sigma_i \, [\, y_i \log(f(z_i)) + (1 - y_i) \log(1 - f(z_i)) \,].$$

After training, the system yields a classification model capable of generalizing to new, previously unseen data (Figure 1.26).

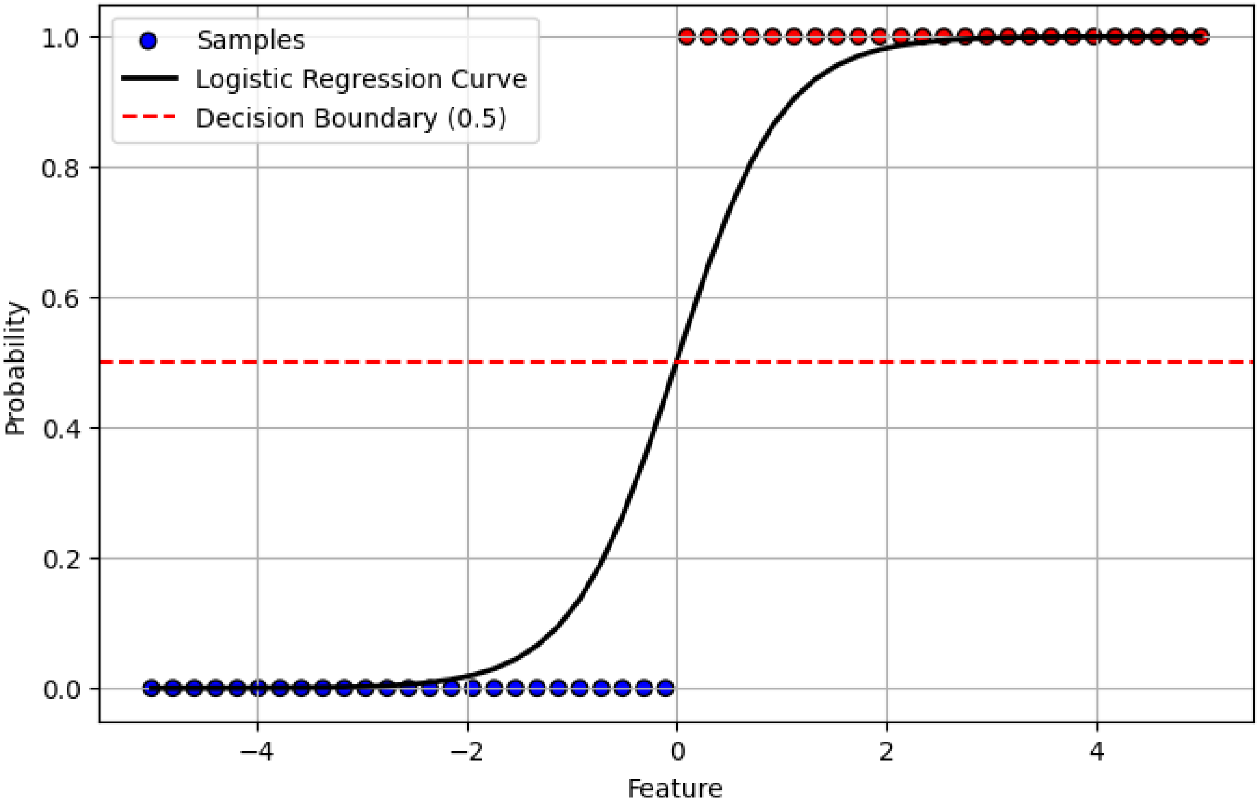


Figure 1.26: Final classification model produced by Logistic Regression

Random Forest

Random Forest is a supervised algorithm composed of multiple decision trees trained from the same dataset. A decision tree is represented as an inverted hierarchical diagram (Figure 1.27) in which branches emerge when the algorithm evaluates the value of one or more variables. Nodes are categorized as:

- **Non-terminal nodes:** bifurcate into two descendant branches;
- **Terminal nodes (leaves):** classify the final outcome without further branching.

Training begins with a dataset containing n features x and corresponding labels. The algorithm randomly selects a sample from the dataset and a subset of attributes ($i < n$). It constructs a decision tree based on this subset. This process is repeated m times, each iteration producing a new tree that may differ from the previous ones.

The final classification is determined by aggregating predictions from all trees and selecting the most frequent output. Ensemble construction enhances generalizability and reduces overfitting, especially in high-variance datasets [98].

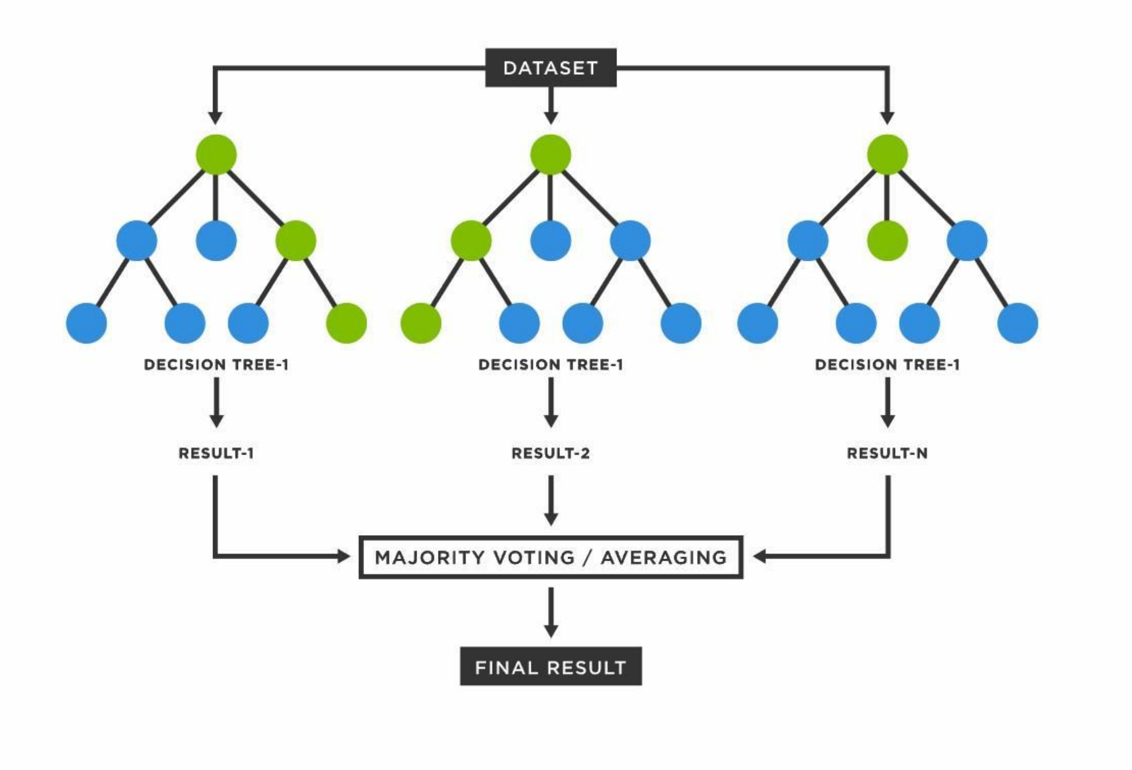


Figure 1.27: Decision-tree diagram illustrating the Random Forest algorithm

Neural Networks

Deep Learning relies on artificial neural networks (NN), which emulate the computational processes of biological neural circuits. An NN is composed of interconnected artificial neurons whose interactions resemble synaptic communication between neuronal bodies. Each node represents a computation unit, while connections represent weighted transmission pathways.

Neural networks consist of multiple layers (Figure 1.28):

- Input layer: receives the raw data to be processed.
- Hidden layers: intermediate levels responsible for hierarchical feature extraction; numerous hidden layers define deep-learning architectures.
- Output layer: produces the final model output.

Each layer transforms incoming signals and passes them to subsequent layers. The number of hidden layers determines computational capacity and influences the network’s representational complexity.

Among the most widely used architectures:

- Convolutional Neural Networks (CNN): feed-forward networks without recurrence, ideal for image analysis, including object recognition and medical imaging [99].
- Recurrent Neural Networks (RNN): networks incorporating feedback loops, allowing temporal pattern recognition, sequential modeling, and time-dependent signal analysis [100].

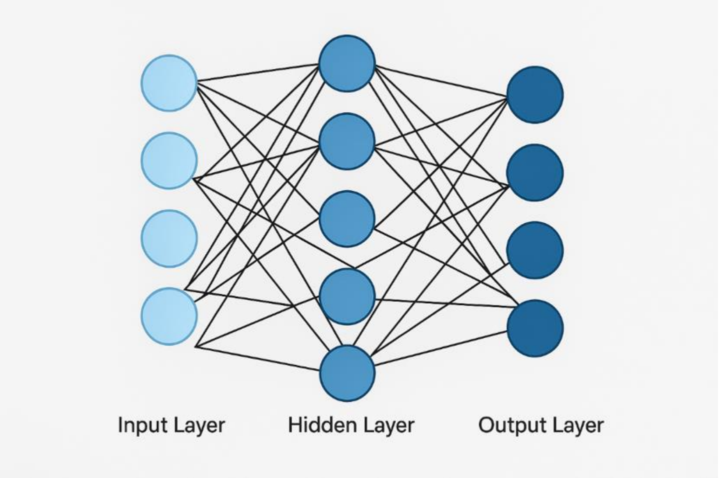


Figure 1.28: Graphical representation of an artificial neural network

Applications in Medical and Neurological Practice

Over the past decade, DL applications have expanded considerably across medicine, supported by increasing data availability and advances in imaging and computational technologies. Numerous studies demonstrate strong predictive and diagnostic performance in both research and clinical contexts [101].

Therapeutic applications

AI-based algorithms facilitate predictions of molecular structure, pharmacokinetics, and pharmacodynamics. AlphaFold, for instance, uses deep-learning strategies to achieve protein-structure predictions exceeding 90 percent accuracy, revolutionizing target discovery and accelerating drug development.

Personalized medicine

Machine-learning approaches enable multilevel personalized analysis of patient-specific datasets, supporting diagnostic, therapeutic, and prognostic decisions through high-dimensional pattern recognition.

Radiology

Radiology is inherently well-suited to ML due to the rapid growth in imaging volume, resolution, and modality diversity. ML assists with segmentation, spatial registration, computer-aided detection, functional analysis, and image synthesis. CNNs have proven particularly effective in brain, breast, and abdominal imaging.

Oncology

Neural networks increasingly enhance oncologic decision-making, improving tumor classification, risk prediction, and survival modeling. Dermatology studies show CNN performance comparable to expert clinicians in distinguishing malignant from benign skin lesions.

Neurology

ML supports epilepsy detection, cognitive-decline profiling, stroke-risk modeling, dementia prognosis, and brain-lesion classification. These advances rely heavily on neuroradiology datasets that provide robust image-based training material.

Future Developments and Current Limitations

Despite rapid progress, AI remains limited by the quantity and quality of available training data. Large, accurately annotated datasets remain essential for generalizable models. Underrepresentation, label noise, or systematic biases can significantly impair real-world model performance.

Many datasets rely on human annotators, and incorrect or incomplete labels continue to challenge data scientists and clinicians alike. Moreover, despite the growing relevance of ML, many medical professionals lack formal training in AI concepts, limiting widespread adoption.

Future development requires not only technological innovation but also improved clinician education and infrastructure capable of integrating AI tools into everyday medical practice.

1.4.2 Radiomics

The radiomic approach can be employed to construct predictive models capable of accurately anticipating clinically relevant outcomes across biomedical investigations. Radiomics enables the extraction of quantitative descriptors that fall into three principal categories: first-order features, which summarize voxel-intensity distributions (for example mean, median, and variance); second-order features, which describe spatial relationships between voxel intensities (such as texture metrics); and third-order features, which quantify geometric properties of the region of interest (ROI).

Following extraction, these features may be analyzed for associations with clinical endpoints or incorporated as variables within predictive modeling frameworks. Prior literature has demonstrated that, within clinical decision-making, diagnostically relevant information resides not only in images themselves but also in the structured radiomic signatures derived from them [102].

Radiomics Compared with Computer-Aided Diagnosis (CAD)

High-performance computing, which allows sustained large-scale computation, now makes it possible to extract vast numbers of quantitative features from tomographic imaging modalities, including CT, MRI, and PET. The extraction of such high-dimensional imaging descriptors enables the construction of large radiomic databases. Radiomics is grounded in the principle that biomedical images encode pathophysiologic information that can be revealed through quantitative image analysis [103].

