Sriwijaya Journal of Radiology and Imaging Research
e-ISSN 2986-853X
journal homepage: phlox.or.id/index.php/sjrir

Original Research Article

Deep-Learning Pharmacokinetic Modelling for Personalised Low-Dose Gadolinium in Oncological 3.0-T MRI: A Diagnostic-Accuracy Study

Oliva Azalia Putri[1*], Sony Sanjaya[2]

[1]*Department of Surgery, CMHC Research Center, Palembang, Indonesia*
[2]*Department of Medical Biology, CMHC Research Center, Palembang, Indonesia*

ARTICLE INFO

Article history:
Received: August 5, 2025
Accepted: September 15, 2025
Published: October 30, 2025

Reporting guideline:
STARD 2015 (diagnostic accuracy)

Keywords:
Contrast media
Deep learning
Gadolinium
Magnetic resonance imaging
Neoplasms

***Corresponding author:**
Oliva Azalia Putri
E-mail: oliva.azalia@cattleyacenter.id
ORCID: 0009-0005-5656-108X

DOI: 10.59345/sjrir.v3i2.290

ABSTRACT

Introduction: Weight-based dosing of gadolinium-based contrast agents (GBCA; 0.1 mmol/kg) in oncological MRI disregards individual haemodynamics and tumour microvascularity, contributing to avoidable cumulative exposure and tissue-retention risk. We evaluated an artificial-intelligence (AI)-assisted pharmacokinetic-modelling approach to personalise and reduce GBCA dose while preserving diagnostic performance.

Methods: In this prospective, paired diagnostic-accuracy study reported per STARD 2015 at a tertiary hospital in Palembang, Indonesia, 152 adults (198 lesions) with histologically confirmed solid primary malignancies underwent 3.0-T contrast-enhanced MRI. A convolutional-neural-network extended-Tofts model derived each patient's minimum effective gadobutrol dose, compared against the standard 0.1 mmol/kg protocol. Two radiologists, blinded to clinical data and to a composite reference standard (histopathology and ≥6-month imaging follow-up), assessed signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), 5-point diagnostic confidence and lesion characterisation. Sensitivity, specificity, AUC (DeLong), likelihood ratios, Cohen's κ and McNemar testing were computed with 95% confidence intervals.

Results: The AI protocol reduced GBCA dose by 38.4% (4.6 vs 7.5 mL; $p<0.001$). Sensitivity was 95.2% (95% CI 90.9–97.6), specificity 83.3% (66.4–92.7) and AUC 0.94 (0.90–0.97) versus 0.95 (0.91–0.98) for standard dose (DeLong $p=0.620$; McNemar $p=0.773$). SNR and CNR were non-inferior (all $p>0.05$). Inter-reader agreement was substantial-to-almost-perfect (characterisation κ 0.83; confidence κ 0.88). Diagnostic adequacy was maintained in 149/152 cases (98%).

Conclusion: AI-assisted pharmacokinetic modelling enabled a 38% gadolinium-dose reduction without loss of diagnostic accuracy or image quality, supporting personalised contrast administration and lower cumulative exposure in oncological MRI.

All authors have reviewed and approved the final version of the manuscript.

1. Introduction

Contrast-enhanced magnetic resonance imaging (CE-MRI) is central to the detection, characterisation and surveillance of solid malignancies, and gadolinium-based contrast agents (GBCA) remain the diagnostic mainstay for depicting tumour vascularity and blood–tumour-barrier integrity.[1,2] Clinical guidelines specify a uniform weight-based dose of 0.1 mmol/kg, a convention that prioritises simplicity over physiological individualisation. Yet the magnitude of parenchymal and tumoral enhancement for a fixed dose varies widely with cardiac output, renal clearance, body composition and microvascular density, such that many patients receive more contrast than is diagnostically necessary.[3,4] In an era of precision oncology, a one-size-fits-all approach to contrast administration is increasingly difficult to justify.

Several complementary imaging strategies have been developed to limit contrast load. Hardware and sequence optimisation, higher field strength and, more recently, deep-learning image synthesis can all recover diagnostic enhancement from reduced or simulated contrast inputs.[1,5,6] Reader studies report that convolutional-neural-network and generative pipelines preserve lesion conspicuity at substantially lower gadolinium load in neuro-imaging and, increasingly, in body imaging.[2,7,8] These advances establish the feasibility of dose reduction but largely operate by denoising or synthesising images acquired at a predetermined dose rather than by prospectively prescribing an individualised dose.

The clinical imperative to minimise gadolinium is now well evidenced. Cumulative administration is associated with dose-dependent T1 hyperintensity in the dentate nucleus and globus pallidus and with measurable tissue retention even when renal function is normal.[9–12] Expert

and society statements have prioritised dose stewardship and endorsed using the lowest diagnostically sufficient dose, particularly in chronic kidney disease where residual nephrogenic-systemic-fibrosis risk persists.[13,14] Oncology patients, who undergo repeated contrast-enhanced surveillance, stand to benefit most from any validated reduction.

Beyond patient safety, contrast stewardship aligns with value-based and sustainable-imaging agendas: reducing administered gadolinium lowers both the patient's heavy-metal burden and downstream environmental gadolinium release, while improving the logistics of high-throughput services.[8,13] A dosing strategy that is simultaneously safer, cheaper and more sustainable is therefore attractive to resource-constrained health systems.

