Histology-informed liver diffusion MRI: relevance in cancer immunotherapy

Francesco Grussu[1]*, Kinga Bernatowicz[1], Marco Palombo[2,3], Irene Casanova-Salas[1], Daniel Navarro-Garcia[1], Ignasi Barba[1,4], Sara Simonetti[1], Garazi Serna[1], Athanasios Grigoriou[1,5], Carlos Macarro[1], Anna Voronova[1,5], 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 Roson[7,8], Manuel Escobar[7,8], Maria Vieito[11], Rodrigo Toledo[1], Paolo Nuciforo[1], Joaquin Mateo[11], Elena Garralda[11] and Raquel Perez-Lopez[1]*

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

[1] Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Spain [2] Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Cardiff, United Kingdom [3]School of Computer Science and Informatics, Cardiff University, Cardiff, UK [4] University of Vic - Central University of Catalonia (UVic-UCC), Vic, Spain [5] Department of Biomedicine, Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, 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), Spain

Abstract

Innovative diffusion Magnetic Resonance Imaging (dMRI) enables the *in vivo* mapping of biologically meaningful properties such as cell size (CS), potential biomarkers in cancer. However, while cancers frequently spread to the liver, dMRI techniques tailored for liver applications and easy to deploy in the clinic are still sought. We tackle this need by delivering a practical dMRI framework for non-invasive liver tumour CS mapping in clinical settings. Through direct comparison of candidate approaches in fixed mouse livers (N = 7) and cancer patients' data (N = 18), we select a dMRI signal model of intra-cellular diffusion fitted to highly diffusion-weighted images, which maximises radiological-histological correlations. We demonstrate the potential of the model for immunotherapy response assessment, by assessing the association between progression-free survival (PFS) and baseline liver tumour dMRI (N = 34). Smaller dMRI CS is associated to faster progression (Cox Hazard Ratio (HR): 0.45 (0.25; 0.81), $P = 0.008$), a finding confirmed by biopsy data (Spearman's $r = 0.60$ between PFS and histological CS, $P = 0.039$, N = 12, with Cox HR: 0.70 (0.47; 1.04), $P = 0.079$), and potentially related to the smaller size of fast-proliferating cells (Spearman's $r = -0.667$ between histological CS and Ki67 staining, $P = 0.049$, N = 7). In conclusion, our study delivers a practical and histologically-meaningful dMRI signal model for CS mapping in liver tumours. Our proposed approach may provide complementary information to routine standard-of-care imaging in applications such as immunotherapy, and thus play a key role in precision oncology.

Introduction

Routine clinical Magnetic Resonance Imaging (MRI) focusses on visualising macroscopic anatomical features, as presence of tumours. Nonetheless, MRI also offers the possibility of measuring biological properties within each pixel of a three-dimensional (3D) scan – known as *voxel*. This approach, referred to as quantitative MRI (qMRI) (1), involves the acquisition of multiple images, each featuring a different contrast, which are then analysed jointly with a mathematical model. qMRI provides promising metrics, which could become quantitative biomarkers complementing the qualitative assessment by the expert radiologist (2). Diffusion MRI (dMRI) is a qMRI approach that sensitises the signal to water diffusion with magnetic field gradients (1). Since diffusion in biological tissues is influenced by the microenvironment where diffusion takes place, dMRI ultimately enables the indirect estimation of properties at the micrometric scale (3), such as the size of cells restricting water random movement (4). dMRI bridges the gap between macroscopic and microscopic imaging, and has found promising applications in oncology (5,6).

Innovative dMRI techniques are also urgently needed in abdominal imaging, as in liver imaging (7,8). The liver is a frequent site for cancer metastasisation (9), and liver tumours are common targets for treatment response assessment in oncology. However, current response criteria such as RECIST (10) have limitations, in that they rely on MRI or computed tomography (CT) merely to measure tumour size, without accounting for changes under therapy at the cellular level. Novel dMRI metrics could 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 in the development of treatment resistance (11,12), and could better stratify patients eligible for treatments such as immunotherapy (13), given the challenge of predicting which patients can benefit from these innovative drugs (14). This would be a major advancement in oncology, as it may allow for personalised treatment planning, sample size reduction in clinical trials, and ultimately improve patient outcomes (15).