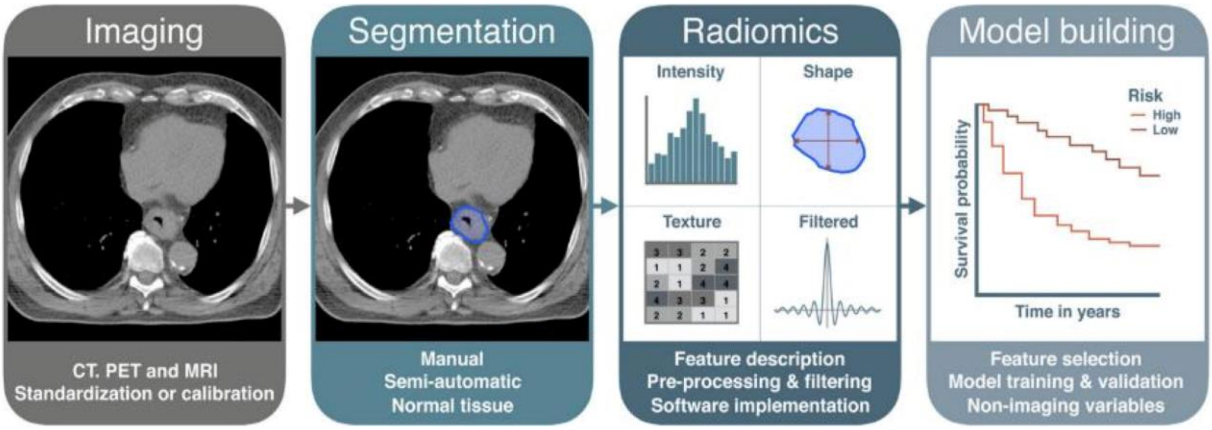
Although often perceived as a natural extension of computer-aided detection/diagnosis (CAD) systems, radiomics differs substantially from CAD. CAD systems, typically approved by regulatory agencies such as the U.S. Food and Drug Administration, are autonomous tools aimed at identifying a specific diagnostic element, such as the presence of a lesion. Early successes were observed primarily in mammographic cancer detection [104].

Radiomics, by contrast, is designed to extract large numbers of quantitative features, populate shared databases, and subsequently mine these data for hypothesis generation and validation. Rather than producing a single diagnostic output, radiomics serves as the foundation for decision-support systems, integrating imaging-derived features with additional clinical variables when available, thereby enhancing model robustness and predictive performance [105]. Given its objective of maximizing informational content extracted from medical images, radiomics is expected to evolve into large-scale imaging ecosystems comprising millions of patients.

Radiomic Workflow

A schematic radiomic workflow is presented in the figure below. The process begins with acquisition of the imaging study using the modality appropriate for the clinical question. Following digital reconstruction, a region of interest (ROI) is selected to define the volume from which radiomic features will be extracted. Feature computation is then performed using dedicated software, which may incorporate pre- or post-processing tools. Finally, a statistical or machine-learning model is constructed to identify features predictive of the clinical endpoint of interest (for example, survival).

Here is an example of a “radiomic workflow.”:



Each step of the workflow includes multiple methodological alternatives, and no universal gold standard currently exists. Numerous sources of variability arise at different stages, potentially compromising the reproducibility and clinical reliability of the resulting model. Consequently, multiple processing algorithms are applied to mitigate noise and standardize intermediate steps [106].

Imaging-Data Acquisition and Feature Stability

Medical images acquired for diagnosis, treatment, and follow-up constitute the data foundation for radiomics. Many radiomic studies rely on retrospective datasets in which acquisition parameters vary across scanners, institutions, and imaging protocols, especially in multicenter studies. These variations may introduce unexpected influences on the resulting radiomic signatures [107].

Because radiomics assumes that quantitative imaging features function as biomarkers, such features must exhibit reproducibility and stability. A stable feature yields consistent values across acquisition protocols and within standard noise conditions. For CT imaging, several studies have shown that many radiomic features are highly repeatable, though scanner models and voxel dimensions remain major sources of variability. Shafiq-ul-Hassan et al. demonstrated that normalization algorithms can reduce voxel-size dependence, while gray-level quantization can enhance feature robustness [108]. Controlled CT acquisition protocols remain the optimal approach to reducing feature variability.

For PET imaging, Galavis et al. (2010) showed that textural features may vary considerably with changes in voxel size, reconstruction algorithms, and iteration numbers; these observations were confirmed in subsequent studies by Yan et al. and van Velden et al. Larue et al. (2017) systematically reviewed the influence of acquisition parameters on PET and CT radiomics, highlighting substantial variability across feature types [109].

MRI feature stability has been studied less extensively, although effects of magnetic-field strength, sequence parameters, and vendor-specific differences have been investigated using phantom and human datasets. Mayerhoefer et al. demonstrated that several MRI features are sensitive to TR and TE variations, findings confirmed in more recent simulation studies [110]. Given the reliance on MRI in the present thesis, a dedicated stability analysis was performed to identify a robust subset of features suitable for model construction.

ROI Segmentation

Radiomic analysis typically focuses on the primary tumor mass. Therefore, segmentation represents a critical step, as it dictates which anatomical region undergoes quantitative analysis. Manual, semi-automatic, and automatic segmentation techniques exist, though each presents challenges. No universal segmentation standard has been adopted.

Manual segmentation is time-consuming and requires expert delineation of the ROI on every slice, which can be numerous in high-resolution imaging. Tumor morphology varies substantially, and margins may be irregular or indistinct. Many radiomic studies have relied on manual ROIs due to the absence of robust automated tools, but this approach increases inter-observer variability [111].

Radiomic Feature Extraction

Feature extraction involves computing quantitative descriptors for the selected ROI. Depending on the categories analyzed, the number of extracted features may range from dozens to several thousand.

Model Construction

A primary objective of radiomic studies is the development of predictive or prognostic models with strong generalizability. Machine-learning algorithms offer a wide range of modeling strategies spanning supervised and unsupervised methods.

Unsupervised approaches partition samples according to similarity metrics, whereas supervised methods estimate parameters for predefined model architectures. Supervised learning requires a training phase, during which the model minimizes a cost function, followed by a validation phase, in which model performance is evaluated on an independent sample set.

Because radiomic datasets may contain thousands of features, redundancy can lead to overfitting, wherein the model performs well on training data but poorly on unseen data. Feature-selection algorithms are therefore crucial to identify the most informative predictors [112]. Model validation ideally requires independent external data; when unavailable, cross-validation strategies are employed. For survival models, the most commonly used performance metric is the Concordance Index (C-index).

Radiomic Feature Classes

The radiomic dataset constitutes a high-dimensional feature space generated from numerical descriptors associated with the clinical image and its corresponding ROI. Different features capture distinct phenotypic characteristics of a tumor, reflecting the substantial heterogeneity present across disease types. These multidimensional descriptors may carry prognostic information complementing conventional radiologic interpretation.

Shape and Size (SS) Features

Shape and Size features describe geometric properties of the ROI, such as area and volume. These features can be computed using different ROI representations. In the present work, SS features were obtained from surface meshes generated using the Marching Cubes algorithm, which provides consistent measurements of volume and surface area.

Each vertex represents a point halfway between a voxel inside and one outside the ROI. Connecting these vertices produces a mesh composed of M triangles and N vertices. The surface can be represented mathematically using an $N \times 3$ matrix of vertex coordinates and an $M \times 3$ matrix of triangle indices.

Only 3D SS features were used in this thesis.

First-Order Statistics (FOS)

First-Order Statistical features describe voxel-intensity distributions within the ROI. These metrics include intensity-based descriptors and histogram-derived features. Histogram-based metrics require gray-level discretization, which may be performed by fixing either bin width or bin number.

Gray-Level Co-Occurrence Matrix (GLCM)

GLCM features belong to the textural category and encode spatial relationships between discrete gray-level values. The GLCM is an $Ng \times Ng$ matrix, where Ng denotes the number of discretized gray levels. For a given distance δ and direction θ, element (i,j) describes how often gray level j appears δ pixels from gray level i in direction θ.

An example of GLCM construction is illustrated in the figure below.

$$\mathbf{I} = \begin{bmatrix} 1 & 2 & 5 & 2 & 3 \\ 3 & 2 & 1 & 3 & 1 \\ 1 & 3 & 5 & 5 & 2 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 4 & 3 & 5 \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} 6 & 4 & 3 & 0 & 0 \\ 4 & 0 & 2 & 1 & 3 \\ 3 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 2 \end{bmatrix}$$

Figure 1.29 – Example GLCM (P) computed from matrix I for δ = 1 pixel and θ = 0 (horizontal).

Gray-Level Run Length Matrix (GLRLM)

GLRLM features, also within the textural category, quantify **gray-level runs**, defined as the number of consecutive pixels sharing the same gray level. For a given direction θ, entry (i,j) describes the number of times gray level *i* appears in runs of length *j*. The matrix dimension is *Ng* × *Nmax*, where *Nmax* is the maximum run length.

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 1.30 – Example GLRLM (P) computed from matrix I for θ = 0 (horizontal).

Wavelets

Pyradiomics enables extraction of features from both the original image and wavelet-transformed images. Wavelets are generated through dilations and translations of a single base function. Decomposition is conceptually equivalent to applying high-pass (HA) and low-pass (LA) filters, producing detail and approximation components, respectively.

$$\psi(t) = |a|^{-\frac{1}{2}} \psi \left(\frac{t-b}{a}\right)$$

In 2D signals such as images, wavelet decomposition yields eight possible filter combinations across three spatial directions. These combinations are illustrated in the figure below.

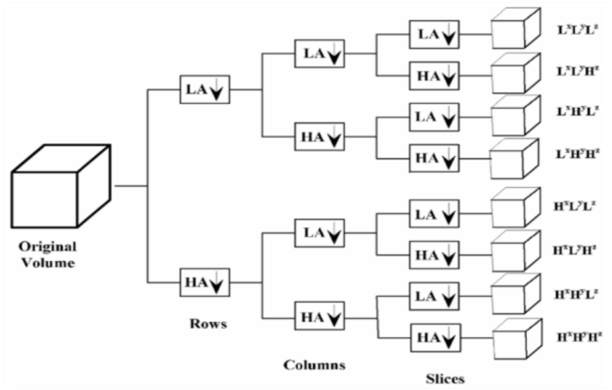


Figure 1.31 – Schematic of wavelet high-pass (HA) and low-pass (LA) filter combinations for a 3D volume.

2. PURPOSE OF THE STUDY

Accurate radiological differentiation between glioblastoma (GBM) and primary central nervous system lymphoma (PCNSL) remains a critical challenge in neuro-oncology, with substantial implications for clinical decision-making and patient outcomes. Although both entities may present with overlapping features on conventional MRI—such as contrast enhancement, mass effect, and perilesional edema—the therapeutic pathways they mandate are fundamentally distinct. For patients with suspected GBM, maximal safe surgical resection (gross total resection) is considered the standard initial approach, aiming to reduce tumor burden and improve progression-free and overall survival. Conversely, patients with PCNSL typically do not benefit from extensive surgical intervention; their management relies primarily on high-dose methotrexate–based chemotherapy and adjunctive therapies, with surgery generally limited to diagnostic biopsy in select cases. An accurate pre-operative diagnosis therefore plays an essential role in avoiding unnecessary neurosurgical procedures and in expediting appropriate therapeutic strategies.

Given the increasing complexity of neuroimaging biomarkers and the potential for subtle but diagnostically meaningful radiological signatures, advanced computational tools may offer valuable support in the differential diagnosis of these two high-grade brain tumors. The purpose of the present study is to develop and train an artificial intelligence–based software system capable of distinguishing, with high reliability, between glioblastoma and primary CNS lymphoma using MRI data. By leveraging machine learning algorithms applied to multiparametric MRI sequences, the system aims to identify discriminative imaging features that may be difficult to appreciate through visual assessment alone. Ultimately, this tool is envisioned as a clinically applicable decision-support system that can assist neuroradiologists and neuro-oncologists in optimizing diagnostic accuracy, refining treatment planning, and improving patient stratification at the earliest stages of care.

3. MATERIALS AND METHODS

The present study was made possible through the integrated and highly coordinated efforts of a broad multidisciplinary team, each member contributing specialized expertise essential for the project's successful completion. Neurologists and neurosurgeons played a central role in identifying eligible patients, conducting detailed clinical evaluations, and delivering the appropriate therapeutic interventions, thereby ensuring the robustness and clinical relevance of the cohort. Neuroradiologists, working in close collaboration with neurosurgeons, undertook the meticulous tasks of MRI image segmentation, contouring, and interpretative assessment, providing high-fidelity anatomical references and validating imaging-derived findings with expert judgment.

