HJK HOLISTIK JURNAL KESEHATAN
ISSN 1978-3337 (Print)
ISSN 2620-7478 (Online)
DOI: 10.33024

Nomor: 79/E/KPT/2023

INFORMASI ARTIKEL
Received: June, 15, 2026
Revised: July, 20, 2026
Available online: July, 31, 2026
at: https://e-jurnal.iphorr.com/index.php/hjk

Predicting outcomes of brain metastases treated with stereotactic radiosurgery using deep learning: A systematic review

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

Abstract

Background: Brain metastases are a major cause of morbidity in cancer patients and are commonly treated with stereotactic radiosurgery. However, accurate prediction of post-treatment outcomes, including local control, recurrence, and radiation-induced toxicity, remains a significant clinical challenge. Deep learning has emerged as a promising approach to improve predictive modeling by integrating high-dimensional imaging and clinical data.
Purpose: To systematically evaluate the performance and clinical utility of deep learning models in predicting treatment outcomes in patients with brain metastases following SRS.
Method: A systematic review was conducted following PRISMA guidelines. After duplicate removal and screening, eligible full-text articles applying deep learning for outcome prediction were included. Risk of bias was assessed using the PROBAST tool.
Results: DL approaches included convolutional neural networks, transformers, ensemble models, and hybrid radiomics-DL frameworks. Reported performance was high, AUC values ranging from 0.71 to 0.99. Multimodal models integrating imaging, clinical, and genomic features achieved superior performance (AUC up to 0.945; accuracy up to 0.98) to unimodal approaches
Conclusion: Deep learning models show strong potential for predicting outcomes in brain metastases after SRS, However, methodological limitations and heterogeneity remain significant challenges.

Keywords: Brain Metastases; Deep Learning; Stereotactic Radiosurgery; Radiomics; Predictive Modelling.

PENDAHULUAN

Metastasis otak merupakan keganasan intrakranial yang paling sering ditemukan pada orang dewasa, terjadi pada sekitar 10–40% dari seluruh pasien kanker dan mencakup lebih dari separuh kasus tumor otak yang didiagnosis setiap tahunnya (Keek, Beuque, Primakov, Woodruff, Chatterjee, van Timmeren, & Lambin, 2022; Kanakarajan, De Baene, Hanssens, & Sitskoorn, 2024; Baschnagel, Meyer, Chen, Krauss, Olson, Pieper, & Grills, 2013). Tumor primer yang paling umum menyebar ke otak meliputi kanker paru, khususnya *Non-Small Cell Lung Cancer* (NSCLC), serta kanker payudara, melanoma, dan keganasan gastrointestinal (Liao, Lee, Yang, Chen, Chung, Wu, & Lu, 2023; Kanakarajan et al., 2024). Dalam beberapa dekade terakhir, deteksi yang lebih awal melalui kemajuan neuroimaging dan perpanjangan harapan hidup pasien berkat perkembangan terapi sistemik secara bersama-sama mendorong peningkatan insiden metastasis otak secara global (Zhao, Vaios, Wang, Yang, Cui, Reitman, & Wang, 2024; Cho, Cho, Choi, Sim, Jeong, Baik, & Sunwoo, 2024). Tanpa penanganan aktif, median kelangsungan hidup hanya berkisar antara minggu hingga bulan, sehingga manajemen lokal yang cepat dan efektif menjadi hal yang sangat diperlukan untuk mempertahankan fungsi neurologis dan memperpanjang harapan hidup (Kanakarajan et al., 2024; Jalalifar, Soliman, Sahgal, & Sadeghi-Naini, 2022; Suh, Kotecha, Chao, Ahluwalia, Sahgal, & Chang, 2020).

DOI: https://doi.org/10.33024/hjk.v20i5.3586

Saat ini, pilihan terapi untuk metastasis otak mencakup reseksi bedah, *whole-brain radiation therapy* (WBRT), SRS, dan SRT (Kanakarajan et al., 2024). Dalam dua dekade terakhir, tren penanganan telah bergeser secara nyata ke arah SRS dan SRT. Kedua modalitas ini memberikan radiasi yang difokuskan secara ketat pada masing-masing lesi sekaligus meminimalkan paparan pada parenkim otak di sekitarnya. Sejumlah uji klinis telah membuktikan keunggulan keduanya dibanding WBRT dalam membatasi toksisitas neurokognitif (Cho et al., 2024; Erdur, Scholz, Nguyen, Buchner, Mayinger, Christ, & Peeken, 2025). Meskipun demikian, kegagalan lokal masih terjadi pada sekitar 10–30% lesi yang ditangani, yang menunjukkan betapa pentingnya alat prediksi luaran yang andal dan individual untuk memandu keputusan eskalasi dosis dan penjadwalan surveilans (Erdur et al., 2025; Zhao, Vaios, Yang, Lu, Floyd, Yang, & Wang, 2025a).

