Communications Medicine

https://doi.org/10.1038/s43856-025-01246-2

Article in Press

Clinically feasible liver tumour cell size measurement through histology-informed in vivo diffusion MRI

Received: 5 December 2024

Accepted: 5 November 2025

Cite this article as: Grussu, F., Grigoriou, A., Bernatowicz, K. *et al.* Clinically feasible liver tumour cell size measurement through histology-informed in vivo diffusion MRI. *Commun Med* (2025). https://doi.org/10.1038/s43856-025-01246-2

Francesco Grussu, Athanasios Grigoriou, Kinga Bernatowicz, Marco Palombo, Irene Casanova-Salas, Daniel Navarro-Garcia, Ignasi Barba, Sara Simonetti, Garazi Serna, Carlos Macarro, Anna Kira Voronova, Valezka Garay, Juan Francisco Corral, Marta Vidorreta, Pablo García-Polo García, Xavier Merino, Richard Mast, Núria Rosón, Manuel Escobar, Maria Vieito, Rodrigo Toledo, Paolo Nuciforo, Joaquin Mateo, Elena Garralda & Raquel Perez-Lopez

Clinically feasible liver tumour cell size measurement through histology-informed *in vivo* diffusion MRI

Francesco Grussu[1]*, Athanasios Grigoriou[1,2], Kinga Bernatowicz[1], Marco Palombo[3,4], Irene Casanova-Salas[1], Daniel Navarro-Garcia[1], Ignasi Barba[1,5], Sara Simonetti[1], Garazi Serna[1], Carlos Macarro[1], Anna Kira Voronova[1,2], Valezka Garay[6], Juan Francisco Corral[7,8], Marta Vidorreta[9], Pablo García-Polo García[10], Xavier Merino[7,8], Richard Mast[7,8], Núria Rosón[7,8], Manuel Escobar[7,8], Maria Vieito[11], Rodrigo Toledo[1], Paolo Nuciforo[1], Joaquin Mateo[11], Elena Garralda[11], Raquel Perez-Lopez[1]*

*Joint corresponding authors. Email: fgrussu@vhio.net (F.G.), rperez@vhio.net (R.P.L.)

Address: Radiomics Group, Vall d'Hebron Institute of Oncology (VHIO), Carrer de Saturnino Calleja 11-13, 08035 Barcelona (Spain).

[1]Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain [2]Department of Biomedicine, Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain [3]Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Cardiff, United Kingdom [4]School of Computer Science and Informatics, Cardiff University, Cardiff, UK [5]University of Vic - Central University of Catalonia (UVic-UCC), Vic, Spain [6]PET/MR Unit, CETIR-ASCIRES, Barcelona, Spain [7]Department of Radiology, Hospital Universitari Vall d'Hebron, Barcelona, Spain [8]Institut de Diagnòstic per la Imatge (IDI), Barcelona, Spain [9]Siemens Healthineers, Madrid, Spain [10]GE HealthCare, Madrid, Spain [11]Medical Oncology Service, Vall d'Hebron Barcelona Hospital Campus, Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain

Abstract

Background: Innovative diffusion Magnetic Resonance Imaging (MRI) models enable the non-invasive measurement of cancer biological properties *in vivo*. However, while cancers frequently spread to the liver, models tailored for liver application and easy to deploy in the clinic are still sought. We fill this gap by delivering a practical, clinically-viable framework for liver tumour diffusion imaging, informing its design through histology.

Methods: We compare MRI and histological data from mice and cancer patients, namely: MRI and hemaotxylin-eosin (HE) stains from N=7 fixed mouse livers; MRI of N=38 patients suffering from liver solid tumours, N=18 of whom with HE biopsies. We study five diffusion models, ranking them according to a total MRI-histology correlation score. Afterwards, we test metrics from the top-ranking model on our cohort, assessing their sensitivity to cell proliferation (Ki-67 staining, N=10), evaluating their association with tumour volume (N=140 tumours), and comparing them across primary cancer types.

Results: We select a dMRI signal model of restricted intra-cellular diffusion with negligible extra-cellular contributions, which maximises radiological-histological correlations (total score: 0.625). The model provides cell size and density estimates that i) correlate with histology (e.g., for cell size: r=0.44, p=0.029), ii) are associated to Ki-67 cell proliferation (for MRI cell density: r=0.80, p=0.006) and tumour volume (r=0.40, $p<10^{-5}$ for tumour volume regression), and iii) that distinguish melanoma (N=8) from colorectal cancer (N=13) (p=0.011 for intra-cellular fraction).

Conclusions: Our biologically meaningful approach may complement standard-of-care radiology, and become a new tool for enhanced cancer characterisation in precision oncology.

Plain language summary

Magnetic Resonance Imaging (MRI) is an imaging tool used frequently to detect and monitor malignant tumours. Notably, the latest MRI technology enables physicians to measure not only the size of a tumour, but even properties related to the cells it contains. This information, like number or size of cells can help oncologists to choose the best treatment. Obtaining this information using this imaging tool in tumors that have spread to the liver, remained challenging. Here we present an MRI method developed specifically to capture the biology of liver tumour cells. We base it on mathematical models that use information gathered from microscopy images – known as histology, where individual cells are visible – to guide the analysis of the scan. We show that the method quantifies tumour cell size and density from a simple, clinically feasible MRI scan and propose this technique for oncological applications.

Introduction

The clinical use of Magnetic Resonance Imaging (MRI) in cancer is based on revealing the presence of tumours within an area of interest, and on measuring their size. Nonetheless, MRI also offers the possibility of obtaining estimates of biological properties within each pixel making up the tumour image. This approach, known as quantitative MRI (qMRI)[1], involves the acquisition of multiple MRI contrasts, which are then analysed jointly with a mathetmatical model. qMRI provides promising metrics, which could become quantitative biomarkers complementing the qualitative assessment by the expert radiologist[2]. Among existing qMRI approaches, diffusion MRI (dMRI) sensitises the MRI signal to water diffusion in biological tissues[1,3,4]. Since the patterns of diffusion are influenced by the microenvironment where diffusion takes place (e.g., by

obstacles such as cell membranes), dMRI ultimately enables the per-voxel estimation of properties related to the size and density of cells[5–7], bridging the gap between radiology and histology.

dMRI has shown promise in multiple areas, including brain[5], spinal cord[8], prostate[6], breast[9], or liver imaging[10,11]. The liver is a frequent site for cancer metastasisation[12], and liver tumours are common targets for treatment response assessment in oncology. However, current image-based response criteria such as RECIST[13] have limitations, in that they rely on MRI or computed tomography merely to measure tumour size, without accounting for changes under therapy at the cellular level. Novel dMRI metrics could instead enable the non-invasive characterisation of cancer microenvironments, shedding light on the composition of tumours that cannot be biopsied. The new readouts could also provide information on tumour heterogeneity, relevant to fight treatment resistance[14,15], and better stratify patients for personalised treatment planning, reducing sample sizes in clinical trials and ultimately improving patient outcomes[16].

In practice, the estimation of the biological properties through dMRI is made possible by mathematical models, which are used to link the observed image contrasts to the biological properties themselves[17–20]. However, despite their promise, advanced liver dMRI models are relegated almost exclusively to research contexts, and seldom used in real-world clinical settings. This is due to two main reasons: i) to the high number of dMRI images (and hence long scan time) required to support the fitting of such models, and ii) to the requirement for specialised dMRI acquisitions[21], beyond those made available by manufacturers in the scanner console.

We aim to fill this clinical-research gap by delivering a practical liver dMRI framework that is truly feasible in hospital settings, i.e., on 1.5T or 3T systems, with scan time that does not exceed 15 minutes, and using vendor-provided dMRI acquisitions. With this objective in mind, we

embraced the latest "histology-informed" dMRI development paradigm, which is based on informing signal model design[22] with co-localised histology. The framework has shown promise in delivering dMRI approaches with unprecedented fidelity to cellular architecture, enhancing the specificity or medical imaging towards the underlying tissue biology[23–25].

Following preliminary abstract dissemination[26], in this article we compared candidate, state-of-the-art approaches using a rich data set of dMRI scans and hematoxylin and eosin (HE)-stained images, from excised mouse livers and patients' liver biopsies. We used these unique data to select the approach maximising the correlation of radiological and histological estimates of metrics such as cell size, corroborating results with computer simulations. To our knowledge, this is the first time that the approach has been used on a rich data set of dMRI scans of liver tumours in patients *in vivo*, featuring multiple types of malignancies. The proposed approach is based on a model of restricted, intra-cellular (IC) diffusion with negligible extra-cellular (EC) signal contributions, which is fitted to highly diffusion-weighted (DW) images. The model provides histologically-meaningful estimates of cell size and density, which are also shown to be associated to cell proliferation within liver tumours, as well as to the tumour size. These results suggest that our dMRI framework may provide histologically-relevant descriptors of tumour characteristics. These may complement standard-of-care readouts in research and clinical practice, equipping oncologists with new tools for precision oncology in hospital settings.

Methods

dMRI models

Common biophysical body dMRI signal models[6,17,18,27–29] describe the signal as arising from three, non-exchanging proton pools: vascular water; restricted, intra-cellular water; hindered,

extra-cellular, extra-vascular water. The dMRI signal for a PGSE measurement at b-value b, gradient duration/separation δ/Δ, and echo time TE is

$$s = s_0\left(f_V\, e^{-\frac{TE}{T2_V}} a_V + (1-f_V)\left(f_I\, e^{-\frac{TE}{T2_I}} a_I + (1-f_I)\, e^{-\frac{TE}{T2_E}} a_E\right)\right). \quad (1)$$

Above, s_0 is the apparent proton density, f_V is the voxel vascular signal fraction, f_I is the tissue intra-cellular signal fraction, $T2_V/T2_I/T2_E$ and $a_V/a_I/a_E$ are compartment-wise T2 and diffusion-weighting factors. a_V captures intra-voxel incoherent motion (IVIM) effects[30]. *In vivo*, the IVIM vascular ADC ranges approximately[31] in [15; 60] $\mu m^2\ ms^{-1}$. For this reason, for b > 100 s/mm^2, the vascular signal vanishes ($a_V \approx 0$), and Eq. (1) reduces to[17]

$$s = s_0\,(1-f_V)\left(f_I\, e^{-\frac{TE}{T2_I}} a_I + (1-f_I)\, e^{-\frac{TE}{T2_E}} a_E\right). \quad (2)$$

A common model for a_I in Eq. 2 is that of restricted diffusion within spheres of diameter L[17,18]:

$$a_I = e^{-b\, ADC_I\left(\delta,\Delta,D_{0,I},L\right)}, \quad (3)$$

where

$$ADC_I = \frac{2}{D_{0,I}\delta^2(\Delta-\delta/3)} \sum_{m=1}^{\infty} \frac{\alpha_m^{-4}}{\alpha_m^2 R^2 - 2} \left(2\delta - \frac{2+e^{-\alpha_m^2 D_{0,I}(\Delta-\delta)} - 2e^{-\alpha_m^2 D_{0,I}\delta} - 2e^{-\alpha_m^2 D_{0,I}\Delta} + e^{-\alpha_m^2 D_{0,I}(\Delta+\delta)}}{\alpha_m^2 D_{0,I}}\right) \quad (4)$$

is the Gaussian phase distribution approximation of the intra-cellular ADC[32]. Above, α_m is the *m*-th root of $\alpha_m R\, J'_{\frac{3}{2}}(\alpha_m R) - \frac{1}{2} J_{\frac{3}{2}}(\alpha_m R) = 0$, $J_{\frac{3}{2}}(x)$ is the Bessel function of the first kind and order 3/2, and $J'_{\frac{3}{2}}(x)$ its first-order derivative ($J'_{\frac{3}{2}}(x) = \frac{d}{dx} J_{\frac{3}{2}}(x)$). ADC_I depends on the intrinsic cytosol diffusivity $D_{0,I}$ and on the cell size $L = 2R$ (R: radius; L: diameter). Noting that dMRI-derived L represents a volume-weighted mean cell size statistics[7,33], we will refer to it as *volume-weighted cell size* (vCS).

Conversely, the extra-cellular, extra-vascular signal may be described in terms of hindered diffusion in a tortuous space[17,28,34]:

$$a_E(b,\Delta) = e^{-b\, ADC_E\left(\Delta, D_{E,\infty}, \beta\right)}, \quad (5)$$

with

$$ADC_E\left(\Delta, D_{E,\infty}, \beta\right) = D_{E,\infty} + \frac{\beta}{\Delta}. \quad (6)$$

In Eq. 6, $D_{E,\infty}$ is the asymptotic[28] ADC_E for $\Delta \to \infty$.

The 5 implementations of the two-compartment model

We investigated 5 implementations of Eq. 2, divided into two families. The first family includes models that do not make assumptions on which of ADC_I/ADC_E is higher:

i. *Diff-in-exTD*: the most general model, relying on the full expression of ADC_E in Eq. 6;

ii. *Diff-in-ex*: a simpler implementation of *Diff-in-exTD* that neglects extra-cellular TD ($\beta = 0$ in Eq. 6).

In the second family of models, we constrain $ADC_E > ADC_I$. It includes

i. *Diff-in-exTDFast*: equivalent to *Diff-in-exTD*, with the lower bound for $D_{E,\infty}$ ensuring $ADC_E > ADC_I$ for any L.

ii. *Diff-in-exFast*: equivalent to *Diff-in-ex*, but again ensuring that $ADC_E > ADC_I$ for any L.

iii. *Diff-in*: a model where the extra-cellular signal is negligible compared to the intra-cellular one, due to ADC_E being much larger than ADC_I, so that Eq. 2 simplifies to

$$s = s_0 (1 - f_V) f_I e^{-\frac{TE}{T2_I}} a_I(b, \delta, \Delta, D_{0,I}, L). \quad (7)$$

In all models we used $T2_I \approx T2_E \doteq T2_T$, given the challenge of resolving accurately multiple T2 constants[18,35].

Fitting

We fitted the 5 models using custom-written Python routines, based on objective function minimisation initialised by a grid search. The objective function was $f_{obj} = -ln(\lambda)$, where λ is the offset-Gaussian likelihood[36]. Fitting provides estimates of vCS and voxel IC signal fraction

$$F = (1 - f_V)\, f_I. \quad (8)$$

We also combined vCS and F into a cell density per unit volume[18]

$$CD = \frac{F}{vCS^3}. \quad (9)$$

Mouse data acquisition

Animals

We obtained data from 7 fixed livers of NOD.Cg-Prkdc^{scid} IL2rg^{tm1WjI}/SzJ mice. These mice develop various liver pathologies when implanted with patient-derived tumour samples. They are routinely used in patient-derived xenograft studies, and represent an ideal setting to validate the performances of new MRI methods for liver imaging. All experimental protocols were

approved and monitored by the Vall d'Hebron Institute of Research Animal Experimentation Ethics Committee (CEEA; registration number 68/20) in accordance with relevant local and EU regulations. Animals were housed at the specific pathogen-free barrier area of the Vall d'Hebron Institute of Oncology following established institutional procedures. A study protocol was defined internally at institutional level, with no registration in external platforms.

We studied six livers from mice implanted with cells derived from biopsies of prostate cancer patients, as part of an ongoing study, plus an additional liver from a mouse without any implantation. We implanted one tumor biopsy core with growth factor-enriched Matrigel (Corning) subcutaneously in the flank of each mice under isoflurane inhalant anaesthesia (3-4% for induction and 1-3% for maintenance). Mice were all males, and they were implanted when 6 weeks old. Author I.C.S. was aware of the status of the mice (implanted vs not-implanted). We derived tissue from the following biopsies: iliac bone metastasis biopsy (metastatic castration-resistant prostate cancer, presenting with bone metastasis and Gleason score 3+4 adenocarcinoma); prostate biopsy (patient with metastatic hormone-sensitive prostate cancer, presenting with bone metastasis and Gleason score 5+4 adenocarcinoma); two liver biopsies (patient with metastatic castration-resistant prostate cancer, presenting with bone and visceral metastasis and Gleason score 4+4 acinar adenocarcinoma; patient with metastatic hormone-sensitive prostate cancer, presenting with bone and liver metastasis and Gleason score 4+4 adenocarcinoma). After implantation, we measured tumour size using calipers and monitored mouse weight weekly, sacrificing animals by cervical dislocation under general anaesthesia (isoflurane inhalant, 3-4%) when tumour volume exceeded 2000 mm^3, to reduce distress. No adverse events were observed in the N=7 animals used in this study. We collected the livers, fixed them overnight in formalin, and transferred them to phosphate-buffered saline (PBS) solution.

205

206 *MRI*

207 We scanned livers on a 9.4T Bruker Avance system at room temperature. Livers were
208 tightened with sewing thread to a histology cassette and placed into a Falcon® tube, filled with
209 PBS solution. A 1-channel birdcage coil was used (excitation/reception). The protocol included a
210 T2-weighted fast spin echo sequence (resolution: 144 μm × 144 μm × 2.216 mm) and PGSE dMRI
211 (Fig. S27; TR = 2700 ms; resolution: 386 μm × 386 μm; matrix size: 86 × 86; 4 slices, 2.216 mm-
212 thick, NEX = 1). The protocol featured: δ = 10 ms, Δ = {15, 30} ms, 10 linearly spaced b-values
213 for each Δ (minimum/maximum nominal *b*: 0/2800 s/mm^2). DW images corresponding to Δ = 15
214 ms were acquired at each of TE = {31, 45, 65} ms, and to Δ = 30 ms at each of TE = {45, 65} ms.
215 We i) denoised dMRI scans with Marchenko-Pastur Principal Component Analysis (MP-PCA)
216 Python denoising[37] (kernel: 7×7×3), ii) mitigated Gibbs ringing (MrTrix3 local sub-voxel shift
217 method[38]), and iii) corrected temporal signal drifts by assessing signal changes in a PBS solution
218 region, accounting for TE (PBS T2: 500 ms).

219 Finally, we fitted the *Diff-in-exTD*, *Diff-in-exTDFast*, *Diff-in-ex*, *Diff-in-exFast* and *Diff-*
220 *in* models voxel-by voxel (tissue parameter bounds: [0; 1] for f_I; [0.8; 2.6] μm^2 ms^{-1} for $D_{0,I}$; [8;
221 40] μm for *vCS*; [0.8; 2.6] μm^2 ms^{-1} for $D_{E,\infty}$ in models *Diff-in-ex* and *Diff-in-exTD* and [1.75; 2.6]
222 μm^2 ms^{-1} in models *Diff-in-exFast* and *Diff-in-exTDFast*; [0; 10] μm^2 for β in models *Diff-in-ex-*
223 *TD* and *Diff-in-exTDFast*). For fitting, we fixed f_V and $T2_T$ to values obtained through a a two-
224 pool vascular-tissue model[39] (fitting bounds: [0; 1] for f_V; [5; 80] ms for $T2_T$). Fitting was
225 performed i) on all images with b > 1000 s/mm^2 (suppressing vascular signals, referred to as *fitting*
226 *on whole image set*); ii) on b > 1000 s/mm^2 images (*high b-value fitting*). In our *ex vivo* data, the

vascular signal captures PBS solution contamination (PBS ADC: roughly 2.4 μm^2 ms^{-1}). For this reason, we adopted a b-value threshold of 1000 s/mm^2 to achieve acceptable PBS signal suppression. We used instead a minimum b-value of 1800 s/mm^2 for high b-value fitting (minimising extra-cellular contributions), given the reduction in intrinsic tissue diffusivity expected *ex vivo*.