The most recent biophysical dMRI techniques describe the extra-vascular liver tissue signal as the sum of contributions from intra-cellular and extra-cellular water (16–18). While these models provide promising readouts (19), their practical use in real-world settings is made unfeasible by i) the high number of dMRI images (and hence long scan time) required to support model fitting, and by ii) the requirement for specialised dMRI acquisitions (20), beyond the standard examinations available by default in the scanner console. In this study we aim to overcome these challenges by delivering a practical liver dMRI signal model 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 protocols. With this objective in mind, we embraced the latest *histology-informed* dMRI development paradigm, which is based on informing signal model design with co-localised histology. The framework has shown promise in delivering dMRI metrics with unprecedented fidelity to cytoarchitecture (21,22), maximising biological specificity.

In this article, we aimed to identify a practical mathematicl model that maximises the agreement of dMRI estimates of metrics of cell morphology to their underlying histological counterparts. We analysed 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 data to select the model maximising radiological-histological correlations, validating results with computer simulations.

Afterwards, we demonstrated the clinical utility of the designed approach in one exemplificatory task, assessing immunotherapy response in patients suffering from advanced solid tumours – an urgent and unmet need, given the lack of robust predictors of treatment response for this class of drugs (14). In summary, we deliver a liver dMRI approach that offers metrics with high fidelity to histopathology, and which is feasible in the clinic. The proposed method, based on a single-compartment model of restricted, intra-cellular diffusion fitted to highly diffusion-weighted (DW) images, enables the stratification of patients' progression under immunotherapy based on pre-treatment dMRI estimates of cell size in liver tumours. The link between cell size and progression, confirmed by clinical-histological data in the same cohort, suggests that our dMRI framework could provide complementary information to standard-of-care imaging, and thus play a key role in oncology research and practice.

Methods

We used co-localised MRI and histology data to select the liver dMRI signal model capable of estimating cell size and intra-cellular fraction with the highest fidelity to histology, given a pool of candidate models. Afterwards, we tested whether voxel-wise maps from such a model offer any added any value compared to routine tumour burden measurement for response assessment in immunotherapy. Fig. S1 illustrates the dMRI signal models and summarises the study design.

dMRI models

We hypothesise that the liver dMRI signal arises from three, non-exchanging proton pools: vascular water; restricted, intra-cellular water; and hindered, extra-cellular, extra-vascular water (4,16,23–25). When the diffusion-weighting is sufficiently high (b > 100 s/mm^2 *in vivo*), the vascular signal vanishes, so that the dMR signal can be written as

$$s = s_0\,(1 - f_V)(f_I\, a_I + (1 - f_I)\, a_E). \quad (1)$$

Above, s_0 is the apparent proton density, f_V is the voxel vascular signal fraction, $F = (1 - f_I)\, f_V$ is the voxel intra-cellular signal fraction, and $a_I = e^{-b\, ADC_I}$ and $a_E = e^{-b\, ADC_E}$ are intra-/extra-cellular diffusion-weighting factors. They depend on the respective intra-/extra-cellular Apparent Diffusion Coefficient (ADC), i.e., ADC_I and ADC_E. We used a common expression for ADC_I based on restricted diffusion within spheres of diameter L (16,17) (see Supplementary Methods). Noting that dMRI-derived estimates of cell size are intrinsically volume-weighted cell size statistics (26), we will refer to L as *volume-weighted cell size* (vCS) index. Conversely, ADC_E models hindered diffusion in a tortuous space (16,24,27), i.e.,

$$ADC_E(\Delta, D_{E,\infty}, \beta) = D_{E,\infty} + \frac{\beta}{\Delta}, \quad (2)$$

where $D_{E,\infty}$ is the asymptotic (24) ADC_E for $\Delta \to \infty$ and Δ the diffusion time.

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. 2;

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

In the second family, we use the constraint $ADC_E > ADC_I$, common in body dMRI (28). This family includes the following models:

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

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

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 $a_E \approx 0$ in Eq. 1.

Fitting the models provides voxel-wise estimates of vCS and F. These can be combined into a cell density per unit volume (17) estimate

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

Preclinical data

Animals

We obtained data from 7 fixed livers of NOD.Cg-Prkdc^{scid} IL2rg^{tm1WjI}/SzJ mice from an ongoing xenograft model development study in prostate cancer, whose experimental protocols were approved and monitored by the Vall d'Hebron Institute of Research Animal Experimentation Ethics Committee (registration number 68/20), in accordance with relevant local/EU regulations. We studied six livers from mice implanted with cells derived from biopsies of prostate cancer patients, plus an additional liver from a mouse without any implantation. After scarification, we collected the livers, fixed them overnight in formalin, and transferred them to phosphate-buffered saline (PBS) solution. Detailed animal procedures are described in the Supplementary Methods.