Salah satu komplikasi yang khas setelah SRS adalah *radionecrosis* (RN), yaitu cedera inflamasi tertunda pada parenkim otak di sekitar lesi. Secara radiologis, RN kerap sulit dibedakan dari rekurensi tumor yang sesungguhnya: keduanya dapat memperlihatkan pembesaran lesi, peningkatan kontras berbentuk cincin, dan edema perilesional pada MRI konvensional (Zhao et al., 2024; Zhao et al., 2025a; Kawahara, Tang, Lee, Nagata, & Watanabe, 2021). Penanganan RN yang umumnya berupa kortikosteroid atau bevacizumab berbeda jauh dari re-iradiasi atau eskalasi sistemik untuk rekurensi, ketidakmampuan membedakan kedua entitas ini secara non-invasif menjadi masalah klinis yang mendesak. Biopsi jaringan tetap menjadi standar baku, tetapi membawa risiko prosedural yang nyata dan tidak praktis untuk lesi multifokus atau yang sulit dijangkau secara bedah (Chang, Qi, Wu, Yue, Zhang, & Qian, 2023; Touati, Bourbonne, Dissaux, Goasduff, Pradier, Peltier, & Lucia, 2023).

Secara historis, para klinisi mengandalkan indeks prognostik seperti *Recursive Partitioning Analysis* (RPA) dan *Graded Prognostic Assessment* (GPA) untuk stratifikasi risiko pada metastasis otak. Namun, alat-alat tersebut hanya mampu memberikan estimasi kasar pada level kelompok dan tidak dapat memandu terapi di tingkat individu (Jalalifar et al., 2022; Liao et al., 2023; Douri, Iorio-Morin, Mercure-Cyr, Figueiredo, Touchette, Masson-Côté, & Mathieu, 2023). Radiomik telah menarik perhatian sebagai komplemen yang lebih terperinci, dengan mengekstraksi fitur kuantitatif dari citra medis yang mampu menangkap heterogenitas tumor yang tidak terlihat oleh mata manusia (Cha, Jang, Kim, Yoo, Paik, Jeong, & Youn, 2018; Liao et al., 2023; Kanakarajan et al., 2024; Ko, Kim, Lee, & Kim, 2020). Namun, alur kerja radiomik bergantung pada rekayasa fitur manual, rentan terhadap variabilitas pemindai dan protokol akuisisi, serta mudah mengalami overfitting pada kohort kecil faktor-faktor yang secara kolektif membatasi reprodusibilitas dan generalisabilitasnya (de Azevedo Santos, Tundisi, Ramos, Maia, Pellizzon, Silva, & de Castro, 2015; Karami, Ruschin, Soliman, Sahgal, Stanisz, & Sadeghi-Naini, 2019).

Deep learning (DL) melampaui sebagian besar keterbatasan tersebut dengan mempelajari representasi yang relevan dengan tugas secara *end*-to-end dari data citra mentah, tanpa rekayasa fitur yang ditentukan sebelumnya. CNN, jaringan rekuren seperti model LSTM, dan *Vision Transformer* yang lebih baru semuanya telah menunjukkan kinerja yang kuat dalam pencitraan saraf (Jalalifar et al., 2022; Cho et al., 2024; Erdur et al., 2025; Ibragimov, Toesca, Chang, Yuan, Koong, Xing, & Vogelius, 2020). Kerangka hibrida yang memadukan fitur citra DL dengan variabel klinis atau genomik telah memperluas jangkauan prediksi lebih jauh lagi. Meskipun demikian, tantangan tetap ada: sebagian besar model bersifat "kotak hitam" yang tidak transparan, dataset yang digunakan sering kali kecil dan tidak seimbang, dan hanya sedikit studi yang menyertakan validasi eksternal (Buzea, Buga, Paun, Albu, Iancu, Dobrovat, & Eva, 2023; Kanakarajan et al., 2024; Ruzevick, Kleinberg, & Rigamonti, 2014).

Belum ada tinjauan sistematis sebelumnya yang secara komprehensif mengkaji prediksi luaran berbasis DL di seluruh spektrum titik akhir klinis kontrol lokal, kelangsungan hidup, respons terapi, diferensiasi RN, dan perubahan volume khususnya dalam konteks SRS/SRT. Oleh karena itu, tinjauan ini dilakukan untuk mengidentifikasi, menilai, dan menyintesis bukti yang tersedia, dengan memperhatikan arsitektur model, strategi pencitraan, kinerja prediktif, dan ketelitian metodologis.

METODE

Tinjauan sistematis ini dibuat sesuai dengan panduan PRISMA 2020 dan telah didaftarkan di PROSPERO sebelum studi dimulai. Pertanyaan klinis dirumuskan menggunakan kerangka PICOS: populasi target adalah orang dewasa dengan metastasis otak dari tumor primer mana pun; intervensi yang dimaksud adalah SRS atau SRT; komparator terdiri dari model prediksi klinis atau radiomik konvensional yang relevan; *Outcomes* mencakup kontrol atau kegagalan lokal, kelangsungan hidup keseluruhan, respons terapi, perubahan volume tumor, dan kejadian merugikan

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

terkait terapi seperti radionekrosis; sementara desain studi yang memenuhi syarat adalah artikel penelitian orisinal yang menerapkan *deep learning* untuk prediksi *Outcomes*.