For comparison, we computed ADC and apparent diffusion excess kurtosis K by fitting

$$s = s_0\, e^{-b\,ADC + \frac{1}{6}K(b\,ADC)^2} \quad (10)$$

to DW images acquired at TE = 45 ms, Δ = 30 ms, with in-house Python code.

Histology

After MRI, samples underwent histology. We cut two 4 μm-thick histological sections at known position, stained them with HE, and digitised them (Hamamatsu C9600-12 slide scanner; 0.227 μm resolution). An experienced pathologist (S.S.) inspected images qualitatively. We then processed them with the automatic cell detection tool of QuPath[40], obtaining per-cell area A and diameter $l = \sqrt{\frac{4}{\pi}A}$. Afterwards, we split images into 386 μm × 386 μm patches (matching the MRI resolution), computing patch-wise histological volume-weighted cell size vCS_{histo}, intra-cellular area fraction F_{histo} and cell density per unit area CD_{histo}[33]. vCS_{histo}, defined as

$$vCS_{histo} = \left(\frac{<l^7>}{<l^3>}\right)^{\frac{1}{4}}, (11)$$

is a more accurate counterpart of dMRI cell size than the arithmetic mean[7,33] $aCS_{histo} = < l >$. We accounted for biases coming from: i) estimating the size of 3D objects from 2D views (bias 1), ii) tissue shrinkage (bias 2), by rescaling vCS_{histo} and CD_{histo}. The final vCS_{histo} estimate was 1.4616 times larger than the value obtained from direct image processing (1.4616 = 1.2732×1.148; 1.2732, derived from the theory of spherical caps, accounts for bias 1; 1.148 accounts for bias 2, and corresponds to a plausible shrinkage of 12.9% following dehydration, clearing and paraffin embedding[41]). The final CD_{histo} estimate was 1.318 times smaller than the value derived from direct image processing, since 1 mm^2 of shrunk tissue corresponds to 1.148×1.148 mm^2 = 1.318 mm^2 of unprocessed tissue (plausible shrinkage 12.9%[41]). Lastly, we co-registered histological maps to MRI[33] using DiPy[42].

Human data acquisition

Cohort

We obtained data from patients suffering from advanced solid tumours, recruited for an ongoing imaging study approved by the Vall d'Hebron University Hospital (VHUH) Ethics committee (PR(AG)29/2020). Patients provided informed written consent to participate in the imaging study. We included 38 patients with liver malignancies, either primary or metastatic, being screened as potentially eligible candidates for various phase I clinical trials at the VHUH. Patients were scanned scanned with either a 1.5T or 3T system. Approximately one week after

MRI, a biopsy from one of the imaged liver tumours was obtained from 18 patients. The histological material was stained for HE, and in 12 cases also through Ki67 IHC, demonstrating cell proliferation. The eligibility criteria to enter this imaging study were: to have liver malignancies of known origin, previously confirmed on a Computer Tomography scan (diameter >10 mm); to meet all standard institutional requirements for having an MRI scan. In those cases where a biopsy was also collected, all standard institutional, requirements for having a biopsy had also to be met.

MRI

We imaged patients at the level of the abdomen. We scanned 11 patients on a 1.5T Siemens Avanto scanner using the vendor 18-channel body coil for detection, and 27 patients on a 3T GE SIGNA Pioneer scanner, using the vendor 48-channel torso coil for signal reception, with 32 channels enabled for detection. The 1.5T system was located at the Vall d'Hebron Hospital radiology department, whereas the 3T at the CETIR Clinic of Barcelona. MRI scans were acquired between October 2020 and June 2023. No software or hardware upgrades were made to the MRI systems during the acquisition of the data. Note that in this study we do not employ any clinical outcome measure related to patient treatment or to the participation in clinical trials.

In the 1.5T Siemens system, the protocol included a T2-weighted fast spin echo scan (resolution: $1.4 \times 1.4 \times 5$ mm^3; 32 slices; TR = 4500 ms; TE = 82 ms; echo train length: 29; NEX = 8; GRAPPA = 2) and fat-suppressed DW TRSE (Fig. S27) EPI (dMRI scan time: 16 minutes). It featured: resolution: $1.9 \times 1.9 \times 6$ mm^3; 32 slices; TR = 7900 ms; bandwidth 1430 Hz/pixel; averaging of 3 orthogonal diffusion directions × 2 signal averages (effective NEX = 6); GRAPPA

factor of 2; 6/8 partial Fourier imaging. The dMRI protocol consisted of b = {0, 50, 100, 400, 900, 1200, 1600} s/mm^2, each for TE = {93, 105, 120} ms. One additional image (b = 0 s/mm^2; TE = 93 ms) was acquired with reversed phase encoding polarity. The gradient timings (Fig. S13B) were: δ_1 = 8.9 ms, δ_2 = 17.6 ms, δ_3 = 20.4 ms, δ_4 = 6.0 ms, $\Delta_{1,2}$ = 17.4 ms and $\Delta_{1,4}$ = 63.9 ms when TE = 93 ms; δ_1 = 13.2 ms, δ_2 = 19.3 ms, δ_3 = 24.8 ms, δ_4 = 7.7 ms, $\Delta_{1,2}$ = 21.7 ms and $\Delta_{1,4}$ = 74.2 ms when TE = 105 ms; δ_1 = 18.9 ms, δ_2 = 21.0 ms, δ_3 = 30.5 ms, δ_4 = 9.5 ms, $\Delta_{1,2}$ = 27.5 ms and $\Delta_{1,4}$ = 87.5 ms when TE = 120 ms. The b-value is

$$b = \gamma^2 G^2 \left(\delta_1^2 (\Delta_{1,2} - \delta_1) + \frac{2}{3}(\delta_1 + \delta_2)^3 + (\delta_1 + \delta_2 - \delta_3)^2 (\Delta_{1,4} - \Delta_{1,2} - \delta_2 - \delta_3) \right). \quad (12)$$

Converely, the protocol implemented on the 3T GE scanner included a respiratory-gated T2-weighted fast spin echo scan (resolution: 1.4 × 1.4 × 6 mm^3; 32 slices; TR = 4615 ms; TE = 52.86 ms; echo train length: 16) and respiratory-gated, fat-suppressed PGSE (Fig. S27) EPI (dMRI scan time: 16 minutes). It featured: resolution: 2.4 × 2.4 × 6 mm^3; 32 slices; TR = 6000 ms; bandwidth 1953 Hz/pixel; averaging of 3 orthogonal diffusion directions × 2 signal averages (effective NEX = 6); ASSET factor of 2. The dMRI protocol consisted of b = {0, 50, 100, 400, 900, 1200, 1500} s/mm^2, each for TE = {75, 90, 105} ms. The gradient timings (Fig. S17A) were: gradient duration δ = {0.0, 3.9, 5.2, 9.2, 15.0, 18.2, 21.0} ms for TE = 75 ms, δ = {0.0, 3.9, 5.2, 9.2, 13.0, 15.8, 18.5} ms for TE = 90 ms and 105 ms; gradient separation Δ = {0.0, 27.8, 29.0, 33.0, 28.7, 31.8, 34.7} ms for TE = 75 ms and Δ = {0.0, 27.8, 29.0, 33.0, 37.0, 39.6, 42.3} ms for TE = 90 ms and TE = 105 ms.

dMRI post-processing consisted of slice-wise Python MP-PCA denoising (kernel: 5×5)[37]; MRTrix3 Gibbs unringing[38]; motion correction via affine co-registration[43]; FSL distortion correction[44] (1.5T data only). An experienced radiologists (R.P.L.) segmented tumours on the T2-w scan, enabling per-patient tumour volume computation. Afterwards, we warped the tumour mask to dMRI using ANTs[45] non-linear co-registration. Two researchers independently reviewed the co-registered masks, and manually edited them to correct for registration inaccuracies (F.G. and K.B.). Lastly, we fitted the 5 dMRI models, fixing again f_V and $T2_T$ to previously computed values[39] (fitting bounds: [0; 1] for f_V; [20; 140] ms for $T2_T$; [0; 1] for f_I; [0.8; 3.0] μm^2 ms^{-1} for $D_{0,I}$; [8; 40] μm for *vCS*; [0.8; 3.0] μm^2 ms^{-1} for $D_{E,\infty}$ in models *Diff-in-ex* and *Diff-in-exTD*, and [1.75; 3.0] μm^2 ms^{-1} in models *Diff-in-exFast* and *Diff-in-exTDFast*; [0; 10] μm^2 for β in models *Diff-in-ex-TD* and *Diff-in-exTDFast*).

We fitted the 5 dMRI models i) on images acquired at a b-value b > 100 s/mm^2, to suppress vascular signals (*fitting to the whole image set*); ii) to b > 900 s/mm^2 images, to also minimize extra-cellular contributions (*high b-value fitting*). For scans performed on the 1.5T Siemens system: i) we used $\Delta_{1,2} + \delta_2$ in place of Δ in Eq. 6 (Fig. S13B), ii) we replaced Eq. 4 with a numerical implementation of restricted diffusion within spheres, based on Radial Basis Function interpolation of synthetic signals generated for DW-TRSE with Monte Carlo simulations[46].

For both scanners, we also computed ADC and excess kurtosis K by fitting Eq. 10 on b > 100 s/mm^2 images (shortest TE), with in-house Python code.

Histology

We performed ultrasound-guided biopsies of one liver tumour at the Barcelona Vall d'Hebron University Hospital (Spain). Biopsies were obtained approximately one week after dMRI. In two patients, the biopsy was obtained after receiving immunotherapy as part of a phase I trial. In those two cases, an additional dMRI scan was also acquired after starting treatment, immediately before the biopsy. The biological material underwent standard processing, HE staining and Ki67 IHC, and final digitalisation on a Hamamatsu C9600-12 slide scanner (resolution: 0.454 μm). An experienced pathologist (S.S.) assessed the images and drew a region-of-interest (ROI) outlining tumour tissue and excluding non-tumour meterial. In parallel, an experienced radiologist (R.P.L.) inspected ultrasound and MR images, outlining the biopsied tumour on the latter. We processed HE data with QuPath and computed per-biopsy vCS_{histo}, F_{histo} and CD_{histo} as previously described. Additionally, we also computed the fraction of tumour labelled for Ki67 (referred to as F_{Ki67}) with in-house routines, in those cases where Ki67 IHC was available (IHC images acquired with the same Hamamatsu scanner used for HE). vCS_{histo} and CD_{histo} were rescaled to account for tissue shrinkage. vCS_{histo} was multiplied by 1.503, where 1.503 = 1.1806×1.2732 accounts for two scaling factors, namely: i) 1.2732 accounts for cell size underestimation due to 2D sectioning, ii) 1.1806 accounts for a plausible tissue shrinkage of 15.3% following fixation, dehydration, clearing and paraffin embedding[41]. The final CD_{histo} estimate was instead 1.3938 times smaller than the value derived from direct image processing, since 1 mm^2 of shrunk tissue corresponds to 1.1806×1.1806 mm^2 = 1.3938 mm^2 of unprocessed tissue for a shrinkage factor of 15.3%[41].

Analyses

dMRI model selection

We carried out model selection independently for each of the two fitting strategies. The MRI-histology *Total Correlation Score* (TCS) selects the model providing the highest Pearson's correlation between vCS_{MRI} and vCS_{histo}, and between F_{MRI} and F_{histo}. It is defined as

$$\mathrm{TCS} = r(vCS_{MRI}, vCS_{histo}) + r(F_{MRI}, F_{histo}), \quad (13)$$

where $r(vCS_{MRI}, vCS_{histo})$ and $r(F_{MRI}, F_{histo})$ are the correlation coefficients of vCS_{MRI} and F_{MRI} with histological vCS_{histo} and F_{histo}. The correlation between CD_{MRI} and CD_{histo} was not included in Eq. 13 since CD_{MRI} is determined analytically from vCS_{MRI} and F_{MRI}. For TCS computation, we pooled together mouse and human data (N = 25, including the two cases who had their biopsy collected after starting treatment, namely immunotherapy).

We remark that the main objective of the study is to use histological information to inform MRI signal model design, through MRI-histology correlation analyses. A sample size of N=25 provides approximately 50% power to detect a Pearson's correlation of |r|=0.4 (moderate correlation) at a significance level α=0.05. Our study size enables the detection of even moderate correlation at a conservative significance level, and was considered appropriate to inform MRI signal modelling. Nonetheless, we also point out that our sample size increases the risk of Type II vs Type I errors, given the statistical power of 50%. In practical terms, this implies that a true

correlation of |r|=0.4 is roughly 10 times more likely to be missed (false negative) than a false correlation of |r|=0.4 detected (false positive).

We also performed model selection using a *Histology Fidelity Criterion* (HFC), and popular *Bayesian Information Criterion* (BIC)[36,47]. HFC rewards the models providing the best accuracy in the numerical estimation of histological cell size and intra-cellular fraction estimation, i.e., minimising

$$\mathrm{HFC} = \frac{|vCS_{MRI} - vCS_{histo}|}{vCS_{histo}} + \frac{|F_{MRI} - F_{histo}|}{F_{histo}}. \quad (14)$$

Information on CD_{MRI} and CD_{histo} was not included in Eq. 14 since CD_{MRI} is not a degree of freedom of the dMRI models (it is determined analytically from vCS_{MRI} and F_{MRI}). BIC selects the model providing the best goodness of fit, penalising complexity, by minimising

$$BIC = Pln(N) - ln(\lambda). \quad (15)$$

λ is the maximised likelihood, and P/N are the number of model parameters/signal measurements. We performed BIC selection voxel-wise, followed by majority voting across voxels.

Analysis of the fitting residuals

For each voxel, fitting strategy and dMRI model, we calculated the fitting residuals as $s_{fit} - s_{measured}$, where $s_{measured}$ is the measured dMRI signal and s_{fit} the signal predicted by the model at the corresponding b-values/diffusion time. Measured signals were normalised according to Eq. 7, removing relaxation effects due to changes in TE, and signal predictions were bound between 0 and 1 (dMRI signal decay with respect to the $b = 0$ signal). For each b-value, diffusion time and fitting strategy, we visualised distributions across all voxels (median and interquartile range), and also scattered $s_{fit} - s_{measured}$ against $s_{measured}$.

Simulated dMRI model selection

We validated results from MRI and histology data through computer simulations. For this experiment, we synthesised signals for each of the three dMRI protocols considered in the study (*ex vivo* PGSE; *in vivo* PGSE; *in vivo* TRSE), fitted the 5 candidate dMRI models, and performed model selection using TCS, HFC and BIC. dMRI signals were synthesised through Monte Carlo simulations, which we performed using code from the freely available MCDC simulator[46] (https://github.com/jonhrafe/MCDC_Simulator_public). We seeded walkers in a substrate made of spherical cells of identical diameter[6,17,20,21] (Fig. S17), controlling the intra-sphere fraction F by adding gaps of increasing size in-between abutting spheres, packed in an ideal cubic lattice. We probed four F values (0.197, 0.323, 0.406, 0.523) and four sphere diameters vCS for each F (8, 16, 22 and 30 μm). We varied intra-/extra-sphere diffusivities (10×10 values; [0.8; 2.6] $\mu m^2\,ms^{-1}$ for the *ex vivo* protocol and [0.8; 3.0] $\mu m^2\,ms^{-1}$ for *in vivo* protocols), for a total of 1600 synthetic voxels. Signals for each synthetic voxel, featuring a unique combination of IC/EC diffusivities

$D_{0,I}/D_{0,E}$, cell diameter vCS and IC fraction F, were generated for 3 levels of trans-cell water permeability, namely: k = 10 μm s^{-1} (low permeability); k = 25 μm s^{-1} (intermediate permeability); k = 50 μm s^{-1} (high permeability)[19]. We corrupted synthetic signals with Rician noise (b = 0 signal-to-noise ratio: 30), and performed model selection according to TCS, HFC and BIC.

To gauge an understanding of the diffusion times for which modelling trans-membrane water exchange becomes relevant, we also studied the random walks of spins seeded in the IC space, and derived their corresponding instantaneous apparent diffusion coefficients $ADC_I(t) = \frac{\langle\, \Delta x^2(t) + \Delta y^2(t) + \Delta z^2(t) \,\rangle}{6t}$ as a function of time t, given spin positions $\mathbf{p}(t) = (x(t), y(t), z(t))$, and having defined $\Delta x(t) = x(t) - x(0)$, $\Delta y(t) = y(t) - y(0)$ and $\Delta z(t) = z(t) - z(0)$, where $\mathbf{p}(0)$ is the initial spin position.

dMRI-histology correlation

We computed mean and standard deviation of all metrics i) within the mouse liver samples, ii) within a mask containing all liver tumours in patients, iii) within the biopsied patients' tumours. We pooled together metrics from mice and patients to calculate Pearson's correlation coefficients r, as described above. ADC was normalised to the ADC of the PBS solutions in mice and to the free water diffusivity in patients (3.0 μm^2 ms^{-1} at 37 °C), given the difference in temperature. The correlation analysis aims to test whether MRI metrics F_{MRI}, vCS_{MRI} and CD_{MRI} are capable of detecting and replicating such contrasts seen across tumours/samples in F_{histo}, vCS_{histo} and CD_{histo}, which originate from a variety of histopathological processes, e.g., necrosis, fibrosis, oedema, or presence of active cancer.

Additionally, we also fitted linear regression model for dMRI metrics F_{MRI}, vCS_{MRI} and CD_{MRI} from the selected dMRI approach to test the specificity of these metrics towards their direct histological counterparts metrics F_{histo}, vCS_{histo} and CD_{histo}. These were:

$$\begin{cases} vCS_{MRI} = \beta_0 + \beta_1 vCS_{histo} + \beta_1 F_{histo} \\ F_{MRI} = \beta_0 + \beta_1 F_{histo} + \beta_2 vCS_{histo} \\ CD_{MRI} = \beta_0 + \beta_1 CD_{histo} + \beta_2 F_{histo} + \beta_3 vCS_{histo} \end{cases} . \quad (16)$$

The models allow us to test whether: (i) the association between vCS_{MRI} and vCS_{histo} is confounded by F_{histo}; (ii) the association between F_{MRI} and F_{histo} is confounded by vCS_{histo}; (iii) the association between CD_{MRI} and CD_{histo} is confounded by either F_{histo} or vCS_{histo}. Metrics were standardised as z-scores before regression, to ensure comparability of the β coefficients. Fitting was performed in python with *statsmodels*[48] (https://www.statsmodels.org).

