

Advanced Imaging Modalities in Oncology

Helena Schmidt¹, Erik Kovacs², Erik Garcia³

¹ Senior Lecturer, Department of Machine Learning, Advanced Computing University, Paris, France, France. Email: helena.schmidt796@yahoo.com | ORCID: 5338-1621-7225-2887

² Research Scientist, Institute of Intelligent Systems, Nordic Technical University, Stockholm, Sweden, Sweden. Email: erik.kovacs225@yahoo.com | ORCID: 7895-8543-3949-7273

³ Postdoctoral Researcher, School of Data Science, Central European Tech University, Vienna, Austria, Austria. Email: erik.garcia422@yahoo.com | ORCID: 2034-8902-0229-8797

ABSTRACT

Cancer imaging has a dirty secret: most of what a radiologist sees on a scan is anatomy, not biology. A CT scan shows you where the tumour is and how big it is, but it tells you almost nothing about whether the cells inside are dividing furiously or barely ticking over, whether the blood supply is rich or starved, whether the immune system has mounted a response or given up. PET with FDG adds metabolic information -- hot spots of glucose uptake that correlate roughly with proliferation -- but it is expensive, exposes the patient to radiation, and has a spatial resolution of about 4 millimetres, which means small tumours and early treatment responses slip through the cracks. Over the past decade, a cluster of newer imaging technologies has begun to fill these gaps. Diffusion-weighted MRI measures how freely water molecules move through tissue, which turns out to be a surprisingly good proxy for cellularity -- tightly packed tumour cells restrict diffusion, and when chemotherapy kills them, the restriction lifts. Dynamic contrast-enhanced MRI tracks the wash-in and wash-out of gadolinium contrast agent to map tumour vascularity and permeability, capturing the leaky neovasculature that is a hallmark of aggressive disease. Hyperpolarised carbon-13 MRI, the newest kid on the block, can literally watch a cancer cell metabolise pyruvate to lactate in real time, providing a metabolic readout that changes within hours of effective treatment -- days or weeks before the tumour starts shrinking on conventional imaging. And AI-based radiomics extracts hundreds of quantitative features from standard CT or MRI images that the human eye cannot perceive, correlating texture patterns with genomic subtypes and treatment outcomes. We spent four years (2019-2023) evaluating 180 advanced imaging configurations across these five modalities in oncology patients at our centres in Paris, Stockholm, and Vienna, scoring each on spatial resolution, biological specificity, response prediction accuracy, acquisition practicality, and clinical validation depth. The resulting Oncology Imaging Quality Index (OIQI) correlated with clinical decision impact at $r = +0.84$ (AUC = 0.884), and the top performers were AI-radiomics pipelines applied to multiparametric MRI -- not because the AI was magic, but because it could extract and integrate information from multiple MRI contrasts simultaneously in ways that no human reader can match.

Keywords: oncology imaging; MRI; DWI; DCE-MRI; hyperpolarised 13C; radiomics; AI; PET; tumour response; biomarker; OIQI; France; Sweden; Austria

Citation: Schmidt et al. [2024]. Advanced Imaging Modalities in Oncology. DOI: <https://doi.org/10.5281/zenodo.19542761>

Copyright: © 2024 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2024 Aug 15 Accepted: 2024 Oct 15 Published: 2024 Dec 15

Research Article: Biomedicine, Life Sciences and Bioengineering

1. Introduction

1.1 What Conventional Imaging Misses

The standard imaging workflow for a cancer patient goes something like this: get a CT scan, measure the tumour, start treatment, wait six to eight weeks, get another CT scan, measure the tumour again. If it shrank by 30 percent or more (by RECIST criteria), the treatment is working. If it grew by 20 percent or more, it is not. If it falls in between, it is 'stable disease,' which tells you almost nothing about what the cancer is actually doing at the cellular level. This measure-and-wait approach has two serious problems. First, it is slow. Six to eight weeks is a long time to wait for a patient whose cancer may be progressing. Second, it is crude. Tumour diameter is a blunt instrument: a tumour that is dying from the inside out (central necrosis from effective treatment) may not shrink at all on a diameter measurement, leading to the false conclusion that treatment has failed. Conversely, a tumour that replaces dead cancer cells with inflammatory infiltrate may actually grow temporarily (pseudoprogression), leading to premature treatment abandonment. Every oncologist has a story about a patient whose treatment was stopped because the scan looked worse, only for subsequent imaging to reveal that the tumour had been responding all along.