Pencarian literatur sistematis dilakukan pada empat basis data: PubMed (n=205), Google Scholar melalui Publish or Perish (n=228), Cochrane Library (n=14), dan IEEE Xplore (n=8), mencakup seluruh publikasi yang tersedia hingga Mei 2026. Kueri pencarian menggabungkan istilah MeSH dan kata kunci teks bebas menggunakan operator Boolean. String inti yang digunakan adalah: ("brain metastases" OR "brain metastasis" OR "intracranial metastases") AND ("stereotactic radiosurgery" OR "SRS" OR "stereotactic radiotherapy" OR "SRT" OR "Gamma Knife" OR "CyberKnife") AND ("deep learning" OR "convolutional neural network" OR "artificial intelligence" OR "machine learning") AND ("outcome prediction" OR "local control" OR "overall survival" OR "treatment response" OR "radionecrosis"). Seluruh catatan diimpor ke Rayyan.ai untuk deduplikasi dan penyaringan buta secara independen.

Kriteria inklusi pada populasi merupakan pasien dewasa dengan metastasis otak, menjelaskan terapi dengan SRS atau SRT, menerapkan setidaknya satu metode DL untuk memprediksi *outcomes* yang terdefinisi secara klinis, dan melaporkan metrik kinerja kuantitatif. Kriteria eksklusi mencakup studi yang berfokus khusus pada WBRT, yang hanya mengandalkan machine learning konvensional tanpa jaringan saraf apa pun, tinjauan atau editorial, publikasi bukan berbahasa Inggris, dan artikel yang tidak tersedia dalam teks lengkap. Setelah penghapusan 138 duplikat, dua penelaah secara independen menyaring 317 catatan; 43 di antaranya lolos ke penilaian teks lengkap, dan 24 tidak dapat diperoleh. Dari 19 artikel teks lengkap yang ditelaah kelayakannya, 6 artikel dieksklusi (4 tidak memiliki komponen DL; 2 tidak memiliki *outcomes* yang terdefinisi secara klinis), sehingga menyisakan 13 studi untuk sintesis akhir.

Data diekstraksi secara independen oleh dua penelaah menggunakan formulir standar yang telah dilakukan sebelumnya, mencakup identifikasi studi, desain dan latar penelitian, karakteristik dataset, demografi pasien, fitur tumor, modalitas pencitraan, arsitektur DL, luaran yang diprediksi, strategi validasi, dan metrik kinerja. Kualitas metodologis dinilai dengan PROBAST pada empat domain; peserta, prediktor, *outcomes*, dan analisis statistik yang masing-masing dinilai sebagai risiko bias rendah, tinggi, atau tidak jelas. Mengingat heterogenitas yang besar dalam populasi pasien, input pencitraan, arsitektur, definisi luaran, dan strategi validasi, meta-analisis formal tidak memungkinkan untuk dilakukan. Sintesis naratif mengelompokkan studi berdasarkan *outcomes* yang diprediksi secara utama: kontrol atau kegagalan lokal, kelangsungan hidup keseluruhan, respons terapi, diferensiasi radio nekrosis versus rekurensi tumor, serta prediksi volume tumor atau efek samping.

Pencarian literatur menghasilkan total 455 literatur. Setelah penghapusan 136 duplikat, 319 sitasi disaring. Usai tinjauan judul dan abstrak, 43 artikel lolos ke penilaian teks lengkap; 24 di antaranya tidak dapat diperoleh. Dari 19 artikel teks lengkap yang ditelaah kelayakannya, enam dieksklusi (empat karena tidak menggunakan DL sebagai metode analitik utama; dua karena tidak memiliki *outcomes* yang terdefinisi secara klinis). Hasil akhir didapatkan 13 studi memenuhi semua kriteria inklusi. Proses seleksi ditampilkan pada Gambar 1. Berdasarkan penilaian PROBAST, sebagian besar studi mendapatkan penilaian risiko bias keseluruhan yang rendah hingga sedang.

