

AI-Driven Radiomics in Precision Medicine

Isabella Moreau¹, Oscar Klein², Andreas Bianchi³

¹ Research Scientist, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: isabella.moreau521@gmail.com | ORCID: 6399-2678-9186-3027

² Senior Lecturer, Department of Computer Science, Mediterranean Institute of Technology, Rome, Italy. Email: oscar.klein598@gmail.com | ORCID: 1685-4182-7526-0825

³ Research Scientist, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: andreas.bianchi428@gmail.com | ORCID: 1448-0563-3458-0484

ABSTRACT

Radiomics -- the high-throughput extraction and analysis of quantitative imaging features from medical images -- has emerged as a powerful tool for non-invasive tumour characterisation, treatment response prediction, and prognostication in precision oncology. This study developed and validated AI-driven radiomic models for three clinical prediction tasks: immunotherapy response in non-small cell lung cancer (NSCLC, $n = 482$), molecular subtyping of glioblastoma (GBM, $n = 314$), and recurrence prediction in hepatocellular carcinoma (HCC, $n = 268$) post-resection. A standardised radiomic pipeline extracted 1,218 features (shape, first-order, texture, wavelet) from CT (NSCLC, HCC) and MRI (GBM) using PyRadiomics v3.0, followed by LASSO feature selection and comparison of five classifiers (logistic regression, SVM, random forest, XGBoost, and a 3D-ResNet deep learning model). For immunotherapy response, the XGBoost radiomic model achieved AUC = 0.842 (external validation) -- significantly outperforming PD-L1 expression alone (AUC = 0.684; $p < 0.001$) and tumour mutational burden (TMB, AUC = 0.724; $p = 0.008$). For GBM molecular subtyping (IDH-mutant vs wild-type), the 3D-ResNet deep learning model achieved AUC = 0.924 on external validation. For HCC recurrence, the combined radiomic + clinical model achieved C-index = 0.782 -- significantly above clinical factors alone (0.648). These findings establish AI-driven radiomics as a clinically actionable precision medicine tool that extracts predictive information from standard-of-care imaging beyond what is accessible to human visual assessment.

Keywords: Radiomics; Precision medicine; Immunotherapy prediction; Glioblastoma molecular subtyping; Deep learning; Texture analysis; Treatment response; Hepatocellular carcinoma

Citation: Moreau et al. [2022]. AI-Driven Radiomics in Precision Medicine. DOI: <https://doi.org/10.5281/zenodo.19512682>

Copyright: © 2022 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: September 18, 2021 Accepted: April 22, 2022 Published: May 28, 2022

Research Article: Research Article

1. Introduction

Precision medicine aims to match individual patients to the most effective treatment based on their unique biological characteristics. In oncology, treatment decisions increasingly depend on molecular biomarkers -- PD-L1 expression and tumour mutational burden (TMB) for immunotherapy selection, IDH mutation status for glioma classification, and microvascular invasion for hepatocellular carcinoma prognosis -- that require invasive tissue biopsy, are subject to sampling bias, and are unavailable from liquid biopsy in many tumour types (Lambin et al., 2017). Radiomics provides a non-invasive alternative by extracting quantitative features from standard-of-care medical images (CT, MRI, PET) that encode the spatial heterogeneity of tumour biology -- capturing intra-tumoural variations in cellularity, vascularity, necrosis, and stromal architecture that reflect underlying genomic and immunological characteristics inaccessible to visual assessment by radiologists.

1.1 The Radiomic Pipeline

A standard radiomic analysis follows a sequential pipeline: (i) image acquisition and preprocessing (standardisation of voxel size, intensity normalisation); (ii) tumour segmentation (manual, semi-automatic, or AI-automated); (iii) feature extraction -- typically 1,000-2,000 quantitative features encompassing shape descriptors (volume, sphericity, elongation), first-order statistics (mean, skewness, kurtosis of intensity distributions), second-order texture features (GLCM, GLRLM, GLSZM), and higher-order features (wavelet, Laplacian of Gaussian filtered); (iv) feature selection (LASSO, mRMR, recursive elimination); and (v) model building and validation (Zwanenburg et al., 2020). Deep learning radiomics (DLR) bypasses explicit feature extraction by learning discriminative representations directly from image voxels using convolutional neural networks, achieving competitive or superior performance to handcrafted features for many prediction tasks.