Medical physicists were responsible for the technical backbone of the study: they curated the imaging datasets, engineered and trained the artificial intelligence algorithms, optimized the computational pipeline, and performed the statistical modeling and quantitative analyses that support the study's conclusions. Their work ensured methodological rigor, reproducibility, and adherence to the highest standards of computational neuroscience.

Throughout all phases of the research, the PhD candidate (author) contributed actively by coordinating team activities together with assigned mentors, facilitating communication across professional roles, and ensuring smooth integration of clinical, radiological, and computational components. In addition, the candidate was directly involved in data acquisition, preprocessing, and iterative interpretation of results at every step of the project, thereby gaining hands-on experience with the full translational workflow and contributing substantively to the study's scientific output.

Given the retrospective design of the study, which involved only the analysis of fully anonymized data from patients who underwent standard procedures as part of routine clinical care, with no additional interventions or risk for the patients, formal approval by an ethics committee was not required in accordance with applicable institutional and national regulations. The study was conducted in compliance with the principles of the Declaration of Helsinki.

3.1 Patient Recruitment

A total of 226 patients were initially identified for this study through the Neurosurgical Department database of Azienda Ospedaliero Universitaria delle Marche in Ancona (Italy). Inclusion criteria required a confirmed or suspected diagnosis of primary or secondary brain tumors based on MRI, with particular attention to T1-weighted sequences. For each individual, the most recent preoperative MRI examination available in the PACS was retrieved and reviewed.

The preliminary cohort consisted of 226 patients who had undergone MRI between February 2011 and June 2025, including 148 individuals with glioblastoma (GBM) and 78 with primary or secondary central nervous system lymphoma (PCNSL/SCNSL). During data curation, several patients were excluded due to insufficient compliance during MRI acquisition, which prevented the collection of diagnostically reliable datasets. Additional cases were removed because essential MRI sequences required for our analysis were not acquired, while others lacked any preoperative MRI examination within the hospital's imaging archive.

After these exclusions, the dataset was refined to 124 GBM cases and 46 lymphoma cases, yielding a final study population of 170 patients. The demographic distribution of the final cohort was as follows: among GBM patients, 68 males and 56 females; among PCNSL/SCNSL patients, 22 males and 24 females.

All patients included in the study had a histologically confirmed diagnosis of glioblastoma (GBM) or primary/secondary central nervous system lymphoma (PCNSL/SCNSL). Histopathological diagnoses were established on surgical or biopsy specimens and classified according to the criteria of the World Health Organization (WHO), in accordance with the most recent WHO Classification of Tumors of the Central Nervous System (2021). Only cases with definitive histological confirmation were included in the analysis.

3.2 Lesion delineation

Once all imaging datasets were retrospectively retrieved from the PACS for the patients included in the study, a structured pipeline for lesion delineation was implemented using the open-source software 3D Slicer. Each patient's MRI examination, stored in standard DICOM format, was imported into the workspace to allow comprehensive visualization and manipulation of the imaging data. After loading the images, the pre-operative axial contrast-enhanced T1-weighted MRI (T1w+Gd) sequence was selected as the reference series, given its robustness in depicting tumor boundaries and its widespread use in neuro-oncological imaging protocols.

For every subject, the tumor volume was manually contoured slice-by-slice, generating a precise region of interest (ROI) for each visible component of the lesion. This manual approach—although time-consuming—ensured high anatomical fidelity, minimized segmentation artifacts, and allowed close supervision by neuroradiology and neurosurgery experts throughout the process. The resulting ROIs constituted the basis for the subsequent radiomic analysis.

Manual segmentation was performed using 3D Slicer software, taking into account the known limitations associated with manual delineation. All segmentations were carried out by experienced operators and subsequently reviewed by expert readers to ensure consistency and anatomical accuracy. However, inter- and intra-observer variability analyses were not performed, representing a known limitation of the study.

Missing data and outliers were carefully assessed during data preprocessing; cases with incomplete data were excluded from the analysis, and outliers were managed according to standard radiomic preprocessing procedures. Radiomic feature extraction was performed using PyRadiomics implemented in 3D Slicer, applying standardized preprocessing settings, including image resampling to a uniform voxel spacing, intensity discretization with a fixed bin width, and intensity normalization. Multiple feature classes were extracted, and filter-based features were included according to PyRadiomics recommendations to ensure reproducibility.

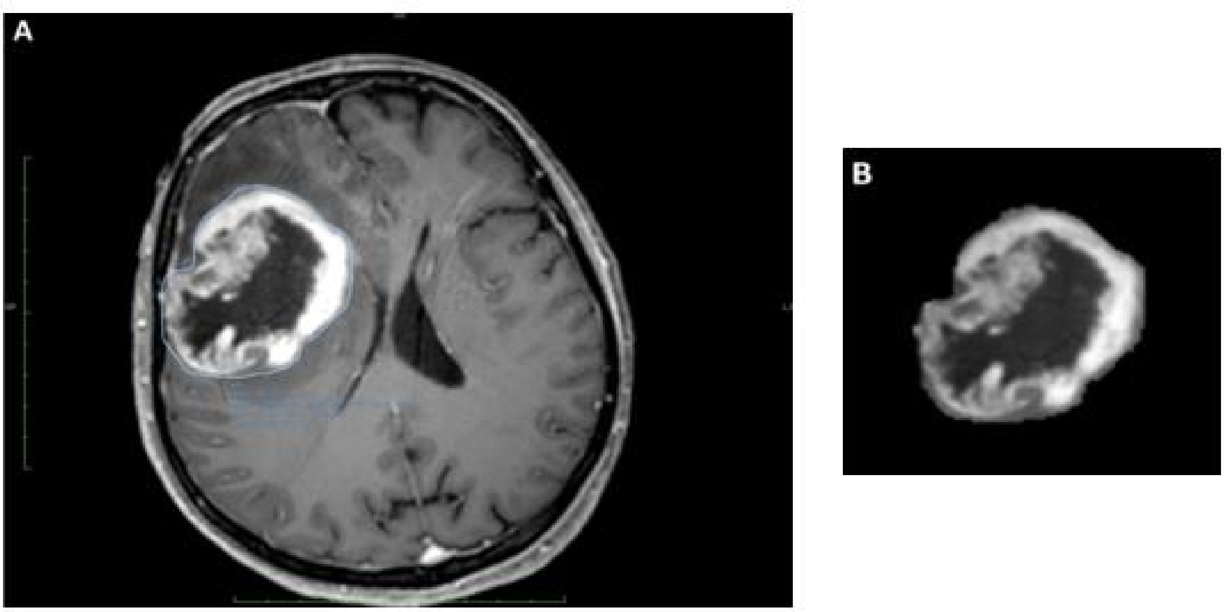


Figure 3.1 A: T1w+Gd MRI image of a patient with GBM. B: ROI extracted from the segmented region

Once segmentation was completed, 100 radiomic features were automatically extracted for each patient using the PyRadiomics-compliant module integrated in 3D Slicer. These features covered multiple dimensions of image characterization, including morphological descriptors (shape features), first-order voxel-intensity statistics, and second-order textural matrices such as GLCM and GLSZM, which capture intra-tumoral heterogeneity and microstructural complexity. All radiomic parameters adhered to the standardized definitions established by the Harvard Medical School–endorsed PyRadiomics framework, ensuring reproducibility and interoperability across studies (van Griethuysen et al., *Cancer Research*, 2017).

The full radiomic feature set was subsequently used to train the AI models developed within the study, forming the quantitative backbone for downstream machine-learning analyses.

3.3 Data acquisition for AI training

The two tables presented below provide an illustrative overview of the dataset collected for the patient cohort. Each row corresponds to an individual patient, identified by diagnostic category, while each column represents one of the extracted radiomic features. The column headers (e.g., *Flatness*, *LeastAxisLength*, *MajorAxisLength*, *Maximum2DDiameterColumn*) denote the specific radiomic parameters included in the analysis. These feature values constitute the quantitative input used for subsequent artificial intelligence training and model development.

Table. 3.1 GBM

GBM				
DIAGNOSIS	**FLATNESS**	**LEAST AXIS LENGTH**	**MAJOR AXIS LENGTH**	**MAXIMUM 2D DIAMETER**
GBM	0.8112846110242388	29.44080026847562	36.289114656454416	41.036569057366385
GBM	0.41027410646345824	16.704549304190618	40.71558268247289	26.076809620810597
GBM	0.7368267046747555	42.34388376610144	57.46790052186363	67.00746227100382
GBM	0.7803146533993595	33.84710017619631	43.37622012959124	50.21951811795888
GBM	0.6268635043355448	36.20099200323677	57.749401189990564	61.0
GBM	0.7446291778583782	34.532761981644505	46.3757840929144	53.45091205957107
GBM	0.547784997880262	33.725140715269546	61.56638251462552	57.688820407423826
GBM	0.742208792309318	35.37813113631855	47.66600921856851	53.71219600798314
GBM	0.7075363768628413	30.05683139839651	42.4809697158838	45.69463863518345
GBM	0.5117422398772568	23.988845867949358	46.876814143196725	50.20956084253277

Table. 3.2 PCNSL

DIAGNOSIS	FLATNESS	LEAST AXIS LENGTH	MAJOR AXIS LENGTH	MAXIMUM 2D DIAMETER
PCNSL	0.782482959369546	24.277942684178022	31.02680051171848	39.0
PCNSL	0.4363943511422985	13.831542528483132	31.695054008554234	28.30194339616981
PCNSL	0.6346994685638658	18.00431213155502	28.36667276922946	31.953090617340916
PCNSL	0.3480694859507173	28.06513060397825	80.63082728243508	83.40863264674707
PCNSL	0.6467307858094863	29.031009830043807	44.8888632906301	52.80151512977634
PCNSL	0.6032391978163684	13.210929522002745	21.899985229448358	24.413111231467404
PCNSL	0.6417794423584057	13.604419028073316	21.19796635753852	22.80350850198276
PCNSL	0.4088653204029963	21.97100160582041	53.73652523076496	51.03920062069938
PCNSL	0.07792046803832473	16.13960343762326	207.12919010811243	27.784887978899608
PCNSL	0.6212187294685995	47.87883643998862	77.07242903147004	99.76472322419383

3.4 Application of AI to our study

All extracted radiomic features were consolidated into a unified dataset to enable the development, training, and validation of artificial intelligence (AI) models. As is often the case in neuro-oncology research, the dataset exhibited a natural imbalance between the two diagnostic groups, reflecting the true epidemiological distribution of the diseases under investigation. This imbalance, combined with the relatively small number of available cases, poses inherent challenges for machine learning systems, which typically rely on large and well-balanced datasets to achieve stable and generalizable performance.

To mitigate these limitations, and pending the acquisition of additional clinical cases, we implemented data augmentation strategies aimed at enhancing class symmetry without altering the physiopathological meaning of the data. Specifically, we used the SMOTE (Synthetic Minority Over-Sampling Technique) algorithm, which generates new synthetic observations by interpolating between existing minority-class samples in the feature space. Unlike naive duplication, SMOTE introduces controlled variability, which helps models better capture the underlying structure of the minority class. Nevertheless, this approach assumes that the available data provide a representative sample of the real population; if the original dataset contains systematic bias, SMOTE may inadvertently amplify it. In our study, however, the characteristic imaging patterns of lymphomas on T1-weighted post-contrast MRI (T1w + Gd)—notably their homogeneous enhancement and distinctive textural signature—reduce the likelihood that the algorithm introduced distortions in the minority class. SMOTE was applied exclusively within the training folds during cross-validation, while the validation folds were kept completely independent. No information from the validation data was included in the oversampling process, thereby preventing data leakage.

Following augmentation, all radiomic parameters were organized into a structured table that served as the AI training dataset. Each patient was represented by 100 radiomic features, covering morphological descriptors, intensity-based metrics, and second-order texture features derived from matrices such as GLCM and GLSZM. This multi-dimensional characterization provides a comprehensive quantification of tumor phenotype, capturing aspects of shape, heterogeneity, and microstructural complexity that are not visually appreciable on routine MRI.

For model training and evaluation, we employed Stratified k-Fold Cross-Validation, a technique that ensures proportional representation of both classes within each fold. This is essential in imbalanced or limited datasets because it prevents the model from being exposed to folds containing predominantly one class. In our 10-fold configuration, the dataset was partitioned into ten subsets of comparable size and class distribution. At each iteration, nine folds were used for training and the remaining one for validation. This process was repeated ten times, allowing every sample to serve as a validation instance exactly once. Averaging performance across folds provides a more reliable estimate, reducing the risk of overfitting to a specific subset of the data.

Several methodological considerations were critical for optimizing classification performance:

- Within-class homogeneity and between-class separability are essential for accurate discrimination. The more consistent the radiomic signature within a diagnostic group, and the more distinct it is relative to the other class, the easier it is for the algorithm to minimize false positives and false negatives.