MRI

We scanned the livers on a 9.4T Bruker Avance system at room temperature. Livers were tightened with sewing thread to a histology cassette and placed into a Falcon® tube, filled with PBS solution. The protocol included a T2-weighted fast spin echo sequence and pulsed-gradient spin echo (PGSE) dMRI (TR = 2700 ms; resolution: 386 µm × 386 µm; matrix size: 86 × 86; 4 slices, 2.216 mm-thick, NEX = 1). dMRI featured δ = 10 ms, Δ = {15, 30} ms, 10 linearly spaced b-values for each Δ (minimum/maximum nominal *b*: 0/2800 s/mm^2). Δ = 15 ms images were acquired at each of TE = {31, 45, 65} ms, and Δ = 30 ms images at each of TE = {45, 65} ms. dMRI data underwent routine processing (MP-PCA denoising (29), Gibbs unringing (30), drift correction), before being

used to fit the *Diff-in-exTD*, *Diff-in-exTDFast*, *Diff-in-ex*, *Diff-in-exFast* and *Diff-in* models voxel-by-voxel. Fitting was performed i) on all images with $b > 1000$ s/mm^2 (suppressing vascular signals, referred to as *fitting on whole image set*); ii) on $b > 1000$ s/mm^2 images (*high b-value fitting*). We also computed routine ADC and apparent diffusion excess kurtosis K by fitting the standard Diffusion Kurtosis Imaging (DKI) representation (31) to DW images acquired at $TE = 45$ ms, $\Delta = 30$ ms. Additional details are reported in the Supplementary Methods.

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 performed automatic cell detection with QuPath (32), deriving maps of volume-weighted cell size vCS_{histo}, intra-cellular area fraction F_{histo} and cell density per unit area CD_{histo} in dMRI space(26). vCS_{histo} is defined as

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

and is a more accurate counterpart of dMRI cell size than the arithmetic mean (26,33) $aCS_{histo} = < l >$. Above, l indicates the generic cell diameter, and $< \cdot >$ performs averaging. We accounted for shrinkage effects due to histological processing (details in the Supplementary Methods).

Clinical data

Cohort

We obtained data from patients suffering from advanced solid tumours, recruited for an ongoing imaging study (Vall d'Hebron University Hospital Ethics committee approval PR(AG)29/2020). We included 38 patients with liver malignancies, who provided informed written consent. Of these 38 patients, 34 received immunotherapy after dMRI, as part of a phase I clinical trial at the Vall d'Hebron Institute of Oncology. Immunotherapy included immune-checkpoint inhibitors (anti-PD1, anti-PDL1, anti-CTLA4, LAG3) in monotherapy or in combination with other drugs, and other cancer immunotherapies as for example bispecific antibodies, T-cell engagers and others (approvals PR(AG)252/2016, PR(AG)321/2018 and PR(AG)70/2018). We also obtained histological data from digitised biopsies from one of the imaged liver tumours in 18 patients. For patients receiving immunotherapy, the Progression Free Survival (PFS) was available. PFS represents the lag between therapy starting date and progression or death (whichever occurs first), with progression determined via RECIST (10), or in case of established clinical worsening.

MRI

We scanned 11 patients on a 1.5T Siemens Avanto scanner, and 27 more patients on a 3T GE SIGNA Pioneer system. In both scanners, the protocol included a abdominal T2-weighted fast spin

echo scan and fat-suppressed dMRI (details in the Supplementary Methods; dMRI sequence block diagram reported in Fig. S2). In short, on the 1.5T system we used a DW TRSE EPI sequence with salient parameters: scan time of 16 minutes, resolution of 1.9 × 1.9 × 6 mm^3, TR = 7900 ms, 2 signal averages, GRAPPA factor of 2, 6/8 partial Fourier imaging, b = {0, 50, 100, 400, 900, 1200, 1600} s/mm^2 at TE = {93, 105, 120} ms, with varying diffusion time. In the 3T system, dMRI was based on PGSE EPI, with salient parameters: scan time of 16 minutes, resolution of 2.4 × 2.4 × 6 mm^3, TR = 6000 ms, 2 signal averages, ASSET factor of 2, b = {0, 50, 100, 400, 900, 1200, 1500} s/mm^2 at TE = {75, 90, 105} ms, with varying gradient duration/separation δ/Δ.