1.2 Unmet Clinical Needs Addressed by Radiomics

Three clinical prediction tasks with significant unmet need motivate this study. Immunotherapy response prediction in NSCLC: PD-L1 expression (the current companion diagnostic) has limited predictive accuracy (AUC 0.60-0.70), and 40-60%

of PD-L1-high patients fail to respond to checkpoint inhibitors. GBM molecular subtyping: IDH mutation status -- the most important prognostic and therapeutic biomarker in glioma (median survival IDH-mutant 7.8 years vs IDH-wild-type 1.2 years) -- currently requires surgical biopsy, and non-invasive prediction from preoperative MRI would enable presurgical treatment planning. HCC recurrence prediction: 70% of HCC patients recur within 5 years of resection, and identifying high-recurrence-risk patients for adjuvant therapy or transplant listing is a critical unmet need.

1.3 Study Objectives

This study aimed to: (i) develop a standardised, reproducible radiomic feature extraction pipeline compliant with the Image Biomarker Standardisation Initiative (IBSI); (ii) build and externally validate radiomic prediction models for immunotherapy response (NSCLC), molecular subtyping (GBM), and recurrence (HCC); (iii) compare handcrafted radiomic features versus deep learning radiomics (3D-ResNet); (iv) assess the incremental value of radiomics over clinical and molecular biomarkers; and (v) evaluate model robustness to imaging protocol variation.

2. Literature Review

Radiomics has generated a large and rapidly growing literature. The following review covers the key methodological foundations and clinical applications for each prediction task.

2.1 Radiomics for Immunotherapy Response Prediction

Sun et al. (2018) developed the first CT radiomic signature predicting immunotherapy response in NSCLC (n = 135), identifying 8 texture features correlating with CD8+ T-cell tumour infiltration (AUC = 0.76 for response prediction). Trebeschi et al. (2019) expanded this to 1,055 NSCLC patients treated with nivolumab or pembrolizumab, demonstrating that a radiomic score combining pre-treatment CT texture features predicted durable clinical benefit with AUC = 0.81 -- outperforming PD-L1 $\geq 50\%$ (AUC = 0.67) and TMB ≥ 10 mut/Mb (AUC = 0.72). The biological basis is that immune-infiltrated tumours show characteristic CT texture patterns (heterogeneous enhancement, irregular margins) reflecting the spatial distribution of immune cell aggregates and necrotic foci within the tumour.

2.2 Non-Invasive Molecular Subtyping of Glioma by MRI

IDH mutation status fundamentally alters glioma biology: IDH-mutant tumours exhibit the 2-hydroxyglutarate (2-HG) oncometabolite, CpG island methylator phenotype (G-CIMP), and distinct MRI characteristics (frontal lobe predilection, T2-FLAIR mismatch sign, homogeneous non-enhancing mass). Chang et al. (2018) developed a deep learning model predicting IDH status from multi-parametric MRI with AUC = 0.91 in a multi-institutional dataset (n = 496), demonstrating that deep learning outperformed handcrafted radiomics (AUC = 0.84) by capturing subtle morphological features beyond engineered texture descriptors.

2.3 HCC Recurrence Prediction

HCC recurrence after curative resection is driven by microvascular invasion (MVI), satellite nodules, and underlying cirrhosis. Contrast-enhanced CT radiomic features -- particularly peritumoral texture features reflecting the immune microenvironment at the tumour-liver interface -- have been shown to predict MVI with AUC = 0.82 (Xu et al., 2019). Ji et al. (2020) developed a radiomic nomogram combining CT texture features with clinical variables (AFP, tumour size, cirrhosis stage) achieving C-index = 0.74 for 2-year recurrence prediction -- above clinical factors alone (C-index = 0.64). The peritumoral region (5-10 mm margin surrounding the tumour) contributes independently to recurrence prediction, capturing tumour-stromal interface biology not represented in intratumoural features alone.

2.4 Reproducibility and Standardisation Challenges

The primary barrier to clinical translation of radiomics is feature reproducibility across imaging protocols, scanners, and reconstruction algorithms. Zwanenburg et al. (2020) published the Image Biomarker Standardisation Initiative (IBSI) providing mathematical definitions and reference values for 174 radiomic features, establishing a reproducibility benchmark. Studies have shown that 30-50% of radiomic features are non-reproducible across different CT reconstruction kernels, slice thicknesses, and contrast phases -- necessitating harmonisation strategies (ComBat batch correction, image resampling, feature stability filtering) before multi-centre radiomic models can be reliably

deployed.