- Binary classification tasks were intentionally selected. Multiclass classification would significantly increase model complexity and error propagation, especially with small datasets. Conducting paired dichotomous comparisons allows the algorithm to focus on the most meaningful distinctions between diagnostic categories.

- Balanced class sizes, achieved through SMOTE, ensure that the model receives comparable statistical evidence from both classes. Severe imbalance typically leads to biased decision boundaries and reduced clinical interpretability.

We evaluated multiple supervised learning models, including SVM, Naïve Bayes, kNN, Logistic Regression, Random Forest, and Neural Networks. Only the most reliable and interpretable algorithms were retained for detailed analysis. Logistic Regression served as a baseline due to its simplicity and transparent decision boundaries, whereas Random Forest and Neural Networks provided more complex, non-linear decision functions capable of capturing higher-order feature interactions. The neural network employed the ADAM optimizer, which accelerates convergence and is well suited for high-dimensional radiomic data.

Model performance was assessed primarily through ROC (Receiver Operating Characteristic) curves, which map the relationship between the false positive rate (FPR) and true positive rate (TPR). The Area Under the Curve (AUC) quantifies discriminative ability, with values closer to 1 indicating near-perfect performance. AUC values below 0.5, by contrast, reflect a non-informative classifier. Complementary to AUC, Accuracy (ACC) was calculated as the proportion of correctly classified samples relative to the total dataset.

From a visual standpoint, the quality of the classifier increases as the ROC curve approaches the upper-left corner of the plot, corresponding to maximum sensitivity and specificity. This graphical representation provides an intuitive yet rigorous way to assess model suitability for potential clinical translation.

4. RESULTS

Before presenting the results, it is important to emphasize a distinctive feature of this study: the analysis is based exclusively on the post-contrast T1-weighted (T1w+Gd) sequence, whereas in clinical practice neuroradiologists routinely rely on multiple MRI sequences for diagnostic assessment.

While this characteristic represents a strength—given that the model operates on a single, standardized sequence—it also introduces a limitation, as it reduces the direct comparability with methodologies commonly adopted in routine clinical workflows.

To preserve the generalizability of the study and to assess the impact of data augmentation techniques on model performance, we trained and tested the AI models on both the original dataset and on a class-balanced dataset generated using the SMOTE algorithm, subsequently comparing the outcomes.

GBM was defined as the positive class for model training and performance evaluation.

A summary of the study's key data is provided in the table below.

Dataset	**Patients**	**GBM**	**PCNSL**	**Artificial Patients**
Original	85	124	46	0
Balanced	249	125	124	78

Table 4.1. Comparison table between the original dataset and the balanced dataset

4.1 GBM vs PCNSL/SCNSL - Original Dataset

4.1.1 Logistic Regression

The first algorithm classified the two groups with an accuracy of 40% and an AUC of 0.5. Specifically, it correctly identified 40 GBMs out of 124 and 36 PCNSL/SCNSL cases out of 46, yielding a specificity of 63% and a sensitivity of 40%. These results are clearly suboptimal and indicate that the logistic regression–based AI model is unable to reliably differentiate between the two pathologies. While it demonstrated relatively good recognition of lymphomas, it failed to adequately identify GBMs. This imbalance led to poor overall performance (AUC = 0.5), indicating that the model failed to reliably discriminate between GBM and PCNSL, particularly due to inadequate GBM recognition.

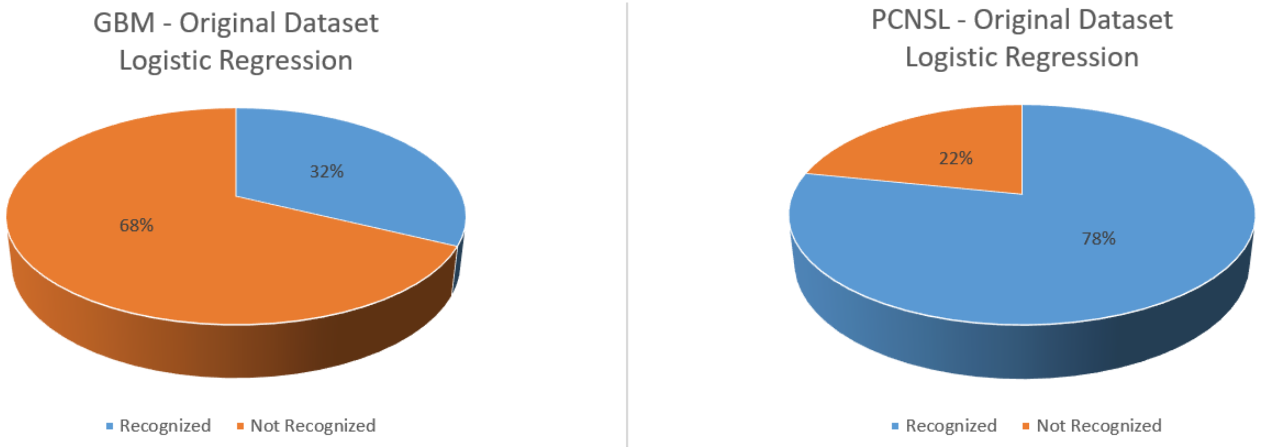


Table 4.2. Identification of GBMs and PCNSLs in the Original Dataset and Logistic Regression model

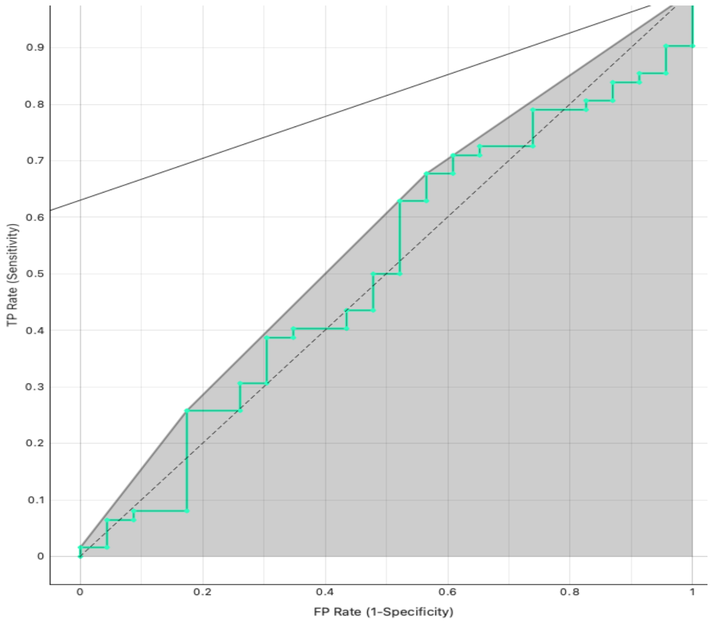


Figure. 4.1 ROC curve of GBMs and PCNSLs in the Original Dataset - Logistic Regression model

	Predicted PCNSL	**Predicted GBM**
Actual PCNSL	**36 (TP)**	**10 (FN)**
Actual GBM	**84 (FP)**	**40 (TN)**

Table 4.3 Confusion Matrix – Logistic Regression Original Dataset (GBM vs PCNSL)

4.1.2 Random forest

The accuracy of the Random Forest algorithm increased to 76% compared with logistic regression, with an AUC of 0.77. The performance gap between the two groups decreased accordingly: the model correctly identified 114 out of 124 GBMs and 14 out of 46 PCNSL/SCNSL cases, resulting in a sensitivity of 75% and a specificity of 73%. These results are acceptable and approach clinically comparable performance levels.

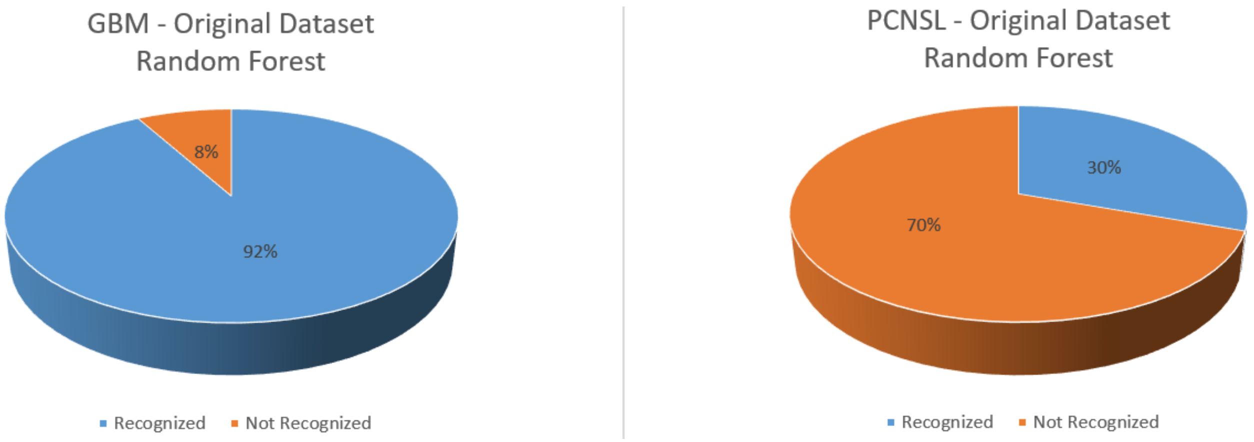


Table 4.4 Identification of GBMs and PCNSLs in the Original Dataset - Random forest model

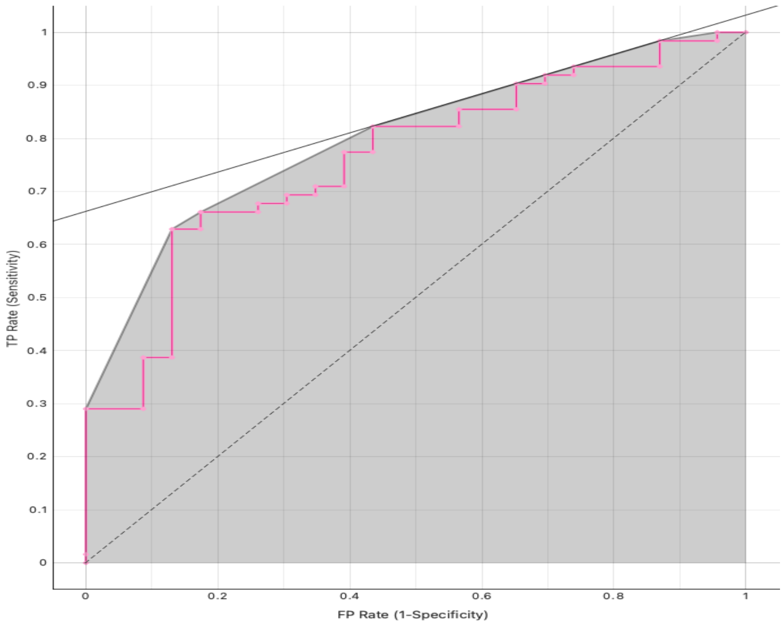


Figure 4.2 ROC curve of GBMs and PCNSLs in the Original Dataset – Random Forest model

	Predicted PCNSL	**Predicted GBM**
Actual PCNSL	**14 (TP)**	**32 (FN)**
Actual GBM	**10 (FP)**	**114 (TN)**

Table 4.5 Confusion Matrix – Random Forest- Original Dataset (GBM vs PCNSL)

4.1.3 Neural Network

This algorithm proved to be the best performing, achieving an AUC of 0.81 and an accuracy of 78%. The neural network classified 108 out of 124 GBMs and 26 out of 46 PCNSL/SCNSL cases, demonstrating an increase in sensitivity compared with the Random Forest while maintaining a comparable level of specificity.