Association with immunohistochemical markers of cell proliferation

We assessed the association between F_{Ki67}, an IHC index assessing the fraction of biopsied tissue stained for Ki67 and demonstrating cell proliferation, with dMRI and HE-derived histological properties. For this purpose, we calculated the Pearson's correlation coefficient between F_{Ki67} and all dMRI and HE-derived metrics. We excluded two cases whose biopsies were obtained after having started receiving immunotherapy as part of a phase I trial.

Association with liver tumour phenotype

Finally, we tested whether metrics from the proposed dMRI approach contribute to explaining the macroscopic tumour phenotype. To this end, we studied the association between

dMRI metrics and tumour volume, and compared their values, alongside histological metricsand patients' age, between CRC and melanoma liver metastases through t-tests. Both tumour volume assessment and CRC-melanoma comparison were performed also for histological metrics, excluding two cases who had entered an immunotherapy phase I trial, and whose biopsy was obtained after receiving treatment (HE and MRI metrics excluded).

Regarding the tumour volume assessment, we computed Pearson's correlation coefficients between mean values of dMRI metrics within each individual liver tumour and the tumour volume (N = 140 liver tumours from 38 patients; 137 metastases, 3 primary cancers). Moreover, we performed experiments in which we attempted to predict liver tumour volume given per-tumour mean values of dMRI metrics. Following a standard 5-fold cross validation design, we split the set of tumours into training and test sets. Afterwards, we used the training set to fit linear statistical models of the form

$$log_{10}(TV) = \beta_0 + \sum_{n=1}^{3} \sum_{k=n-1}^{n} \beta_{k,n}\, x^k\, y^{3-n} \quad .(17)$$

Above, TV is the tumour volume in mm^3, modelled as a 3rd order polynomial function of x and y, which is linear in the (unknown) β coefficients of the polynomial terms. x and y in Eq. 16 represent in turn different pairs of dMRI metrics, as for example DKI ADC and K, or pairs of metrics from the *Diff-in* model (we tested explicitly vCS_{MRI} and F_{MRI}; $D_{0,I}$ and F_{MRI}; and $D_{0,I}$ and vCS_{MRI}). Afterwards, we deployed the trained statistical models on the test data, calculating the Pearon's correlation coefficient between ground truth and predicted $log_{10}(TV)$ values, pooling together predictions from all 5 cross-validation folds. Note that the $\log_{10}$ of the volume was studied

instead of the volume itself, since the latter spans almost 4 orders of magnitude in our data set, and is therefore difficult to handle numerically.

Regarding the CRC-melanoma comparison instead, we harmonised dMRI metrics across scanners with a custom-written ComBat[49] implementation, rescaling metrics obtained on the 1.5T system to the 3T range. This was done before performing the CRC-melanoma comparions to account for the fact that CRC scanned almost entirely with just one MRI system.

Statistics and reproducibility

The following statistical analyses were performed. (i) Evaluation of correlations between MRI and histology data. This test was based on a sample size of N = 25 paired MRI-histological data points with a significance level of 0.05 (two-sided Pearson's correlation). P-values were unadjusted for multiple comparisons. Additional correlations models were also fitted accounting for the confounding effect that other MRI metrics can have in the correlation between each MRI index and its direct histological pair. These were based on linear regression models on the same sample. P-values were also unadjusted for multiple comparisons. (ii) Evaluation of correlations between synthetic MRI signals and simulated histological properties. This test was based on a sample size of N = 1600 synthetic tissue environments with a significance level of 0.05 (two-sided Pearson's correlation). P-values were unadjusted for multiple comparisons. (iii) Evaluation of the correlation between MRI metrics and histological Ki-67 stains, visualising cell proliferation. This test was based on a sample size of N = 10 paired MRI-histology data points with a significance level of 0.05 (two-sided Pearson's correlation). P-values were unadjusted for multiple comparisons. (iv) Tumour volume prediction from mean tumour-wise MRI metrics. This analysis was based on the analysis of N = 140 tumours from N = 38 patients. Tumour volume was predicted

through logistic regression in 5-fold cross-validation. Predicted and ground truth tumour volumes were correlated through two-sided Pearson's correlation tests (sample size N = 140). P-values were unadjusted for multiple comparisons. (v) MRI and histological metric comparison between CRC and melanoma cases. This analysis relied on a cohort of N = 13 CRC and N = 8 melanoma cases for MRI metrics; N = 6 CRC and N = 4 melanoma cases for histological HE metrics; N = 4 CRC and N = 2 melanoma cases for Ki-67 immunostain fraction. The between-group comparison was based on two-sided t-tests. P-values were unadjusted for multiple comparisons. No imputation methods are used in any of the analyses above in case of missing data (sample sizes already report the exact number of data points available).

We carried out all our analyses in *python* 3.12.2 (*Anaconda* distribution), using the following third-party packages: *scipy* version 1.13.1; *pandas* version 2.2.2; *numpy* version 1.26.4; *matplotlib* version 3.9.2; *seaborn* version 0.13.2; *scikit-learn* version 1.5.2; *statsmodels* version 0.14.4. The python scripts written to perform the analyses reported in the main manuscript have been included as supplementary data. Supplementary Data S6 performs MRI-histology correlation analysis and model selection (analysis (i) above); Supplementary Data S7 relates MRI metrics to Ki-67 cell proliferation immunostains (analysis (iii) above); Supplementary Data S8 evaluates the association with tumour volume (analysis (iv) above); Supplementary Data S9 compares MRI and histology metrics across CRC and melanoma groups (analysis (v) above).

Ethics

All experimental protocols in animals and PDX generation were approved and monitored by the Vall d'Hebron Institute of Research Animal Experimentation Ethics Committee (CEEA;

registration number 68/20 and 70/20) in accordance with relevant local and EU regulations. The imaging study and sampe collection from cancer patients for histopathology as well as PDX generation were approved by the Vall d'Hebron University Hospital Ethics committee (PR(AG)29/2020, PR(AG)5248), Barcelona, Spain. Patients provided informed written consent to participate in the PDX study and in the imaging study.

Results

Overview: mouse and human data for histology-informed dMRI signal modelling

We designed and demonstrated our dMRI approach using data obtained in fixed mouse livers (*mouse data*) and in cancer patients *in vivo* (*human data*) (Fig. 1). Mouse data consists of pulsed gradient spin echo (PGSE) DW MRI scans of seven fixed mouse livers, performed *ex vivo* on a 9.4T Bruker system. It also includes whole-organ HE-stained sections, obtained at known radiographic position. We obtained livers of mice sacrificed as part of ongoing xenograft model development studies in prostate cancer. Six had been implanted with biopsies of prostate cancer patients, while one had not had any implantation. While the livers from the implanted mice did not grow any tumours, they feature a variety of pathologies, with three unique histopathological phenotypes (Fig. S1). The liver from the mouse with no implantation features normal liver structures, and we will refer to it as *Control*. Of the six implanted cases, two also show normal liver tissue, with normal representation of all hepatic structures. We will refer to these cases as Pat_{NA1} and Pat_{NA2} (patient biopsy implantation, but normal appearing). Another case exhibits generalised necrosis and diffuse acute and chronic inflammation surrounding necrotic areas, with presence of occluded thrombotic vessels. This specimen will be identified as Pat_{nec} (patient biopsy

implantation, with necrosis). Finally, three specimens feature an immature, lymphoproliferative process, with various degrees of infiltration of small, lymphoid, atypical cells with abundant mitosis, which infiltrate portal vessels and sinusoidal capillaries, but without producing tumours. These will be referred to as Pat_{inf1} to Pat_{inf3} (patient biopsy implantation, with lymphoid cell infiltration).

We obtained human data on cancer patients suffering from advanced solid tumours, participating in an imaging study at the Barcelona Vall d'Hebron Institute of Oncology (VHIO, Spain), before being enrolled in a number of clinical trials. We included data from 38 patients with liver malignancies (mean/standard deviation (SD) of age: 61.88/11.90 years; median/inter-quartile range (IQR) of age: 62.8/14.7 years; 19 male, 19 female), of which 3 suffered from primary hepatocellular carcinoma (HCC), while 35 had liver metastases from different primary cancers (10 colon, 8 melanoma, 4 breast, 4 gastric, 3 rectal, 2 ovarian, 1 renal, 1 endometrial, 1 ureteral, 1 thyroidal). dMRI was based on diffusion-weighted (DW) echo planar imaging (EPI). We also obtained digitised HE-stained biopsies from one of imaged liver tumours approximately one week after MRI in 18 cases (mean/SD of age: 61.1/12.0 years; median/IQR of age: 62.1/14.4 years; 8 male, 10 female). Of these 18 patients, 2 suffered from primary HCC, while 16 had liver metastases (5 from colon cancer, 4 from melanoma, 3 from rectal cancer, 2 from ovarian cancer, 1 from breast and endometrial cancer). In 10 cases, Ki67 imunnohistochemsitry (IHC), demonstrating cell proliferation, was also available beyond routine HE stains.

Overview: dMRI signal models for cell size and density measurement

We studied five biophysical dMRI models, grouped into two families (Fig. 2A). All models describe the dMRI signal as originating from a combination of IC and EC contributions[6,17,20], making different assumptions. The first family, which is more general, includes:

i. *Diff-in-exTD*: a model accounting for restricted IC diffusion within spherical cells[6], and hindered diffusion in the EC space, with diffusion time dependence[34] in both IC and EC signals[50]. The diffusion time quantifies the time during which water molecules can sense cellular barriers, before the MR signal is measured.

ii. *Diff-in-ex*: similar to model *Diff-in-exTD*, but neglecting diffusion time dependence in the EC space. Popular body dMRI techniques such as IMPULSED[17] or VERDICT[18] are essentially implementations of this model.

Conversely, the second family explicitly assumes that the EC apparent diffusion coefficient (ADC) is higher in the EC than in the IC space, i.e., that $ADC_E > ADC_I$, and assumption used in certain body dMRI techniques[51]. This family includes:

i. *Diff-in-exTDFast*: equivalent to *Diff-in-exTD*, but ensuring that $ADC_E > ADC_I$.

ii. *Diff-in-exFast*: equivalent to *Diff-in-ex*, but ensuring that $ADC_E > ADC_I$.

iii. *Diff-in*: a model where it is hypothesised that the EC extra-cellular signal is negligible with respect to the IC one due to its much larger ADC.

All models enable the estimation of *volume-weighted mean cell size (CS)* (vCS_{MRI}, expressed in µm) and *IC signal fraction* (F_{MRI}, dimensionless). These can be combined into an apparent *cell density per unit volume* ($CD_{MRI} = F_{MRI}/vCS_{MRI}^3$, expressed in cell/mm^3)[18]. All models were fitted twice: to the complete set of images, or to highly diffusion-weighted (DW) images only (i.e., to

high *b*-values), to minimse signal contributions from the EC space. For comparison, we also considered routine ADC (in $\mu m^2/ms$) and apparent diffusion excess kurtosis K (dimensionless) from diffusion kurtosis imaging (DKI)[52], since these are popular dMRI indices sensitive to cancer cellularity and easy to compute from short acquisitions[18,53].

We processed HE-stained histological data with automatic cell detection[40] to derive histological counterparts of the dMRI metric. The two were compared to inform the design of a dMRI signal model maximising radiological-histological correlations (Fig. 2B) . The histological metrics were: *histological volume-weighted mean cell size* (vCS_{histo}, in µm), *histological intra-cellular area fraction* (F_{histo}, dimensionless), and *histological cell density per unit area* (CD_{histo}, in cell/mm^2). In patients for which Ki-67 ICH was available, the fraction of histological material stained for Ki-67 within the biopsy was also computed through custom routines (metric F_{Ki67}).

Supplementary Data S1 contains a CSV file with the MRI and histology data used for model selection, while Supplementary Data S2 contains a CSV file with MRI and histology data for model selection based on high b-value images only. Supplementary Data S3 contains a CSV file with MRI and histology metrics for each individual tumour. Additionally, Supplementary Data S4 and Supplementary Data S5 contain a CSV file with cohort information, histological and mean MRI metrics across all tumour voxels (inter-scanner harmonisation performed in Supplementary Data S5).

A model of restricted intra-cellular diffusion with negligible extra-cellular contributions maximises radiological-histological correlations

We compared dMRI models through an MRI-Histology *Total Correlation Score* (TCS). TCS measures the overall correlation between histological and radiological readouts of cell size and intra-cellular fraction, and is obtained by summing Pearson's correlation coefficients between MRI and histological cell size and IC fraction (i.e., vCS_{MRI} and vCS_{histo}, and F_{MRI} and F_{histo}). Negative correlations are penalised, since the model providing the highest TCS should be preferred. TCS values unequivocally suggest that models where the IC ADC is constrained to be smaller than the EC ADC provide higher correlations with histology. Among all dMRI implementations, fitting a model of IC restricted diffusion with negigilble EC signal contributions to highly DW images (referred to as model *Diff-in*) provides the highest TCS values, and hence the highest correlation to histological readouts (Fig. 3). Fig. 3A summarises the different dMRI models, while Fig. 3B and Fig. 3C report actual TCS values. In the figure, models where $ADC_E > ADC_I$, shown in violet shades, provide consistently higher TCS values than models that do not make such an assumption (orange shades), with the highest TCS observed for model *Diff-in* fitted to highly DW images.

Additional model selection criteria used to compare dMRI models confirm findings from TCS-based model selection. The additional criteria were: an *Histology Fidelity Criterion* (HFC), measuring the sum of absolute errors in F and vCS estimation via dMRI, compared to histology; and the *Bayesian Information Criterion* (BIC)[47]. BIC is a common criterion for dMRI model selection[36,54]: it quantifies the model fitting quality, penalising model complexity. Note that model performance increases as both HFC and BIC decrease. Fig. S2 reports the number of times, in percentage terms, that a model provides the lowest HFC and BIC across our N = 25 MRI-histology cases. TCS ranking is confiremd: models hypothesising $ADC_E > ADC_I$ are selected more frequently than models that do not do, according to HFC. The model *Diff-in* is the most selected model according to both BIC and HFC (fig. S3.B; fitting to high b-value images). Fig. S3 splits

HFC and BIC rankings depending on the MRI scanner. In all cases, models with $ADC_E > ADC_I$ (*Diff-in*, *Diff-in-exFast*, *Diff-in-exFastTD*) are selected more frequently than models *Diff-in-ex* and *Diff-in-ex-TD*. When fitting is performed only on high b-value images, *Diff-in* is the most selected model according to both BIC and HFC.

Supplemetnary Figures S4 to S10 show examples of signal model fitting to dMRI measurements obtained from exemplificatory voxels. Supplementary Figures S4, S5, S6 refer to *ex vivo* mouse data (examples of voxels from the three main histological phenotypes); Supplementary Figures S7 to S10 refer instead to *in vivo* cancer patient (active tumour, necrosis, and other tumour regions-of-interest, from both 3T and 1.5T systems). On visual inspection, all models capture the main trends of the signal as a function of the b-value and of the diffusion time, e.g., low diffusion attenuation in highly cellular areas; high diffusion attenuation in necrosis; increasing dMRI signal for increasing diffusion time.

The analysis of fitting residuals reveals differences among dMRI models

We also studied fitting residuals as a function of the acquisition parameters, as patterns in the distribution of the residual can provide insight onto the features of the signal that are best – or wosrt – estimated by a model. Supplementary Figures S11, S12 and S13 show median and interquartile of residual (i.e., signal predicted by a dMRI model, minus the measured signal) distributions as a function of the b-value and of the diffusion time, for both fitting strategies (Supplementary Figures S11: distributions across *ex vivo* mouse tissue voxels; Supplementary Figures S12: distributions across *in vivo* cancer patients voxels, 3T scans; Figures S13: distributions across *in vivo* cancer patients voxels, 1.5T scans). In general, all models tend to

overestimate the signal at longer diffusion times, compared to shorter diffusion times acquired at the same b-value. This is potentially compatible with the fact that as the diffusion time increases, the measured signal becomes more sensitivie to IC/EC water exchange – a biological phenomenon not considered by our models, and leading to an increase in apparent diffusivities (and hence, to stronger signal decay) compared to ideal diffusion restriction[19]. The underestimation of the signal is more apparent for models that assume faster EC diffusion, i.e., *Diff-in*, *Diff-in-exFast* and *Diff-in-exTDFast*, compared to the models that do not make assumptions on which is faster between IC and EC diffusion.

These trends are confirmed by voxel-wise scatter plots in Supplementary Figures S14, S15 and S16, in which residuals are scattered against the corresponding measured dMRI signals (Supplementary Figures S14: *ex vivo* mouse tissue voxels; Supplementary Figures S15: *in vivo* cancer patients voxels, 3T scans; Supplementary Figures S16: *in vivo* cancer patients voxels, 1.5T scans). Another important aspects revelead by all plots in Supplementary Figures S11 to S16 is that, for a given dMRI measurement (i.e., fixed *b* and diffusion time), residuals are smaller when fitting is performed only on high b-value images, rather than on all images. This finding remarks that simplified analytical models with 3-5 parameters, as those considered here, may not always capture the full complexity of the dMRI signal over large acquisition spaces due to anaccounted biological processes, as water exchange and intra-compartmental relaxation.

Computer simulations give insight into histology-informed model selection

We related the model selection performed on mouse and human dMRI data to computer simulations of diffusion in virtual substrates made of packed spheres, representing cells, a common

tissue representation in body dMRI[6,17,19–21] (Fig. S17). Simulations consisted in: i) synthesing dMRI signals for the three acquisition protocols used in this study, through Monte Carlo simulations; ii) fitting all candidate models to the synthetic dMRI signals; iii) selecting the best approach according to the TCS (correlation ground truth-estimated parameters), HFC (accuracy in parameter estimation) and BIC (quality of signal fit, penalising model complexity) criteria introduced above. Our signal synthesis explicitly accounted for IC/EC water exchange (3 membrane permeability levels, i.e., low, intermediate and high), and probed a wide range of IC fractions, cell sizes and IC/EC diffusivities.