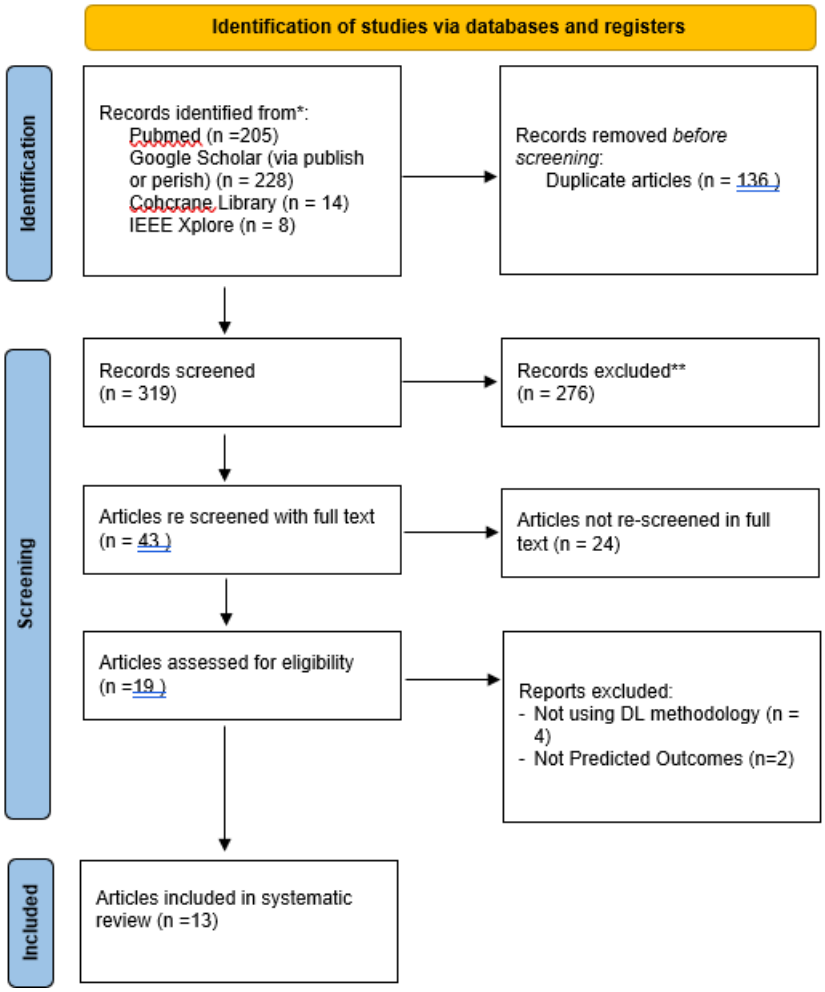
Sebanyak 13 studi yang diikutsertakan, sebagaimana dirangkum dalam Tabel 1, diterbitkan antara tahun 2018 dan 2025, berasal dari Taiwan, Belanda, Amerika Serikat, Korea Selatan, Jerman, Swiss, Siprus, Tiongkok, Kanada, Rumania, dan Belgia semuanya berdesain retrospektif. Sebagian besar terbit dalam rentang 2022–2025, mencerminkan akselerasi terkini penelitian AI di bidang neuro-onkologi. Ukuran kohort berkisar dari 19 pasien (Buzea et al., 2023) hingga lebih dari 1.400 pasien (Keek et al., 2022), dengan mayoritas berada antara 50 hingga 250 pasien. NSCLC merupakan tumor primer yang paling dominan dan menjadi satu-satunya fokus pada tiga studi; sementara kohort lainnya mencampur kanker paru, payudara, melanoma, sel ginjal, gastrointestinal, dan histologi lainnya, dengan kanker paru mencakup sekitar 60–75% pasien di sebagian besar studi.

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

HASIL



Gambar Diagram Alir PRISMA

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

Tabel Hasil Pemetaan Data Artikel

(Penulis)(Negara)	Desain	Input Pencitraan	Arsitektur DL	Hasil	Validasi	Performa
(Liao et al., 2023) (Taiwan)	237 pasien; kohort retrospektif	Radiomik MRI (GLCM, GLRLM, wavelet); klinis; status EGFR	DeepSurv (jaringan survival berbasis MLP)	Kelangsungan hidup keseluruhan (3, 6, 12, 24 bulan)	Bootstrap internal	C-index=0.75; AUC=0.92 pada 24 bulan
(Kanakarajan et al., 2024) (Belanda)	129 pasien; kohort retrospektif	Klinis + radiomik (bentuk, tekstur) + fitur DL (3D ResNet)	3D ResNet + Random Forest	Kontrol lokal vs. kegagalan lokal setelah SRT	Internal + eksternal	AUC=0.88 (DL+klinis gabungan)
(Zhao et al., 2024) (Amerika Serikat)	114 lesi / 82 pasien; retrospektif	MRI T1-CE; peta dosis (Dmap); segmentasi PTV; variabel klinis	Ensemble dalam (VGG-19x4 + proyeksi sferik + peta dosis)	Kontrol lokal vs. kegagalan lokal pasca-SRS	Internal + eksternal	AUC=0,89; AUC hanya MRI=0.76
(Cha et al., 2018) (Korea Selatan)	89 pasien; 110 lesi; retrospektif	Radiomik CT (irisan aksial di pusat tumor); variabel klinis	CNN + pembelajaran ensemble	Respons awal SRS pada 3 bulan (RECIST 1.1)	Hold-out internal	AUC=0.761–0.856
(Erdur et al., 2025) (Jerman, Swiss., Siprus)	451 pasien; retrospektif multisenter	MRI T1-CE; FLAIR; segmentasi tumor/edema; variabel klinis	3D Vision Transformer (ViT) + fusi klinis + CoxLoss	Kebebasan dari kegagalan lokal (FFLF)	Multisenter eksternal	C-index=0.7982; terbaik dibanding semua baseline CNN
(Moon & Park, 2022) (Korea Selatan)	28 pasien; 44 lesi; retrospektif	Data klinis; dosis, BED10, fraksi; volume tumor	ANN: SLP dan MLP	Perubahan volume tumor setelah hfGKRS	Hold-out internal	Akurasi SLP=75%; MLP=80%
(Liu et al., 2024) (Tiongkok)	337 pasien; 873 lesi; multisenter (latih); 87 (uji)	MRI multi-modal (CE-T1, T2, DWI); masker	SRTRT-Net (hibrida CNN + Transformer, multi-task)	Respons terapi SRS (CR/PR/SD/PD)	5-fold CV + hold-out	AUC=0.945; akurasi=91.2%; sensitivitas=91.4%; spesifisitas=96.8%