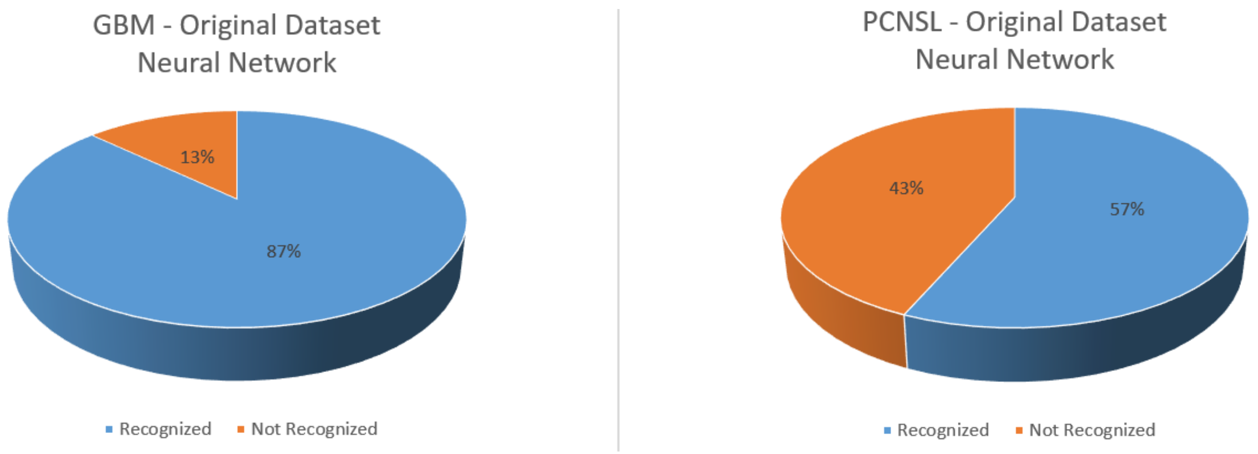


Table 4.6 Identification of GBMs and PCNSLs in the Original Dataset – Neural Network model

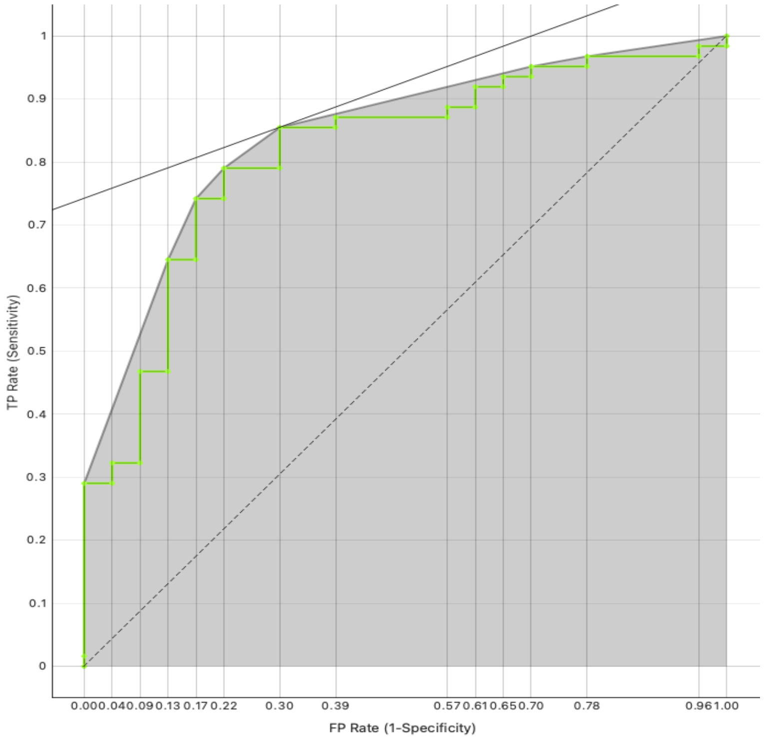


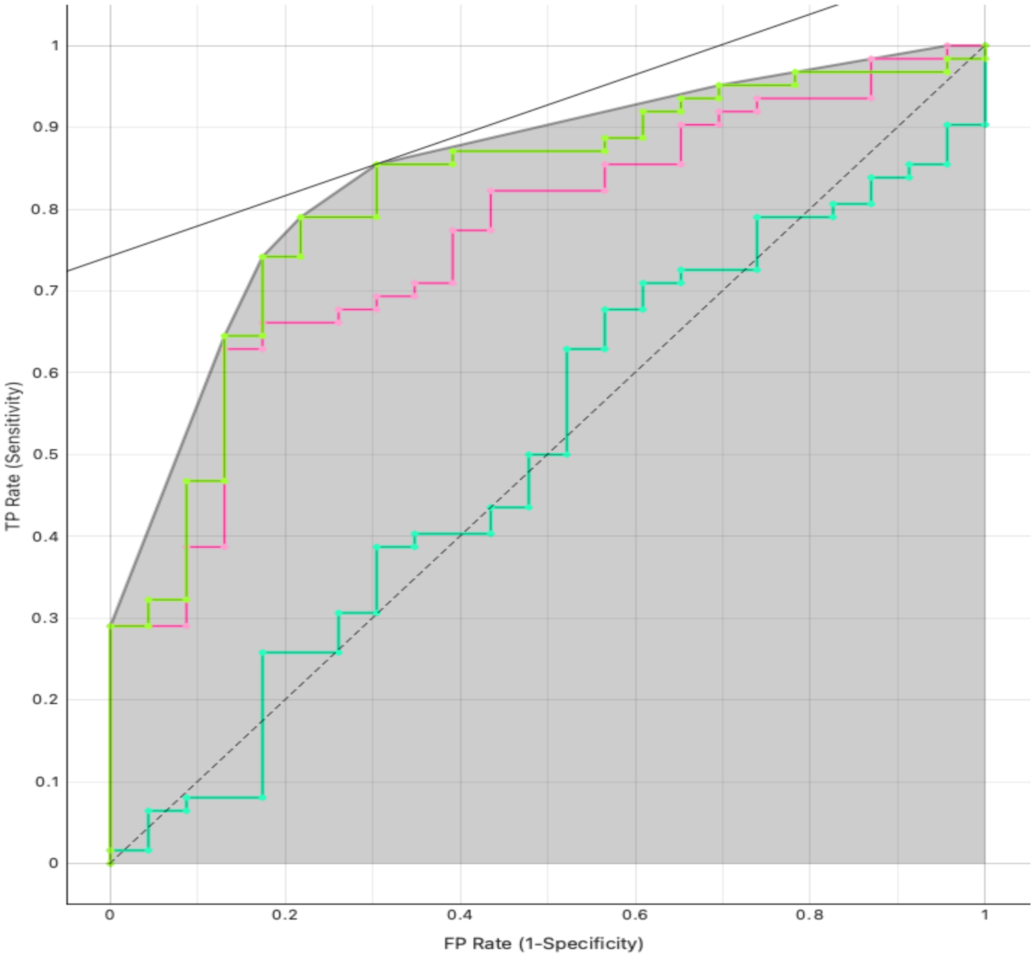
Figure 4.3. ROC curve of GBMs and PCNSLs in the Original Dataset – Neural Network model

	Predicted PCNSL	**Predicted GBM**
Actual PCNSL	**26 (TP)**	**20 (FN)**
Actual GBM	**16 (FP)**	**108 (TN)**

Table 4.7 Confusion Matrix – Neural Network- Original Dataset (GBM vs PCNSL)

4.1.4 Comparison between the ROC curves

In the graph below, a comparison of the three ROC curves is presented. Specifically, the ROC curve corresponding to the Logistic Regression algorithm is shown in blue, the curve associated with the Neural Network in green, and the Random Forest curve is depicted in pink.



Algorithm	AUC	95% CI (AUC)
Logistic Regression	0.50	0.402 – 0.598
Random Forest	0.77	0.698 – 0.842
Neural Network	0.81	0.745 – 0.875

Table 4.8 AUC with 95% Confidence Intervals (Hanley & McNeil)

4.2 GBM vs PCNSL/SCNSL - Balanced Dataset

4.2.1 Logistic Regression

The first algorithm classified the two groups with an accuracy of 60% and an AUC of 0.635. It correctly identified 40 out of 124 GBMs and 110 out of 124 PCNSL/SCNSL cases, resulting in a sensitivity of 60% and a specificity of 65%. These results are not satisfactory, as they fall below the expected clinical performance standards.

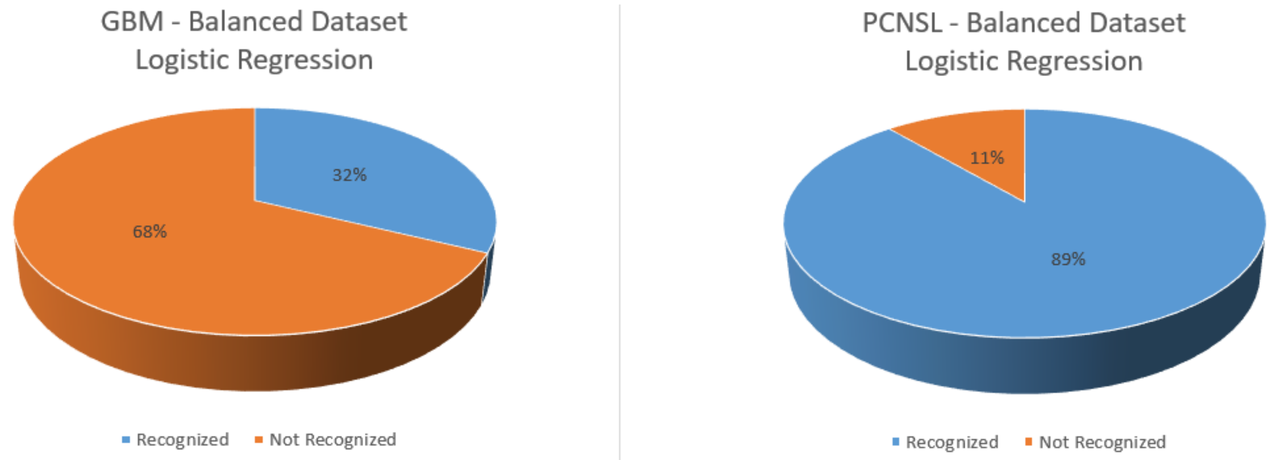


Table 4.9 Identification of GBMs and PCNSLs in the Balanced Dataset – Logistic regression model

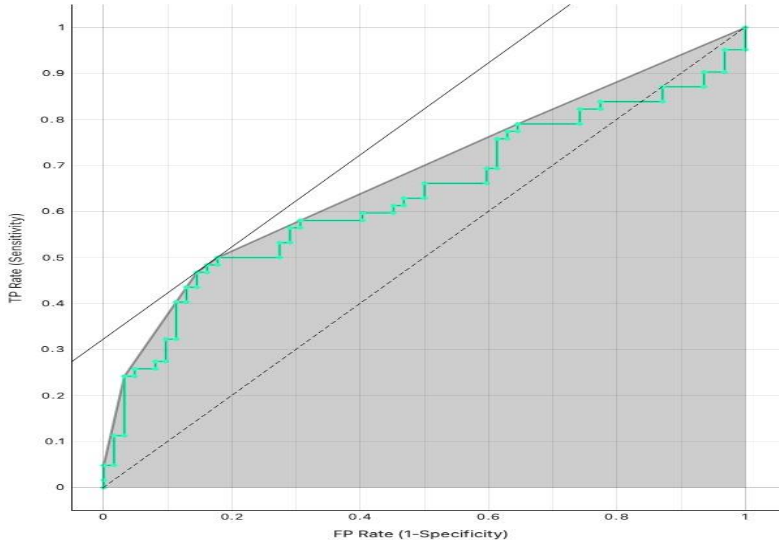


Figure 4.4 ROC curve of GBMs and PCNSLs in the Balanced Dataset – Logistic Regression model

	Predicted PCNSL	**Predicted GBM**
Actual PCNSL	**110 (TP)**	**14 (FN)**
Actual GBM	**85 (FP)**	**40 (TN)**

Table 4.10 Confusion Matrix – Logistic Regression- Balanced Dataset (GBM vs PCNSL)

4.2.2 Random Forest

The accuracy of the Random Forest algorithm increased to 86.3% compared with logistic regression, with an AUC of 0.93. The performance gap between the two groups narrowed accordingly: the model correctly identified 112 out of 124 GBMs and 102 out of 124 PCNSL/SCNSL cases, yielding a sensitivity of 86% and a specificity of 85%. These results are acceptable and comparable to clinical diagnostic performance.