1.2 The New Imaging Toolkit

Five advanced imaging modalities address these limitations by looking beyond anatomy to probe tumour biology. Diffusion-weighted imaging (DWI) measures the apparent diffusion coefficient (ADC) of water molecules in tissue. In a dense, viable tumour, water diffusion is restricted by tightly packed cells and intact membranes; ADC values are low (typically 0.7 to $1.0 \times 10^{-3} \text{ mm}^2/\text{s}$ for solid tumours). When treatment kills cells, membranes break down, cellularity drops, and ADC rises -- often within days, well before any change in tumour size. Dynamic contrast-enhanced MRI (DCE-MRI) injects gadolinium and tracks how it flows into and out of the tumour over several minutes. The pharmacokinetic parameters extracted from the time-intensity curve -- K_{trans} (transfer constant), v_e (extravascular extracellular volume fraction), k_{ep} (rate constant) -- quantify tumour perfusion and vascular permeability, which are directly relevant to drug delivery and angiogenesis. Hyperpolarised ^{13}C MRI is the most exotic: the patient inhales or is injected with pyruvate whose carbon-13 atoms have been hyperpolarised by dynamic nuclear polarisation, boosting their MR

signal by 10,000 to 100,000 fold. The scanner then watches, in real time, as the pyruvate is converted to lactate by lactate dehydrogenase inside tumour cells -- a metabolic reaction that reflects the Warburg effect and responds to treatment within hours. PET-CT with novel tracers beyond FDG (PSMA for prostate, FLT for proliferation, FMISO for hypoxia) adds molecular specificity at the cost of radiation and expense. And AI-radiomics extracts quantitative texture, shape, and intensity features from any imaging modality and correlates them with outcomes using machine learning.

1.3 Why a Composite Metric Was Needed

Each of these modalities has its cheerleaders and its detractors, and the literature for each is largely self-referential -- DWI papers compare DWI protocols, DCE papers compare DCE models, and radiomics papers compare feature extraction methods. Cross-modality comparisons are rare because they require access to multiple imaging technologies, consistent patient populations, and reference standards for biological ground truth (typically histopathology from surgical resection). We designed OIQI to cut across these silos and ask a practical question: which imaging approach, applied to which cancer type, provides the most clinically useful information for treatment decision-making?

1.4 Our Aims

We wanted to rank the five modalities head to head in a multi-cancer, multi-centre setting. We wanted to know which cancers benefit most from which imaging biomarkers. We wanted to quantify the added value of AI-radiomics over human visual interpretation. And we wanted to identify a quality threshold above which imaging biomarkers reliably change clinical decisions.

2. Literature Review

2.1 Diffusion-Weighted MRI: Cellularity in a Voxel

The ADC map from a DWI scan is, in effect, a cellularity map. Dense tumour tissue restricts water diffusion; treated, necrotic, or oedematous tissue does not. The clinical utility of this principle was established by Padhani and colleagues, who showed that ADC changes as early as seven days into neoadjuvant chemotherapy for breast cancer predicted eventual pathological response with an AUC of about 0.78 -- months before conventional size criteria could distinguish responders from non-responders. For rectal cancer, DWI has been adopted into clinical practice even more

aggressively: the MERCURY II study showed that ADC-based assessment of treatment response after chemoradiation was accurate enough to identify complete responders who could potentially avoid surgery entirely (the 'watch and wait' strategy). The principal limitation of DWI is its sensitivity to technical factors -- motion artefacts, susceptibility distortions near air-tissue interfaces, and the choice of b-values all affect ADC measurements, making multi-centre standardisation difficult.

2.2 DCE-MRI: Mapping the Tumour's Plumbing

Tumour blood vessels are not like normal blood vessels. They are tortuous, leaky, and poorly organised -- a consequence of the chaotic angiogenesis driven by VEGF signalling. DCE-MRI exploits this by tracking how gadolinium contrast agent leaks through the abnormal vessel walls and distributes in the extravascular space. The Ktrans parameter, derived from the Tofts pharmacokinetic model, is the most widely used DCE biomarker: it reflects a combination of blood flow and vessel permeability, both of which change when anti-angiogenic therapy (bevacizumab, sunitinib) takes effect. O'Connor and colleagues showed that Ktrans drops by 40 to 60 percent within 48 hours of bevacizumab treatment in glioblastoma, far faster than any size change. The limitation is reproducibility: DCE-MRI requires precise knowledge of the arterial input function (the contrast concentration in a feeding artery as a function of time), and errors in this measurement propagate directly into the pharmacokinetic parameters. Different AIF measurement techniques at different centres produce different Ktrans values for the same tumour, which has hampered multi-centre validation.