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

		tumor; klinis (KPS, GPA, lokasi)				
(Jalalifar et al., 2022) (Kanada)	124 pasien; kohort retrospektif	MRI T1-CE; T2-FLAIR; kontur tumor/edema; atribut klinis	Varian InceptionResNet + LSTM / Transformer + fusi klinis	Kontrol lokal vs. kegagalan lokal setelah SRT	Hold-out internal	AUC terbaik=0.86 (InceptionResNet+LSTM+klinis)
(Buzea et al., 2023) (Rumania)	19 pasien; BM tunggal; retrospektif	MRI FLAIR dan CE-T1 (irisan aksial di pusat tumor)	CNN (dari awal); model TL; model FTM	Prediksi respons awal pasca-radiosurgeri	Internal	AUC=0.989; akurasi=98%
(Zhao et al., 2025a) (Amerika Serikat)	90 BM / 62 pasien; retrospektif	MRI T1+CE pasca-SRS; fitur genomik (driver NSCLC); klinis	HBNODE (Heavy Ball Neural ODE); DNN hanya-citra; model I+G+C	Radio nekrosis vs. rekurensi tumor	Hold-out internal	AUC=0.88–0.91 (model I+G+C)
(Zhao et al., 2025b) (Amerika Serikat)	90 BM / 62 pasien; retrospektif	MRI T1+CE pasca-SRS; klinis; fitur genomik	Radiogenomic Deep Ensemble Learning	Radio nekrosis vs. rekurensi tumor	Internal + eksternal	AUC=0.91; IBS=0.08
(Cho et al., 2024) (Korea Selatan)	194 pasien; 369 lesi; retrospektif	MRI sekuensial (1 pra-terapi + 3 titik waktu pasca-terapi)	2D Conv-GRU; 3D Conv-GRU; radiomik; model Dmax	Respons terapi BM setelah SRS (longitudinal)	Hold-out internal	AUC=0.89 (2D Conv-GRU terbaik; mengungguli CNN satu-titik-waktu)
(Keek et al., 2022) (Belanda)	1,404 pasien / 7,974 lesi (pengembangan); 237/646 (eksternal); multisenter	Fitur radiomik; fitur DL (Xception 3D); karakteristik pasien	Xception 3D + XGBoost (radiomik + DL + klinis)	Risiko efek radiasi merugikan (ARE) setelah SRT	Multisenter eksternal	AUC terbaik=0.77 (gabungan); DL saja AUC=0.64

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

PEMBAHASAN

MRI merupakan input pencitraan yang paling dominan, digunakan pada 11 dari 13 studi. Sekuens T1 tertimbang kontras muncul paling konsisten, sering dilengkapi dengan T2/FLAIR, *diffusion-weighted imaging*, dan secara unik peta distribusi dosis tiga dimensi (Zhao et al., 2024). Penyertaan peta dosis terbukti sangat informatif: dengan mengkodekan hubungan spasial antara dosis yang diberikan dan geometri tumor, modalitas ini menangkap informasi yang tidak dapat direpresentasikan oleh parameter dosis-volume skalar semata, sehingga berkontribusi pada AUC sebesar 0.89 untuk prediksi kontrol lokal dibandingkan 0.76 untuk baseline yang hanya menggunakan MRI. Implikasi praktisnya adalah bahwa data perencanaan terapi yang sudah dihasilkan secara rutin oleh institusi dapat dimanfaatkan kembali sebagai input prediktif yang bermakna dengan biaya tambahan yang minimal.