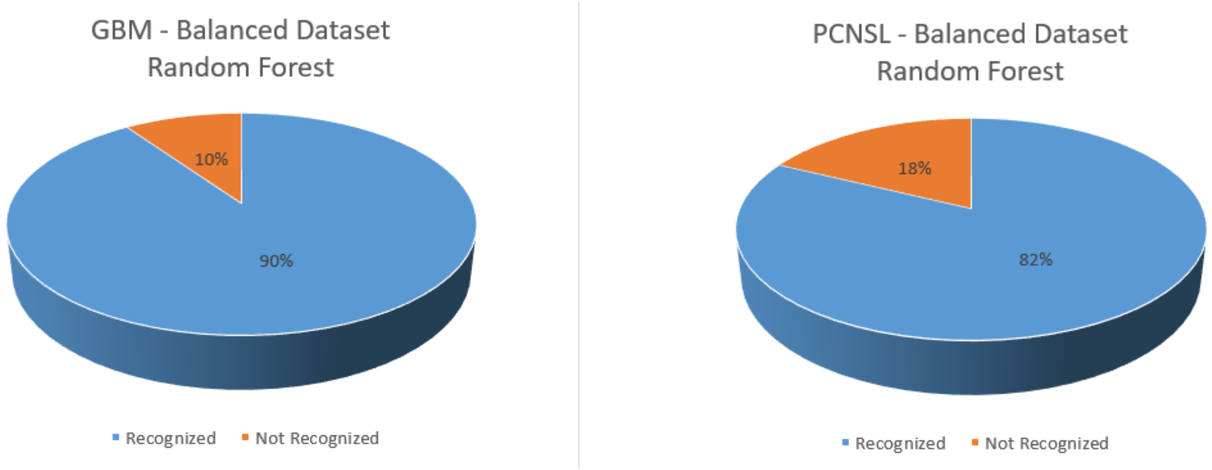


Table 4.11 Identification of GBMs and PCNSLs in the Balanced Dataset – Random Forest model

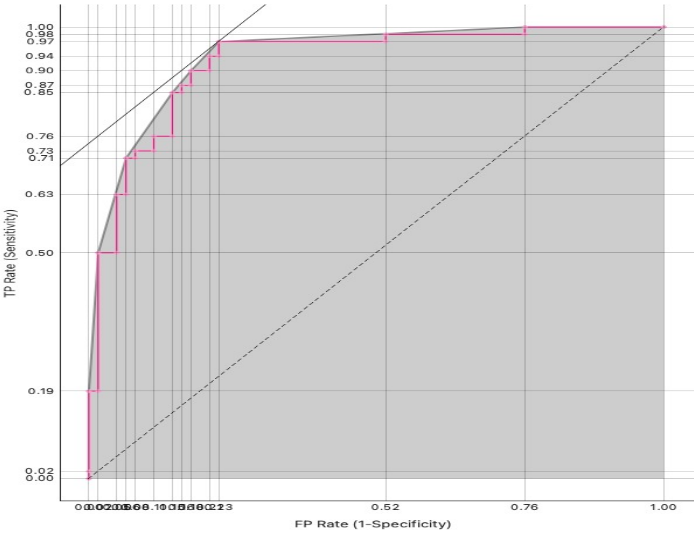


Figure 4.5 ROC curve of GBMs and PCNSLs in the Balanced Dataset – Random Forest model

	Predicted PCNSL	**Predicted GBM**
Actual PCNSL	**102 (TP)**	**22 (FN)**
Actual GBM	**13 (FP)**	**112 (TN)**

Table 4.12 Confusion Matrix – Randon Forest- Balanced Dataset (GBM vs PCNSL)

4.2.3 Neural Network

This algorithm demonstrated the highest overall performance, achieving an accuracy of 88% and an AUC of 0.92. The neural network correctly classified 102 out of 124 GBMs and 116 out of 124 lymphomas, with improved specificity and sensitivity of 89% and 88%, respectively.

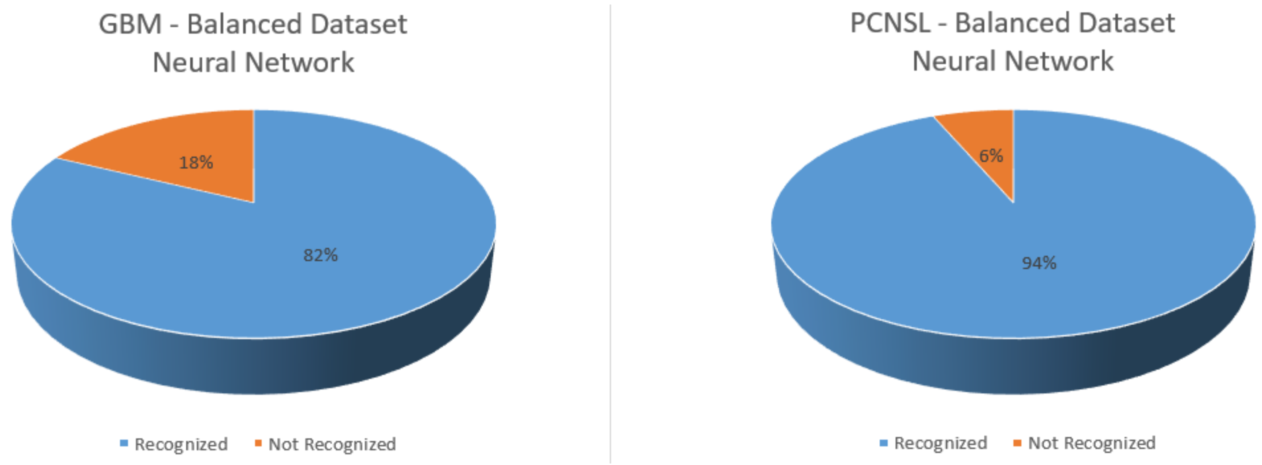


Table 4.13 Identification of GBMs and PCNSLs in the Balanced Dataset – Neural Network model

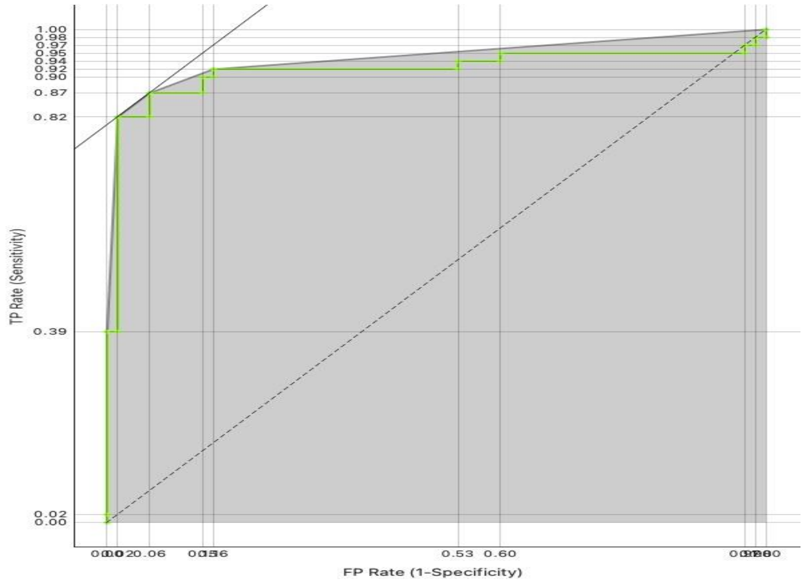


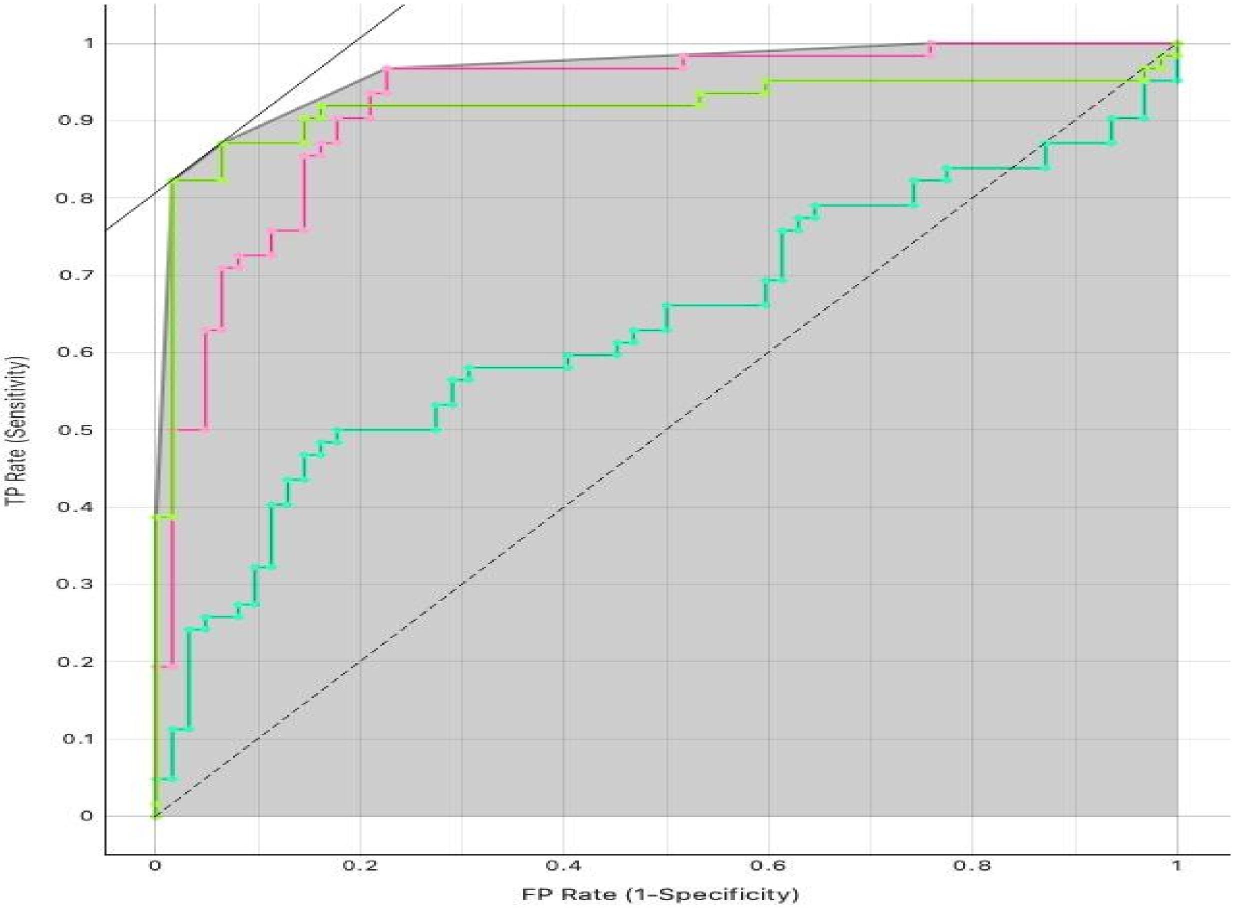
Figure 4.6 ROC curve of GBMs and PCNSLs in the Balanced Dataset – Neural Network model

	Predicted PCNSL	**Predicted GBM**
Actual PCNSL	**116 (TP)**	**8 (FN)**
Actual GBM	**23 (FP)**	**102 (TN)**

Table 4.14 Confusion Matrix – Randon Forest- Balanced Dataset (GBM vs PCNSL)

4.2.4 Comparison between the ROC curves

In the graph below, the three ROC curves derived from the balanced dataset are compared. The ROC curve corresponding to the Logistic Regression algorithm is shown in blue, the curve associated with the Neural Network in green, and the Random Forest curve is depicted in pink.



Model	AUC	95% CI
Logistic Regression	0.635	0.566 – 0.704
Random Forest	0.930	0.897 – 0.963
Neural Network	0.920	0.884 – 0.956

Table 4.15 AUC with 95% Confidence Intervals (Hanley & McNeil)

5. DISCUSSION

Distinguishing glioblastoma (GBM) from primary or secondary central nervous system lymphoma (PCNSL/SCNSL) is a critical task in neuro-oncology, as the therapeutic implications are profound and immediate. A correct preoperative diagnosis determines whether a patient should undergo maximal safe surgical resection—standard of care for GBM—or instead be managed primarily with medical therapy, as required for CNS lymphoma, where aggressive surgical removal confers no benefit and may increase morbidity. The present study aimed to evaluate whether machine-learning (ML) and deep-learning (DL) models trained exclusively on post-contrast T1-weighted (T1w+Gd) MRI could reliably differentiate GBM from CNS lymphoma, thereby offering clinicians a potential decision-support tool during the diagnostic workup.

A distinctive feature of this study is its reliance on a single MRI sequence (T1w+Gd) for model training. While this reduces heterogeneity and simplifies the analysis pipeline, it also deviates from conventional neuroradiological practice, where multiparametric MRI—including T1w pre- and post-contrast, T2w, FLAIR, DWI/ADC, SWI, and perfusion imaging—is typically used to differentiate GBM from lymphoma [113–116]. These additional sequences provide essential physiological and microstructural information, such as diffusion restriction (commonly seen in lymphoma), vascular proliferation (frequently noted in GBM), and susceptibility effects [114-115]. Therefore, the single-sequence approach adopted in this work represents both a methodological constraint and a strength: it tests whether fundamental radiological differences between these tumors can be detected using only core anatomical sequences, which could enhance the generalizability and reproducibility of ML pipelines across centers.