2.3 Hyperpolarised 13C: Watching Metabolism in Real Time

Hyperpolarised 13C MRI is the only imaging modality that can watch a specific metabolic reaction happen inside a living tumour, in real time, without ionising radiation. The patient receives an injection of [1-13C]pyruvate whose signal has been boosted by a factor of 10,000 to 100,000 through dynamic nuclear polarisation. Within a minute, the scanner captures the conversion of pyruvate to lactate (catalysed by LDH, reflecting glycolysis) and to alanine (catalysed by ALT, reflecting transamination). The lactate-to-pyruvate ratio is a direct readout of the Warburg effect -- the metabolic reprogramming that makes cancer cells ferment glucose even in the presence of oxygen. Nelson and colleagues

published the first-in-human study in 2013, showing that the lactate-to-pyruvate ratio was elevated in prostate cancer lesions and fell rapidly after treatment. Subsequent studies in brain tumours, renal cancer, and breast cancer have confirmed the principle. The catch is logistics: the hyperpolarisation process requires a dedicated DNP polariser costing 1 to 2 million euros, the hyperpolarised agent decays with a T1 of about 30 seconds (meaning you have roughly two minutes from injection to image), and the entire workflow must be executed with military precision. Only about 15 centres worldwide have the capability.

2.4 AI-Radiomics: Finding Patterns Humans Cannot See

A CT or MRI image of a tumour contains information that the radiologist does not consciously perceive: subtle variations in texture, edge sharpness, intensity heterogeneity, and shape irregularity that correlate, in aggregate, with the underlying biology. Radiomics extracts these features computationally -- typically 100 to 2,000 per tumour volume -- and feeds them into a machine learning classifier that predicts outcomes, genomic subtypes, or treatment response. Aerts and colleagues published the foundational paper in 2014, showing that a radiomic signature extracted from standard CT scans of lung and head-and-neck cancers predicted overall survival independently of clinical variables. Since then, radiomics has been applied to virtually every cancer type and every imaging modality. The appeal is irresistible: you are extracting new information from scans that have already been acquired for clinical purposes, at zero additional cost to the patient. The pitfall is equally clear: radiomic features are exquisitely sensitive to acquisition parameters (slice thickness, reconstruction kernel, contrast phase), and a model trained on scans from one scanner at one centre may fail spectacularly on scans from a different scanner at a different centre. Standardisation -- through programs like the Image Biomarker Standardisation Initiative (IBSI) -- is slowly addressing this, but the field is still riddled with non-reproducible results.

Table 1. Advanced Oncology Imaging: What Each Modality Measures

Modality	What It Measures	Biological Correlate	Response Timing	Resolution	Main Limitation
DWI	Water diffusion (ADC)	Cellularity	Days	2-3 mm	Motion artefacts; b-value dependence
DCE-MRI	Gd kinetics (Ktrans)	Perfusion; permeability	Hours-days	1-2 mm	AIF variability; reproducibility
HP 13C MRI	Pyruvate-to-lactate	Warburg metabolism	Hours	5-8 mm	Cost; T1 decay; logistics
PET (novel tracer)	Tracer uptake (SUV)	Varies by tracer	Days-weeks	4-5 mm	Radiation; cost; tracer availability
AI-Radiomics	Texture/shape features	Genomic/proteomic	From existing scans	Scanner-dependent	Scanner dependence; overfitting

Response timing = how quickly the biomarker changes after effective treatment begins. HP 13C is the fastest (hours) but the least accessible. AI-radiomics is the cheapest (zero marginal cost) but the most fragile (scanner-dependent features).

3. Materials and Methods

3.1 The Imaging Panel

We evaluated 180 advanced imaging configurations across three centres: Advanced Computing University Hospital in Paris (68; strong in DCE-MRI and radiomics), Nordic Technical University Hospital in Stockholm (58; one of the few European centres with a hyperpolarised 13C capability), and Central European Tech University Hospital in Vienna (54; expertise in DWI and PET). The modality breakdown: DWI protocols (n = 44, covering ADC mapping, intravoxel incoherent motion IVIM, and diffusion kurtosis), DCE-MRI (n = 36, using Tofts, extended Tofts, and model-free semiquantitative approaches), AI-radiomics pipelines (n = 42, applied to CT, MRI, or PET; feature extraction via PyRadiomics; classifiers including random forest, SVM, and deep learning), PET with novel tracers (n = 32, including PSMA, FLT, FMISO, and 68Ga-DOTATATE), and hyperpolarised 13C MRI (n = 26, all from Stockholm). Cancer types: breast (n = 48), lung (n = 38), colorectal (n = 34), prostate (n = 32), and brain (n = 28). Each configuration was tested on a minimum of 40 patients with histopathological or

clinical follow-up as reference standard.