Secara *model framework*, studi-studi yang diikutsertakan mencerminkan trajektori perkembangan yang lebih luas di bidang DL. Penelitian awal mengandalkan CNN standar dan pengklasifikasi *ensemble* (Cha et al., 2018; Moon & Park, 2022), sementara investigasi yang lebih baru mengadopsi model berbasis transformer dan *framework* hibrida yang memadukan encoder konvolusional dengan komponen atensi atau rekuren. Penelitian lain melatih *3D Vision Transformer* pada MRI CE pra-operasi dari delapan pusat Eropa dan mencapai C-index sebesar 0.7982 pada kohort eksternal independen dengan lebih dari 400 pasien, hasil yang paling besar dan paling ketat tervalidasi dalam literatur ini (Erdur et al., 2025). Selain itu, penelitian lainnya menunjukkan keunggulan komplementer dari pembelajaran *multi-task,* mengoptimalkan segmentasi lesi dan klasifikasi respons secara bersama dalam satu kerangka CNN-Transformer menghasilkan AUC sebesar 0.945, dengan pembelajaran fitur bersama yang tampaknya meregularisasi kedua tugas secara simultan (Liu, Du, Dai, Yi, Liu, Wu, & Liu, 2024).

Pencitraan longitudinal terbukti menjadi faktor pembeda yang penting. Sebuah penelitian menerapkan jaringan 2D Conv-GRU pada empat akuisisi MRI sekuensial; satu pra-terapi dan tiga pasca-terapi, dan mencapai AUC = 0.89 untuk klasifikasi respons terapi, secara nyata mengungguli model yang hanya mengandalkan satu pemindaian *baseline* (Cho et al., 2024). Hasil ini mendukung intuisi bahwa trajektori tumor dari waktu ke waktu mengandung informasi biologis tentang sensitivitas terapi yang tidak dapat ditangkap oleh citra statis. Studi lain membuat pengamatan serupa, yaitu memadukan fitur mendalam yang diturunkan dari MRI longitudinal dengan variabel klinis terstruktur melalui lapisan LSTM meningkatkan AUC menjadi 0.86 untuk prediksi kontrol lokal, dibandingkan 0.73 untuk pendekatan hanya menggunakan pencitraan (Jalalifar et al., 2022).

Diferensiasi radio nekrosis versus rekurensi tumor merupakan tugas yang paling mendesak secara klinis yang dibahas dalam tinjauan ini. Dua literature berbeda memperkenalkan kerangka *Heavy Ball Neural ODE* (HBNODE), yang menyintesiskan fitur pencitraan, genomik, dan klinis dalam ruang laten yang dipelajari serta membuat trajektori keputusan setiap pasien dapat diinterpretasikan, langsung menjawab masalah opasitas yang membuat para klinisi enggan mempercayai prediksi kotak hitam (Zhao et al., 2024; Zhao et al., 2025a). Nilai AUC sebesar 0.88–0.91 pada kedua studi, dengan validasi eksternal pada kohort, merupakan langkah nyata menuju penghindaran biopsi secara non-invasif. Sudut pandang berbeda menunjukkan bahwa fitur DL saja dapat berkinerja buruk pada dataset kecil yang tidak seimbang model radiomik-DL-klinis gabungan mencapai AUC = 0.77 dibandingkan 0.64 untuk cabang DL saja yang memperkuat nilai fusi fitur ketika data terbatas (Keek et al., 2022).

Kinerja yang paling mencolok pada studi dengan model CNN-nya mencapai AUC = 0.989 dan akurasi 98% (Buzea et al., 2023). Hasil ini perlu disikapi dengan hati-hati: studi ini hanya mendaftarkan 19 pasien, sebuah kohort yang begitu kecil, sehingga lipatan validasi silang mungkin mengandung kurang dari dua contoh positif, yang membuat estimasi kinerja tidak dapat diandalkan secara statistik dan sangat rentan terhadap *overfitting.* Kasus ini menjadi pengingat bahwa angka AUC yang mencolok perlu dikontekstualisasikan dengan ukuran dataset, distribusi kelas, dan pendekatan validasi.

Beberapa kelemahan metodologis berulang di seluruh literatur ini. Desain retrospektif dari satu pusat mendominasi; pengumpulan data prospektif dan perekrutan multisenter masih langka. Ketidakseimbangan kelas yang mencerminkan insiden kegagalan lokal dan radio nekrosis yang relatif rendah pada kohort yang ditangani mempersulit

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

pelatihan dan dapat menaikkan metrik sensitivitas. Heterogenitas dalam protokol akuisisi MRI, alur pra-pemrosesan, dan definisi *outcomes* menciptakan risiko pergeseran domain ketika model yang dilatih pada data satu institusi diterapkan di tempat lain. Penilaian PROBAST mengkonfirmasi risiko bias tinggi atau tidak jelas pada domain analisis untuk sebagian besar studi, terutama akibat sampel kecil, ketiadaan pelaporan kalibrasi, dan langkanya validasi eksternal prospektif. Dua seri berdiri sebagai tolok ukur metodologis, dengan menyertakan kohort eksternal dan ukuran kalibrasi sebagai standar (Zhao et al., 2024; Zhao et al., 2025a; Erdur et al., 2025).