In our synthetic data, TCS confirms that model *Diff-in* fitted to highly DW images provides the best overall performances in terms of correlation between ground truth and estimated CS and IC fraction, when trans-membrane permeability is low (Supplementary Table S1). In this scenario, neglecting EC signal contribution at high *b* provides a compact signal representation that captures salient contrasts in terms of CS and IC fraction. Nontheless, Supplementary Table S1 also shows that the performances of such a model degrade as permeability levels grows. As IC/EC water exchange increases, the IC and EC signals become more and more intermixed, and models that account explicitly for EC diffusion provide increasingly better performances. Rankings based on HFC (Supplementary Table S2) show that while *Diff-in* excels in ground truth-estimated parameter correlations, it is outperformed by other models in terms of accuracy of the estimation, regardless of the permeability level. This finding suggests that while a simplified signal representations such as *Diff-in* suffices to capture CS and IC fraction contrasts, it may provide biases microstructural estimates, due to accounted diffusion processes. Finally, BIC-based rankings (Supplementary Table S3) show that *Diff-in*, despite being a simplified signal representations that by design neglects signal contributions from specific water pools (i.e., EC), suffices to capture salient

patterns of the dMRI signal. According to BIC, *Diff-in* provides excellent goodness of fit, as it systematically ranks first for almost all protocols and all fitting strategies. .

To more clearly visualise how trans-membrane water exchange affects diffusion, and hence dMRI signal modelling, we plotted the instantaneous diffusion coefficient for spins seeded inside cells ($ADC_I(t)$) for different CS, intrinsic diffusivities and cell permeability levels (Supplementary Fig. S18). The figure shows that IC/EC water exchange affects the temporal profile of $ADC_I(t)$, as compared to curves obtained for impermeable cells. For example, in presence of membrane permeability, $ADC_I(t)$ is in general higher than when obtained for impermeable cells, for any t. Moreover, it does not exhibit the characteristic decreasing behaviour approaching $ADC_I \rightarrow 0$ typical of restricted diffusion: while firstly it decreases sharply, due to the restricting effect of cell membranes, it then increases with time, as more and more spins leave the cells and explore the EC space with their random walks. However, the plots also show that exchange-induced deviations of ADC_I from the ideal diffusion restriction case are the strongest for diffusion times that are unlikely to be probed with clinical systems in vivo (t > 40 ms), and that they are small for large cell diameters (22-30 μm), as those that can be expected in normal liver or in certain liver tumours.

Cell size and density estimates from the proposed dMRI model correlate with histology

In view of the results obtained from the analysis of *in silico*, *ex vivo* and *in vivo* MRI data, our recommended liver tumour dMRI approach is the *fitting of model "Diff-in" – a one-compartment model of restricted, IC diffusion within spherical cells, with negligible EC signal contributions – to high b-values images* ($\gtrapprox$ 1800 s/mm^2 *ex vivo*, $\gtrapprox$ 900 s/mm^2 *in vivo*).

The model provides estimates of IC fraction, cell size and cell density that are weakly, moderately and strongly correlated to the underlying histology. Table 1 shows Pearson's correlation between *Diff-in* metrics F_{MRI} (intra-cellular fraction), vCS_{MRI} (cell size index) and CD_{MRI} (cell density) with their histology counterparts F_{histo}, $vCS_{histo}$and CD_{histo}. We observe a weak correlation between F_{MRI} and F_{histo} (r = 0.19, p = 0.370), moderate between vCS_{MRI} and vCS_{histo} (r = 0.44, p = 0.029), and strong between CD_{MRI} and CD_{histo} (r = 0.70, p = 0.0001). The weak correlation between F_{MRI} and F_{histo} is likely due, at least in part, to the fact that F_{MRI} is a signal fraction, rather than an actual volume/area fraction. This implies that F_{MRI} also encodes T2/T1 differences between IC and EC signals, unlike F_{histo} [35]. Moreover, F_{MRI} values may have also been biased by the exchange between intra- and extra-cellular water[19,55], especially in presence of small cells[19] – a phenomenon that is is not accounted for in none of the studied signal models. Finally, we emphasise that in the N = 18 human observations, dMRI metrics from an entire tumour have been compared to a small sliver of biopsied tissue. This implies that the correlation between dMRI and histology may have been underestimated, especially for the IC fraction, for example when the biopsy misses large areas of necrosis or fibrosis, which affect F_{MRI} values. Supplementary Table S4 shows results from linear regression models that test the specificity of vCS_{MRI}, F_{MRI} and CD_{MRI} to their histologica counterparts vCS_{histo}, F_{histo} and CD_{histo}. Results suggest that $vCS_{histo}$is the main correlate of vCS_{MRI} (even considering the confounding effect of F_{histo}), that $F_{histo}$is the main correlate of F_{MRI} (even considering the confounding effect of vCS_{histo}), and that CD_{histo} is the main correlate of CD_{MRI} (even considering the confounding effect of vCS_{histo} and F_{histo}). Note that collinearity among the regressors may cause β coefficients be larger than 1 even if working with z-scores (i.e., for the model featuring CD_{MRI} as dependent variable). In general, the association between F_{histo} and F_{MRI} is confired to be weak; between

741 vCS_{histo} and vCS_{MRI} moderate; and between CD_{histo} and CD_{MRI} to be strong. These results give
742 additional confidence about the biological specificity of the proposed dMRI approach.

743 Table 1 also reports correlation coefficients for DKI ADC and kurtosis K. The correlations
744 are in line with previous studies[56,57]. For example, both ADC and K exhibit significant, moderate
745 correlations with histological properties, i.e., a negative correlation with cell density CD_{histo} for
746 ADC, and a positive correlation with CD_{histo} for K (r = –0.47 and 0.43 respectively). Significant
747 correlations are also seen between DKI metrics and F_{histo} (r = 0.40, p = 0.048 for K).

748 Fig. S19 and Fig. S20 show Pearson's correlation coefficients for all possible pairs of
749 metrics, in the form of correlation matrices. Supplementary Data S1 and S2 contain all
750 MRI/histology metric values used to generate such matrices. In general, metrics from dMRI
751 models where $ADC_E > ADC_I$ show stronger correlations with their histological counterparts than
752 models *Diff-in-exTD* and *Diff-in-ex*. Specifically, we observe the strongest dMRI-histology
753 correlations for model *Diff-in* fitted to high b-value images. Correlations among pairs of dMRI
754 metrics are also seen, e.g., a strong negative correlation between CD_{MRI} and vCS_{MRI} (r = –0.81 for
755 model *Diff-in* fitted to high b-value images). This finding, indicating that tighter cell packings per
756 unit volume are achieved with smaller cells, is mirrored by histological CD_{histo} and vCS_{histo} (r =
757 –0.88 between these two), and thus appears biophysically plausible.

758

759 **The proposed dMRI model reveals histologically-meaningful tumour characteristics**

760 The proposed model *Diff-in* is demonstrated in 7 fixed livers of mice suffering from a variety of
761 pathologies, as well as in a cohort of 38 patients suffering from advanced solid tumours of the liver
762 (both primary and metastatic). Voxel-wise *Diff-in* parameteric maps reveal intra-tissue and inter-

763 tissue contrasts that are histologically meaningful, being compatible with histological features
764 observed in HE stains. These include, for example, presence of small cells in areas characterised
765 by lymphoproliferative processes in mouse livers, or reductions in cell density in necrosis or
766 fibrosis in human tumours.

767 Fig. 4 shows *Diff-in* and histological maps in 3 mouse livers, representative of the 3
768 phenotypes seen in our mouse data (*Control*, for normal liver structures; Pat_{inf1}, for small cell
769 infiltration; Pat_{nec}, for necrosis). Visually, we observe excellent co-localisation between MRI
770 slices and HE sections. The histological details reveal higher cellularity in sample Pat_{inf1} compared
771 to *Control*, due to packing of small cells in between larger hepatocytes, or an alternation of areas
772 with lower/higher cell density in sample Pat_{nec}. These qualitative trends are confirmed in the
773 histological maps F_{histo}, vCS_{histo}, CD_{histo}, with values in physiologically plausible ranges, as for
774 example intra-cellular fractions around 0.75 and cell sizes of the order of 20 μm[58,59]. Maps F_{MRI},
775 vCS_{MRI} and CD_{MRI} replicate the contrasts seeing in their histological counterparts F_{histo}, vCS_{histo}
776 and CD_{histo}. Fig. S21 shows standard DKI ADC and kurtosis excess K in the same mouse livers.
777 Visual trends highlight that the higher cell density of sample Pat_{inf1} translates to remarkably
778 reduced ADC and increased K compared to the *Control*. Lastly, Fig. S22 and S23 show additional
779 dMRI metrics, including metrics from model *Diff-in-exFast* (i.e., the one providing the higher TCS
780 when fitting is performed to the whole dMR image set). The figures highlight that overall, spatial
781 trends seen in maps from the selected model *Diff-in* agree with those seen in *Diff-in-exFast*, but
782 metrics from the latter appear noisier. Metrics $D_{0,I}$ and $D_{E,\infty}$ show limited between-sample
783 contrast, and are difficult to validate histologically. Table S5 report qualitative per-sample mean
784 and standard deviation of all MRI and histology metrics in mice (see also Supplementary Data S1
785 and S2). F_{MRI} slightly underestimates F_{histo}, while vCS_{MRI} slightly overestimates vCS_{histo}. We

speculate that the discrepancies may be due, at least in part, to unaccounted factors such as variability in intrinsic cell shape/cytosol diffusivity[33] or water exchange[19], and by the difficulty of relating accurately 2D histology to 3D MRI[60].

Fig. 5 shows F_{MRI}, vCS_{MRI} and CD_{MRI} maps in patients, alongside biopsies. Table S6 reports salient demographic, MRI and histology information of the cohort. Histopathological assessment highlights the variety of characteristics that can coexist within advanced solid tumours, e.g.: areas of fibrosis; localised areas of tightly packed cancer cells, sourrounded by stromal fibres; necrosis. dMRI F_{MRI}, vCS_{MRI} and CD_{MRI} show contrasts that are plausible with such histopathological features. For example, in a breast cancer liver metastasis in Fig. 5, we observe a core of low intra-cellular fraction F_{MRI} and low cell density CD_{MRI}, compatible with necrosis. In a HCC case instead, we see areas of high F_{MRI} and high CD_{MRI}, sourrounded by lower F_{MRI} and lower CD_{MRI}, potentially indicating the alternation of high cell densities with fibrotic tissue. Fig. S24 shows routine dMRI ADC and K in the same tumours. These show plausible contrasts given the known tumour histology, e.g., a core of high ADC and low K are seen in the necrotic core of the breast cancer tumour. Supplementary Fig. S25 shows F_{MRI}, vCS_{MRI} and CD_{MRI} from model *Diff-in-exFast*. Image contrasts match visually those seen in *Diff-in* metrics (the proposed approach), although *Diff-in-exFast* metrics appear noisier on visual inspection. Fig. S26 shows intra-cellular cytosol diffusivity $D_{0,I}$ asymptotic ADC_E ($D_{E,\infty}$). Their speckled appearance suggests that these metrics are difficult to measure accurately *in vivo*[17,61].

Metrics from the proposed dMRI model correlate with cell proliferation in liver tumours

We tested the potential utility of the proposed model for non-invasive liver tumour phenotyping. To this end, we studied associations between dMRI metrics an cell proliferation within patients' biopsied liver tumours, quantified through Ki67 immunohistochemistry (IHC). The maing finding of this analysis is that dMRI cell size (vCS_{MRI}) and density (CD_{MRI}) are respectively negatively and positively associated to the fraction of Ki67 staining (F_{Ki67}). This results, confirmed by histological cell size and density (vCS_{histo} and CD_{histo}), suggests that tumours featuring high cell proliferation are made, on average, by smaller and more densely packed cells, compared to tumours with lower proliferation. Fig. 6 visualises the association between F_{Ki67} and MRI/histological cell size and density through scatter plots, reporting Pearons's correlation coefficients. There is a strong, positive correlation between CD_{MRI} and F_{Ki67} (r = 0.80, p = 0.006), which is confirmed by a positive correlation between CD_{histo} and F_{Ki67} (r = 0.66, p = 0.038). Regarding cell size, there is a moderate, negative correlation between vCS_{MRI} and F_{Ki67} (r = –0.53, p = 0.111). This correlation, albeit not statistically significant owing to the small sample size (N = 10), is in the same direction of the correlation observed between histological vCS_{histo} and F_{Ki67} (r = –0.68, p = 0.031). Concerning other MRI metrics, weaker correlations are seen between DKI indices and F_{Ki67} (r = 0.17, p = 0.633 between ADC and F_{Ki67}; r = 012, p = 0.744 between K and F_{Ki67}), or between F_{MRI} and $D_{0,I}$ from the proposed dMRI model and F_{Ki67} (r = –0.35, p = 0.319 between F_{MRI} and F_{Ki67}; r = 0.07, p = 0.857 between $D_{0,I}$ and F_{Ki67}).

Fig. 6 also visualises examples of biopsies from two liver tumours, featuring respectively strong and weak Ki67 staining for cell proliferation. These are: a metastatic ovarian cancer (Fig. 6F), whose F_{Ki67} is of 0.197, and a primary HCC (Fig. 6E), which features a F_{Ki67} of 0.008. The former tumour, with stronger Ki67 immunostain, is characterised by smaller and more densely packed cells than the latter, as suggested by both MRI and histology (for histology: vCS_{histo} of

23.41 µm against 27.50 µm, and CD_{histo} of $5.24 \cdot 10^3$ cell/mm^2 against $2.44 \cdot 10^3$ cell/mm^2; for MRI: vCS_{MRI} of 25.15 µm against 29.39 µm, and CD_{MRI} of against $2.08 \cdot 10^5$ cell/mm^3 against $0.58 \cdot 10^5$ cell/mm^3).

Metrics from the proposed dMRI model are associated with macroscopic tumour characteristics

We investigated whether the microscopic composition of each individual liver tumour assessed through dMRI can explain its macroscopic phenotype, e.g., salient characteristics such as its volume. For this analysis, we studied correlations between tumour-wise dMRI indices and the corresponding tumour volume, using data from N = 140 liver tumours from the 38 patients. We also performed cross-validation experiments in which we predicted liver tumour volume through statistical models that used dMRI metrics as predictors. The main finding of this analysis is that *Diff-in* metrics are significantly correlated with tumour volume, as is the case, for example, for MRI-derived cell size vCS_{MRI}. Statistical models that rely on cell size vCS_{MRI} as a regressor enable meaningful predictions of tumour volumes, which are significantly correlated with ground truth values, and outperforming predictions from statistical models based on DKI metrics.

Fig. 7A reports Pearson's correlation coefficients between tumour volume and dMRI metrics from both *Diff-in* and DKI dMRI. The strongest, statistically significant associations are seen for *Diff-in* metrics $D_{0,I}$ and vCS_{MRI} (r = –0.305, $p = 2.50 \cdot 10^{-4}$ and r = 0.256, $p = 2.26 \cdot 10^{-3}$ respectively), while no significant associations are seen for DKI ADC (r = –0.092, p = 0.277) and kurtosis K (r = –0.035, p = 0.680). Fig. 7B shows an example of two tumours, one larger (a primary hepatocellular carcinoma, with a volume of circa $431 \cdot 10^3$ mm^3) and one considerably smaller (an

ovarian cancer liver metastasis, with a volume of approximately $21 \cdot 10^3$ mm^3) (same tumours shown in Fig. 6). Histology shows that the former contains, on average, larger cells than the latter, a contrast that is detected by dMRI vCS_{MRI}. Fig. 7C to Fig. 7F report scatter plots of different pairs of dMRI metrics coloured by tumour volume, visualising the latter as a bi-dimensional function of the diffusion metrics. Trends are seen in the colouring of the plots, e.g., in panel 7F, where the dependence of tumour volume on $D_{0,I}$ and vCS_{MRI} is visualised (Fig. 7F). In Fig. 7F, the tumour volume does not appear to be monotonic in the ($D_{0,I}$, vCS_{MRI}) plane, as it peaks for ($D_{0,I}$, vCS_{MRI}) around (1.8 μm^2/ms; 28 μm). Finally, Fig.7G to Fig. 7J show results from the 5-fold cross-validation experiments, by scattering ground truth and predicted tumour volume values for different statistical models. The highest correlation between ground truth and predicted volumes is obtained for a model where $D_{0,I}$ and vCS_{MRI} are used as predictors, i.e., a moderate but significant Pearson's r = 0.396 ($p = 1.26 \cdot 10^{-6}$; range of r across the 5 folds of (0.146; 0.524)). This surpsasses values obtained when other dMRI metrics are used, e.g., DKI ADC and kurtosis K (r = 0.244, $p = 3.68 \cdot 10^{-3}$; range of r across the 5 folds of (0.226; 0.403)). Supplementary Data S3 reports all liver tumour values used for these analyses.

The proposed dMRI model distinguishes colorectal cancer from melanoma liver metastases

Lastly, we tested whether the proposed dMRI approach may complement standard diffusion metrics, such as ADC, to aid the interpretation of between-tumour differences. For this purpose, we compared the composition of liver metastases from colorectal cancer (CRC) and melanoma, the two most frequent primary cancers in our cohort. We tested for differences in dMRI metrics between the two groups, verifying results with histology. These two cancers feature distinct

874 cellular composition: while CRC metastases are characterised by the presence of large luminal
875 spaces, melanoma metastases consist of tightly packed malignant melanocytes (Fig. 8, panels A
876 and B)[62,63]. The proposed *Diff-in* model provides estimates of IC fraction F_{MRI} that are significantly
877 larger in melanoma than in CRC metastases (p = 0.011, N = 21; Fig. 8, panels A, B and K), a
878 finding that is consistent with the presence of EC water trapped within CRC luminal spaces. This
879 result is confirmed by the histological IC fraction F_{histo}, which is also lower in CRC than in
880 melanoma (p = 0.034, N = 12; Fig. 8, panels A, B and E). More standard diffusion indices also
881 detects differences between the two cancer types: DKI ADC is higher in CRC than in melanoma
882 metastases (p = 0.038, N = 21; Fig. 8, panels A, B and I), also consistent with presence of free
883 diffusion in the lumina. No statistically significant differences between CRC and melanoma were
884 detected for any other metric, including patient's age.

885

886 **Discussion**

887 The latest liver dMRI models aim to disentangle IC and EC water contributions to the total
888 signal[6,17,20,21]. This powerful approach enables the estimation of innovative tissue property maps,
889 but its clinical deployment is hampered by the high number of unknown tissue parameters to
890 estimate, which requires impractically long dMRI acquisitions[17,64,65]. With this challenge in mind,
891 this paper delivers a practical implementation of a two-compartment dMRI signal model, tailored
892 for liver imaging, and truly feasible in the clinic. Through histology-informed model selection, we
893 design a compact framework consisting of fitting a model of restricted IC diffusion to highly DW
894 images, with negligible EC signal contributions. This provides cell size and density estimates that

correlate with histology, and which may provide additional information to standard-of-care imaging for non-invasive tumour phenotyping.