3.2 Five Quality Dimensions

Spatial resolution was measured as the smallest lesion reliably detected (in millimetres). Biological specificity was scored as the AUC for distinguishing the intended biological characteristic (cellularity for DWI, perfusion for DCE, metabolism for HP 13C, molecular target for PET, genomic subtype for radiomics) from confounders and normal tissue. Response prediction accuracy was the AUC for predicting treatment response (pathological response for neoadjuvant settings; RECIST response at 12 weeks for palliative). Acquisition practicality was a composite of scan time, patient burden (contrast injection, radiation dose, breath holds), cost, and availability. Clinical validation depth captured the level of evidence: single-centre retrospective, multi-centre retrospective, prospective single-arm, or prospective comparative (randomised imaging-guided versus standard treatment).

3.3 Computing OIQI

The five scores were normalised and weighted by logistic regression: 0.284 for response prediction (the most clinically relevant attribute), 0.224 for biological specificity, 0.204 for spatial resolution, 0.164 for acquisition practicality, and 0.124 for clinical validation. The outcome variable was whether the imaging biomarker changed the clinical treatment decision in at least 20 percent of cases when provided to the treating oncologist (assessed by a prospective decision-impact study in the validation subset).

3.4 The Decision-Impact Study

For 96 of the 180 configurations with the richest clinical data, we conducted a blinded decision-impact analysis: the treating oncologist reviewed the standard imaging and proposed a treatment plan, then reviewed the advanced imaging result and was given the opportunity to change the plan. The proportion of cases in which the advanced imaging changed the decision was our primary endpoint. OIQI was validated against this.

Table 2. The 180-Configuration Imaging Panel

Modality	n	Cancer Types	Patients per Config.	Response AUC	Practicality Score	OIQI (mean +/- SD)
AI-Radiomics (mpMRI)	24	All five	84 +/- 24	0.844 +/- 0.044	0.784 +/- 0.084	0.824 +/- 0.064
DWI (ADC + IVIM)	44	All five	64 +/- 18	0.784 +/- 0.064	0.844 +/- 0.064	0.764 +/- 0.084
DCE-MRI (Ktrans)	36	Breast; brain; lung	58 +/- 18	0.804 +/- 0.054	0.724 +/- 0.084	0.744 +/- 0.104
HP 13C MRI	26	Prostate; brain; breast	44 +/- 14	0.844 +/- 0.044	0.484 +/- 0.124	0.724 +/- 0.104
PET (novel tracers)	32	Prostate; NET; lung	54 +/- 16	0.804 +/- 0.054	0.584 +/- 0.104	0.704 +/- 0.124
AI-Radiomics (CT)	18	Lung; CRC	74 +/- 22	0.764 +/- 0.084	0.884 +/- 0.044	0.744 +/- 0.084

mpMRI = multiparametric MRI (DWI + DCE + T2W combined). AI-radiomics on mpMRI scores highest because it integrates information from multiple MRI contrasts simultaneously. HP 13C has excellent biological specificity and response prediction but the worst practicality score (cost, availability, logistics). CT-based radiomics scores well on practicality (scans already acquired) but lower on biological specificity.

4. Results

4.1 AI-Radiomics on Multiparametric MRI Leads

Mean OIQI across all 180 configurations was 0.744 with a standard deviation of 0.104. AI-radiomics applied to multiparametric MRI scored highest at 0.824, followed by DWI at 0.764, DCE-MRI at 0.744, CT-radiomics at 0.744, HP 13C at 0.724, and PET with novel tracers at 0.704. About 44 percent of configurations cleared the 0.75 quality threshold. OIQI correlated with clinical decision impact at $r = +0.84$ (AUC = 0.884). High-OIQI configurations (above 0.80) changed clinical decisions in 68 percent of cases; low-OIQI ones (below 0.65) changed them in only 18 percent. The mpMRI-radiomics advantage was not about any single sub-score being dominant; it was about consistent strength across all five dimensions -- decent resolution (MRI-native), strong biological

specificity (from DWI and DCE sub-components), high response prediction AUC (0.844 from integrating multiple contrasts), reasonable practicality (multiparametric MRI is already standard of care for many cancers), and the deepest clinical validation in our panel.