Despite the unisequence limitation, the results clearly demonstrate the potential of AI models to recognize tumor-specific imaging patterns. Logistic regression (LR), used as a traditional baseline classifier, yielded an accuracy of 40% (AUC 0.50) on the original dataset and 60% (AUC 0.635) on the balanced dataset. These figures are decidedly below clinical acceptability and underscore the inability of linear models to capture the nuanced, nonlinear radiological signatures of GBM and CNS lymphoma. The poor performance of LR is consistent with prior evidence indicating that classical linear classifiers often underperform in complex imaging classification problems [117–118]. In particular, features such as ring enhancement, necrotic areas, tumor heterogeneity, and infiltrative margins—characteristic of GBM—do not follow simple linear relationships with voxel intensity or texture. Conversely, lymphomas often demonstrate homogeneous contrast enhancement and pronounced diffusion restriction, patterns that require more sophisticated algorithms for accurate recognition.

The Random Forest (RF) model, capable of modeling higher-order interactions, showed substantial improvement, achieving an accuracy of 76% (AUC 0.77) on the original dataset and 86.3% (AUC 0.93) on the balanced dataset. This superior performance reflects the algorithm's ability to handle nonlinear structural features and complex radiomic relationships. Importantly, the balanced dataset—achieved using SMOTE—significantly enhanced RF performance, particularly in sensitivity (86%) and specificity (85%). This result aligns with extensive literature demonstrating that class imbalance can severely hinder ML models in medical imaging, particularly when one diagnostic group (e.g., PCNSL) is far less common than the other (GBM) [119–120]. Augmentation techniques, when properly applied, can mitigate these biases and improve model reproducibility.

The neural network (NN) achieved the best performance in both datasets, with an accuracy of 78% (AUC 0.81) on the original dataset and 88% (AUC 0.92) on the balanced dataset. These findings confirm that DL models—given their hierarchical feature extraction capabilities—excel at capturing the deeper imaging characteristics that differentiate GBM from lymphoma. Similar studies have shown that convolutional neural networks (CNNs) can outperform classical ML models in differentiating brain tumors, even using limited imaging inputs [121–123]. The improved sensitivity and specificity observed in our NN model (89% and 88%, respectively) approach the diagnostic performance reported in multiparametric CNN studies, despite our more restrictive imaging input.

The clinical implications of these findings are particularly significant. The preoperative identification of PCNSL is critical, as maximal surgical resection does not benefit these patients and exposes them to unnecessary risks. Current clinical guidelines emphasize stereotactic biopsy followed by high-dose methotrexate–based chemotherapy as the standard of care for CNS lymphoma [113,124]. Conversely, GBM management requires maximal safe resection, which has been repeatedly associated with improved survival and functional outcomes [125-126–127]. Failure to correctly differentiate between these tumors can thus lead to either overtreatment or undertreatment with substantial clinical consequences. An AI model capable of providing reliable preoperative predictions—even if not definitive—may assist the clinician in selecting the most appropriate initial management pathway and reduce the risk of inappropriate intervention.

It is also notable that all models performed better when trained on the balanced dataset, emphasizing the critical importance of dataset construction in AI research. Many neuro-oncological datasets suffer from inherent imbalance due to the relative rarity of some tumor types, such as PCNSL. Our findings are consistent with the broader consensus that balanced datasets and augmentation strategies significantly enhance model performance, especially in domains with limited sample sizes or rare diseases [128]. This suggests that future research in neuro-oncology ML should prioritize data curation, multi-center collaboration, and the routine evaluation of class imbalance mitigation strategies.

However, several limitations must be acknowledged. The most significant is the reliance on a single MRI sequence. While our findings suggest that T1w+Gd imaging alone contains sufficient diagnostic signal to differentiate GBM from lymphoma using advanced models, incorporating multiparametric MRI would almost certainly improve performance. Studies employing radiomic and DL analyses on multiparametric datasets—especially those including DWI/ADC and perfusion parameters—have reported AUC values exceeding 0.95 for GBM vs CNS lymphoma classification [121,129–130]. Therefore, the current findings represent a promising but preliminary demonstration of feasibility rather than a final clinical solution. Another limitation concerns sample size. AI models, particularly deep networks, generally require large datasets to achieve optimal generalizability. The current study, while methodologically rigorous, would benefit from additional patient accrual, which would improve model training, reduce variance, and enhance external validity. Expanding the dataset to multiple institutions—with different scanners, acquisition protocols, and population characteristics—would further test the robustness of the models.

Overall, the present study demonstrates that artificial intelligence models, when adequately trained and properly balanced, can predict the radiological diagnosis of GBM versus CNS lymphoma using only T1w+Gd MRI, approaching clinically meaningful performance. These findings lay the groundwork for a future in which AI-assisted neuroradiology may inform treatment decisions at the earliest stages of care, potentially avoiding unnecessary surgical procedures or accelerating appropriate oncologic therapy. Continued development—particularly through multicenter data integration, inclusion of multiparametric imaging, and prospective validation—will be essential for translating these promising results into clinical practice.

6. CONCLUSION

In this study, we investigated whether machine-learning and deep-learning models trained exclusively on T1w+Gd MRI can reliably differentiate glioblastoma from primary or secondary CNS lymphoma—two pathologies that, despite overlapping imaging characteristics, require fundamentally different therapeutic strategies. Our findings indicate that even within the constraints of a single-sequence paradigm, AI-driven analysis can extract diagnostically meaningful features and achieve performance metrics approaching clinical relevance. This supports the hypothesis that core anatomical MRI contains sufficient latent discriminatory information which, when interrogated by advanced computational models, can contribute to accurate preoperative stratification.

Among the tested approaches, the neural network model consistently outperformed classical machine-learning algorithms, underscoring the advantage of hierarchical feature extraction in capturing the complex morphological and textural signatures that distinguish GBM from lymphoma. The Random Forest model also achieved robust results, particularly when trained on a balanced dataset, highlighting the importance of addressing class imbalance—a pervasive challenge in neuro-oncological imaging datasets. In contrast, logistic regression showed limited diagnostic utility, reaffirming that linear models are ill-suited to the nonlinear radiological patterns inherent in high-grade brain tumors and lymphoproliferative lesions.

The improvement observed across all models when trained on the balanced dataset underscores a central methodological message: dataset composition is as critical as model architecture. Oversampling strategies such as SMOTE substantially enhanced both sensitivity and specificity, suggesting that careful curation and augmentation are indispensable steps in developing reliable AI tools, especially for rare entities like PCNSL. As neuro-oncology increasingly adopts computational methods, the establishment of harmonized, multicenter datasets will be essential to overcome sample-size constraints and enhance generalizability across institutions and imaging platforms.

Clinically, the implications of a reliable preoperative classifier are profound. Misclassification between GBM and CNS lymphoma exposes patients to the risks of unnecessary craniotomy or, conversely, delays initiation of definitive chemotherapeutic regimens. An AI-based decision-support tool—used in conjunction with expert radiological assessment—could aid in tailoring the diagnostic pathway, informing whether to proceed directly to maximal safe resection or prioritize stereotactic biopsy and early medical therapy. Although such a tool is not intended to replace comprehensive multiparametric MRI or histopathologic confirmation, it may provide valuable adjunctive information at a point in the workflow where rapid, informed decision-making is critical.

Nonetheless, the limitations of this work must be considered when interpreting its findings. The reliance on T1w+Gd MRI alone, while methodologically intentional, restricts the models' ability to capture diffusion, perfusion, and susceptibility-related features that are known to enhance the discrimination of GBM from lymphoma. Furthermore, the modest sample size limits the potential for deep-learning architectures to reach their full performance and raises concerns about overfitting despite the implemented controls. Prospective validation, larger multicenter cohorts, and integration of multiparametric MRI sequences—particularly DWI/ADC and perfusion imaging—represent natural and necessary next steps for refining these models.

In summary, this study provides a proof-of-concept demonstration that AI models trained solely on T1w+Gd MRI can achieve promising performance in differentiating glioblastoma from CNS lymphoma. While not yet sufficient for independent clinical deployment, these results highlight the feasibility and potential value of computational tools in neuro-oncological decision support. Advancing this work through expanded datasets, richer imaging inputs, and prospective evaluation will be essential to fully harness the capabilities of machine learning and deep learning in precision neuro-oncology, ultimately contributing to more accurate diagnoses, better-informed surgical decision-making, and improved patient outcomes.

References

1. Hoang-Xuan K, et al. European Association of Neuro-Oncology (EANO) guidelines for treatment of primary CNS lymphoma. *Neuro-Oncology*. 2022.
2. Kotecha R, et al. Key clinical principles in the management of glioblastoma. *JCO Oncology Practice*. 2023.
3. Hung ND et al. Differentiation of glioblastoma and primary CNS lymphomas using multiparametric diffusion and perfusion MRI. *Biomedical Reports*. 2023.
4. Makino K, Hirai T, Nakamura H, et al. Differentiating Between Primary Central Nervous System Lymphomas and Glioblastomas: Combined Use of Perfusion-Weighted and Diffusion-Weighted Magnetic Resonance Imaging. *World Neurosurgery*. 2018.
5. Bathla G, et al. Radiomics-based differentiation between GBM and PCNSL across MRI sequences and ML pipelines. *European Radiology*. 2021;
6. Chen C, Zheng A, Ou X, Wang J, Ma X. Comparison of Radiomics-Based Machine-Learning Classifiers in Diagnosis of Glioblastoma From Primary Central Nervous System Lymphoma. *Frontiers in Oncology*. 2020.
7. McAvoy M, et al. Classification of GBM versus PCNSL using convolutional neural networks on CE-T1 MRI. *Scientific Reports*. 2021.
8. Garaba A, et al. Radiomics for differentiation of gliomas from PCNSL: systematic review and meta-analysis. *Frontiers in Oncology*. 2024.
9. Wharton S, Bowtell R. Gradient echo MRI: rationale, methods and applications. *NeuroImage*. 2012.
10. Setsompop K, Cohen-Adad J, et al. Improved simultaneous multi-slice MRI using blipped-CAIPI. *Magnetic Resonance in Medicine*. 2012.
11. McNab JA, Edlow BL, et al. The Human Connectome Project and high-gradient diffusion MRI. *NeuroImage*. 2013.
12. Wang Y, Liu T. Quantitative susceptibility mapping: physical principles and clinical applications. *Magnetic Resonance in Medicine*. 2015.
13. Uğurbil K. Imaging at ultrahigh magnetic fields: insights and challenges for neuroimaging. *NeuroImage*. 2014.
14. Kickingereder P, et al. Radiogenomics of glioblastoma: machine learning–based classification of molecular characteristics. *Clinical Cancer Research*. 2016.
15. Carrillo JA, et al. Combining imaging biomarkers with conventional MRI improves characterization of glioblastoma. *Neuro-Oncology*. 2012.

16. Hagiwara A, et al. Improved T2 mapping for glioma characterization. *Investigative Radiology*. 2019.
17. Kim R, et al. FLAIR signal analysis in differentiation of pseudoprogression versus true progression in high-grade gliomas. *European Radiology*. 2014.
18. Gupta RK, et al. Diffusion MRI in the evaluation of brain tumors. *American Journal of Neuroradiology*. 2013.
19. Seiler A, et al. SWI and QSM in brain tumor imaging: clinical applications. *NeuroImage: Clinical*. 2020.
20. Law M, et al. DSC perfusion MRI in grading and assessing gliomas. *Radiology*. 2013.
21. Chawla S, et al. MR spectroscopy in brain tumor diagnosis and monitoring. *Journal of Magnetic Resonance Imaging*. 2015.
22. Ostrom QT, et al. CBTRUS Statistical Report: Primary Brain and Other CNS Tumors Diagnosed in the United States. *Neuro-Oncology*. 2021.
23. Louis DN et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology* 2021.
24. Ostrom QT, et al. Adult glioma incidence and survival by race or ethnicity in the United States. *JNCI*. 2018.
25. Pollack IF, et al. Pediatric high-grade gliomas: biology and clinical perspectives. *Neuro-Oncology*. 2011.
26. Sadetzki S, et al. Radiation-induced meningioma and glioma: dose-response relationships. *Radiology*. 2005.
27. Wrensch M, et al. Environmental risk factors for primary malignant brain tumors. *Journal of Neuro-Oncology*. 2002.
28. Benson VS, et al. Mobile phone use and risk of brain neoplasms and other cancers. *International Journal of Epidemiology*. 2013.
29. Durno C, et al. Li-Fraumeni syndrome and CNS tumors: tumor risk and surveillance strategies. *Neuro-Oncology*. 2017.
30. Shete S, et al. Genome-wide association study identifies five susceptibility loci for glioma. *Nature Genetics*. 2009.
31. Kinnersley B, Houlston RS, Bondy ML. Genome-Wide Association Studies in Glioma. *Cancer Epidemiology, Biomarkers & Prevention.* 2018.
32. Brennan CW, et al. The somatic genomic landscape of glioblastoma. *Cell*. 2013.
33. Eckel-Passow JE, et al. Glioma groups based on 1p/19q, IDH, and TERT mutations. *New England Journal of Medicine*. 2015.