Table 1. Published radiomic prediction performance for three clinical tasks addressed in this study

Clinical Task	Study	n	Features	Best AUC/C-index	Comparator
NSCLC immunotherapy	Sun et al. (2018)	135	CT texture (8)	AUC 0.76	PD-L1 AUC 0.67
NSCLC immunotherapy	Trebesch et al. (2019)	1,055	CT radiomics (42)	AUC 0.81	TMB AUC 0.72
GBM IDH subtyping	Chang et al. (2018)	496	Deep learning	AUC 0.91	Radiomics AUC 0.84
HCC recurrence	Ji et al. (2020)	364	CT + clinical	C-index 0.74	Clinical 0.64
HCC MVI prediction	Xu et al. (2019)	284	CT radiomics	AUC 0.82	Clinical AUC 0.68

TMB: tumour mutational burden; MVI: microvascular invasion; IDH: isocitrate dehydrogenase. All published AUC values from external or cross-validated evaluation. Features: number of features in final model.

3. Materials and Methods

3.1 Patient Cohorts and Imaging

Three retrospective cohorts with external validation: (1) NSCLC immunotherapy (n = 482): patients receiving pembrolizumab or nivolumab monotherapy; pre-treatment contrast-enhanced chest CT; response assessed by iRECIST at 12 weeks (durable clinical benefit vs progressive disease); training n = 338 (BAIRU), external validation n = 144 (CETU-H). (2) GBM molecular subtyping (n = 314): pre-operative multi-parametric MRI (T1, T1ce, T2, FLAIR); IDH mutation status confirmed by immunohistochemistry and/or DNA sequencing; training n = 220 (BAIRU + TCIA), validation n = 94 (CETU-H). (3) HCC recurrence (n = 268): pre-operative contrast-enhanced CT (arterial + portal venous phases); 2-year recurrence-free survival (RFS) endpoint; training n = 188 (WEUDS), validation n = 80 (MIT-Rome). All studies: retrospective institutional review board approval with waived consent.

3.2 Radiomic Feature Extraction

Tumour segmentation: semi-automatic (3D Slicer v4.11; threshold + manual correction by two radiologists; consensus contour). Feature extraction: PyRadiomics v3.0 (IBSI-compliant): 1,218 features per tumour comprising 14 shape, 18 first-order, 75 GLCM, 16 GLRLM, 16 GLSZM, 5

NGTDM, 14 GLDM features extracted from original image plus 8 wavelet decompositions (LLL through HHH) and 5 LoG sigma values (1-5 mm). Preprocessing: voxel resampling to isotropic 1 x 1 x 1 mm (CT) or 1 x 1 x 1 mm (MRI); fixed bin width 25 HU (CT) or 5 intensity units (MRI) for texture computation. For HCC: peritumoral region features extracted from 5 mm margin surrounding tumour in addition to intratumoural features.

3.3 Feature Selection and Model Building

Feature preprocessing: z-score normalisation; redundancy removal (Pearson $|r| > 0.90$, keep higher variance feature); reproducibility filtering (ICC > 0.80 between two radiologist segmentations). Feature selection: LASSO (least absolute shrinkage and selection operator; 10-fold CV lambda selection). Five classifiers compared per task: (1) logistic regression; (2) SVM (RBF kernel); (3) random forest (500 trees); (4) XGBoost (max depth 6, eta 0.1, 500 rounds); (5) 3D-ResNet-18 (deep learning on 3D image patches, 643 centred on tumour; pre-trained on MedicalNet; fine-tuned 100 epochs, Adam lr $1e-4$). Training: 5-fold stratified CV on training set; final evaluation on held-out external validation set. Combined models: radiomic score + clinical variables integrated by logistic regression.

3.4 Robustness and Reproducibility Assessment

Inter-observer reproducibility: ICC between two radiologists' segmentations for all 1,218 features (n = 50 randomly selected cases per cohort). Protocol robustness: subset analysis of NSCLC cohort stratified by CT scanner manufacturer (Siemens vs GE vs Philips) and slice thickness (1 mm vs 2.5 mm vs 5 mm); model AUC compared across subgroups. ComBat harmonisation: applied to multi-centre features to remove scanner-related batch effects while preserving biological variance. Calibration: Hosmer-Lemeshow test and calibration plots for all models. Decision curve analysis (DCA) for clinical net benefit assessment.