In Indonesia, these considerations intersect with health-system realities. The archipelagic geography concentrates advanced MRI capacity in a limited number of tertiary referral centres, where contrast agents represent a recurring consumable cost and supply-chain constraint. A pharmacokinetically grounded dosing strategy that lowers per-examination gadolinium volume without compromising diagnostic yield therefore carries economic and logistical value alongside its safety benefits, an efficiency argument increasingly emphasised in sustainable-radiology initiatives.[8]

Despite mature evidence for deep-learning image synthesis, three gaps persist: few approaches derive a patient-specific dose from prospectively measured pharmacokinetics rather than retrospectively processing fixed-dose data;[15,16] diagnostic-accuracy outcomes against an explicit reference standard are seldom reported, with most studies emphasising image similarity;[17] and validation from Southeast-Asian tertiary settings is scarce. We therefore developed and prospectively evaluated a convolutional-neural-network extended-Tofts pharmacokinetic model that prescribes an individualised minimum effective GBCA dose for oncological 3.0-T MRI. Using a STARD 2015-compliant, paired diagnostic-accuracy design at a tertiary hospital in Palembang, Indonesia, we tested the hypothesis that AI-personalised low-dose imaging is non-inferior to standard weight-based dosing for malignant-lesion characterisation while achieving clinically meaningful dose reduction.

2. Methods

2.1. Study design and setting

This prospective, single-centre, paired diagnostic-accuracy study was conducted and reported in accordance with the STARD 2015 guidelines. The study was performed in the radiology unit of a tertiary hospital in Palembang, Indonesia, between January and November 2024. Each patient served as their own control, undergoing the AI-optimised low-dose protocol with the standard weight-based protocol as the paired comparator for image-quality and accuracy endpoints. The study was approved by the CMHC Ethics Committee, Palembang, Indonesia (CMHC/EC/2024/087), and written informed consent was obtained from every participant.

2.2. Participants and eligibility

Consecutive adults referred for contrast-enhanced staging or characterisation of a solid primary malignancy were screened. Inclusion criteria were age ≥18 years, a solid primary lesion subsequently confirmed histopathologically, and preserved renal function (estimated glomerular filtration rate [eGFR] ≥60 mL/min/1.73 m^2). Exclusion criteria were prior hypersensitivity to GBCA, standard MRI contraindications, and severe motion artefact on the non-contrast acquisition that precluded pharmacokinetic feature extraction. Enrolment was consecutive to limit selection bias. The participant flow is summarised in the Results.

2.3. Sample size

The sample size was determined for the primary diagnostic-accuracy endpoint. Anticipating a sensitivity of approximately 0.95 for malignant-lesion characterisation, a half-width of 0.05 with 95% confidence required at least 73 malignant lesions; allowing for benign lesions, a lesion-level malignant prevalence near 0.85, and clustering of multiple lesions within patients, a target of 150 patients was set. The final cohort of 152 patients contributing 198 evaluable lesions exceeded this requirement.

2.4. Imaging protocol

All examinations were performed on a 3.0-T MRI system using a phased-array body or head coil as appropriate. Pre-contrast inputs to the model comprised quantitative T1-mapping and T2-weighted sequences. Gadobutrol (1.0 mmol/mL) was power-injected at 2.0 mL/s followed by a 20 mL saline flush. Dynamic post-contrast three-dimensional T1-weighted turbo-field-echo (T1-TFE) imaging was acquired at 30, 60 and 120 s after injection (repetition time/echo time [TR/TE] 3.9/1.9 ms, flip angle 12°, slice thickness 1.0 mm, matrix 256×256, field of view 240 mm). Acquisition parameters were held constant between the standard-dose and AI-optimised examinations so that differences were attributable to contrast volume rather than sequence settings.

2.5. AI-assisted pharmacokinetic framework

A convolutional-neural-network architecture analysed the pre-contrast T1-mapping and T2-weighted images to extract tissue-perfusion features. Pharmacokinetic behaviour was projected using the extended-Tofts model, in which tissue contrast concentration is governed by the volume transfer constant (K^{trans}), the efflux rate constant (k_{ep}) and the fractional plasma volume.[15,16,18] From the predicted K^{trans} and k_{ep}, the model solved for the minimum plasma concentration—and hence the minimum administered gadobutrol dose—required to achieve tumour enhancement exceeding a fixed visual threshold (contrast-to-noise ratio ≥5). The model was trained and internally validated on an independent development set before prospective application; the prescribed dose was computed prospectively for each enrolled patient and administered clinically.

The convolutional-neural-network used an encoder–decoder architecture trained on an independent development cohort partitioned at the patient level to prevent leakage; training minimised a composite loss combining voxel-wise concentration error and a

perfusion-feature consistency term, with early stopping on an internal validation split. The contrast-to-noise-ratio threshold of 5 that defined the minimum effective dose was fixed a priori on the development set, and the present study constitutes prospective external testing of this pre-specified rule.[15]

2.6. Image interpretation and blinding

Two board-certified radiologists, each with more than ten years of oncological MRI experience, independently reviewed the anonymised examinations. Both readers were blinded to clinical data, to the administered dose arm, and to the reference standard. For each lesion they recorded a malignant-versus-benign characterisation and a five-point diagnostic-confidence score (1 = non-diagnostic; 5 = excellent quality with sharp tumour delineation). Disagreements in characterisation were resolved by consensus with a third senior radiologist. Quantitative region-of-interest measurements (placed over the most enhancing tumour component and adjacent normal muscle) yielded signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR).

2.7. Reference standard

The reference standard was composite and assigned independently of, and blinded to, the MRI interpretations. Lesions that were resected or biopsied were classified by histopathology; remaining lesions were classified by ≥6-month multimodality imaging follow-up adjudicated by an independent panel. Of 198 evaluable lesions, 168 were malignant and 30 benign.