dMRI post-processing consisted of MP-PCA denoising (29), Gibbs unringing (30), motion correction (34), EPI distortion correction (35) (1.5T data only). An experienced radiologist (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 (36) non-linear co-registration, and fitted the 5 dMRI models object of this study. Fitting was performed: 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*). Finally, we also computed standard ADC and excess kurtosis K on b > 100 s/mm^2 images (shortest TE), with in-house DKI fitting routines.

Histology

We performed ultrasound-guided biopsies of one liver tumour at the Barcelona Vall d'Hebron University Hospital. The biological material underwent standard processing, followed by routine HE staining, Ki67 immunohistochemical labelling (demonstrating cell replication), and digitalisation (Hamamatsu C9600-12 slide scanner; resolution: 0.454 μm). An experienced pathologist (S.S.) assessed the images and outlined tumour tissue. 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 to estimate per-biopsy vCS_{histo}, F_{histo} and CD_{histo} (detailed pipeline in the Supplementary Methods), and Ki67 images with custom-written code to estimate the fraction of tumour labelled for Ki67 (referred to as F_{Ki67}).

Statistical analyses

dMRI model selection

We ranked dMRI models fitted according to the two fitting strategies to find the approach yielding estimates of cell size (vCS_{MRI}) and intra-cellular fraction (F_{MRI}) that most closely match histological vCS_{histo} and F_{histo}. For this purpose, we introduced an MRI-histology *Total Correlation Score* (TCS), as well as an *Histology Fidelity Criterion* (HFC). TCS combines the correlation between vCS_{MRI} and vCS_{histo} and between F_{MRI} and F_{histo}. Higher TCS point towards stronger histological-radiological correlation; negative correlations reduce TCS, so they are penalised. Conversely, HFC measures the accuracy in using vCS_{MRI} and F_{MRI} as estimates of vCS_{histo} and F_{histo}. For comparison, we also ranked dMRI models using the popular *Bayesian Information Criterion* (BIC) (37), i.e., according to the goodness of fit, while penalising model complexity. Additional details on model selection are included in the Supplementary Methods.

Simulated dMRI model selection

We validated the histology-informed dMRI model selection with computer simulations. Briefly, we synthesised dMRI signals via Monte Carlo diffusion random walks for all protocols of this study, using the MCDC simulator (38). We seeded walkers in a substrate made of spherical cells of identical diameter (4,16,19,20) (Fig. S3), varying the intra-sphere fraction F and the sphere diameter L. Afterwards, we performed model selection on signal corrupted with Rician noise (b = 0 signal-to-noise ratio: 30) according to the TCS, HFC and BIC criteria, thus comparing ground truth and estimated values of F and L (additional details in the Supplementary Methods).

dMRI-histology correlation analysis

We computed mean and standard deviation of all dMRI and HE-derived 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 among all possible pairs of metrics, visualing results in terms of correlation matrices.

Response assessment in immunotherapy

We studied mean values of *in vivo* baseline dMRI metrics within liver tumours (in the form of z-scores) to stratify the risk of progression under immunotherapy (N = 34). We fitted multi-variate proportional hazard Cox regression models, where one dMRI metric was included in turn, while also controlling for age, sex and total baseline tumour volume. We also fitted two additional multi-variate models similar to the above, but including i) both DKI metrics ADC and K, or ii) vCS_{MRI} and CD_{MRI} from model *Diff-in* fitted on high b-value data (i.e., the proposed approach; additional details are included in the Supplementary Methods).

Finally, we also studied the relationship between histological properties and PFS, thus corroborating dMRI results. We calculated the Spearman's correlation coefficient between PFS and baseline vCS_{histo}, CD_{histo}, F_{histo} and F_{Ki67}, and fitted univariate Cox regression models (F_{histo} and F_{Ki67} expressed in percetange terms). Spearman's correlation coefficients between F_{Ki67} and each of vCS_{histo}, CD_{histo} and F_{histo} were also evaluated.