To find the optimal dMRI signal implementation, we analysed co-localised dMRI and histology data (N = 25) from fixed mouse livers and from human tumours of the liver. We compared 5 signal models, each fitted according to two distinct strategies, and ranked them for their ability to estimate histological IC fraction and cell size (the uknown tissue parameters in the dMRI signal models). Rankings unequivocally suggest the highest radiological-histological agreement is obtained by fitting a model of restricted diffusion within spherical cells, with negligible EC signal contributions – a model here referred to as *Diff-in* –, to images acquired with b-values higher than approximately 900 s/mm^2 *in vivo* and 1800 s/mm^2 *ex vivo*. Interestingly, our central result, points towards the use of simpler models of diffusion if these are deployed in appropriate measurement regimes, and is consistent with recent estimates of the EC liver ADC, as high as 2.5 μm^2/ms [17]. Such a high ADC_E implies that the EC signal would decay to roughly 5% or less of its non-DW value even for b-values of 1200 s/mm^2 ($exp(–b\ ADC_E) \approx 0.05$ for b = 1.2 ms/μm^2 and ADC_E = 2.5 μm^2 ms$^{–1}$), justifying the hypothesis of negligible EC signal contributions[7,29,61]. Nevertheless, we also point out that our main modelling result, should be interpreted with care, since some unaccounted biological processes may also contribute to explaining the performances of the simplified *Diff-in* model (one-pool model of restricted intra-cellular diffusion). Notably, none of the tested signal models accounts for water exchange across cell membranes, a these are assumed to be perfectly impermeable. It is therefore possible that the good performances of the *Diff-in* model, or, more generally, of models featuring faster EC diffusion may be, at least in part, a consequence of non-negligible membrane permeability, and hence a manifestation of fast IC/EC water exchange.

To corroborate dMRI model selection on *ex vivo* and *in vivo* MRI, we performed model selection on simulated dMRI data. These were generated using virtual cellular models made of packed spherical cells, exhibiting a variety of packing densities, IC/EC diffusivities, and cell membrane permeabilities. Results from simulations essentially confirm rankings from *ex vivo* and *in vivo* imaging in presence of low cell permeability, as model *Diff-in* fitted to high b-value images provides among the best performance in terms of ground truth-parameter estimate correlation (high ranking in TCS). However, as permeability increases, the performances of *Diff-in* degrade, and models that account explicitly for EC signals (*Diff-in-ex* and *Diff-in-exTD*), without imposing constraints of IC vs EC diffusion, provide better performances – a finding that is in line with the lower biases in signal estimation of these models seen on actual *ex vivo* and *in vivo* dMRI measurements. Despite the degradation in performances measured by TCS and HFC as exchange increases, the model *Diff-in* is still capable of delivering acceptable signal fitting quality for all permeability levels *in silico*, as demonstrated by the fact that it ranks high on BIC (recall that BIC essentially corrects fitting mean squared errors, penalising model complexity). Nonetheless, while ranking high on BIC, our simulated HFC model selection also reveals that *Diff-in* is outperformed by other models in terms of accuracy of the parameter estimation. This result is consistent with the fact that the estimation of dMRI parameters can be biased by unaccounted biological processes, as, in this case, residual EC signals at high *b* as well as water exchange. All in all, these finding confirm that *Diff-in*, while being a simplified dMRI signal representation that is more prone to biases than more complex models, is a compact and potentially useful representation for liver tumour imaging when deployed in appropriate measurement regimes, that is, with negligible EC signal contributions for sufficiently high b-values, and slow IC/EC water exchange with respect to the measurement diffusion times.

Our proposed dMRI modelling approach provides metrics that visualise tissue heterogeneity over a wide field of view, beyond what can be routinely sampled in the clinic with biopsies. *Diff-in* metrics point, for example, to areas of high cell densities in some of the fixed mouse livers, or areas of necrosis within the core of large tumours in patients *in vivo* – facts that are all confirmed by histology. Notably, spatial trends from the dMRI model *Diff-in* agree with metrics from other candidates dMRI signal implementation, e.g., *Diff-in-exFast*, the model selected when fitting is performed on the whole image data set (Fig. S22 and S25). On the one hand, this gives confidence on the robustness of our signal modelling routines. On the other hand, the noisier appearance of metrics *Diff-in-exFast* compared to *Diff-in* is consistent with ill-posedness of complex dMRI models, when they include several unknowns to estimate[61].

We investigated the potential clinical utility of the proposed *Diff-in* dMRI approach by testing whether its metrics can serve as non-invasive biomarkers of biologically relevant processes in cancer. For this, we conducted three analyses, namely: i) we studied the association between *Diff-in* metrics and cell proliferation; ii) we investigated the relationship between *Diff-in* metrics and macroscopic tumour characteristics, as their volume; iii) we tested whether *Diff-in* metrics reveal between-tumour differences that are histologically meaningful. Regarding the comparison with Ki67 immunostains demonstrating cell proliferation, we found that dMRI cell size and density are sensitive to Ki67 stains of cell proliferation, being respectively moderately and strongly correlated with the staining fraction F_{Ki67}. This finding, confirmed by histological cell size and density from HE-stained biopsies, may indicate that tumours featuring higher proliferation levels contain smaller and more densely packed cells. This is in line with previous *in vitro* measurements, as those obtained, for example, in colon carciomas[66], in which the time required for malignant cell populations to double in number was shorter for small cells with large nuclei, compared to larger

964 cells but with small nuclei. The main takeaway of this analysis is that cell morphology assessment
965 through our dMRI model may offer sensitivity to key biological processes of cancer biology, as
966 cell proliferation[67], which are directly linked to clinical outcomes. The reported association
967 between *in vivo* cell size measurements and Ki67 proliferation is, to the best of our knowledge, a
968 novel finding. However, we acknowledge that its exact biological implications deserve additional
969 elucidation in larger cohorts, given the small size of our sample.

970 Regarding the analysis on tumour volume instead, our key result is that metrics from the
971 proposed *Diff-in* approach are associated to tumour volume, albeit only moderately, in a data set
972 of N = 140 liver tumours. This fact allowed us to develop simple regression models capable of
973 predicting tumour volume given characteristic per-tumour dMRI metric values. The best
974 prediction was obtained using cell size vCS_{MRI} and cell diffusivity $D_{0,I}$ – statistical properties of
975 the tumour cell ensemble – as regressors (correlation $r = 0.396$, $p = 1.26 \cdot 10^{-6}$ between ground
976 truth and tumour volume predictions), surpassing correlations obtained with DKI ADC and K,
977 which only provides $r = 0.244$, $p = 3.68 \cdot 10^{-3}$.

978 Taken as a whole, these findings may suggest that biological properties directly related to
979 the morphology of cancer cells may contribute to explaining, at least in part, the macroscopic
980 phenotype of individual tumours, as for example their size. Therefore, our proposed dMRI
981 approach may offer value to complement standard-of-care radiological assessments, which focus
982 on macroscopic tumour descriptors, providing information that is more closely related to the
983 tumour biology. Nevertheless, we acknowledge again that further validation in larger cohorts is
984 required to confirm these findings, since we were not able to validate the dependence of tumour
985 volume on dMRI metrics histologically, given the limited size of the sample of histological data
986 (Supplementary Table S7). Future work will focus on studying areas of active tumour, since

necrosis – common in the large tumours contained in our advanced patients' cohort – may confound in part the associations, being CS values from both automatic histology processing and dMRI less reliable in highly necrotic areas (see, for example, Fig. 5).

As a final demonstration of the potential utility of our approach, we compared *Diff-in* dMRI metrics across melanoma and CRC liver metastases, being these the two most frequent cancers in our advanced cohort. The analysis showed that *Diff-in* model pins down the exact microstructural property underlying the higher ADC seen in CRC, i.e., the lower IC fraction, being F_{MRI} higher in CRC than in mealona, in agreement with F_{histo}. This finding points once again towards the higher biological specificity of the proposed approach against standard dMRI techniques such as DKI: these, while sensitive to tissue cellulear properties, provide surrogate markers that are more difficult to interpret, and that pool several tissue properties into one number[65].

In our data, ADC and K do offer value beyond standard-of-care tumour size assessment for cancer characterisation, for example showing correlation with cancer cellularity and cell size, a fact that agrees with known literature[57,65], and reveal inter-tumour differences, as discussed boave for CRC and melanoma. However, we point out that histology-informed, biophysical dMRI techniques as the one proposed in this article attempt to provide truly quantitative indices, with a direct biological interpretation, and thus overcome the issues related to the interpretaibility and biological specificity of DKI. Contrarily to our proposed metrics, ADC and K are semi-quantitative and protocol-dependent, and their numerical value can change as function of factors as the dMRI sequence effective diffusion time[65,68]. This may be one of the reason contributing, for example, to the disappointing association between DKI metrics and Ki67 cell proliferation, a fact that confirming the need for developing novel dMRI approaches with high biological specificity for liver tumour imaging – one of the key objectives of this work.

We would like to acknowledge the following potential limitations. Firstly, our sample size is relatively small. This paper provides a first demonstration of the potential utility of metrics from the a practical liver dMRI model, e.g., as markers of cell proliferation or of other biological phenomena that drive tumour phenotypes. Nonetheless, while works proposing related dMRI techniques relied on similar[21,35] (if not even smaller[17,18]), sample sizes, we acknowledge that our exploratory findings require further confirmation in larger cohorts. Confirmation from larger samples is needed to ensure the generalisability of the proposed "*Diff-in*" approach, as well as to rule out additional confounding effects. The limited statistical power of our study (sample of N = 25 MRI-histology acquisitions) implies that additional data from different MRI manufacturers and scanner models are required to better characterise the performances the effect of diffusion time. These could enable us to test, for example, whether the same model selection results would be obtained with shorter diffusion times through oscillating gradient diffusion encoding *in vivo*[17,69], or had different malignancies been included in the *in vivo* data set. These could have included, for example, lymphomas, being this a class of malignancies that involve the liver and in which malignant lymphocites infiltrate the liver parenchyma[70]. Lymphocites are much smaller compared to hepatic cells[17,19]. As a consequence, the dMRI signal from the former is more likely to be sensitive to IC/EC water exchange than the latter, potentially requiring more complex dMRI signal models than those used here, as discussed in more detail below. Similarly, additional data from other animals, beyond mice, would be needed to confirm the generalisbility of the proposed dMRI findings across species.

Secondly, we point out that results from any dMRI-histology comparison should always be taken with care. Here we related dMRI metrics obtained *in vivo* to histological indices from biopsies. While we were able to identify the tumours from which the biopsies were taken, we could

not identify exactly the tumour area that was biopsied. This implies that a biopsy may not be fully representative of the characteristics of an entire tumour, so that the true extent of the associations between dMRI and histology may have been underestimated. Also, histology has its own limitations *per se*, since it provides readouts that may not be fully accurate (a fact that may explain, for example, the discrepancy in HFC rankings seen when comparying in silico and actucal in vivo/*ex vivo* dMRI experiments). For example, routine HE histology is an inherently 2D technique, unlike 3D MRI. Moreover, it is affected by artifacts (e.g., due to dehydration, paraffin embedding, imperfect staining, cutting, etc[41]), and the automatic processing of large fields-of-view requires trading off between sensitivity and specificity. We took steps to mitigate these issues, e.g., by accounting for biases due to tissue shrinkage. Nonetheless, our histology-derived estimates of cell properties are likely biased versions of the true figures, and are also subject to sampling bias compared to full 3D volumetric imaging in MRI. While it is well known that acquiring high-quality, co-localised dMRI and histology data is challenging, especially when dMRI scans are acquired *in vivo* in humans (e.g., prior to tumour excision), future studies are required to confirm the findings presented in this article, and ensure its generalisability across MRI manufactures, scanner models, dMRI encoding schemens, as well as tumour types.

We would also like to acknowledge that the proposed dMRI approach neglects other potentially relevant microstructural properties, such as water exchange between intra-/extra-cellular spaces[19,55], presence of cell size/cytosolic diffusivity distributions[33,69], or intra-compartmental T2 or T1[35]. On the one hand, ignoring these properties may have biased the estimation of F_{MRI} and vCS_{MRI}[19,55]. F_{MRI}, for example, is known to be underestimated when cell permeability is neglected[19,55], and is also confounded by TE changes across imaging sessions or across images from the same session acquired at varying TE[71]. This latter aspects makes F_{MRI}

inherently protocol-dependent, potentially hindering the adoption of the metric as a quantitative biomarker. On the other hand, these properties, overlooked in our modelling framework, may be relevant markers of cellular stress *per se*. In future, we plan to incorporate these properties in our models, while ensuring the clinical feasibility of the dMRI protocols required to fit them. Incorporating IC/EC water exchange modelling and intra-compartmental relaxation effect may be especially relevant to increase the disappointing MRI-histology correlations observed for the IC fraction F for all dMRI models and fitting strategies, being the metric known to be biased by both permeability and relaxation effects[19,55,71].

Conclusions

To conclude, *this study delivers a practical liver dMRI signal model consisting of restricted diffusion within spherical cells, with negligible EC signals, which should be fitted to b-values higher than, approximately, 900 s/mm^2* in vivo. This model offers estimates of cell size and density that are correlated to the underlying histology, and which may provide complementary information to routine standard-of-care imaging, as for example for the characterisation of cell proliferation, or for non-invasive tumour phenotyping *in vivo*. The proposed approach is appropriate in regimes where IC/EC water exchange is slow compared to the measurement diffusion time, as for example to image liver tumours containing predominantly large cells, e.g., 20-30 μm in diameter for diffusion times not exceeding 40 ms. In applications where high cell permeability is expected, as, for example, to image tumours rich in small cells, as lymphocites[19] (e.g., in immunotherapy or in lymphomas), care would be needed in interpreting metrics from the proposed model, due to potential biases coming from unaccounted IC/EC water exchange, and non-negligible EC signals.

Supplementary Information

This article includes Supplementary Figures S1 to S27, Supplementary Tables S1 to S7, and Supplementary Data S1 to S9. Supplementary Data S1: CSV file with MRI and histology data for model selection. Supplementary Data S2: CSV file with MRI and histology data for model selection based on high b-value images only. Supplementary Data S3: CSV file with MRI and histology metrics for each individual tumour. Supplementary Data S4: CSV file with cohort information, histological and mean MRI metrics across all tumour voxels. Supplementary Data S5: CSV file with cohort information, histological and mean ComBat-harmonised MRI metrics across all tumour voxels. Supplementary Data S6: zipped folder containing the code generating Table 1 and Figure 3. Supplementary Data S7: zipped folder containing the code generating Figure 6. Supplementary Data S8: zipped folder containing the code generating Figure 7. Supplementary Data S9: zipped folder containing the code generating Figure 8.

Data availability

Raw MRI and histological images from the mouse livers are available online through the Radiomics Group web site[72] . They can also be accessed by contacting the corresponding authors directly.

Raw MRI and histological images of patients cannot be made freely available due to ethical restrictions at this stage (the release of the linked MRI-histology data was not contemplated in the current institutional project protocol, ethical approval and patient's consent form). Scientists

interested in accessing them for research purposes can contact the corresponding authors to establish appropriate data sharing agreements at institutional level. Requests should be sent via e-mail to Dr Raquel Pérez López (rperez@vhio.net). They should include: (i) a short of description of the reasons for requesting the data (project description, hypotheses to be tested, expected contribution of the data towards hypothesis testing); (ii) contact details of the researcher(s) in charge of the data analysis, of their direct principal investigator/line manager, and of a representative of the legal team of the institution. A first response will be provided in reasonable time (approximately 10 working days), from which legal steps may be coordinated for the stipulation of an official data sharing agreement. Requests will be assessed on a case-by-case basis and may be denied if not deemed of satisfactory scientific quality, unclear or unethical, or for any other relevant reason. For approved data sharing, the data must be used solely for the purposes stipulated in the data sharing agreement and may not be used for commercial purposes. Requests will only be considered from not-for-profit scientific or research institutions. The submission of a request shall be considered the initiation of a formal collaboration between the requesting institution and VHIO.

Spreadsheets with per-patient and per-tumour values of MRI and histological metrics, jointly with salient demographical information, are included as Supplementary Data S1 to S5. Supplementary Data S6 contains the code used to generate Table 1 and Figure 3. Supplementary Data S7 contains the code used to generate Figure 6. Supplementary Data S8 contains the code used to generate Figure 7. Supplementary Data S9 contains the code used to generate Figure 8.

Code availability

Python routines enabling the computation of the diffusion MRI and histological metrics presented in this article are freely available online[73] (command line tools *pgse2sphereinex.py*, *mri2micro_dictml.py* and *getPatchMapFromQuPath.py*). Python scripts reproducing Fig. 3 and Fig. 6 to 8 are included as Supplementary Data S6 to S9. All other scripts used to pre-process the MR and histological images can be obtained by contacting the corresponding authors.

Author contributions

Conceptualization: F.G., R.P.L., K.B., M.P., E.G., R.T., P.N., J.M.. **Methodology:** F.G., R.P.L., K.B., M.P., A.G, D.N.G.. **Investigation:** F.G., R.P.L., K.B., I.C.S., I.B., S.S., G.S., A.G., D.N.G., C.M., V.G., J.F.C., X.M., R.M., N.R., M.E., M.Vie., R.T., P.N., J.M., E.G.. **Resources:** R.P.L., F.G., P.N., J.M., E.G., N.R., M.E., V.G., M.Vid., P.G.P.G, I.B.. **Formal analysis:** F.G.. **Visualization:** F.G.. **Validation:** F.G.. **Software:** F.G., K.B., A.G., D.N.G., C.M.. **Data curation:** F.G., R.P.L., K.B., A.V., G.S., I.C.S., A.G, D.N.G.. **Project administration:** F.G., R.P.L., K.B., E.G., P.N., R.T., J.M., I.C.S.. **Funding acquisition:** R.P.L., E.G., R.T., P.N., J.M., F.G., K.B., I.C.S.. **Supervision:** F.G., R.P.L., E.G., R.T., P.N., J.M.. **Writing—original draft:** F.G., R.P.L., K.B., M.P.. **Writing—review & editing:** all authors.

Competing interests

This study received funding from AstraZeneca. M.Vid. works for Siemens Healthineers. P.G.P.G. works for GE HealthCare. K.B. worked as a researcher at the Vall d'Hebron Institute of Oncology (Barcelona), and is now an employee of AstraZeneca. AstraZeneca, Siemens and General Electric did not influence the acquisition and analysis of the data, the interpretation of the

results, or the decision to submit the manuscript in its current form for publication. All other authors declare no competing interests.