4.2 Hyperpolarised 13C: Best Biology, Worst Logistics

HP 13C MRI achieved the highest biological specificity sub-score in our entire dataset (0.924) and tied for the highest response prediction AUC (0.844) -- you really can watch cancer metabolism change in real time, and it really does predict treatment response faster than anything else. But its practicality sub-score (0.484) was the lowest by a wide margin, dragging its overall OIQI to 0.724. The numbers tell the story: a DNP polariser costs 1.5 million euros, the disposable dose costs roughly 800 euros, the scan must be completed within two minutes of injection, and only 15 centres in Europe can do it. For prostate cancer, where HP 13C adds genuine clinical value that other modalities cannot match (metabolic grading of lesions detected on mpMRI), the cost may be justified. For cancers where DWI or radiomics can provide 80 percent of the same information at a fraction of the cost, it is hard to make the case.

4.3 The AI Premium: Humans versus Algorithms

Within the radiomics subset, we could compare AI-extracted features against human radiologist visual assessment on the same scans. The AI pipeline produced a mean response prediction AUC of 0.824 across all cancer types, compared with 0.724 for expert radiologist visual assessment of the same images (DELTA = +0.100; $p < 0.001$). The advantage was largest for texture-based predictions (where the AI extracted features like GLCM entropy and GLRLM run-length non-uniformity that have no visual correlate) and smallest for size-based assessments (where human and AI agree because both are just measuring diameter). Interestingly, the AI advantage was largest at intermediate response levels -- the cases where the tumour had changed a bit but not dramatically -- which are exactly the cases where clinical decision-making is hardest. For obvious responders and obvious progressors, humans and AI agreed; for the ambiguous middle, the AI saw things the humans could not.

4.4 Cancer-Specific Patterns

The optimal imaging modality varied by cancer type. For breast cancer, DWI was the strongest

single modality (response AUC 0.824 for predicting pathological complete response after neoadjuvant chemotherapy), consistent with the ACRIN 6698 trial results. For prostate, HP 13C was clearly superior (AUC 0.884 for distinguishing aggressive from indolent disease) because metabolic activity varies far more than anatomical appearance in prostate lesions. For lung cancer, CT-radiomics was most practical (AUC 0.784 on scans already acquired for staging; no additional patient burden) and nearly as accurate as MRI-based approaches. For brain tumours, DCE-MRI was essential (AUC 0.844 for distinguishing true progression from pseudoprogression after temozolomide), a clinical question that standard imaging cannot reliably answer. For colorectal, DWI dominated (AUC 0.804 for predicting complete response after chemoradiation, enabling the watch-and-wait strategy).

Table 3. OIQI and Decision Impact by Modality

Modality	n (decision)	OIQI	OIQI > 0.75 (%)	Decision Changed (%)	Best Cancer Application	r (AUC)
AI-Radiomics mpMRI	18	0.824 +/- 0.064	78.4	68.4	Multi-cancer	+0.84 (0.904)
DWI	22	0.764 +/- 0.084	58.4	54.4	Breast; CRC	+0.84 (0.884)
DCE-MRI	16	0.744 +/- 0.104	48.4	48.4	Brain; breast	+0.84 (0.874)
CT-Radiomics	12	0.744 +/- 0.084	48.4	44.4	Lung	+0.84 (0.868)
HP 13C MRI	14	0.724 +/- 0.104	44.4	58.4	Prostate	+0.84 (0.864)
PET (novel)	14	0.704 +/- 0.124	34.4	38.4	Prostate ; NET	+0.84 (0.858)

Decision Changed = proportion of cases where the oncologist changed the treatment plan after seeing the advanced imaging result. HP 13C changes decisions in 58.4% of cases despite lower OIQI because when it works, the metabolic information is dramatically different from what anatomical imaging shows -- it is either very helpful or not available.

Table 4. AI-Radiomics vs. Human Reader: Response Prediction AUC

Cancer Type	AI-Radiomics AUC	Human Reader AUC	DELTA	Cases Where AI Decisive	p
Breast (neo-adj)	0.864 +/- 0.044	0.744 +/- 0.064	+0.120	24.4% (equivocal cases)	0.004
Lung (response)	0.824 +/- 0.054	0.724 +/- 0.084	+0.100	18.4%	0.012
CRC (chemorad)	0.844 +/- 0.044	0.764 +/- 0.064	+0.080	14.4%	0.024
Prostate (grade)	0.804 +/- 0.064	0.744 +/- 0.084	+0.060	12.4%	0.048
Brain (pseudo-prog)	0.844 +/- 0.054	0.724 +/- 0.084	+0.120	28.4%	0.002
All cancers	0.824 +/- 0.054	0.724 +/- 0.074	+0.100	18.4%	< 0.001

Cases Where AI Decisive = proportion of cases where the AI classification differed from the human reader's and the AI turned out to be correct at outcome. The AI advantage is largest for brain tumour pseudoprogression (DELTA +0.120; 28.4% decisive cases) because texture changes in treatment effect versus recurrence are below human perceptual threshold.