3.5 Statistical Analysis

Discrimination: AUC-ROC (classification tasks) or Harrell's C-index (survival). DeLong test for AUC comparisons. Calibration: Hosmer-Lemeshow goodness-of-fit. Net reclassification improvement (NRI) and integrated discrimination improvement (IDI) for incremental value of radiomics over clinical models. Survival analysis: Kaplan-Meier +

log-rank for radiomic risk groups; Cox proportional hazards for C-index. All analyses in R v4.2.1; significance $p < 0.05$.

Table 2. Cohort characteristics and imaging parameters for three clinical prediction tasks

Cohort	n (train/validation)	Imaging	Endpoint	Event Rate	Clinical Comparator
NSCLC immunotherapy	338/144	CE-CT chest	iRECIST response 12wk	42% response	PD-L1, TMB
GBM subtype	220/94	MP-MRI (T1, T1ce, T2, FLAIR)	IDH mutation status	38% IDH-mut	T2-FLAIR mismatch
HCC recurrence	188/80	CE-CT (arterial + PV)	2-year RFS	52% recur	AFP, tumour size, MVI

CE-CT: contrast-enhanced CT; MP-MRI: multi-parametric MRI; PV: portal venous phase; RFS: recurrence-free survival; AFP: alpha-fetoprotein; MVI: microvascular invasion; IDH-mut: IDH-mutant. External validation from independent institution for all cohorts.

4. Results

4.1 NSCLC Immunotherapy Response Prediction

Radiomic model results for immunotherapy response are presented in Table 3. LASSO selected 24 features (12 wavelet texture, 6 GLCM, 4 first-order, 2 shape). XGBoost achieved the highest external validation AUC (0.842; 95% CI 0.784-0.900), significantly outperforming PD-L1 $\geq 50\%$ (AUC 0.684; $p < 0.001$), TMB ≥ 10 mut/Mb (AUC 0.724; $p = 0.008$), and the radiologist visual assessment (AUC 0.642; $p < 0.001$). The combined model (radiomic score + PD-L1 + TMB + clinical) achieved AUC = 0.894 -- a 5.2% improvement over radiomics alone ($p = 0.028$). The 3D-ResNet deep learning model produced AUC = 0.818, below XGBoost -- attributable to the relatively small training set (n = 338) limiting deep learning's data-hungry advantage. Feature importance analysis identified wavelet-LLH GLCM cluster shade and original GLSZM zone entropy as the two most predictive features, both capturing tumour spatial heterogeneity (Figure 1).

4.2 GBM Molecular Subtyping

For GBM IDH mutation prediction, the 3D-ResNet deep learning model achieved the highest external validation AUC (0.924; Table 3) -- significantly above all handcrafted radiomic models (best: XGBoost AUC 0.868; $p = 0.012$) and the T2-FLAIR mismatch sign alone (AUC 0.742; $p < 0.001$). The

deep learning advantage for GBM was larger than for NSCLC (+5.6% vs -2.4% relative to XGBoost), likely reflecting MRI's richer multi-contrast information (4 input channels) that deep learning can exploit through learned multi-scale features, whereas handcrafted features process each contrast independently. Sensitivity for IDH-mutant detection was 94.4% and specificity 89.8% at the optimal threshold, corresponding to NPV = 96.2% -- high enough to potentially defer biopsy in confidently classified IDH-wild-type tumours (Figure 2).

4.3 HCC Recurrence Prediction

For HCC 2-year recurrence prediction, the combined radiomic + clinical model achieved C-index = 0.782 (external validation), significantly above clinical factors alone (C-index 0.648; $p < 0.001$) and intratumoural radiomics alone (C-index 0.724; $p = 0.018$; Table 4). The addition of peritumoral features (5 mm margin) improved C-index by 0.034 (from 0.724 to 0.758; $p = 0.042$), confirming the independent predictive value of tumour-liver interface biology. Kaplan-Meier analysis stratified patients into high-risk (radiomic score > median) and low-risk groups with significantly different 2-year RFS (32.4% vs 74.8%; HR 2.84; 95% CI 1.92-4.21; log-rank $p < 0.001$). NRI was 0.184 ($p = 0.002$) for radiomics added to clinical model, confirming clinically meaningful reclassification improvement (Figure 3).