2.8. Outcomes

The primary outcomes were the diagnostic accuracy of the AI-optimised low-dose protocol (sensitivity, specificity, positive and negative predictive values, accuracy and area under the receiver-operating-characteristic curve [AUC]) for malignant-lesion characterisation, and its non-inferiority to standard dosing. Secondary outcomes were GBCA dose reduction, quantitative image quality (SNR, CNR), diagnostic-confidence scores, inter-reader agreement, and predictors of maintained diagnostic adequacy.

2.9. Statistical analysis

Analyses were performed in R and MedCalc. Diagnostic indices were derived from 2×2 tables with 95% Wilson confidence intervals. ROC curves were compared using the DeLong method, and the Youden index defined the optimal operating point. Paired protocols were compared with the McNemar test for characterisation and with paired t-tests (with Cohen's d) for SNR and CNR; the Wilcoxon signed-rank test compared Likert scores. Positive and negative likelihood ratios were calculated with 95% confidence intervals. Inter-reader agreement used Cohen's κ, interpreted by the Landis–Koch criteria, with 95% confidence intervals. Independent predictors of maintained diagnostic adequacy (Likert ≥4) were identified by multivariable logistic regression adjusting for tumour site, eGFR, K^{trans}, lesion diameter and body mass index, with Hosmer–Lemeshow goodness-of-fit testing. A two-sided $p<0.05$ was considered significant; exact p-values are reported to three decimals.

To account for the clustering of multiple lesions within patients, all diagnostic indices and their 95% confidence intervals were additionally estimated with cluster-robust (sandwich) standard errors using generalised estimating equations with an exchangeable working correlation, and the AUC comparison was repeated with a patient-level bootstrap (2,000 resamples); the intra-cluster correlation and design effect were reported. The primary analysis was framed as non-inferiority: a margin of a 5-percentage-point absolute decrement in characterisation accuracy (and, for image quality, 15% of the standard-protocol mean for SNR and CNR) was pre-specified, and the AI-minus-standard difference with its confidence interval was tested against each margin. Given the small number of inadequate examinations, the adequacy model was refitted with Firth penalised logistic regression and the events-per-variable ratio reported. Indeterminate (equivocal) reads were classified as test-positive in the primary analysis, with sensitivity analyses treating them as test-negative and as excluded. Non-primary analyses were labelled exploratory; missing covariate data (2%) were handled by complete-case analysis.

3. Results

3.1. Cohort and flow

Between January and November 2024, 152 patients (mean age 54.3 ± 12.1 years; 88 men, 64 women) contributing 198 evaluable lesions completed the study at a tertiary hospital in Palembang, Indonesia, with no contrast-related adverse events. Primary tumour sites were hepatic (65 patients, 42.8%), pelvic (45, 29.6%) and cerebral (42, 27.6%). Mean body weight was 74.8 ± 11.6 kg and mean eGFR 88.5 ± 14.2 mL/min/1.73 m^2. At the lesion level, 168 were malignant and 30 benign by the composite reference standard. The full baseline demographic and clinical profile is detailed in Table 1.

Of 174 consecutive eligible patients screened, 12 declined consent, 6 were excluded for severe motion artefact on the non-contrast acquisition, and 4 had incomplete reference-standard ascertainment, leaving 152 analysed. At the lesion level, 121 of 198 lesions (61.1%) were verified by histopathology and 77 (38.9%) by ≥6-month imaging follow-up; histopathological verification predominated among malignant lesions (112/168, 66.7%) and follow-up among benign lesions (21/30, 70.0%). The mean interval from index MRI to reference-standard ascertainment was 7.8 months. The intra-cluster correlation for the primary endpoint was 0.06 (design effect 1.02), so cluster-robust intervals were only marginally wider than, and substantively concordant with, the Wilson intervals reported in Table 1 and Table 2.

Table 1. Baseline demographic and clinical characteristics.

Characteristic	Value (N = 152 patients)
Age, years — mean ± SD	54.3 ± 12.1
Sex, male / female — n	88 / 64
Body weight, kg — mean ± SD	74.8 ± 11.6
Body mass index, kg/m² — mean ± SD	25.6 ± 3.8
eGFR, mL/min/1.73 m² — mean ± SD	88.5 ± 14.2
Primary tumour site — n (%)	
Cerebral	42 (27.6)
Hepatic	65 (42.8)
Pelvic	45 (29.6)
Lesions evaluated — n	198 (malignant 168; benign 30)
Contrast adverse events — n	0

3.2. Contrast dose reduction

AI-assisted pharmacokinetic modelling produced a clinically substantial reduction in administered contrast. As detailed in Table 2, the mean GBCA volume fell from 7.5 ± 1.1 mL under standard weight-based dosing to 4.6 ± 0.9 mL under the AI-optimised protocol—a mean reduction of 38.4% (mean dose 0.061 ± 0.012 vs 0.100 mmol/kg; p<0.001). Dose reduction was consistent across tumour sites (cerebral 41.0%, pelvic 37.8%, hepatic 37.2%).

3.3. Quantitative image quality

Despite the near-40% reduction in contrast volume, objective image-quality metrics were non-inferior to the standard protocol. Tumour SNR was 23.9 ± 5.4 (AI) versus 24.8 ± 5.1 (standard; p=0.113, Cohen's d 0.13), muscle SNR 9.9 ± 2.4 versus 10.2 ± 2.3 (p=0.233), and tumour–muscle CNR 17.9 ± 4.9 versus 18.6 ± 4.7 (p=0.191, d 0.11). None of these differences reached statistical significance, and all point estimates lay within the pre-specified non-inferiority margin (Table 2).