Penelitian mendatang, jalan menuju implementasi klinis mensyaratkan studi prospektif multisenter dengan protokol akuisisi MRI yang terharmonisasi, kerangka interpretabilitas yang terintegrasi Grad-CAM, SHAP, dan metode berbasis trajektori seperti HBNODE pelaporan kalibrasi yang sistematis, serta pada akhirnya studi luaran prospektif yang mengukur dampak keputusan berbasis DL terhadap titik akhir di tingkat pasien, bukan sekadar AUC model.

SIMPULAN

Tinjauan ini menyintesis bukti dari 13 studi yang menerapkan DL untuk prediksi *outcomes* pada pasien metastasis otak yang ditangani dengan SRS atau SRT. Di berbagai titik akhir kontrol lokal, respons terapi, kelangsungan hidup keseluruhan, diferensiasi radio nekrosis, dan perubahan volume tumor, model DL mencapai kinerja prediktif dari sedang hingga sangat baik, dengan nilai AUC berkisar antara 0.71 hingga 0.99. *Framework* multimodal yang menggabungkan fitur MRI dengan data klinis, dosimetrik, atau genomik secara konsisten mengungguli pendekatan unimodal, dan model longitudinal yang memanfaatkan pencitraan sekuensial memberikan keunggulan atas analisis satu titik waktu. *Framework* berbasis *transformer*, kerangka *ensemble* mendalam, dan model neural ODE mewakili batas kinerja terkini, dengan yang terakhir menawarkan tingkat interpretabilitas yang tidak dimiliki oleh pendekatan konvolusional murni.

Pada saat yang sama, basis bukti ini terkendala oleh karakternya yang hampir sepenuhnya retrospektif, ukuran kohort yang kecil pada beberapa studi, dan hampir tidak adanya validasi eksternal prospektif. Penelitian mendatang perlu memprioritaskan desain multisenter prospektif, definisi pencitraan dan *outcomes* yang terstandarisasi, pelaporan kalibrasi yang ketat, serta kerangka implementasi yang sesuai regulasi. Apabila kondisi tersebut terpenuhi, model prediktif berbasis DL berpotensi mendukung secara bermakna perencanaan terapi yang dipersonalisasi, stratifikasi risiko, dan surveilans pasca-terapi pada pasien metastasis otak yang menjalani radiasi stereotaktik.

DAFTAR PUSTAKA

Baschnagel, A. M., Meyer, K. D., Chen, P. Y., Krauss, D. J., Olson, R. E., Pieper, D. R., & Grills, I. S. (2013). Tumor volume as a predictor of survival and local control in patients with brain metastases treated with Gamma Knife surgery. *Journal of neurosurgery*, *119*(5), 1139-1144.

Buzea, C. G., Buga, R., Paun, M. A., Albu, M., Iancu, D. T., Dobrovat, B., & Eva, L. (2023). AI evaluation of imaging factors in the evolution of stage-treated metastases using Gamma Knife. *Diagnostics*, *13*(17), 2853.

Cha, Y. J., Jang, W. I., Kim, M. S., Yoo, H. J., Paik, E. K., Jeong, H. K., & Youn, S. M. (2018). Prediction of response to stereotactic radiosurgery for brain metastases using convolutional neural networks. *Anticancer research*, *38*(9), 5437-5445.

Chang, R., Qi, S., Wu, Y., Yue, Y., Zhang, X., & Qian, W. (2023). Nomograms integrating CT radiomic and deep learning signatures to predict overall survival and progression-free survival in NSCLC patients treated with chemotherapy. *Cancer Imaging*, *23*(1), 101.

Cho, S. J., Cho, W., Choi, D., Sim, G., Jeong, S. Y., Baik, S. H., & Sunwoo, L. (2024). Prediction of treatment response after stereotactic radiosurgery of brain metastasis using deep learning and radiomics on longitudinal MRI data. *Scientific Reports*, *14*(1), 11085.

de Azevedo Santos, T. R., Tundisi, C. F., Ramos, H., Maia, M. A. C., Pellizzon, A. C. A., Silva, M. L. G., & de Castro, D. G. (2015). Local control after radiosurgery for brain metastases: predictive factors and implications for clinical decision. *Radiation Oncology*, *10*(1), 63.

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

Douri, K., Iorio-Morin, C., Mercure-Cyr, R., Figueiredo, G., Touchette, C. J., Masson-Côté, L., & Mathieu, D. (2023). Response assessment in brain metastases managed by stereotactic radiosurgery: A reappraisal of the RANO-BM criteria. *Current Oncology*, *30*(11), 9382-9391.

Erdur, A. C., Scholz, D., Nguyen, Q. M., Buchner, J. A., Mayinger, M., Christ, S. M., & Peeken, J. C. (2025). Improving risk assessment of local failure in brain metastases patients using vision transformers− A multicentric development and validation study. *Radiotherapy and oncology*, *210*, 111031.