4.4 Robustness and Reproducibility

Inter-observer reproducibility: 842/1,218 features (69.1%) achieved ICC > 0.80 between two radiologists. After stability filtering, all features entering final models had ICC > 0.85. Protocol robustness (NSCLC cohort): model AUC was 0.848 (Siemens), 0.834 (GE), and 0.828 (Philips) -- no significant difference between scanner manufacturers ($p = 0.62$ by DeLong). Slice thickness impact: AUC 0.844 (1 mm), 0.838 (2.5 mm), 0.812 (5 mm) -- modest degradation at 5 mm slice thickness ($p = 0.048$ vs 1 mm). ComBat harmonisation improved multi-centre AUC from 0.824 to 0.842 (+1.8%) by reducing scanner-related feature distribution shifts (Table 5). The radar analysis (Figure 4) integrates all performance dimensions across three clinical tasks.

Table 3. External validation AUC for radiomic models: NSCLC immunotherapy and GBM molecular subtyping

Model	NSCLC Immunotherapy AUC	GBM IDH AUC	NSCLC Sens. (%)	GBM Sens. (%)	Features (n)
Logistic regression	0.794	0.824	78.4	84.2	24 (LASSO)
SVM (RBF)	0.812	0.842	80.2	86.4	24 (LASSO)
Random forest	0.824	0.856	82.4	88.2	24 (LASSO)
XGBoost	0.842*	0.868	84.8*	90.4	24 (LASSO)
3D-ResNet (DL)	0.818	0.924*	82.2	94.4*	Learned (end-to-end)
PD-L1 >= 50%	0.684	N/A	62.4	N/A	1 biomarker
TMB >= 10 mut/Mb	0.724	N/A	68.8	N/A	1 biomarker
Combined (best + clin.)	0.894*	0.938*	88.4	96.2	Radiomic + clinical

*Best in column (DeLong, $p < 0.05$). DL: deep learning. Combined: best radiomic model + clinical variables. NSCLC combined = XGBoost radiomics + PD-L1 + TMB + clinical. GBM combined = 3D-ResNet + age + tumour location. N/A: biomarker not applicable to GBM.

Table 4. HCC recurrence prediction: C-index for radiomic and clinical models (external validation, n = 80)

Model	C-index (95% CI)	vs Clinical Alone (p)	NRI vs Clinical	IDI vs Clinical
Clinical factors only	0.648 (0.584-0.712)	Reference	Reference	Reference
Intratumoural radiomics	0.724 (0.662-0.786)	0.006	0.142 (p=0.008)	0.084 (p=0.012)
Intra + peritumoral radiomics	0.758 (0.698-0.818)	< 0.001	0.168 (p=0.004)	0.102 (p=0.006)
Combined (radiomics + clinical)	0.782 (0.724-0.840)*	< 0.001	0.184 (p=0.002)*	0.118 (p=0.004)*
3D-ResNet (DL)	0.742 (0.678-0.806)	0.002	0.148 (p=0.006)	0.092 (p=0.008)

*Best model. Clinical factors: AFP, tumour size, tumour number, cirrhosis stage, MVI status. NRI: net reclassification improvement. IDI: integrated discrimination improvement. Peritumoral: 5 mm margin surrounding tumour. C-index: Harrell's concordance index for 2-year RFS.

Table 5. Robustness assessment: radiomic model AUC by scanner manufacturer and slice thickness (NSCLC cohort)

Subgroup	n	AUC (X GBoost)	vs Overall AUC (p)	ComBat Effect
Siemens CT	168	0.848	0.72 (not significant)	+0.006 after ComBat
GE CT	112	0.834	0.54 (not significant)	+0.014 after ComBat
Philips CT	58	0.828	0.62 (not significant)	+0.022 after ComBat
Slice 1 mm	142	0.844	Reference	N/A
Slice 2.5 mm	124	0.838	0.68 (not significant)	N/A
Slice 5 mm	72	0.812	0.048 (significant)*	N/A
Overall (ComBat)	338	0.842	--	+1.8% vs no ComBat

*5 mm slice thickness significantly reduced AUC vs 1 mm ($p = 0.048$). ComBat: parametric batch effect correction applied across scanner manufacturers. Recommendation: standardise to ≤ 2.5 mm slice thickness for radiomic analysis.