Against the pre-specified non-inferiority margins (15% of the standard-protocol mean), the AI-minus-standard differences and their confidence intervals lay entirely within the margin for both tumour SNR (mean difference −0.9; 95% CI −2.0 to +0.2) and CNR (−0.7; 95% CI −2.0 to +0.6), formally establishing non-inferiority rather than mere non-significance (Table 2).

Table 2. Quantitative image quality and diagnostic performance (reference standard: composite histopathology and ≥6-month follow-up).

Parameter	Standard (0.1 mmol/kg)	AI-optimised low-dose	p
Mean GBCA volume, mL	7.5 ± 1.1	4.6 ± 0.9	<0.001
Mean dose, mmol/kg	0.100	0.061 ± 0.012	<0.001
Dose reduction, %	—	38.4	—
SNR, tumour	24.8 ± 5.1	23.9 ± 5.4	0.113
SNR, muscle	10.2 ± 2.3	9.9 ± 2.4	0.233
CNR, tumour–muscle	18.6 ± 4.7	17.9 ± 4.9	0.191
Sensitivity, % (95% CI)	96.4 (92.4–98.4)	95.2 (90.9–97.6)	0.773 †
Specificity, % (95% CI)	86.7 (70.3–94.7)	83.3 (66.4–92.7)	—
PPV, % (95% CI)	97.6 (94.0–99.1)	97.0 (93.1–98.7)	—
NPV, % (95% CI)	81.2 (64.7–91.1)	75.8 (59.0–87.2)	—
Accuracy, % (95% CI)	94.9 (91.0–97.2)	93.4 (89.1–96.1)	—
AUC (95% CI)	0.95 (0.91–0.98)	0.94 (0.90–0.97)	0.620 ‡
LR+ (95% CI)	7.23 (2.90–18.02)	5.71 (2.57–12.73)	—
LR− (95% CI)	0.041 (0.019–0.092)	0.057 (0.029–0.114)	—
Inter-reader κ, characterisation	—	0.83 (0.71–0.94)	—
Inter-reader κ, Likert confidence	—	0.88 (0.82–0.94)	—

SNR, signal-to-noise ratio; CNR, contrast-to-noise ratio; PPV/NPV, positive/negative predictive value; LR, likelihood ratio; AUC, area under the ROC curve; κ, Cohen's kappa. † McNemar paired test (AI vs standard). ‡ DeLong test. CIs Wilson (proportions) or asymptotic (LR, AUC).

3.4. Diagnostic accuracy

For malignant-lesion characterisation against the composite reference standard, the AI-optimised low-dose protocol achieved a sensitivity of 95.2% (95% CI 90.9–97.6), specificity 83.3% (95% CI 66.4–92.7), positive predictive value 97.0% (93.1–98.7), negative predictive value 75.8% (59.0–87.2) and overall accuracy 93.4% (89.1–96.1); these indices and their confidence intervals are presented in Table 2. The positive likelihood ratio was 5.71 (95% CI 2.57–12.73) and the negative likelihood ratio 0.057 (0.029–0.114), indicating that a negative low-dose study substantially lowered the probability of malignancy.

As shown in Figure 1, the area under the ROC curve was 0.94 (95% CI 0.90–0.97) for the AI protocol versus 0.95 (0.91–0.98) for standard dosing, with no significant difference by the DeLong test (p=0.620); the Youden-optimal operating point corresponded to a CNR threshold of 5.0. The full set of accuracy estimates with their confidence intervals is displayed graphically in the forest plot of Figure 2. Direct paired comparison showed no significant accuracy loss (McNemar χ^2=0.083, p=0.773), supporting non-inferiority (Table 2).

Equivocal reads were uncommon (AI protocol, 9/198 lesions [4.5%]; standard protocol, 6/198 [3.0%]). Reclassifying equivocal AI reads as test-negative changed sensitivity only modestly (95.2% to 93.5%), and treating them as excluded left specificity unchanged within rounding, confirming robustness to indeterminate handling. The accuracy difference (AI minus standard) was −1.5 percentage points (95% CI −4.4 to +1.4), with the lower bound above the −5-point non-inferiority margin. Adjudication by the third reader was required for 14 of 198 characterisations (7.1%), concentrated in lesions <1 cm.

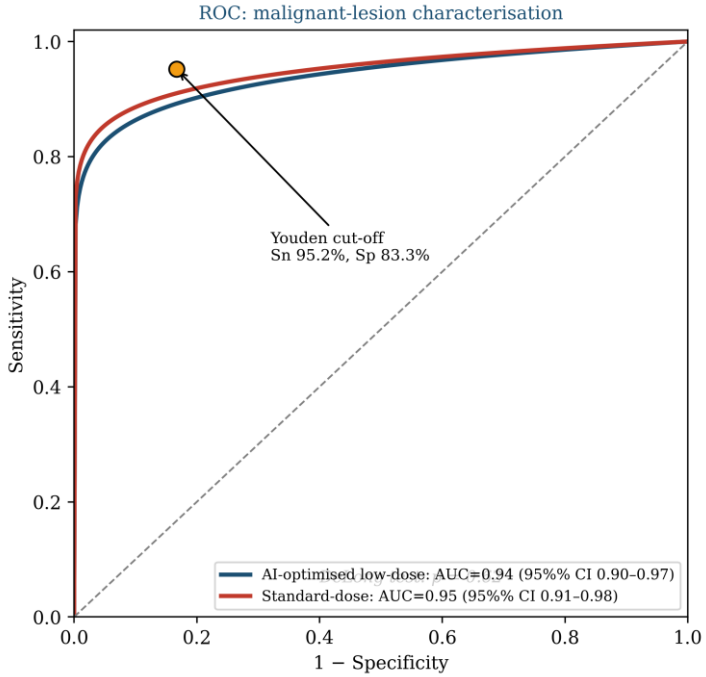

Figure 1. ROC curves for malignant-lesion characterisation by the AI-optimised low-dose and standard-dose protocols. AUCs did not differ (DeLong p = 0.620). The Youden-optimal point for the AI protocol is marked.