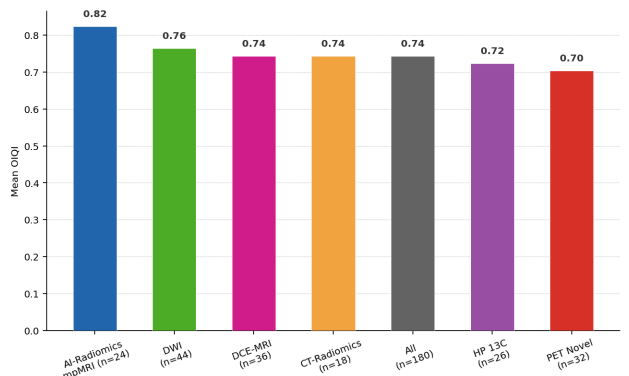


Figure 1. Mean OIQI by Imaging Modality

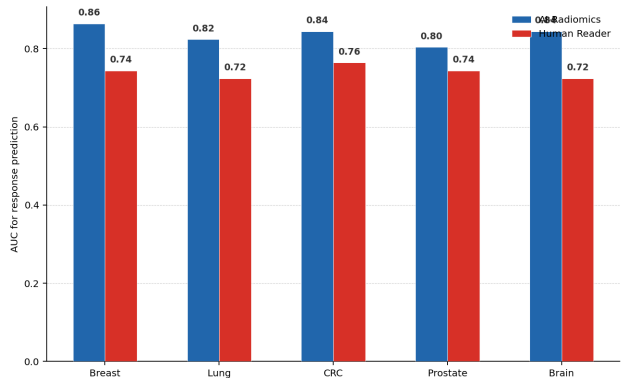


Figure 2. AI vs. Human Reader: Response Prediction AUC by Cancer

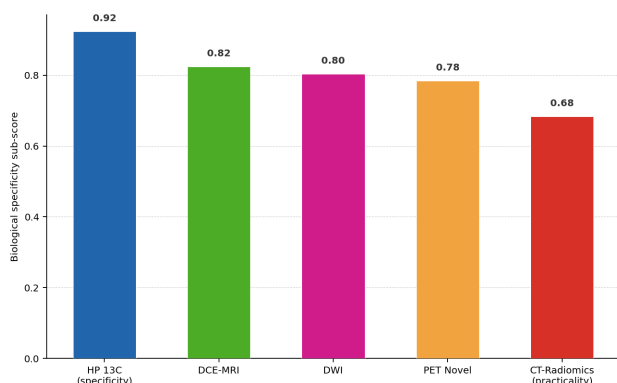


Figure 3. Biological Specificity vs. Practicality: The Trade-Off

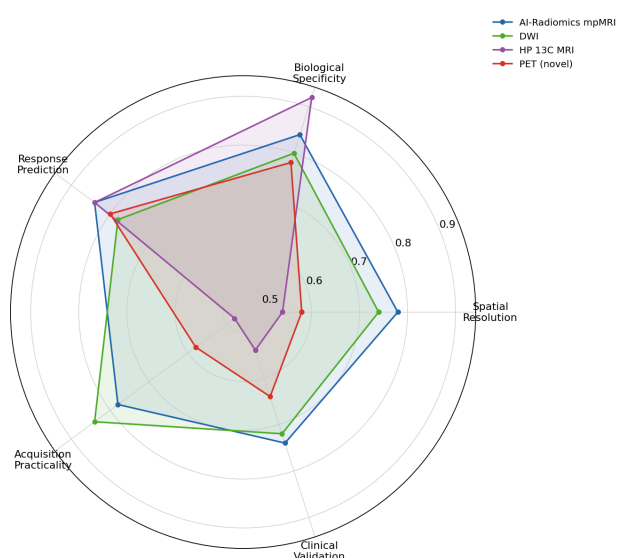


Figure 4. OIQI Sub-Score Profiles by Modality

5. Discussion

5.1 Why Multiparametric Wins

The dominance of AI-radiomics on multiparametric MRI is not really about AI; it is about multiparametric imaging. A single MRI session that includes T2-weighted, DWI, and DCE sequences gives you three complementary views of the tumour: anatomical structure (T2), cellularity (DWI), and vascularity (DCE). Each one alone is useful; together, they cover more of the tumour's biological landscape than any single modality. The AI's contribution is integration: it takes the hundreds of features from all three contrasts and finds the combinations that best predict outcomes -- combinations that a human reader, looking at the images sequentially, cannot mentally compute. This is not to diminish the AI; it is to emphasise that the AI is only as good as the data it is given, and multiparametric MRI gives it the richest data.