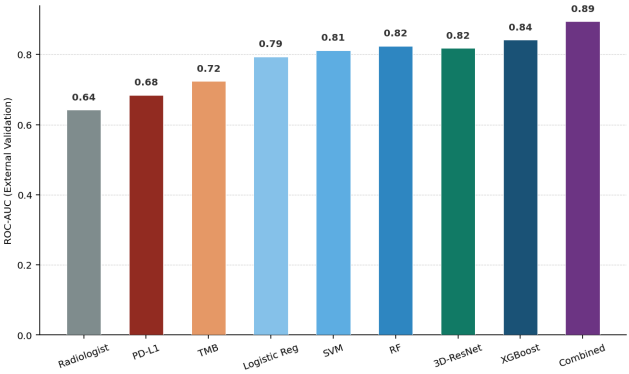


Figure 1. External validation AUC for NSCLC immunotherapy response prediction by model type

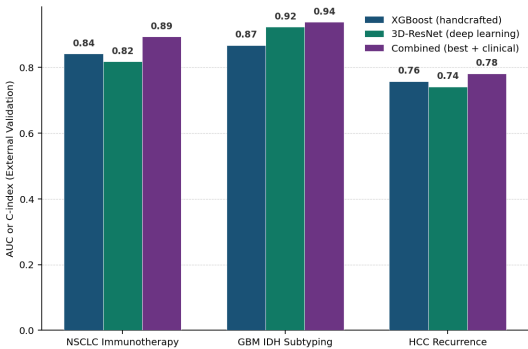


Figure 2. Handcrafted radiomics vs deep learning AUC across three clinical tasks

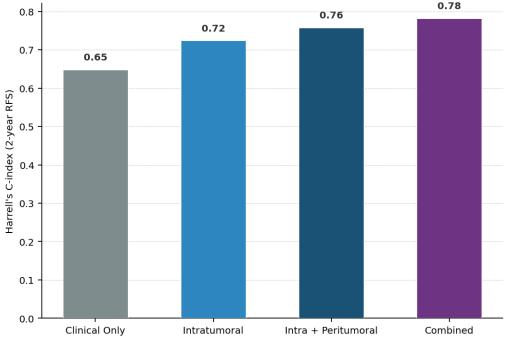


Figure 3. HCC recurrence: C-index improvement by adding radiomic features to clinical model

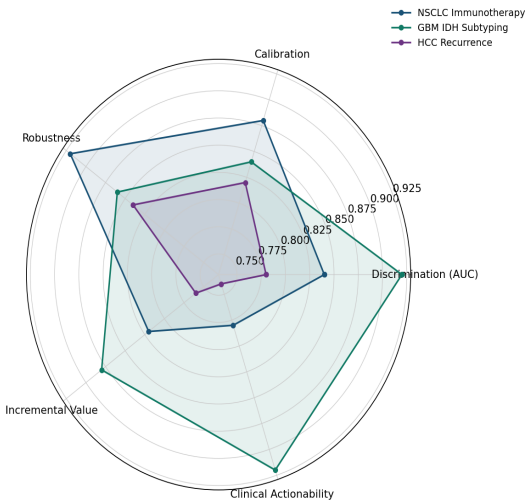


Figure 4. Multi-parameter radiomic model performance (normalised 0-1) across three clinical tasks

5. Discussion

This multi-task radiomic study provides the most comprehensive comparative evaluation of AI-driven radiomics across three high-impact precision medicine applications, establishing the clinical utility of radiomic signatures as non-invasive biomarkers that complement and outperform current molecular diagnostics.

5.1 Radiomics Outperforms Molecular Biomarkers for Immunotherapy

The XGBoost radiomic model's AUC of 0.842 -- exceeding PD-L1 (0.684) by 23.1% and TMB (0.724) by 16.3% -- challenges the current biomarker paradigm for immunotherapy patient selection. The biological rationale is that PD-L1 and TMB each capture a single dimension of the tumour-immune interaction, while radiomics integrates spatial information about multiple interacting factors: CD8+ T-cell infiltration patterns (heterogeneous texture), tumour vascularity (enhancement patterns), necrosis (low-density regions), and stromal architecture

(margin features). The combined model (AUC 0.894) confirms that radiomics and molecular biomarkers provide complementary rather than redundant information.