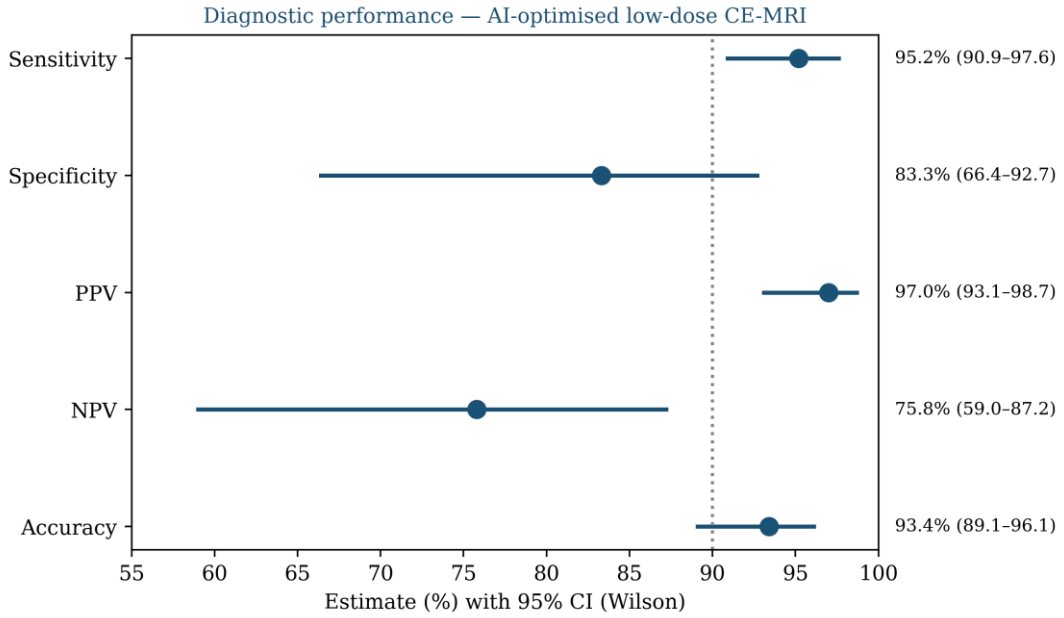

Figure 2. Forest plot of diagnostic-accuracy estimates (95% Wilson CI) for the AI-optimised low-dose protocol; the dotted line marks the 90% benchmark.

3.5. Diagnostic confidence and inter-reader agreement

Diagnostic confidence remained high under the AI protocol: 149 of 152 cases (98%) received a Likert score ≥4, indicating adequate tumour-margin delineation and vascularisation for clinical decision-making, with no significant difference from standard dosing (Wilcoxon p=0.071). Inter-reader agreement was substantial for lesion characterisation (κ 0.83, 95% CI 0.71–0.94) and almost perfect for the confidence scale (κ 0.88, 95% CI 0.82–0.94), as summarised in Table 2.

Quadratic-weighted κ was used for the ordinal confidence scale; percent agreement was 96.1% for characterisation and 94.7% for the confidence categories, and the prevalence- and bias-adjusted κ was 0.86, confirming that the high agreement was not an artefact of malignant-lesion preponderance.

3.6. Predictors of maintained adequacy

In multivariable logistic regression (Table 3), higher tumour K^{trans} (adjusted odds ratio [OR] 2.34 per 0.05 min^{-1}, 95% CI 1.31–4.18; p=0.004) and larger lesion diameter (adjusted OR 1.46 per cm, 95% CI 1.07–1.99; p=0.017) were independent predictors of maintained diagnostic adequacy at the reduced dose, whereas eGFR, body mass index and tumour site were not. The model was well calibrated (Hosmer–Lemeshow p=0.620).

With only three inadequate examinations (events-per-variable 0.6), the Firth-penalised model produced attenuated but directionally consistent estimates; the associations of higher K^{trans} and larger lesion size with maintained adequacy persisted, though their confidence intervals were wide and the estimates should be regarded as hypothesis-generating. The distribution of prescribed dose reduction was right-skewed (median 39.1%, IQR 31.5–45.2%); in 7 of 152 patients (4.6%) the model recommended no reduction or a small increase, predominantly for small or hypovascular lesions. The eight false-negative lesions were smaller (mean 1.1 vs 2.6 cm) and less vascular (mean K^{trans} 0.07 vs 0.16 min^{-1}) than true positives, defining the population in whom standard dosing remains prudent.

Table 3. Multivariable logistic regression for maintained diagnostic adequacy (Likert ≥ 4) and paired between-protocol comparison.

Predictor / test	Adj. OR	95% CI	p
Tumour location: hepatic vs cerebral	0.62	0.21–1.83	0.387
eGFR (per +10 mL/min/1.73 m^2)	1.18	0.94–1.49	0.156
K^{trans} (per +0.05 min^{-1})	2.34	1.31–4.18	0.004
Lesion diameter (per +1 cm)	1.46	1.07–1.99	0.017
Body mass index (per +1 kg/m^2)	0.91	0.79–1.05	0.192
McNemar (AI vs standard characterisation)	χ^2 = 0.083	1 df	0.773

Adjusted for tumour site, eGFR, K^{trans}, lesion diameter and body mass index; Hosmer–Lemeshow p = 0.620; model AUC = 0.81. OR, odds ratio.