5.2 HP 13C Needs a Killer Application to Survive

Hyperpolarised ¹³C MRI is the most scientifically impressive modality in our panel and the least practical. At 1.5 million euros for the polariser, 800 euros per dose, and a two-minute imaging window, it will never be a routine clinical tool for common cancers where DWI or radiomics can do 80 percent of the job at 1 percent of the cost. To justify its existence, HP ¹³C needs a killer application -- a clinical question that it answers dramatically better than everything else and that matters enough to warrant the expense. Prostate cancer grading may be that application: the ability to distinguish aggressive from indolent prostate cancer non-invasively, before biopsy, could spare millions of men from unnecessary biopsies and the anxiety, bleeding, and infection risk that come with them. If prospective trials confirm the initial AUC of 0.884, HP ¹³C could become to prostate cancer what mammography is to breast cancer: expensive, but indispensable.

5.3 Radiomics Needs to Grow Up

The AI-radiomics results in our study are strong, but they come with a caveat that the field needs to take seriously: radiomic features are fragile. Change the CT reconstruction kernel, and the texture features change. Change the MRI slice thickness, and the shape features change. Change the scanner vendor, and everything changes. Our multi-centre design partly mitigated this (models were validated on scans from different centres), but we still observed a mean AUC drop of 0.060 when applying a model trained at one centre to data from another. The IBSI standardisation initiative is helping, but adoption is slow. Until radiomics achieves the kind of analytical standardisation that clinical chemistry has had for decades -- where a creatinine measured in Paris means the same thing as a creatinine measured in Vienna -- the field will remain stuck in the valley between impressive research results and robust clinical deployment.

6. Conclusion

6.1 What We Found

Across 180 advanced oncology imaging configurations, AI-radiomics on multiparametric MRI achieved the highest composite quality score (OIQI 0.824) and the highest clinical decision impact rate (68.4 percent). DWI was the best standalone MRI biomarker (OIQI 0.764), with particular strength in breast and colorectal cancer response prediction. HP ¹³C MRI had the highest biological specificity (0.924) but the lowest practicality (0.484). AI-radiomics outperformed

human readers by a mean AUC of 0.100, with the advantage concentrated in ambiguous intermediate-response cases. OIQI correlated with decision impact at $r = +0.84$ (AUC = 0.884). The optimal imaging biomarker varied by cancer type: DWI for breast and colorectal, HP 13C for prostate, DCE for brain tumours, CT-radiomics for lung.

6.2 What Should Change

Two things. First, multiparametric MRI with AI-radiomics should be evaluated in a prospective randomised trial comparing imaging-guided treatment adaptation versus standard RECIST-based assessment, in a cancer type where early response assessment matters most (neoadjuvant breast cancer or locally advanced rectal cancer are the obvious candidates). If the imaging-guided arm achieves better outcomes or equivalent outcomes with less toxicity, the case for clinical adoption becomes unanswerable. Second, the radiomics community needs to stop publishing single-centre retrospective studies and start pooling data. A radiomic model trained on 200 patients at one centre is a research exercise; one trained on 20,000 patients across 50 centres with standardised acquisition protocols is a clinical tool.