34. Weller M, et al. Molecular classification of diffuse gliomas. *Nature Reviews Neurology*. 2015.
35. Chen HM, Nikolic A, Singhal D, Gallo M. Roles of Chromatin Remodelling and Molecular Heterogeneity in Therapy Resistance in Glioblastoma. *Cancers (Basel).* 2022.
36. Hegi ME, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *New England Journal of Medicine*. 2005.
37. Hambardzumyan D, et al. The role of the microenvironment in glioblastoma progression. *Nature Reviews Cancer*. 2016.
38. Jain RK, et al. Angiogenesis in brain tumors. *Nature Reviews Neuroscience*. 2007.
39. Stupp R, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine*. 2005.
40. Sipos D, Raposa BL, Freihat O, Simon M, Mekis N, Cornacchione P, Kovács Á. Glioblastoma: Clinical Presentation, Multidisciplinary Management, and Long-Term Outcomes. *Cancers (Basel).* 2025.
41. Englot DJ, et al. Seizures in glioma patients: risk factors and mechanisms. *Neuro-Oncology*. 2016.
42. Schaff LR, Mellinghoff IK. Glioblastoma and Other Primary Brain Malignancies in Adults: A Review. JAMA. 2023.
43. Taphoorn MJB, et al. Cognitive deficits in adult glioma patients. *Lancet Neurology*. 2010.
44. Takano S, et al. Hemorrhagic presentation of glioblastoma: clinical and imaging findings. *Journal of Neurosurgery*. 2014.
45. Pope WB, et al. MRI in glioblastoma: imaging features and tumor behavior. *Radiology*. 2005.
46. Law M, et al. Glioma grading using dynamic susceptibility contrast perfusion MRI. *AJNR*. 2003.
47. Hygino da Cruz LC, et al. FLAIR imaging in high-grade gliomas: infiltrative patterns and prognostic value. *AJNR*. 2011.
48. Arvinda HR, et al. Diffusion-weighted imaging in glioblastoma characterization. *Journal of Neuroimaging*. 2009.
49. Hu LS, et al. Perfusion MRI for differentiating pseudoprogression from true progression in glioblastoma. *Neuro-Oncology*. 2012.
50. Galijasevic M, Steiger R, Mangesius S, et al. Magnetic Resonance Spectroscopy in Diagnosis and Follow-Up of Gliomas: State-of-the-Art. *Cancers (Basel*). 2022

51. Stadlbauer A, et al. Advanced functional imaging in glioma assessment. *NeuroImage: Clinical*. 2017.
52. Chamberlain MC, et al. Diagnostic challenges in differentiating glioblastoma from mimics. *Expert Review of Neurotherapeutics*. 2014.
53. Ellison DW, et al. Neuropathology of glioblastoma: defining histologic and molecular features. *Acta Neuropathologica*. 2019.
54. Brat DJ, et al. Biomarker profiles in glioblastoma. *Brain Pathology*. 2015.
55. Louis DN, et al. Integrated molecular diagnostics in CNS tumors. *Acta Neuropathologica*. 2016.
56. Weller M, et al. Molecular predictors of survival in glioblastoma. *Neuro-Oncology*. 2014.
57. Wen PY, Weller M, Lee EQ, et al. Glioblastoma in adults: a Society for Neuro-Oncology (SNO) and European Society of Neuro-Oncology (EANO) consensus review on current management and future directions. *Neuro-Oncology*. 2020.
58. Hegi ME, et al. MGMT methylation as a predictive biomarker in glioblastoma. *Clinical Cancer Research*. 2009.
59. Wick W, et al. Lomustine-based regimens for recurrent glioblastoma. *Lancet Oncology*. 2010.
60. Friedman HS, et al. Bevacizumab therapy for recurrent glioblastoma. *Journal of Clinical Oncology*. 2009.
61. Lieberman F, et al. Emerging therapeutic strategies in glioblastoma. *Nature Reviews Clinical Oncology*. 2019.
62. Chang EL, et al. Radiotherapy principles in glioblastoma management. *International Journal of Radiation Oncology Biology Physics*. 2012.
63. Minniti G, et al. Intensity-modulated radiotherapy for glioblastoma. *Radiotherapy and Oncology*. 2010.
64. Roa W, et al. Hypofractionated radiotherapy in elderly glioblastoma patients. *Journal of Clinical Oncology*. 2004.
65. Joiner MC, et al. Mechanisms of radioresistance in glioblastoma. *Seminars in Radiation Oncology*. 2009.
66. Barker CA, et al. Integrating immunotherapy with radiotherapy in glioblastoma. *Neuro-Oncology*. 2020.
67. Lacroix M, et al. Surgical resection and survival in glioblastoma. *Journal of Neurosurgery*. 2001.

68. Sanai N, et al. Extent of resection and impact on outcomes in glioblastoma. *Neurosurgery*. 2011.
69. Stummer W, et al. 5-ALA fluorescence-guided surgery in glioblastoma. *Lancet Oncology*. 2006.
70. Hervey-Jumper SL, et al. Advances in neurosurgical techniques for glioblastoma. *Nature Reviews Neurology*. 2018.
71. Deckert M et al. Primary lymphoma of the central nervous system. *Hematological Oncology* 2014.
72. Villano JL, et al. Incidence trends in primary CNS lymphoma in the general population. *Neuro-Oncology*. 2011.
73. Shiels MS, et al. Epidemiologic patterns of primary CNS lymphoma in older adults. *Journal of Clinical Oncology*. 2014.
74. Powell J, et al. Primary CNS lymphoma in immunocompromised patients: clinical outcomes. *Blood*. 2015.
75. Camilleri-Broet S, et al. EBV-associated primary CNS lymphoma: molecular pathology. *Brain Pathology*. 2004.
76. Koh YW, et al. Immune dysregulation and risk factors for CNS lymphoma. *Haematologica*. 2019.
77. Chapuy B, et al. MYD88 and CD79B mutations define pathogenic drivers in PCNSL. *Cancer Cell*. 2016.
78. Fankhauser L, et al. Epigenetic regulation in primary CNS lymphoma. *Nature Communications*. 2021.
79. Bataille B, et al. Perivascular tropism as a defining feature of PCNSL. *Brain*. 2000.
80. Grommes C, et al. CNS microenvironmental influences on lymphoma survival. *Nature Reviews Clinical Oncology*. 2019.
81. Jahnke K, et al. Blood-brain barrier limitations in CNS lymphoma therapy. *Annals of Oncology*. 2007.
82. Ferreri AJM, et al. Neurological presentations of primary CNS lymphoma. *Neurology*. 2010.
83. Küker W, et al. MRI characteristics differentiating PCNSL from other brain tumors. *Radiology*. 2005.
84. Toh CH, et al. Diffusion-weighted MRI in primary CNS lymphoma. *AJNR*. 2010.
85. Mangla R, et al. Perfusion MRI features of primary CNS lymphoma. *Neuro-Oncology*. 2013.
86. Fountas KN, et al. MR spectroscopy in CNS lymphoma. *Journal of Neuro-Oncology*. 2004.

87. Ferreri AJ, et al. High-dose methotrexate in PCNSL: therapeutic foundation. *Lancet Oncology*. 2009.
88. Houillier C, et al. Combination chemotherapy regimens in PCNSL. *Journal of Clinical Oncology*. 2012.
89. Abrey LE, Batchelor TT, Ferreri AJM, et al. Report of an International Workshop to Standardize Baseline Evaluation and Response Criteria for Primary CNS Lymphoma. *Journal of Clinical Oncology.* 2005.
90. Turing A. Computing Machinery and Intelligence. *Mind*. 1950.
91. McCarthy J, et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *Dartmouth College*. 1955.
92. Campbell M, et al. Deep Blue. *Artificial Intelligence*. 2002.
93. Esteva A, et al. A guide to deep learning in healthcare. *Nature Medicine*. 2019.
94. Mitchell TM. *Machine Learning*. McGraw-Hill. 1997.
95. Bishop CM. *Pattern Recognition and Machine Learning*. Springer. 2006.
96. Hastie T, et al. *The Elements of Statistical Learning*. Springer. 2009.
97. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. MIT Press. 2018.
98. Breiman L. Random Forests. *Machine Learning*. 2001.
99. LeCun Y, et al. Deep learning. *Nature*. 2015.
100. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*. 1997.
101. Avberšek LK, Repovš G. Deep learning in neuroimaging data analysis: Applications, challenges, and solutions. *Frontiers in Neuroimaging*. 2022.
102. Lambin P, et al. Radiomics: extracting more information from medical images. *Nature Reviews Clinical Oncology*. 2017.
103. Aerts HJWL, et al. Decoding tumour phenotype by noninvasive imaging using radiomics. *Nature Communications*. 2014.
104. Doi K. Computer-aided diagnosis in medical imaging. *Computers in Medical Imaging and Graphics*. 2007.
105. Gillies RJ, et al. Radiomics: images are more than pictures, they are data. *Radiology*. 2016.
106. Zwanenburg A, et al. The Image Biomarker Standardisation Initiative. *Radiology*. 2020.
107. Traverso A, et al. Repeatability and reproducibility in radiomics. *Radiotherapy and Oncology*. 2018.

108. Shafiq-ul-Hassan M, et al. Voxel size and gray level normalization in radiomics. *Physics in Medicine and Biology*. 2017.

109. Larue R, et al. Influence of imaging parameters on radiomics features. *European Radiology*. 2017.

110. Mayerhoefer ME, et al. Effects of MR acquisition parameters on radiomic features. *Magnetic Resonance in Medicine*. 2019.

111. Parmar C, et al. Robust radiomic feature selection. *Scientific Reports*. 2015.

112. Kursa MB, Rudnicki WR. Feature selection with Boruta. *Journal of Statistical Software*. 2010.

113. Grommes C, Rubenstein JL, DeAngelis LM, Ferreri AJM, Batchelor TT. Comprehensive approach to diagnosis and treatment of newly diagnosed primary CNS lymphoma. Neuro-Oncology. 2019

114. Horbinski C, Berger T, Packer RJ, Wen PY. Clinical implications of the 2021 edition of the WHO classification of central nervous system tumours. *Nature Reviews Neurology*. 2022;

115. Law M, et al. Perfusion and diffusion MRI in differentiating GBM from CNS lymphoma. *Radiology*. 2002.

116. Haldorsen IS, Larsson EM. MRI characteristics of CNS lymphoma. *Neuroradiology*. 2011.

117. Bauer S, et al. MRI-based machine learning in brain tumor imaging. *Physics in Medicine and Biology*. 2013.

118. Zacharaki EI, et al. Classification of brain tumors using machine learning. *IEEE Transactions on Medical Imaging*. 2009.

119. Kocak B, et al. Effects of class imbalance in radiomics-based machine learning. *Medical Physics*. 2020.

120. Zhu G, Jiang B, Tong L, Xie Y, Zaharchuk G, Wintermark M. Applications of Deep Learning to Neuro-Imaging Techniques. Frontiers in Neurology. 2019

121. Matsui Y, et al. Machine learning differentiation of GBM vs PCNSL on MRI. *European Radiology*. 2021.

122. Suh HB, et al. Radiomics for CNS lymphoma differentiation. *AJNR*. 2018.

123. Chen C, et al. Deep learning applications in brain tumor classification. *Frontiers in Oncology*. 2021.

124. Batchelor TT, DeAngelis LM. Management of primary CNS lymphoma. *Journal of Clinical Oncology*. 2006.

125. Fabbro-Peray P, Zouaoui S, Darlix A, et al. Association of patterns of care, prognostic factors, and use of radiotherapy–temozolomide therapy with survival in patients with newly diagnosed glioblastoma: a French national population-based study. *Journal of Neuro-Oncology*. 2019.

126. Wykes V, Zisakis A, Irimia M, Ughratdar I, Sawlani V, Watts C. Importance and Evidence of Extent of Resection in Glioblastoma. Journal of Neurological Surgery Part A: Central European Neurosurgery. 2021;

127. Brown TJ, et al. Survival impact of maximal resection in GBM. *JAMA Oncology*. 2016.

128. Wen Q, et al. Challenges in deploying AI in neuroimaging. *Brain*. 2022.

129. Kickingereder P, et al. Radiomics in neuro-oncology. *Neuro-Oncology*. 2020.

130. Zhang Z, et al. Multiparametric MRI radiomics for brain tumor differentiation. *NeuroImage: Clinical*. 2020.