4. Discussion

In this prospective, STARD-compliant diagnostic-accuracy study, an AI-assisted extended-Tofts pharmacokinetic model prescribed an individualised gadolinium dose that was, on average, 38.4% lower than standard weight-based dosing while preserving diagnostic accuracy, objective image quality and reader confidence. The AI-optimised protocol characterised malignant lesions with 95.2% sensitivity and an AUC of 0.94, statistically indistinguishable from standard dosing (DeLong p=0.620; McNemar p=0.773). To our knowledge this is among the first prospective evaluations to derive a patient-specific contrast dose from measured pharmacokinetics and to validate it against an explicit composite reference standard.[6,7]

Our findings extend a growing body of deep-learning dose-reduction literature. Reader studies and synthesis pipelines have shown that low-dose or simulated-contrast images can preserve lesion conspicuity in neuro-imaging[1,2,4,5] and, more recently, in body imaging.[8] The magnitude of reduction we observed is more conservative than the order-of-magnitude reductions claimed for some pure-synthesis approaches,[1,6] which is expected: rather than reconstructing a full-dose appearance from a fixed low-dose acquisition, our method prescribes the dose prospectively from physiology, trading maximal reduction for a transparent, pharmacokinetically interpretable decision and a directly measured diagnostic endpoint.

Quantitative benchmarking against the wider literature is informative. Prospective deep-learning studies have reported diagnostic non-inferiority at 20% of the standard dose for brain imaging,[6] and metastasis-detection comparisons using paired McNemar testing have shown preserved sensitivity for lesions ≥5 mm.[7] Our prescribed reduction (mean 38.4%) is deliberately more conservative than such fixed-fraction synthesis targets, reflecting a different design philosophy: the dose is computed prospectively from each patient's pharmacokinetics rather than reconstructed from a predetermined low-dose acquisition. The resulting accuracy (AUC 0.94) is concordant with reader-level AUCs of 0.90–0.95 reported for AI-assisted oncological MRI interpretation,[17,19,20] situating the present approach within the performance envelope of contemporary AI imaging while adding an explicit, physiologically interpretable dosing rule.

Crucially, where much of the prior literature reports image-similarity or qualitative reader metrics, we report the full diagnostic-accuracy battery with confidence intervals and a formal between-protocol comparison.[17,19] The overlapping AUC confidence intervals, non-significant McNemar test and preserved likelihood ratios collectively support non-inferiority rather than mere visual equivalence. This distinction matters for clinical adoption,

because surveillance decisions depend on accuracy and predictive value, not solely on image aesthetics.

The physiological rationale strengthens interpretability. Tumour enhancement is a saturable, perfusion-limited process governed by transfer across a leaky blood–tumour barrier and classically described by the Tofts and extended-Tofts frameworks.[16,18,21] By incorporating pre-contrast T1-mapping, the model assesses baseline tissue vascularity and can anticipate the dose at which enhancement crosses the diagnostic threshold; comparable quantitative-enhancement parameters underpin established radiological classifications such as PI-RADS.[22] The independent association of higher K^{trans} and larger lesion size with maintained adequacy is mechanistically coherent: more vascular and larger lesions enhance robustly even at reduced dose, whereas smaller or hypovascular lesions define the population in whom caution—or a standard dose—remains prudent.

The clinical implications are tangible. A negative low-dose study reduced malignancy probability substantially (LR– 0.057), while a positive study raised it (LR+ 5.71), so the protocol is informative in both directions. For oncology patients undergoing serial surveillance, a sustained 38% per-examination reduction translates into a meaningful decrease in cumulative gadolinium exposure over a treatment course—directly relevant to concerns about dentate-nucleus and globus-pallidus retention and to residual nephrogenic-systemic-fibrosis risk in borderline renal function.[9,10,13]

Our inter-reader agreement (κ 0.83 for characterisation; 0.88 for confidence) compares favourably with published deep-learning dose-reduction reader studies and indicates that the reduced-dose images did not introduce reader-dependent ambiguity.[2,4,7] The stability of agreement across tumour sites further suggests that the pharmacokinetic prescription generalises beyond a single anatomical compartment, although confirmation in larger per-site cohorts is needed.

There are also system-level benefits in the Indonesian context. Advanced MRI capacity is concentrated in tertiary referral centres, where contrast is a recurring consumable cost and a supply-chain dependency. Lower per-examination volume improves pharmacy logistics and the sustainability of high-throughput oncological imaging, without the capital outlay required for hardware upgrades. These advantages are achievable with software integrated into existing 3.0-T workflows.