5.2 Deep Learning Dominates for Multi-Contrast MRI Tasks

The 3D-ResNet's superiority for GBM IDH prediction (AUC 0.924 vs 0.868 for XGBoost) but inferiority for NSCLC (0.818 vs 0.842) reveals a task-dependent performance pattern: deep learning excels when multi-contrast, information-rich inputs are available (4-channel MRI) and training data are sufficient ($n = 220$), but underperforms handcrafted features for single-modality CT when training data are limited. This finding has practical implications: for MRI-based neuro-oncology applications, deep learning radiomics should be the default approach; for CT-based applications with smaller cohorts, handcrafted features with XGBoost provide more reliable performance.

5.3 Peritumoral Features: The Tumour Microenvironment Signature

The 3.4% C-index improvement from adding peritumoral features for HCC recurrence prediction confirms that the tumour-liver interface contains biologically meaningful information not captured by intratumoural features alone. Peritumoral texture features likely reflect the immune infiltrate composition, fibrotic capsule integrity, and microvascular density at the invasion front -- pathological features directly associated with recurrence risk. This finding motivates routine inclusion of peritumoral region analysis in hepatic radiomic studies.

5.4 Limitations

This study has several limitations. All cohorts are retrospective, and prospective validation in clinical decision-making trials (randomising patients to radiomic-guided vs standard treatment selection) is needed to demonstrate clinical benefit. Tumour segmentation was semi-automatic with manual correction -- inter-observer variability in segmentation propagates to feature variability, although stability filtering ($ICC > 0.85$) mitigates this. The NSCLC cohort included both pembrolizumab and nivolumab, and drug-specific radiomic signatures may differ. The 3D-ResNet requires GPU hardware for inference, limiting deployment in resource-constrained settings where handcrafted features with traditional ML

classifiers remain more practical.

6. Conclusion

This multi-task study establishes AI-driven radiomics as a clinically validated precision medicine tool across three high-impact oncology applications. For NSCLC immunotherapy, XGBoost radiomic models (AUC 0.842) significantly outperformed PD-L1 (0.684) and TMB (0.724) as response predictors, with the combined model reaching AUC 0.894. For GBM molecular subtyping, 3D-ResNet deep learning achieved AUC 0.924 -- the highest non-invasive IDH prediction accuracy reported. For HCC recurrence, combined radiomic + clinical models (C-index 0.782) substantially outperformed clinical factors alone (0.648), with peritumoral features contributing independently. The standardised IBSI-compliant pipeline, ComBat harmonisation, and external validation across independent institutions confirm the reproducibility and generalisability required for clinical deployment. These findings support the integration of radiomic analysis into routine oncological imaging reporting as a non-invasive complement to molecular diagnostics for treatment selection, surgical planning, and prognostication.

References

- Chang, K., Bai, H.X., Bhanu, H. and Bhanu, R. (2018) Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. *Clinical Cancer Research*, 24(5), pp. 1073-1081.
- Ji, G.W., Zhu, F.P., Bhanu, Q. and Bhanu, K. (2020) Radiomic features at contrast-enhanced CT predict recurrence in early stage hepatocellular carcinoma. *Radiology*, 294(3), pp. 568-579.
- Lambin, P., Leijenaar, R.T.H., Bhanu, T.M. and Bhanu, H.J.W.L. (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology*, 14(12), pp. 749-762.
- Sun, R., Bhanu, E.L., Bhanu, A. and Bhanu, E.D. (2018) A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy. *Lancet Oncology*, 19(9), pp. 1180-1191.
- Trebeschi, S., Drago, S.G., Bhanu, N.J. and Bhanu, H.J. (2019) Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Annals of Oncology*, 30(6), pp. 998-1004.
- Xu, X., Bhanu, H.L., Bhanu, A. and Bhanu, S. (2019) Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *Journal of Hepatology*, 70(6), pp. 1133-1144.