Acknowledgments

We thank the whole medical oncology, radiology, pathology, molecular biology, clinical trial, and IT teams at the Vall d'Hebron University Hospital and at the Vall d'Hebron Institute of Oncology in Barcelona (Spain), without whom this study would not have been possible. We are also thankful to the Vall d'Hebron Radiology department and to the ASCIRES CETIR clinical team for their assistance, and to past and present members of the Radiomics group for useful discussion and advice. Finally, we would like to express our sincere gratitude to all patients and their families for dedicating their time to research. VHIO would like to acknowledge: the State Agency for Research (Agencia Estatal de Investigación) for the financial support as a Center of Excellence Severo Ochoa (CEX2020-001024-S/AEI/10.13039/501100011033), the Cellex Foundation for providing research facilities and equipment and the CERCA Programme from the Generalitat de Catalunya for their support on this research. This study has been funded by Instituto de Salud Carlos III (ISCIII) through the project "PI21/01019" and co-funded by the European Union. Data acquisition was also supported by PREdICT, sponsored by AstraZeneca. This study has been co-funded by the European Regional Development Fund/European Social Fund 'A way to make Europe' (to R.P.L.), and by the Comprehensive Program of Cancer Immunotherapy & Immunology (CAIMI), funded by the Banco Bilbao Vizcaya Argentaria Foundation Foundation (FBBVA, grant 89/2017). R.P.L is supported by the "la Caixa" Foundation CaixaResearch Advanced Oncology Research Program, the Prostate Cancer Foundation (18YOUN19), ISCIII (PI18/01395), a CRIS Foundation Talent Award (TALENT19-05), the FERO Foundation through

the XVIII Fero Fellowship for Oncological Research, the Asociación Española Contra el Cancer (AECC) (PRYCO211023SERR) and the Generalitat de Catalunya Agency for Management of University and Research Grants of Catalonia (AGAUR) (2023PROD00178). The project that gave rise to these results received the support of a fellowship from ”la Caixa” Foundation (ID 100010434). The fellowship code is “LCF/BQ/PR22/11920010" (funding F.G. and A.V.) and “LCF/BQ/PI20/11760033” (funding I.C.S). I.C.S. also receives the support of the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 847648. This research has received support from the Beatriu de Pinós Postdoctoral Program from the Secretariat of Universities and Research of the Department of Business and Knowledge of the Government of Catalonia, and the support from the Marie Sklodowska-Curie COFUND program (BP3, contract number 801370; reference 2019 BP 00182) of the H2020 program (to K.B.). M.P. is supported by the UKRI Future Leaders Fellowship MR/T020296/2. A.G. is supported by a Severo Ochoa PhD fellowship (PRE2022-102586). C.M. is funded by the Asociación Española Contra el Cancer (AECC) (PRYCO211023SERR). Funding supported personnel, data acquisition and processing, and dissemination. Funders did not influence the content of the manuscript and the decision to submit it in its current form for publication.

References

1. Weiskopf, N., Edwards, L. J., Helms, G., Mohammadi, S. & Kirilina, E. Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. *Nature Reviews Physics* **3**, 570–588 (2021).
2. Jarrett, A. M. *et al.* Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting. *Nat. Protoc.* **16**, 5309–5338 (2021).

3. Kiselev, V. G. Fundamentals of diffusion MRI physics. *NMR Biomed.* **30**, (2017).
4. Novikov, D. S., Fieremans, E., Jespersen, S. N. & Kiselev, V. G. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR Biomed.* **32**, e3998 (2019).
5. Le Bihan, D. Looking into the functional architecture of the brain with diffusion MRI. *Nat. Rev. Neurosci.* **4**, 469–480 (2003).
6. Panagiotaki, E. *et al.* Noninvasive quantification of solid tumor microstructure using VERDICT MRI. *Cancer Res.* **74**, 1902–1912 (2014).
7. Veraart, J. *et al.* Noninvasive quantification of axon radii using diffusion MRI. *Elife* **9**, e49855 (2020).
8. Cohen-Adad, J. *et al.* Generic acquisition protocol for quantitative MRI of the spinal cord. *Nat. Protoc.* **16**, 4611–4632 (2021).
9. Baxter, G. C., Graves, M. J., Gilbert, F. J. & Patterson, A. J. A Meta-analysis of the Diagnostic Performance of Diffusion MRI for Breast Lesion Characterization. *Radiology* **291**, 632–641 (2019).
10. Donato, H., França, M., Candelária, I. & Caseiro-Alves, F. Liver MRI: From basic protocol to advanced techniques. *Eur. J. Radiol.* **93**, 30–39 (2017).
11. Tapper, E. B. & Lok, A. S.-F. Use of Liver Imaging and Biopsy in Clinical Practice. *N. Engl. J. Med.* **377**, 756–768 (2017).
12. Tsilimigras, D. I. *et al.* Liver metastases. *Nat Rev Dis Primers* **7**, 27 (2021).
13. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
14. Buikhuisen, J. Y., Torang, A. & Medema, J. P. Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges. *Oncogenesis* **9**, 66 (2020).

15. Vitale, I., Shema, E., Loi, S. & Galluzzi, L. Intratumoral heterogeneity in cancer progression and response to immunotherapy. *Nat. Med.* **27**, 212–224 (2021).
16. O'Connor, J. P. B. *et al.* Imaging biomarker roadmap for cancer studies. *Nat. Rev. Clin. Oncol.* **14**, 169–186 (2017).
17. Jiang, X., Xu, J. & Gore, J. C. Mapping hepatocyte size in vivo using temporal diffusion spectroscopy MRI. *Magn. Reson. Med.* **84**, 2671–2683 (2020).
18. Panagiotaki, E. *et al.* Microstructural Characterization of Normal and Malignant Human Prostate Tissue With Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours Magnetic Resonance Imaging. *Invest. Radiol.* **50**, 218 (2015).
19. Gardier, R. *et al.* Cellular Exchange Imaging (CEXI): Evaluation of a diffusion model including water exchange in cells using numerical phantoms of permeable spheres. *Magn. Reson. Med.* (2023) doi:10.1002/mrm.29720.
20. Hoffmann, E. *et al.* Profiling specific cell populations within the inflammatory tumor microenvironment by oscillating-gradient diffusion-weighted MRI. *J Immunother Cancer* **11**, e006092 (2023).
21. Jiang, X. *et al.* MRI of tumor T cell infiltration in response to checkpoint inhibitor therapy. *J Immunother Cancer* **8**, (2020).
22. Novikov, D. S., Kiselev, V. G. & Jespersen, S. N. On modeling. *Magn. Reson. Med.* **79**, 3172–3193 (2018).
23. Ye, Z. *et al.* Diffusion Histology Imaging Combining Diffusion Basis Spectrum Imaging (DBSI) and Machine Learning Improves Detection and Classification of Glioblastoma Pathology. *Clin. Cancer Res.* **26**, 5388–5399 (2020).

24. Howard, A. F. *et al.* Joint modelling of diffusion MRI and microscopy. *Neuroimage* **201**, 116014 (2019).
25. Zhu, S. *et al.* Imaging the structural connectome with hybrid MRI-microscopy tractography. *Med. Image Anal.* **102**, 103498 (2025).
26. Grussu, F. *et al.* Histology-informed biophysical diffusion MRI model selection for enhanced liver cancer immunotherapy assessment. in *Proceedings of the 2024 annual meeting of the International Society for Magnetic Resonance in Medicine* 0649 (2024).
27. Xu, J. *et al.* Magnetic resonance imaging of mean cell size in human breast tumors. *Magn. Reson. Med.* **83**, 2002–2014 (2020).
28. Reynaud, O. *et al.* Pulsed and oscillating gradient MRI for assessment of cell size and extracellular space (POMACE) in mouse gliomas. *NMR Biomed.* **29**, 1350–1363 (2016).
29. Reynaud, O. Time-Dependent Diffusion MRI in Cancer: Tissue Modeling and Applications. *Frontiers in Physics* **5**, (2017).
30. Le Bihan, D. *et al.* MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* **161**, 401–407 (1986).
31. Li, Y. T. *et al.* Liver intravoxel incoherent motion (IVIM) magnetic resonance imaging: a comprehensive review of published data on normal values and applications for fibrosis and tumor evaluation. *Quant. Imaging Med. Surg.* **7**, 59–78 (2017).
32. Balinov, B., Jonsson, B., Linse, P. & Soderman, O. The NMR Self-Diffusion Method Applied to Restricted Diffusion. Simulation of Echo Attenuation from Molecules in Spheres and between Planes. *J. Magn. Reson. A* **104**, 17–25 (1993).

33. Grussu, F. *et al.* Diffusion MRI signal cumulants and hepatocyte microstructure at fixed diffusion time: Insights from simulations, 9.4T imaging, and histology. *Magn. Reson. Med.* **88**, 365–379 (2022).
34. Xu, J. *et al.* Diffusion time dependency of extracellular diffusion. *Magn. Reson. Med.* **89**, 2432–2440 (2023).
35. Palombo, M. *et al.* Joint estimation of relaxation and diffusion tissue parameters for prostate cancer with relaxation-VERDICT MRI. *Sci. Rep.* **13**, 2999 (2023).
36. Panagiotaki, E. *et al.* Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *Neuroimage* **59**, 2241–2254 (2012).
37. Veraart, J. *et al.* Denoising of diffusion MRI using random matrix theory. *Neuroimage* **142**, 394–406 (2016).
38. Kellner, E., Dhital, B., Kiselev, V. G. & Reisert, M. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn. Reson. Med.* **76**, 1574–1581 (2016).
39. Jerome, N. P. *et al.* Extended T2-IVIM model for correction of TE dependence of pseudo-diffusion volume fraction in clinical diffusion-weighted magnetic resonance imaging. *Phys. Med. Biol.* **61**, N667–N680 (2016).
40. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
41. Boonstra, H., Oosterhuis, J. W., Oosterhuis, A. M. & Fleuren, G. J. Cervical tissue shrinkage by formaldehyde fixation, paraffin wax embedding, section cutting and mounting. *Virchows Arch. A Pathol. Anat. Histopathol.* **402**, 195–201 (1983).
42. Garyfallidis, E. *et al.* Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.* **8**, 8 (2014).

43. Ourselin, S., Roche, A., Subsol, G., Pennec, X. & Ayache, N. Reconstructing a 3D structure from serial histological sections. *Image Vis. Comput.* **19**, 25–31 (2001).
44. Andersson, J. L. R., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* **20**, 870–888 (2003).
45. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* **12**, 26–41 (2008).
46. Rafael-Patino, J. *et al.* Robust Monte-Carlo Simulations in Diffusion-MRI: Effect of the Substrate Complexity and Parameter Choice on the Reproducibility of Results. *Front. Neuroinform.* **14**, 8 (2020).
47. Schwarz, G. Estimating the Dimension of a Model. *aos* **6**, 461–464 (1978).
48. Seabold, S. & Perktold, J. Statsmodels: Econometric and statistical modeling with python. in *Proceedings of the Python in Science Conference* 92–96 (SciPy, 2010).
49. Fortin, J.-P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* **167**, 104–120 (2018).
50. Neuman, C. H. Spin echo of spins diffusing in a bounded medium. *J. Chem. Phys.* **60**, 4508–4511 (1974).
51. Conlin, C. C. *et al.* A Multicompartmental Diffusion Model for Improved Assessment of Whole-Body Diffusion-weighted Imaging Data and Evaluation of Prostate Cancer Bone Metastases. *Radiol Imaging Cancer* **5**, e210115 (2023).
52. Jensen, J. H., Helpern, J. A., Ramani, A. & Lu, H. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *in Medicine: An …* (2005).

53. Hectors, S. J. *et al.* Advanced Diffusion-weighted Imaging Modeling for Prostate Cancer Characterization: Correlation with Quantitative Histopathologic Tumor Tissue Composition-A Hypothesis-generating Study. *Radiology* **286**, 918–928 (2018).
54. Ferizi, U. *et al.* A ranking of diffusion MRI compartment models with in vivo human brain data. *Magn. Reson. Med.* **72**, 1785–1792 (2014).
55. Jiang, X., Devan, S. P., Xie, J., Gore, J. C. & Xu, J. Improving MR cell size imaging by inclusion of transcytolemmal water exchange. *NMR Biomed.* **35**, e4799 (2022).
56. Rosenkrantz, A. B. *et al.* Assessment of hepatocellular carcinoma using apparent diffusion coefficient and diffusion kurtosis indices: preliminary experience in fresh liver explants. *Magn. Reson. Imaging* **30**, 1534–1540 (2012).
57. Tang, W.-J. *et al.* Evaluation of the Effects of Anti-PD-1 Therapy on Triple-Negative Breast Cancer in Mice by Diffusion Kurtosis Imaging and Dynamic Contrast-Enhanced Imaging. *J. Magn. Reson. Imaging* **56**, 1912–1923 (2022).
58. Martin, N. C. *et al.* Functional analysis of mouse hepatocytes differing in DNA content: volume, receptor expression, and effect of IFNgamma. *J. Cell. Physiol.* **191**, 138–144 (2002).
59. Zhou, Z., Xu, M.-J. & Gao, B. Hepatocytes: a key cell type for innate immunity. *Cell. Mol. Immunol.* **13**, 301–315 (2016).
60. Morawski, M. *et al.* Developing 3D microscopy with CLARITY on human brain tissue: Towards a tool for informing and validating MRI-based histology. *Neuroimage* **182**, 417–428 (2018).
61. Jelescu, I. O., Veraart, J., Fieremans, E. & Novikov, D. S. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR Biomed.* **29**, 33–47 (2016).
62. Höppener, D. J. *et al.* The relationship between primary colorectal cancer histology and the histopathological growth patterns of corresponding liver metastases. *BMC Cancer* **22**, 911 (2022).

63. Barnhill, R. *et al.* Replacement and desmoplastic histopathological growth patterns in cutaneous melanoma liver metastases: frequency, characteristics, and robust prognostic value. *J. Pathol. Clin. Res.* **6**, 195–206 (2020).
64. Nilsson, M. *et al.* Mapping prostatic microscopic anisotropy using linear and spherical b-tensor encoding: A preliminary study. *Magn. Reson. Med.* **86**, 2025–2033 (2021).
65. Fokkinga, E. *et al.* Advanced Diffusion-Weighted MRI for Cancer Microstructure Assessment in Body Imaging, and Its Relationship With Histology. *J. Magn. Reson. Imaging* **60**, 1278–1304 (2024).
66. Derenzini, M. *et al.* Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. *J. Pathol.* **191**, 181–186 (2000).
67. Warth, A. *et al.* Tumour cell proliferation (Ki-67) in non-small cell lung cancer: a critical reappraisal of its prognostic role. *Br. J. Cancer* **111**, 1222–1229 (2014).
68. Lee, H.-H., Papaioannou, A., Novikov, D. S. & Fieremans, E. In vivo observation and biophysical interpretation of time-dependent diffusion in human cortical gray matter. *Neuroimage* **222**, 117054 (2020).
69. Xu, J. *et al.* MRI-cytometry: Mapping nonparametric cell size distributions using diffusion MRI. *Magn. Reson. Med.* **85**, 748–761 (2021).
70. Anthony, P. P., Sarsfield, P. & Clarke, T. Primary lymphoma of the liver: clinical and pathological features of 10 patients. *J. Clin. Pathol.* **43**, 1007–1013 (1990).
71. Jiang, X. *et al.* Joint estimation of compartment-specific T2 relaxation and tumor microstructure using multi-TE IMPULSED MRI. *Magn. Reson. Med.* **93**, 96–107 (2025).
72. Perez-Lopez., R. Released data bases. *VHIO Radiomics group* https://radiomicsgroup.github.io/data (2025).

73. Grussu, F. BodyMRITools. *Github repository* https://github.com/fragrussu/bodymritools (2024).

Figures

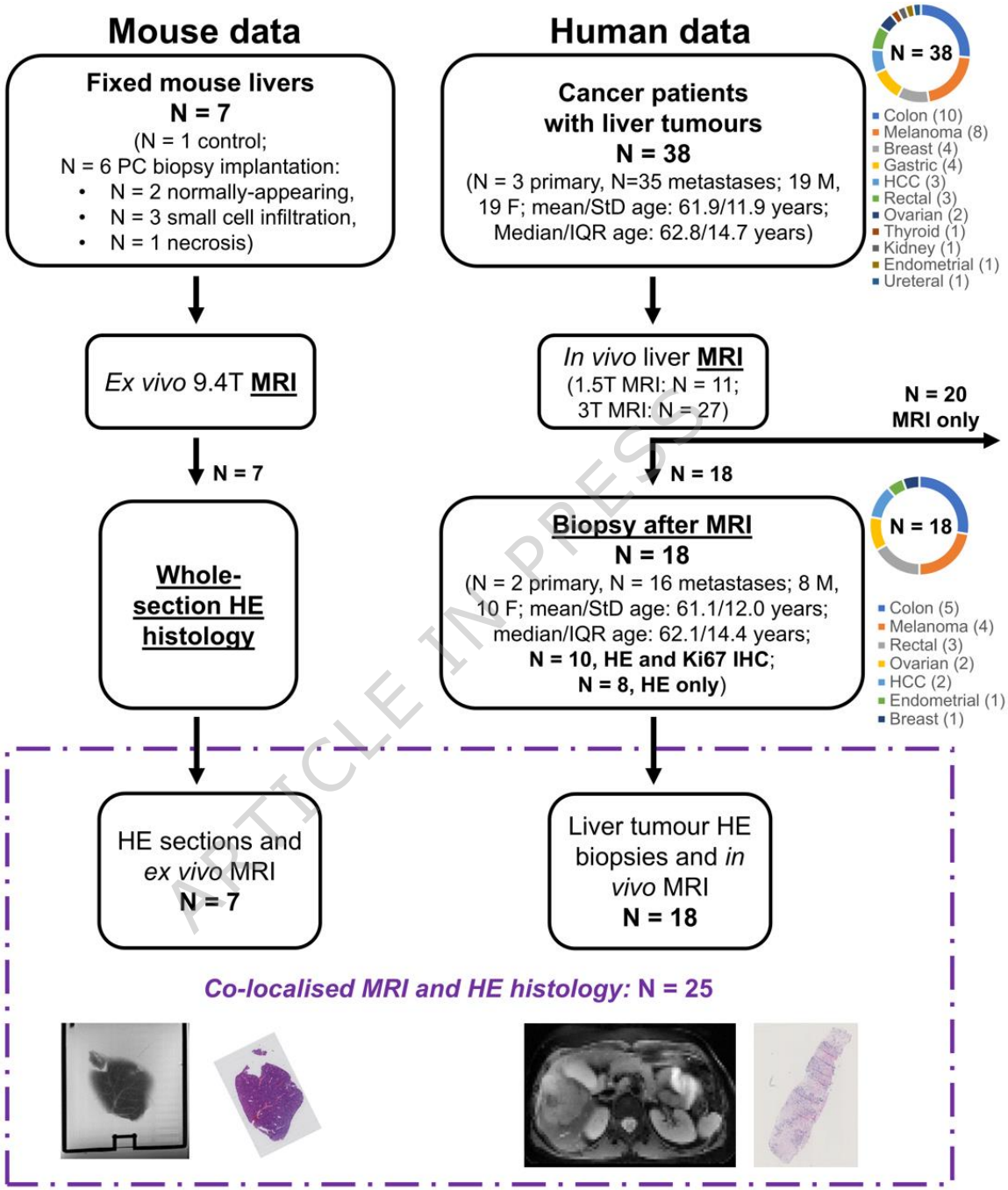


Fig. 1. Illustration of the liver MRI and histology data used in the study. Our data set consisted of preclinical mouse data and clinical human data. The mouse data encompasses dMRI scans of seven fixed livers from mice (six implanted with tissue from biopsies of patients suffering from

prostate cancer (PC); one without any implantation). We scanned the livers *ex vivo* on a 9.4T system, and obtained HE histological sections at known position. The clinical data includes *in vivo* liver dMRI scans performed on 38 patients suffering from advanced liver tumours. Scans were performed on clinical 1.5T and 3T MRI systems, and a HE-stained biopsy from one of the images tumours was collected after MRI in 18 cases. In 10 out of these 18, Ki67 immunohistochemsitry (IHC), demonstrating cell proliferation, was available beyond routine HE. In the figure, PC stands for prostate cancer; HCC for hepatocellular carcinoma; IHC for immunohistochemsitry.