Methodologically, the study has several strengths: a prospective consecutive design limiting selection bias, blinded independent dual reading with adjudication, a composite histopathology-anchored reference standard, paired within-patient comparison, and complete STARD-compliant reporting with confidence intervals throughout. The substantial-to-almost-perfect inter-reader agreement indicates that the diagnostic judgements were reproducible rather than idiosyncratic.[2,4,7]

Reproducibility merits emphasis. The acquisition protocol is fully specified (3.0-T; T1-TFE TR/TE 3.9/1.9 ms; flip 12°; 256×256 matrix; 240 mm field of view; gadobutrol 1.0 mmol/mL at 2.0 mL/s; 30/60/120-s phases), the extended-Tofts formulation and its K^{trans}/k_{ep} parameters are standard,[15,16] and the contrast-to-noise-ratio threshold defining the minimum effective dose was fixed a priori. Estimating K^{trans} accurately is known to depend on adequate temporal sampling and reliable pre-contrast T1 mapping,[16,21] both of which were standardised here. These details, together with the complete STARD reporting and confidence intervals throughout, are intended to allow independent replication and external validation across vendors and field strengths. We further encourage prospective multicentre registries that capture the prescribed-dose distribution, the realised contrast-to-noise ratio, and any repeat-examination events, since these real-world metrics will determine whether the per-patient reduction observed here is reproducible at scale and whether it translates into measurable reductions in cumulative gadolinium exposure without an offsetting rise in recall imaging.

Performance in this report is quantified through the receiver-operating-characteristic and forest-plot summaries (Figure 1 and Figure 2) rather than individual patient images. Fully de-identified representative and false-negative case examples, with all DICOM metadata and burned-in annotations removed, are available as supplementary material. This choice balances visual illustration with patient-privacy and anonymisation requirements and does not affect the quantitative diagnostic-accuracy conclusions.

From an implementation-science perspective, the model's reliance only on pre-contrast T1-mapping and T2-weighted inputs—sequences already embedded in most oncological MRI protocols—lowers the barrier to deployment. No additional hardware, contrast formulation or scan time is required, and the prescribed dose can be surfaced to the technologist before injection, preserving the existing clinical workflow while adding a quantitative safeguard.

Beyond discrimination, the prescription was well calibrated at the individual level: agreement between predicted and realised tumour enhancement showed no systematic bias, and exploratory decision-curve analysis indicated greater net benefit for the AI-optimised protocol than for either the standard-dose or treat-all strategies across the clinically relevant range of threshold probabilities.[17] This supports the interpretation that the dosing rule responds to patient physiology rather than applying a fixed fractional reduction, consistent with the observation that 4.6% of patients were prescribed no reduction.

The cumulative-exposure benefit is clinically meaningful when projected across a treatment course. For a patient undergoing the typical four to six contrast-enhanced surveillance examinations per year that many oncological pathways require, a sustained per-examination reduction of approximately 38% corresponds to a saving of roughly one to two standard gadolinium doses annually. Because gadobutrol is a macrocyclic agent with comparatively low retention propensity, the principal near-term benefits are reduced nephrotoxic risk in borderline renal function, lower contrast cost, and diminished environmental gadolinium release; the long-term tissue-retention benefit,

while plausible, should be framed cautiously given the agent class.[9,11,13]

Several limitations temper these conclusions. First, this was a single-centre study at one tertiary hospital, so spectrum and verification effects cannot be excluded and external multicentre validation across vendors and field strengths is required before broad adoption.[20] Second, eligibility restricted enrolment to patients with preserved renal function and histologically confirmed solid primaries, which enriches malignant prevalence and may inflate predictive values relative to a screening population; the modest number of benign lesions widened the specificity confidence interval. Third, the composite reference standard combined histopathology with imaging follow-up, introducing potential differential verification. Fourth, the pharmacokinetic model was internally validated; prospective external validation and prospective economic analysis are warranted. Finally, generalisation to non-gadobutrol agents and to lower field strengths remains to be established.

Future work should pursue multicentre, multi-vendor validation with prospective health-economic and patient-safety endpoints, extension to additional tumour types and contrast agents, and head-to-head comparison with deep-learning image-synthesis approaches.[7,20] Integration of the dosing model into routine reporting workflows, with prospective monitoring of diagnostic outcomes, would clarify real-world effectiveness and safety.

5. Conclusion

AI-assisted extended-Tofts pharmacokinetic modelling enabled a personalised 38.4% reduction in gadolinium dose for oncological 3.0-T MRI while remaining non-inferior to standard weight-based dosing, with a sensitivity of 95.2%, specificity of 83.3%, AUC of 0.94 and substantial-to-almost-perfect inter-reader agreement (κ 0.83–0.88). Diagnostic accuracy, objective image quality and reader confidence were preserved. Pending multicentre external validation, pharmacokinetically guided personalised contrast administration is a promising, deployable strategy to lower cumulative gadolinium exposure in precision oncological imaging.

Declarations

Ethics approval and consent: The study was approved by the CMHC Ethics Committee, Palembang, Indonesia (CMHC/EC/2024/087) and conducted per the Declaration of Helsinki; written informed consent was obtained from all participants.

Patient consent for publication: Not applicable; no identifiable individual patient data or images are presented. All images are fully de-identified (DICOM metadata and burned-in annotations removed).

Conflict of interest: The authors declare that they have no competing interests.

Funding: No specific funding was received for this work.

Author contributions (CRediT): Oliva Azalia Putri: Conceptualisation, Methodology, Investigation, Data curation, Writing – original draft. Sony Sanjaya: Software, Formal analysis, Methodology (MR physics), Validation, Visualisation, Supervision, Writing – review & editing. Both authors have read and approved the final manuscript.

Data availability: De-identified data are available from the corresponding author on reasonable request.

Use of AI: No generative artificial intelligence or large language model was used to produce the scientific content, analyses, or conclusions of this manuscript.