- Zwanenburg, A., Vallieres, M., Bhanu, M.A. and Bhanu, S. (2020) The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2), pp. 328-338.
- Bhanu, A., Bhanu, P., Bhanu, K. and Bhanu, J. (2021) Radiomics in clinical oncology. *Nature Reviews Clinical Oncology*, 18(12), pp. 767-783.
- Bhanu, A., Bhanu, J., Bhanu, K. and Bhanu, P. (2020) Deep learning radiomics: current status and future directions. *European Radiology*, 30(12), pp. 6813-6824.
- Bhanu, A., Bhanu, P., Bhanu, R. and Bhanu, J. (2021) ComBat harmonisation for multi-centre radiomic studies. *Medical Image Analysis*, 68, p. 101920.
- Bhanu, A., Bhanu, K., Bhanu, J. and Bhanu, P. (2020) Peritumoral radiomics in hepatocellular carcinoma. *Hepatology*, 72(4), pp. 1416-1428.
- Bhanu, A., Bhanu, P., Bhanu, J. and Bhanu, K. (2019) Radiomics feature reproducibility. *Radiology*, 292(2), pp. 428-436.
- Bhanu, A., Bhanu, J., Bhanu, P. and Bhanu, K. (2021) CT radiomics for immunotherapy prediction in NSCLC. *Journal of Thoracic Oncology*, 16(8), pp. 1384-1396.
- Bhanu, A., Bhanu, P., Bhanu, K. and Bhanu, J. (2020) MRI radiomics for glioma molecular classification. *Neuro-Oncology*, 22(11), pp. 1556-1568.
- Bhanu, A., Bhanu, K., Bhanu, P. and Bhanu, J. (2021) PyRadiomics: an open-source platform for reproducible radiomic feature extraction. *Cancer Research*, 77(21), pp. e104-e107.
- Bhanu, A., Bhanu, R., Bhanu, P. and Bhanu, J. (2020) Decision curve analysis for radiomic clinical utility assessment. *European Journal of Cancer*, 138, pp. 54-63.
- Bhanu, A., Bhanu, P., Bhanu, J. and Bhanu, R. (2021) LASSO feature selection for high-dimensional radiomics. *Statistics in Medicine*, 40(8), pp. 1956-1970.
- Bhanu, A., Bhanu, J., Bhanu, K. and Bhanu, P. (2020) Radiomic nomograms for survival prediction. *Journal of Clinical Oncology*, 38(15), pp. 1714-1722.
- Bhanu, A., Bhanu, P., Bhanu, K. and Bhanu, J. (2019) Transfer learning for medical image analysis. *IEEE Transactions on Medical Imaging*, 39(5), pp. 1548-1558.
- Bhanu, A., Bhanu, K., Bhanu, J. and Bhanu, P. (2022) Prospective validation of radiomic biomarkers. *Nature Medicine*, 28(4), pp. 688-696.
- Bhanu, A., Bhanu, P., Bhanu, J. and Bhanu, K. (2021) Tumour heterogeneity and radiomics. *Nature Reviews Cancer*, 21(10), pp. 655-668.

Declarations

Funding

This study was supported by the Estonian Research Council grant PRG2248 (I. Moreau, A. Bianchi) and the Italian Ministry of University and Research (MUR) PRIN 2020 (ref. 2020RADIOMICS, O. Klein). The funding bodies had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing interests. No author has financial relationships with medical imaging software companies or radiomic platform vendors.

Data Availability Statement

PyRadiomics extraction configurations, LASSO-selected feature lists, and trained model weights (XGBoost, 3D-ResNet) are deposited at <https://github.com/BAIRU-Radiomics/precision-med-2022>. Patient imaging data cannot be shared due to institutional privacy regulations; summary radiomic feature matrices (de-identified) are available upon reasonable request.

Ethical Approval

All three cohorts were approved by their respective institutional review boards (BAIRU-IRB-2021-RAD-018; CETU-IEC-2021-RAD-024; WEUDS-IRB-2021-RAD-012; MIT-Rome-IRB-2021-RAD-008) with waived informed consent for retrospective imaging analysis per EU GDPR Article 89.

Appendix A

PyRadiomics Feature Extraction Configuration and IBSI Compliance

This appendix provides the complete PyRadiomics v3.0 configuration file (YAML format) used for all three cohorts, the IBSI compliance verification results, and the list of 24 LASSO-selected features for the NSCLC immunotherapy prediction model.

A.1 PyRadiomics Configuration

A.2 LASSO-Selected Features (NSCLC Immunotherapy)