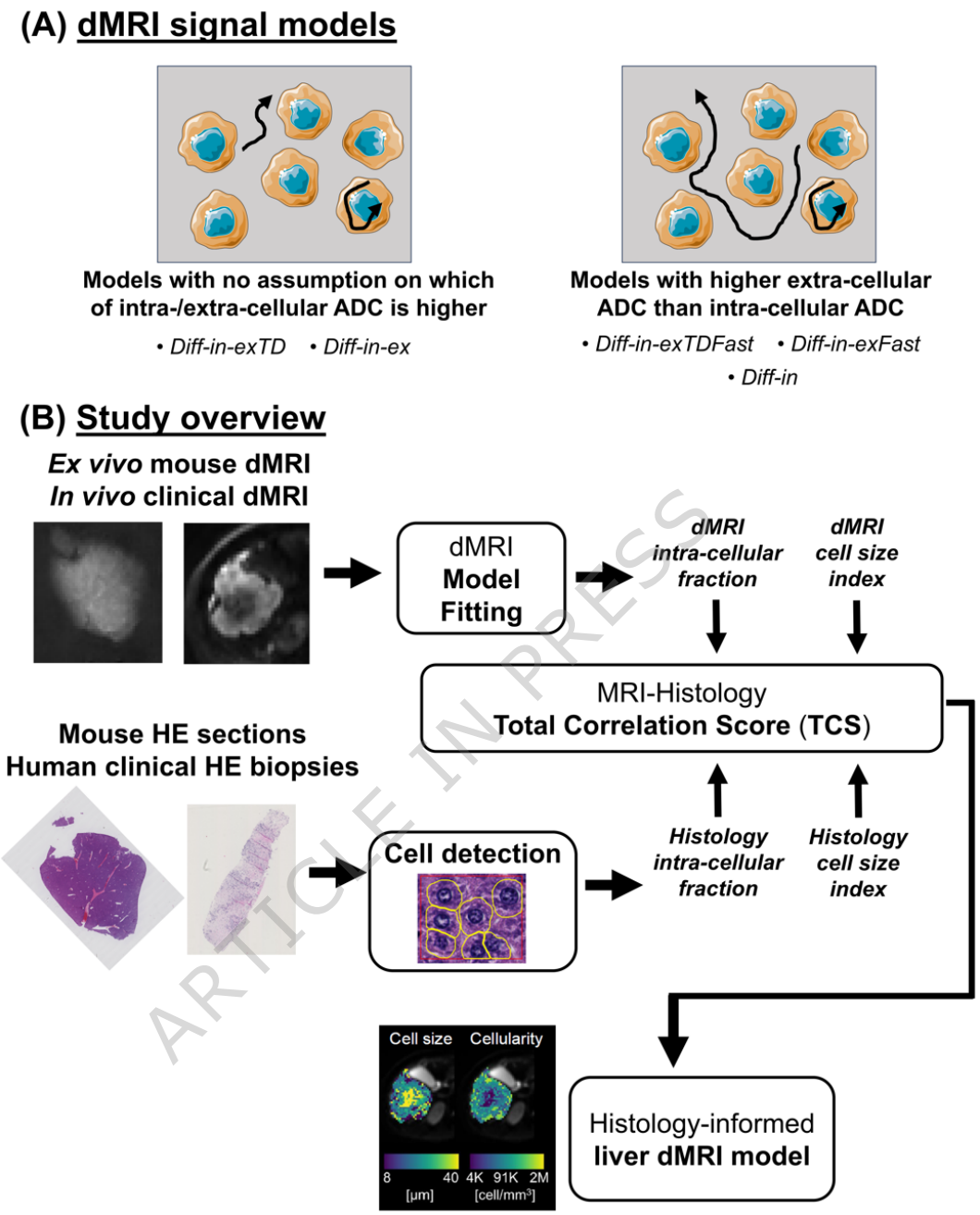


Fig. 2. Description of the dMRI signal models and study overview. (**A**), top: cartoon illustrating the two families of dMRI models considered in this study, consisting of 1) models with no assumption of which of intra/extra-cellular ADC is higher, and 2) models where the extra-cellular ADC is hypothesised to be higher than the intra-cellular ADC. (**B**), bottom: study overview. We analysed dMRI data from fixed mouse livers and from cancer patients imaged *in vivo* to derive

estimates of intra-cellular fraction and of cell size. In parallel, we processed histological material from the same tissues (whole-liver sections for the preclinical mouse data; biopsies from one of the imaged tumours for the clinical data), and derived the histological counterparts of such dMRI metrics. We compared dMRI and histological cell size and intra-cellular fraction to select the dMRI model featuring the best fidelity to histology. In Fig. 2, pictures from Servier Medical Art have been used. Image(s) provided by Servier Medical Art (https://smart.servier.com), licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

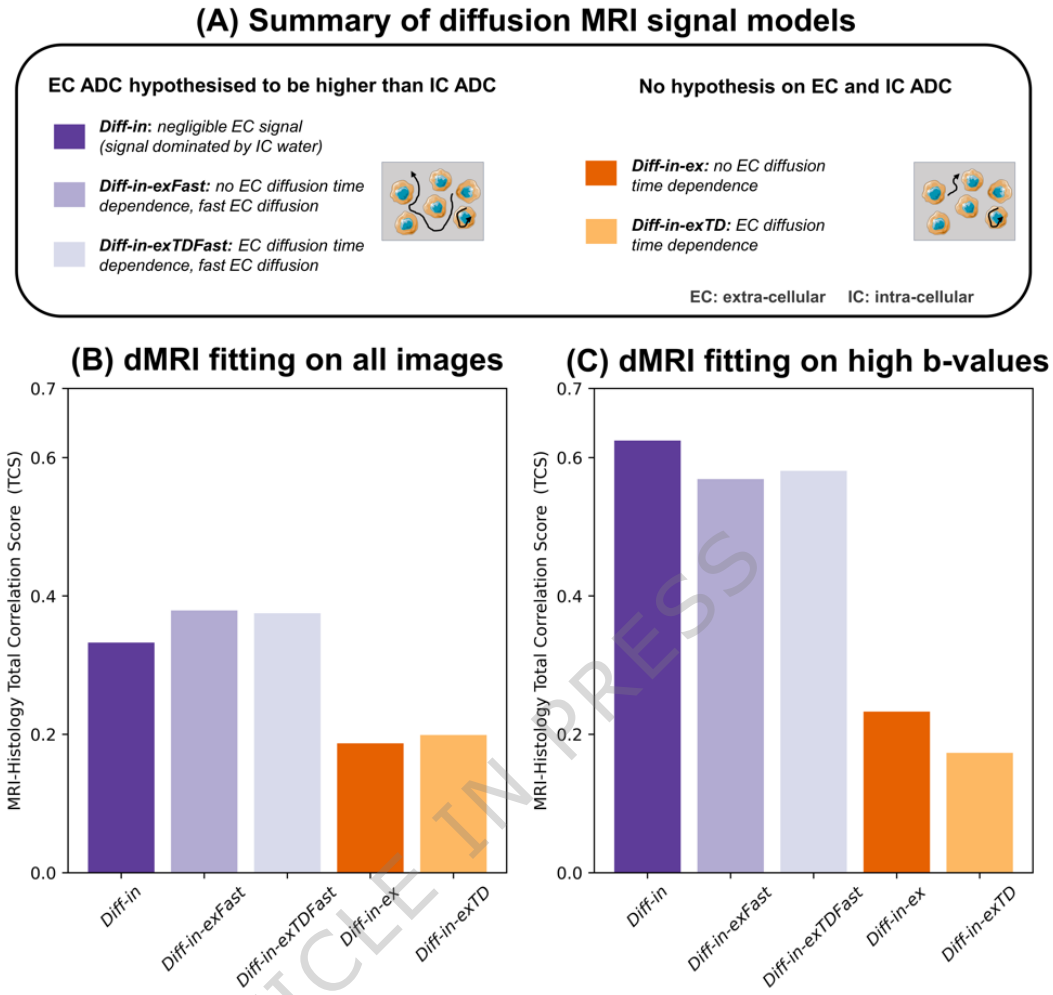


Figure 3. Biophysical dMRI signal model selection based on the MRI-histology Total Correlation Score (TCS). (A): panel summarising the salient differences between the biophysical dMRI models compared in this study. Models can be divided in two families, i.e.: i) models where it is hypothesised that the extra-cellular ADC is higher than the intra-cellular ADC, and ii) models with no hypothesis on which, between intra-/extra-cellular ADC is higher. Violet shades are used for the first family (models *Diff-in*, *Diff-in-exFast* and *Diff-in-exTDFast*), while orange shades for the second family (models *Diff-in-ex* and *Diff-in-exTD*). **(B)**: values of TCS for all models, as obtained by fitting them on the whole set of images featuring negligible vascular signal contribution (b > 1000 s/mm^2 in the fixed mouse livers; b > 100 s/mm^2 *in vivo*). **(C)**: TCS values

for all models when these are fitted only on high b-value images (b > 1800 s/mm^2 in the fixed mouse livers; b > 900 s/mm^2 *in vivo*). TCS was computed using a sample size of N = 25 MRI-histology data points. In this figure, pictures from Servier Medical Art have been used. Image(s) provided by Servier Medical Art (https://smart.servier.com), licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

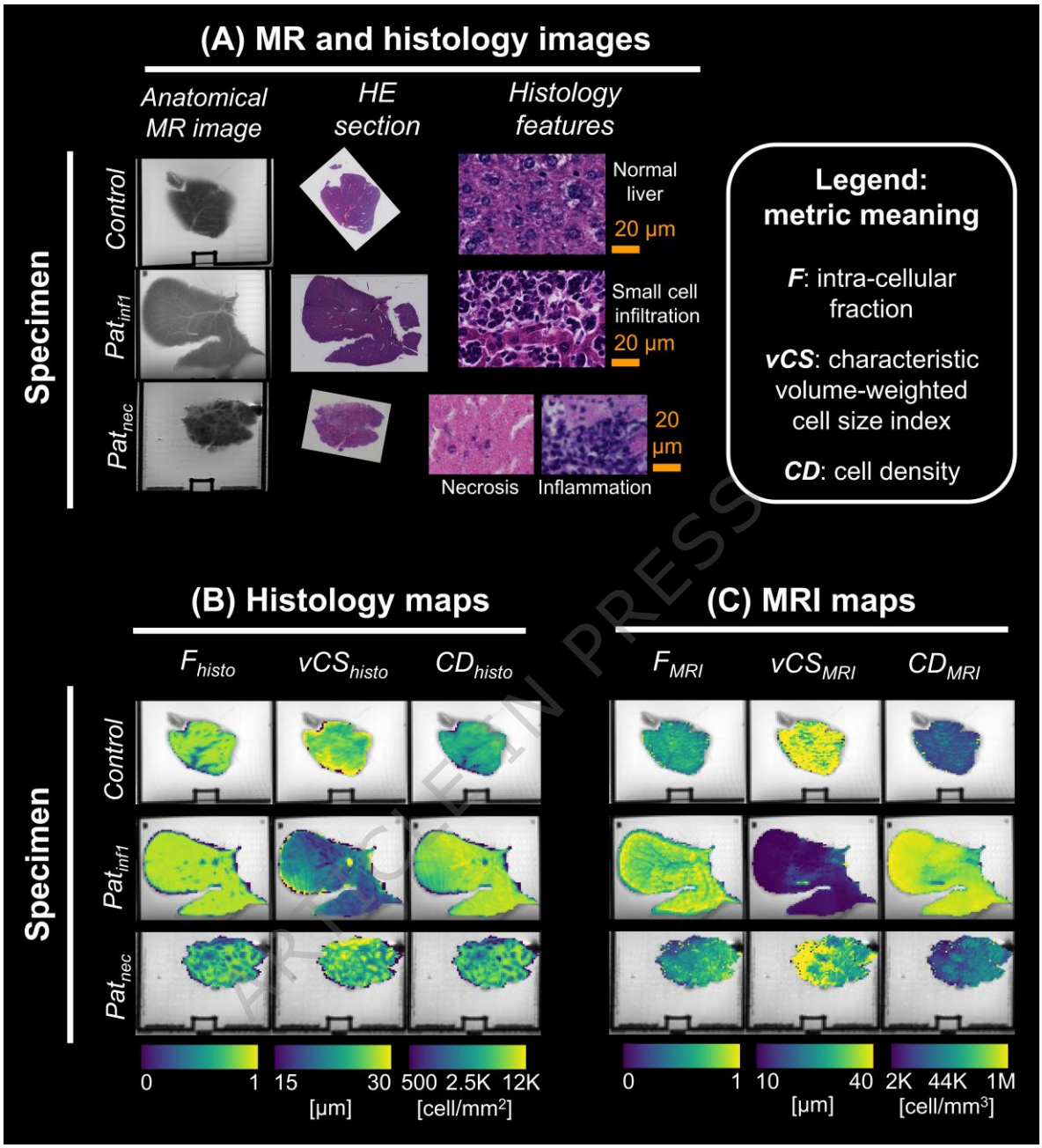


Fig. 4. Maps from the selected dMRI model *Diff-in* with their histological counterparts in the fixed mouse livers scanned at 9.4T *ex vivo*. The figure reports MRI and histology data for 3 specimens, representative of the 3 microstructural phenotypes observed in our *ex vivo* data set, namely: normal liver structures (*Control* case); a proliferative process, characterized by infiltration of small cells (Pat_{inf1} case); necrosis and inflammation (Pat_{nec} case). For all specimens, the

following is shown. **(A)**, *top left*: a high-resolution T2-w anatomical scan is shown next to the corresponding HE section, with histological details. **(B)**, *bottom left*: histological maps warped to the dMRI space (intra-cellular patch area fraction F_{histo}; volume-weighted mean cell size index vCS_{histo}; cell density per unit patch area CD_{histo}). **(C)**, *bottom right*: dMRI maps F_{MRI}, vCS_{MRI} and CD_{MRI} from the selected dMRI signal model (model *Diff-in*, fitted to high *b*-value images, i.e., $b > 1800$ s/mm^2).

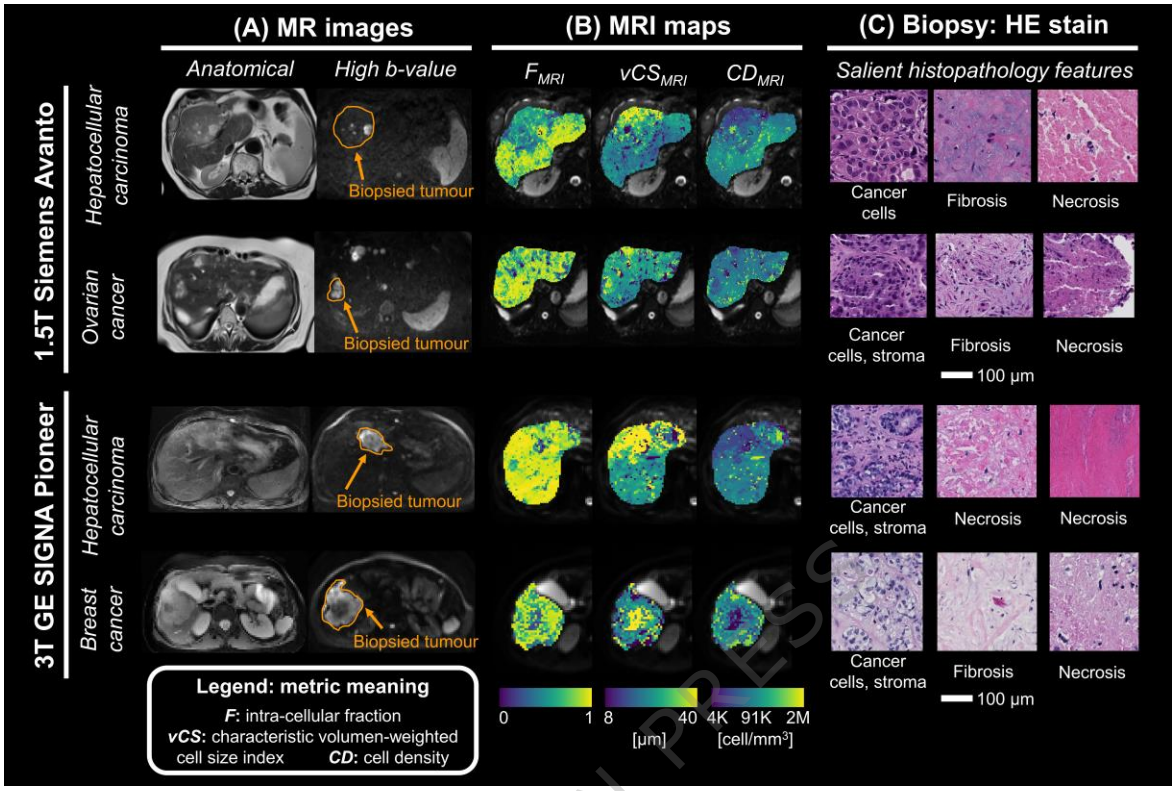


Fig. 5. Examples of maps from the proposed dMRI model *Diff-in* in liver tumours of patients scanned at 1.5T and 3T *in vivo*, with co-localised biopsies. MRI maps are shown in a biopsied liver tumour in two patients for each MRI scanner, arranged along rows. **(A)**: examples of slices from the high-resolution anatomical T2-w image and from a high b-value image, with biopsied tumour outlined. **(C)**: maps from the selected model (*Diff-in*, fitted to high b-value images $b > 900$ s/mm^2). From left to right: intra-cellular signal fraction F_{MRI}; volume-weighted mean cell size index vCS_{MRI}; cell density per unit volume CD_{MRI}. **(C)**: histological details from the HE-stained biopsy. For the 1.5T Siemens scanner (first and second rows from top) we report: patient 6 (primary hepatocellular carcinoma) and patient 3 (liver metastases from ovarian cancer). For the 3T GE scanner (third and fourth rows from top) we report: patient 24 (primary hepatocellular carcinoma (HCC)) and patient 30 (liver metastases from breast cancer).