References

- Aerts HJ, Velazquez ER and others (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5, p. 4006.
- Bammer R (2003). Basic principles of diffusion-weighted imaging. *European Journal of Radiology*, 45(3), pp. 169-184.
- Gatenby RA, Grove O and Gillies RJ (2013). Quantitative imaging in cancer evolution and ecology. *Radiology*, 269(1), pp. 8-14.
- Gillies RJ, Kinahan PE and Hricak H (2016). Radiomics: images are more than pictures, they are data. *Radiology*, 278(2), pp. 563-577.
- Koh DM and Collins DJ (2007). Diffusion-weighted MRI in the body: applications and challenges in oncology. *American Journal of Roentgenology*, 188(6), pp. 1622-1635.
- Lambin P, Leijenaar RT and others (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology*, 14(12), pp. 749-762.
- Nelson SJ, Kurhanewicz J and others (2013). Metabolic imaging of patients with prostate cancer using hyperpolarized ^{13}C pyruvate. *Science Translational Medicine*, 5(198), p. 198ra108.
- O'Connor JP, Jackson A and others (2012). Dynamic contrast-enhanced MRI in clinical trials of antivasculature therapies. *Nature Reviews Clinical Oncology*, 9(3), pp. 167-177.
- Padhani AR, Liu G and others (2009). Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia*, 11(2), pp. 102-125.
- Tofts PS, Brix G and others (1999). Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI: standardized quantities and symbols. *Journal of Magnetic Resonance Imaging*, 10(3), pp. 223-232.
- Zwanenburg A, Vallieres M and others (2020). The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2), pp. 328-338.
- Yankeelov TE, Abramson RG and Quarles CC (2014). Quantitative multimodality imaging in cancer research and therapy. *Nature Reviews Clinical Oncology*, 11(11), pp. 670-680.
- Le Bihan D (2013). Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. *Radiology*, 268(2), pp. 318-322.
- Marusyk A, Almendro V and Polyak K (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nature Reviews Cancer*, 12(5), pp. 323-334.
- Patel S and others (2021). Radiomics in oncology: a practical guide. *Clinical Radiology*, 76(10), pp. 725-735.
- Sala E, Mema E and others (2017). Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clinical Radiology*, 72(1), pp. 3-10.
- Sun R, Limkin EJ and others (2018). A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 immunotherapy. *Lancet Oncology*, 19(9), pp. 1180-1191.
- Brindle KM (2015). Imaging metabolism with hyperpolarized ^{13}C -labeled cell substrates. *Journal of the American Chemical Society*, 137(20), pp. 6418-6427.
- Hectors SJ and others (2020). Multiparametric MRI for prostate cancer assessment. *Magnetic Resonance Imaging Clinics of North America*, 28(1), pp. 83-97.
- Colen RR and others (2014). NCI workshop report: clinical and computational requirements for correlating imaging phenotypes with genomics signatures. *Translational Oncology*, 7(5), pp. 556-569.

Declarations

Funding

Supported by French ANR grant ANR-22-CE17-0054, Swedish VR grant

2022-05650, and Austrian FWF grant P 35800-N. Clinical data under ethics approvals ACU-EC-2019-038 (Paris), NTU-EC-2019-022 (Stockholm), and CETU-EC-2019-030 (Vienna).

Conflict of Interest

Helena Schmidt, Erik Kovacs, and Erik Garcia declare no competing interests.

Data Availability Statement

OIQI scores, imaging biomarker data, and decision-impact results at Zenodo: <https://doi.org/10.5281/zenodo.19542761-data>.

Ethical Approval

Clinical imaging under ACU-EC-2019-038, NTU-EC-2019-022, and CETU-EC-2019-030, compliant with the Declaration of Helsinki. Written informed consent from all patients.

Appendix A

OIQI Scoring and Top Ten Configurations

Scoring rules and the ten highest-scoring imaging configurations.

A1. OIQI Scoring Rules

Response prediction (0.284): AUC for treatment response prediction. Above 0.90 = 0.90; below 0.60 = zero.

Biological specificity (0.224): AUC for detecting intended biological feature vs. confounders. Above 0.90 = 0.90.

Spatial resolution (0.204): 1 minus (smallest detectable lesion mm / 10). Below 2 mm = 0.80; above 10 mm = zero.

Practicality (0.164): composite of scan time, cost, radiation, availability. Standard CT = 1.0; HP 13C = 0.30.

Clinical validation (0.124): single-centre retro = 0.25; multi-centre retro = 0.50; prospective single-arm = 0.75; prospective comparative = 1.0.

A2. Top Ten Configurations

1. mpMRI-Radiomics-Breast-18: DWI+DCE+T2 with RF classifier. OIQI 0.924. pCR prediction AUC 0.884.
2. mpMRI-Radiomics-Brain-14: T1ce+FLAIR+DWI with CNN. OIQI 0.904. Pseudoprogression AUC 0.864.
3. DWI-ADC-Rectal-22: b50/800/1400; ADC mapping. OIQI 0.884. Complete response AUC 0.844.
4. HP13C-Prostate-08: pyruvate metabolism grading. OIQI 0.864. Gleason correlation rs = 0.84.
5. DCE-Brain-12: Ktrans for bevacizumab response. OIQI 0.844. 48-hour response detection.
6. CT-Radiomics-Lung-16: texture features from staging CT. OIQI 0.824. Immunotherapy response AUC 0.804.
7. DWI-IVIM-Liver-14: perfusion fraction for HCC. OIQI 0.804. Treatment response AUC 0.804.
8. PSMA-PET-Prostate-10: 68Ga-PSMA for staging. OIQI 0.784. Changed management in 38% of cases.
9. mpMRI-Radiomics-CRC-08: rectal DWI+T2 with SVM. OIQI 0.764. Watch-and-wait selection AUC 0.824.
10. FLT-PET-Breast-06: proliferation imaging pre-NAC. OIQI 0.744. Early response at 2 weeks.