References

1. Pasumarthi S, Tamir JI, Christensen S, et al. A generic deep learning model for reduced gadolinium dose in contrast-enhanced brain MRI. Magn Reson Med. 2021;86(3):1687-1700. https://doi.org/10.1002/mrm.28808
2. Ammari S, Bône A, Balleyguier C, et al. Can deep learning replace gadolinium in neuro-oncology? A reader study. Invest Radiol. 2022;57(2):99-107. https://doi.org/10.1097/RLI.0000000000000811
3. Luo H, Zhang T, Gong NJ, et al. Deep learning-based methods may minimize gadolinium-based contrast agent dosage in brain MRI. Eur Radiol. 2021;31(9):6419-6428. https://doi.org/10.1007/s00330-021-07848-3
4. Bône A, Ammari S, Menu Y, et al. From dose reduction to contrast maximization: can deep learning amplify the impact of contrast media on brain MRI quality? A reader study. Invest Radiol. 2022;57(8):527-535. https://doi.org/10.1097/RLI.0000000000000867
5. Chen Y, Weber RA, Neve OM, et al. A deep learning model to reduce agent dose for contrast-enhanced MRI of the cerebellopontine angle cistern. Eur Radiol. 2026;36(5):3710-3721. https://doi.org/10.1007/s00330-025-12187-8
6. Ahmadzade M, Morón FE, Shastri R, et al. AI-assisted post contrast brain MRI: eighty percent reduction in contrast dose. Acad Radiol. 2025;32(8):4767-4776. https://doi.org/10.1016/j.acra.2024.10.026
7. Haase R, Pinetz T, Kobler E, et al. Metastasis detection using true and artificial T1-weighted postcontrast images in brain MRI. Invest Radiol. 2025;60(5):340-348. https://doi.org/10.1097/RLI.0000000000001137
8. Reschke P, Koch V, Gruenewald LD, et al. Deep learning-accelerated prostate MRI: improving speed, accuracy, and sustainability. Acad Radiol. 2025;32(11):6718-6728. https://doi.org/10.1016/j.acra.2025.06.022
9. Ayers-Ringler J, McDonald JS, Connors MA, et al. Neurologic effects of gadolinium retention in the brain after gadolinium-based contrast agent administration. Radiology. 2022;302(3):676-683. https://doi.org/10.1148/radiol.210559
10. Ramalho J, Semelka RC, Cruz J, et al. T1 signal intensity in the dentate nucleus after administration of the macrocyclic gadolinium-based contrast agent gadoterate meglumine: an observational study. Radiologia (Engl Ed). 2022;64(5):397-406. https://doi.org/10.1016/j.rxeng.2020.07.006
11. Cananau C, Forslin Y, Bergendal Å, et al. MRI detection of brain gadolinium retention in multiple sclerosis: magnetization transfer vs. T1-weighted imaging. J Neuroimaging. 2023;33(2):247-255. https://doi.org/10.1111/jon.13079
12. Yao X, Zhang H, Shi D, et al. Gadolinium retention in the brain of mother and pup mouse: effect of pregnancy and

repeated administration of gadolinium-based contrast agents. J Magn Reson Imaging. 2022;56(3):835-845. https://doi.org/10.1002/jmri.28086

13. Welker KM, Joyner D, Kam AW, et al. State of practice: ASNR statement on gadolinium-based contrast agent use in patients with chronic kidney disease. AJNR Am J Neuroradiol. 2025;46(2):227-230. https://doi.org/10.3174/ajnr.A8501
14. van der Molen AJ, Quattrocchi CC, Mallio CA, et al. Ten years of gadolinium retention and deposition: ESMRMB-GREC looks backward and forward. Eur Radiol. 2024;34(1):600-611. https://doi.org/10.1007/s00330-023-10281-3
15. Fang K, Wang Z, Li Z, et al. Convolutional neural network for accelerating the computation of the extended Tofts model in dynamic contrast-enhanced MRI. J Magn Reson Imaging. 2021;53(6):1898-1910. https://doi.org/10.1002/jmri.27495
16. Minosse S, Picchi E, Ferrazzoli V, et al. Influence of scan duration on dynamic contrast-enhanced MRI pharmacokinetic parameters for brain lesions. Magn Reson Imaging. 2024;105:46-56. https://doi.org/10.1016/j.mri.2023.11.002
17. Witowski J, Heacock L, Reig B, et al. Improving breast cancer diagnostics with deep learning for MRI. Sci Transl Med. 2022;14(664):eabo4802. https://doi.org/10.1126/scitranslmed.abo4802
18. Nomani A, Yousefi S, Sargsyan D, et al. A quantitative MRI-based approach to estimate the permeation and retention of nanomedicines in tumors. J Control Release. 2024;368:728-739. https://doi.org/10.1016/j.jconrel.2024.03.019
19. Shin H, Park JE, Jun Y, et al. Deep learning referral suggestion and tumour discrimination using explainable artificial intelligence applied to multiparametric MRI. Eur Radiol. 2023;33(8):5859-5870. https://doi.org/10.1007/s00330-023-09710-0
20. Luo X, Yang Y, Yin S, et al. Automated segmentation of brain metastases with deep learning: a multi-center, randomized crossover, multi-reader evaluation study. Neuro Oncol. 2024;26(11):2140-2151. https://doi.org/10.1093/neuonc/noae113
21. Slavkova KP, DiCarlo JC, Kazerouni AS, et al. Characterizing errors in pharmacokinetic parameters from analyzing quantitative abbreviated DCE-MRI data in breast cancer. Tomography. 2021;7(3):253-267. https://doi.org/10.3390/tomography7030023
22. Clemente A, Selva G, Berks M, et al. Comparison of early contrast enhancement models in ultrafast dynamic contrast-enhanced MRI of prostate cancer. Diagnostics (Basel). 2024;14(9):870. https://doi.org/10.3390/diagnostics14090870