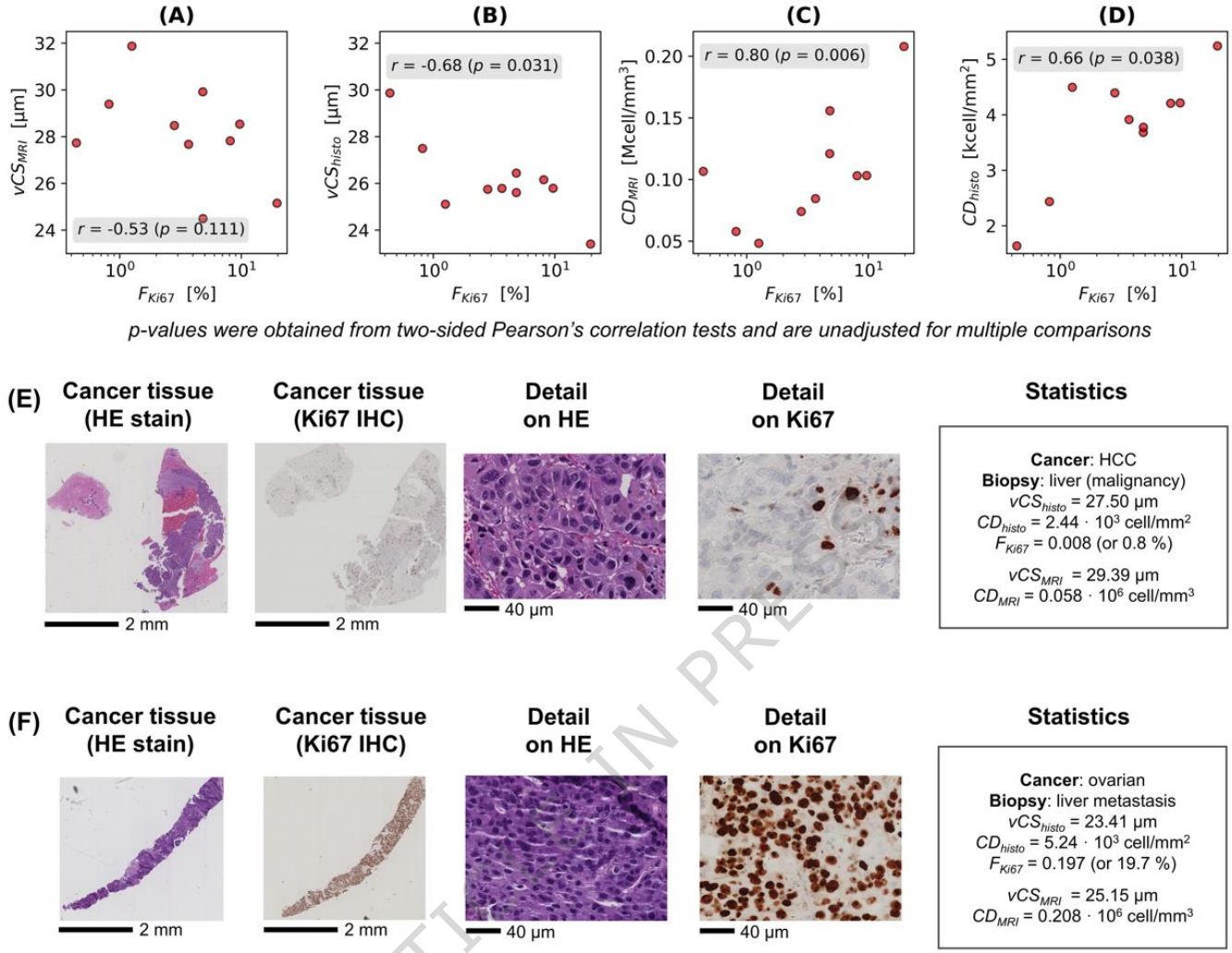
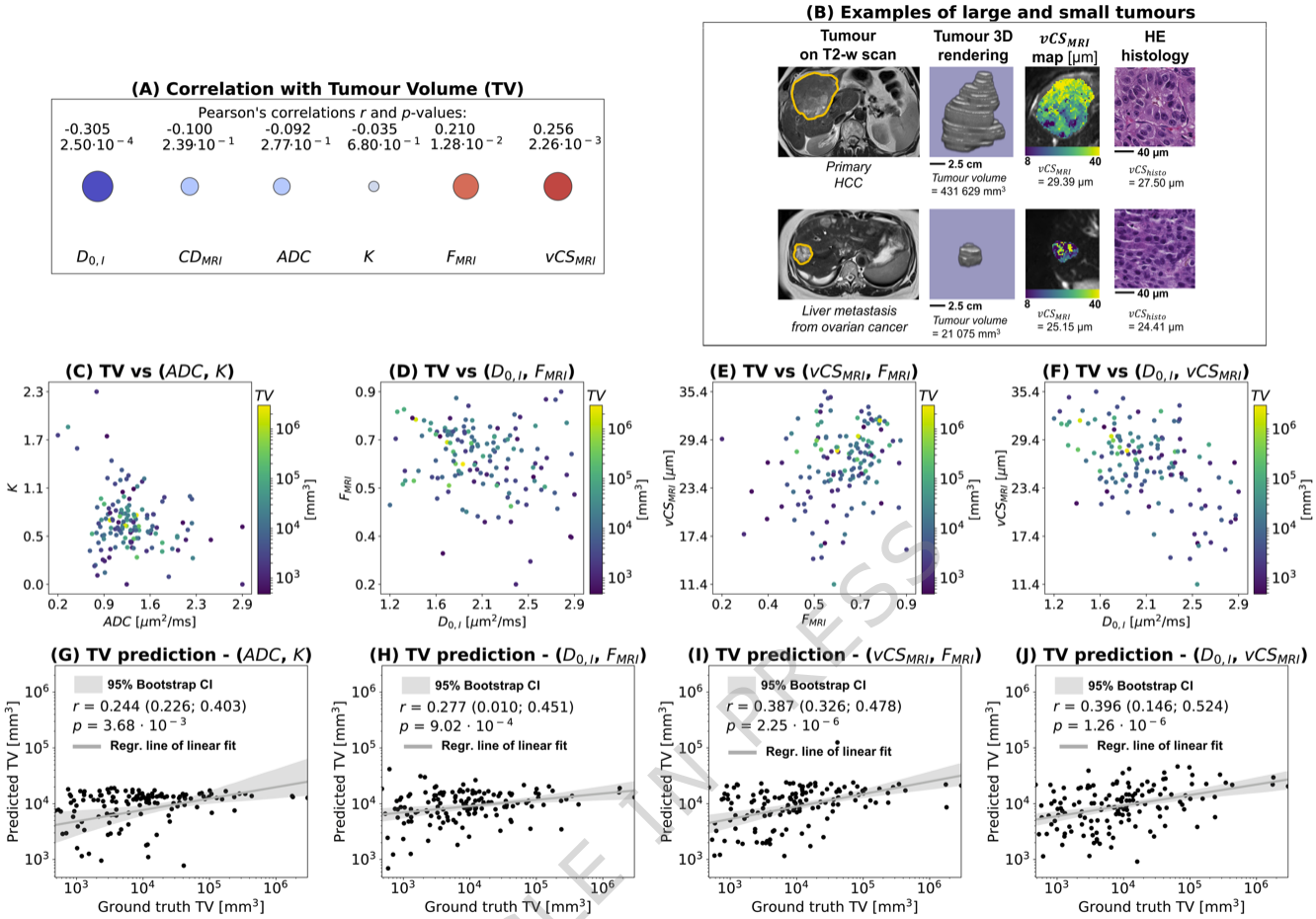


Fig. 6. Association between dMRI metrics and cell proliferation within the biopsied tumours in patients. Top row: scatter plots and Pearson's correlation coefficients between F_{Ki67} (strength of the IHC stain demonstrating proliferation within the biopsied tumour, here expressed in percentage terms) and indices of cell size and density from both MRI and HE histology (sample size: N = 10). From left to right, **(A):** scatter plot and correlation between vCS_{MRI} and F_{Ki67}; **(B):** between vCS_{histo} and F_{Ki67}; **(C):** between CD_{MRI} and F_{Ki67}; **(D):** between CD_{histo} and F_{Ki67}. In the scatter plots, p-values were obtained from a two-sided Pearson's correlation test, and are unadjusted for multiple comparisons. **(E)**, middle row: example of HE and Ki67 stains from a slowly-proliferating tumour (primary hepatocellular carcinoma,

F_{Ki67} of 0.8%). **(F)**, bottom row: example of HE and Ki67 stains from a fast-proliferating tumour (ovarian cancer liver metastasis, F_{Ki67} of 19.7%), with a summary of salient MRI and histological metrics. MRI maps of the two patients are shown in Fig. 5 and Fig. 7 (both patients were scanned on the 1.5T system). In **(E)** and **(F)**, length bars are reported (2mm for the HE and IHC tissue biopsies; 40 μm for the zoomed details visualising cells).



p-values are unadjusted for multiple comparisons and were obtained from two-sided Pearson's correlation tests

Fig. 7. Association between dMRI metrics and liver tumour volume in patients. (A): Pearson's correlation coefficients between liver tumour volume and per-tumour mean values of *Diff-in* and DKI metrics. **(B)**: examples of large and small tumours from a primary hepatocellular carcinoma (large) and a metastatic ovarian cancer (small), with corresponding vCS_{MRI} dMRI metric of cell size and HE histology. **(C) to (E)**: visualisation of tumour volume as a bi-dimensional function of different pairs of dMRI metrics, namely: DKI ADC and K in **(C)**; *Diff-in* $D_{0,I}$ and F_{MRI} in **(D)**; *Diff-in* vCS_{MRI} and F_{MRI} in **(E)**; *Diff-in* $D_{0,I}$ and vCS_{MRI} in **(F)**. **(G) to (J)**: prediction of tumour volume based on different pairs of input dMRI, in 5-fold cross validation experiments. These were: DKI ADC and K in **(G)**; *Diff-in* $D_{0,I}$ and F_{MRI} in **(H)**; *Diff-in* vCS_{MRI} and F_{MRI} in **(I)**; *Diff-in* $D_{0,I}$ and vCS_{MRI} in **(J)**. In panels (G) to (J), scatter plots of ground truth

tumour volumes against predicted tumour volumes are obtained by pooling together predictions from all 5 cross-validation folds. The figures also report the Pearson's correlation coefficient between ground truth and predicted tumour volume values, with the linear regression line in grey. This lines visualises the fitting relating ground truth and predicted tumour volumes through a linear relationship ($x_{predicted} = a + b\ x_{ground\text{-}truth}$). The line visualises slope/off-set of the fit, which was performed on the sample made of predictions and ground truth volumes for all N = 140 tumours from 38 patients included in this experiment. The grey interval around the regression line represents the 95% confidence interval over 1000 boostrap repetitions of the tumour volume regression. The p-values refer to a two-sided Pearson's correlation test between ground truth and predicted tumour volumes (in logarithmic scale), and are unadjusted for multiple comparisons. The range for the correlation coefficient is obtained across cross-validation folds. Note that in these experiments, the $\log_{10}$ of the tumour volume was studied, since the tumour volume spans several orders of magnitude and is therefore challenging to handle numerically. The results shown in this figure are based on a sample size of N = 140 tumours from 38 patients.

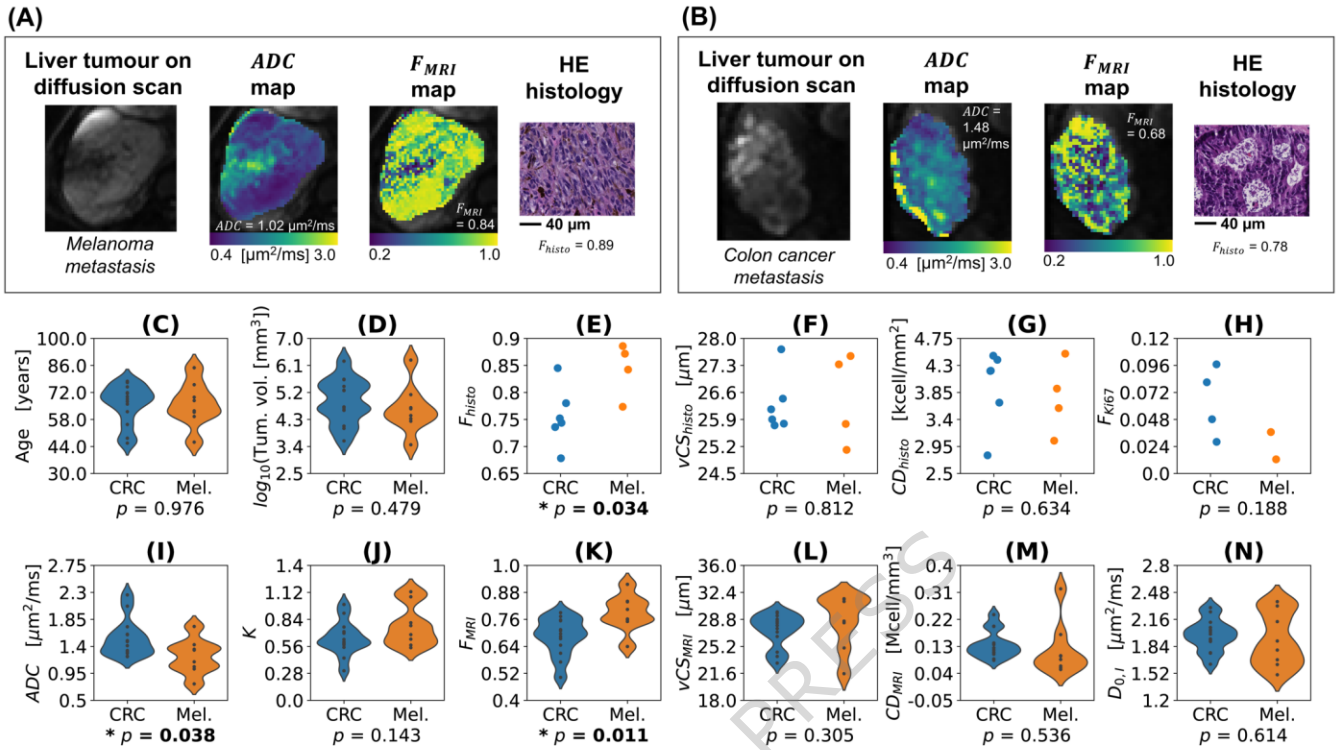


Fig. 8. Differences in dMRI and histological metrics between liver metastases from melanoma and colorectal cancer (CRC). Visualisation of illustrative parametric maps in a liver metastasis from melanoma and CRC, and differences between the two groups across dMRI and histological metrics. Differences are assessed through two-sided t-tests, and p-values are reported as unadjusted for multiple comparisons. **(A)**: data from a melanoma liver metastasis (from left to right: DW image, ADC map and F_{MRI}map from dMRI; HE histology, with quantitative IC fraction F_{histo}). **(B)**: same data as in (A), but for a colon cancer liver metastasis. **(C):** age differences between the CRC and melanoma groups (N = 13 CRC; N = 8 melanoma). **(D)**: tumour volume differences (N = 13 CRC; N = 8 melanoma). **(E):** histological IC fraction F_{histo} differences (N = 8 CRC; N = 4 melanoma). **(F)**: histological cell size vCS_{histo} differences (N = 6 CRC; N = 4 melanoma). **(G)**: histological cell density CD_{histo} differences (N = 6 CRC; N = 4 melanoma). (H): histological Ki67 staining fraction F_{Ki67} differences (N = 4 CRC; N = 2 melanoma). **(I)**: DKI *ADC* differences (N = 13 CRC; N = 8 melanoma). **(J)**: DKI kurtosis *K* differences

(N = 13 CRC; N = 8 melanoma). (K): *Diff-in* model IC fraction F_{MRI} differences (N = 13 CRC; N = 8 melanoma). **(L)**: *Diff-in* model cell size vCS_{MRI} differences (N = 13 CRC; N = 8 melanoma). **(M)**: *Diff-in* model cell density CD_{MRI} differences (N = 13 CRC; N = 8 melanoma). **(N)**: *Diff-in* model IC diffusivity $D_{0,I}$ differences (N = 13 CRC; N = 8 melanoma).

Tables

Table 1. Correlation between metrics from the selected biophysical dMRI approach with their direct histological counterparts. The table reports Pearson's correlation coefficients *r* and corresponding p-values *p* for dMRI metrics F_{MRI} (intra-cellular fraction), vCS_{MRI} (volume-weighted cell size index) and CD_{MRI} (cell density per unit volume) of model *Diff-in*, fitted to high b-value images. The correlations are reported between each MRI metric and its direct histological counterpart (F_{histo} for F_{MRI}; vCS_{histo} for vCS_{MRI}; and CD_{histo} for CD_{MRI}). For comparison, the table also reports correlation coefficients between routine ADC and *K* from DKI and each of F_{histo}, vCS_{histo} and CD_{histo}. The sample size was N = 25, implying that $p < 0.05$ if $|r| > 0.396$. $p < 0.05$ is flagged by bold font. The p-values were obtained from two-sided Pearson's correlation tests, and are unadjusted for multiple comparisons.

	Histology F_{histo}	**Histology vCS_{histo}**	**Histology CD_{histo}**
Model *Diff-in* fitted to high *b*-value images	With F_{MRI} $r = 0.19$; $p = 0.370$	With vCS_{MRI} **$r = 0.44$; $p = 0.029$**	With CD_{MRI} **$r = 0.70$; $p = 0.00011$**
DKI *ADC*	$r = -0.28$; $p = 0.180$	**$r = 0.49$; $p = 0.014$**	**$r = -0.47$; $p = 0.017$**
DKI *K*	**$r = 0.40$; $p = 0.048$**	$r = -0.31$; $p = 0.130$	**$r = 0.43$; $p = 0.033$**

Ed Summary :

Grussu et al. use co-localised MRI and histology data to design a practical MRI technique for cell size and density measurement in liver tumours in vivo. The method provides non-invasive proxies of histological properties that are associated to cell proliferation, that explain tumour volume and that distinguish liver tumour types.

Peer review information:

Communications Medicine thanks Thierry Metens and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.