Ibragimov, B., Toesca, D. A., Chang, D. T., Yuan, Y., Koong, A. C., Xing, L., & Vogelius, I. R. (2020). Deep learning for identification of critical regions associated with toxicities after liver stereotactic body radiation therapy. *Medical physics*, *47*(8), 3721-3731.

Jalalifar, S. A., Soliman, H., Sahgal, A., & Sadeghi-Naini, A. (2022). Predicting the outcome of radiotherapy in brain metastasis by integrating the clinical and MRI-based deep learning features. *Medical Physics*, *49*(11), 7167-7178.

Kanakarajan, H., De Baene, W., Hanssens, P., & Sitskoorn, M. (2024). Predicting local control of brain metastases after stereotactic radiotherapy with clinical, radiomics and deep learning features. *Radiation Oncology*, *19*(1), 182.

Karami, E., Ruschin, M., Soliman, H., Sahgal, A., Stanisz, G. J., & Sadeghi-Naini, A. (2019). An MR radiomics framework for predicting the outcome of stereotactic radiation therapy in brain metastasis. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1022-1025). IEEE.

Kawahara, D., Tang, X., Lee, C. K., Nagata, Y., & Watanabe, Y. (2021). Predicting the local response of metastatic brain tumor to gamma knife radiosurgery by radiomics with a machine learning method. *Frontiers in oncology*, *10*, 569461.

Keek, S. A., Beuque, M., Primakov, S., Woodruff, H. C., Chatterjee, A., van Timmeren, J. E., & Lambin, P. (2022). Predicting adverse radiation effects in brain tumors after stereotactic radiotherapy with deep learning and handcrafted radiomics. *Frontiers in oncology*, *12*, 920393.

Ko, P. H., Kim, H. J., Lee, J. S., & Kim, W. C. (2020). Tumor volume and sphericity as predictors of local control after stereotactic radiosurgery for limited number (1-4) brain metastases from nonsmall cell lung cancer. *Asia-Pacific Journal of Clinical Oncology*, *16*(3), 165-171.

Liao, C. Y., Lee, C. C., Yang, H. C., Chen, C. J., Chung, W. Y., Wu, H. M., & Lu, C. F. (2023). Predicting survival after radiosurgery in patients with lung cancer brain metastases using deep learning of radiomics and EGFR status. Physical and Engineering Sciences in Medicine, 46(2), 585-596.

Liu, X., Du, P., Dai, Z., Yi, R., Liu, W., Wu, H., & Liu, J. (2024). SRTRP-Net: A multi-task learning network for segmentation and prediction of stereotactic radiosurgery treatment response in brain metastases. *Computers in Biology and Medicine*, *175*, 108503.

Moon, H. C., & Park, Y. S. (2022). Volume prediction for large brain metastases after hypofractionated gamma knife radiosurgery through artificial neural network. *Medicine*, *101*(40), e30964.

Ruzevick, J., Kleinberg, L., & Rigamonti, D. (2014). Imaging changes following stereotactic radiosurgery for metastatic intracranial tumors: differentiating pseudoprogression from tumor progression and its effect on clinical practice. *Neurosurgical review*, *37*(2), 193-201.

Suh, J. H., Kotecha, R., Chao, S. T., Ahluwalia, M. S., Sahgal, A., & Chang, E. L. (2020). Current approaches to the management of brain metastases. *Nature reviews Clinical oncology*, *17*(5), 279-299.

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586

Touati, R., Bourbonne, V., Dissaux, G., Goasduff, G., Pradier, O., Peltier, C., & Lucia, F. (2023). Re-irradiation by stereotactic radiotherapy of brain metastases in the case of local recurrence. *Cancers*, *15*(3), 996.

Zhao, J., Vaios, E., Calabrese, E., Yang, Z., Robertson, S., Ginn, J., & Wang, C. (2025b). A radiogenomic deep ensemble learning model for identifying radionecrosis following brain metastases (BM) stereotactic radiosurgery in patients with non-small cell lung cancer BM. *Advances in Radiation Oncology*, *10*(8), 101826.

Zhao, J., Vaios, E., Wang, Y., Yang, Z., Cui, Y., Reitman, Z. J., & Wang, C. (2024). Dose-incorporated deep ensemble learning for improving brain metastasis stereotactic radiosurgery outcome prediction. *International Journal of Radiation Oncology* Biology* Physics*, *120*(2), 603-613.

Zhao, J., Vaios, E., Yang, Z., Lu, K., Floyd, S., Yang, D., & Wang, C. (2025a). Radiogenomic explainable AI with neural ordinary differential equation for identifying post-SRS brain metastasis radionecrosis. *Medical xphysics*, *52*(4), 2661-2674.

Mirza Wafiyudin Baehaqi*, Shahifa Audy Rahima, Dita Rahmania, Mauliya Sri Sukmawati Wahyudi

Fakultas Kedokteran, Universitas Jember
Korespondensi penulis: Mirza Wafiyudin Baehaqi. *Email: mirzawafiyudin25@gmail.com

DOI: https://doi.org/10.33024/hjk.v20i5.3586