We performed all statistical analyses in Python, using the *SciPy* and *lifelines* packages. Data S4 includes Python scripts performing immunotherapy response assessment.

Results

Data set overview

Fig. 1 illustrates the data used in this study. We will refer to data obtained in fixed mouse livers as *preclinical*, while to data obtained in cancer patients as *clinical*.

The preclinical data includes dMRI and histology data from fixed livers of seven mice, of which six had been implanted with biopsies of prostate cancer patients, while one had no implantation, and features normal liver structures (we will refer to it as *Control*). The livers of the six implanted cases did not grow any tumours, and feature three unique histopathological phenotypes (Fig. S4). 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 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 also obtained clinical data on N = 38 cancer patients suffering from advanced liver malignancies (Fig. 1). The mean/standard deviation of age was 61.88/11.90 years. 19 patients were male and 19 female. 3 patients 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 kidney, 1 endometrial, 1 ureteral, 1 thyroid). 34 out of 38 patients received immunotherapy after dMRI (primary cancers: 8 colon, 7 melanoma, 4 breast, 4 gastric, 3 HCC, 2 rectal, 2 ovarian, 1 thyroid, 1 kidney, 1 endometrial, 1 ureteral). Their mean/standard deviation of age were 61.84/12.41 years (16 male, 18 female), and the median/interquartile range of PFS were 65/61.75 days. Biopsies from one of the imaged liver tumours were also available in 18 patients. Of these 18 patients, 4 did not receive immunotherapy; 12 receieved immontherapy after the collection of the biopsy; and 2 received immunotherapy before the collection of the biopsy (in these cases, an additional dMRI scan immediately before the biopsy was acquired).

dMRI model selection: a one-pool model of intra-cellular diffusion maximises the MRI-histology Total Correlation Score (TCS)

Fig. 2A summarises the different dMRI models, while Fig. 2B and Fig. 2C report the MRI-Histology (fitting performed on the whole image set in Fig. 2B, or in high high b-value images in Fig. 2C). dMRI models where $ADC_E > ADC_I$, shown in violet, provide consistently higher TCS than models that do not make such an assumption (orange shades). We observe the highest TCS for model *Diff-in* fitted to high b-value images. Fitting on high b-value images provides higher TCS values than fitting to the whole image set, especially for models where $ADC_E > ADC_I$ (violet shades). When fitting is performed on the whole image set, *Diff-in-exFast* provides the highest TCS (Fig. 2B), depsite being lower than *Diff-in* TCS (Fig. 2C).

Fig. S5 reports rankings according to HFC and BIC. It reports the number of times, in percentage terms, that a model provides the lowest HFC and lowest BIC (both indicative of the best performance) across N = 25 MRI-histology cases. Results essentially confirm rankings seen on TCS: models hypothesising $ADC_E > ADC_I$ are selected more frequently than models that do not. The model *Diff-in* is the most selected model according to both BIC and HFC (fig. S5.B; fitting to high b-value images). Fig. S6 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*.

Computer simulations confirm model selection from MRI measurements

Supplementary Tables S1, S2, and S3 report TCS, HFC and BIC rankings from simulated signals. Results confirm that model *Diff-in* enables the best estimation of cell size and intra-cellular fraction in the vast majorities of cases, thus confirming *ex vivo* and *in vivo* MRI results.

In view of all rankings reported above, our recommended modelling approach is the *fitting of a one-compartment model of restricted, intra-cellular diffusion to high b-values images* (here, $\gtrapprox$ 1800 s/mm^2 *ex vivo*, $\gtrapprox$ 900 s/mm^2 *in vivo*) – referred to as model *Diff-in*.

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

Table 1 shows that Pearson's correlation between metrics from the selected *Diff-in* dMRI model and from DKI with histology. We consider correlations to be weak, moderate, and strong when $|r| < 0.4$, $|r| \geq 0.4$ but $|r| < 0.6$, and $|r| \geq 0.6$. Given our N = 25, $P < 0.05$ if $|r| > 0.3961$.

Diff-in F_{MRI} (intra-cellular fraction), vCS_{MRI} (cell size index) and CD_{MRI} (cell density) are respectively weakly ($r = 0.19$ between F_{MRI} and F_{histo}), moderately ($r = 0.44$ between vCS_{MRI} and vCS_{histo}) and strongly ($r = 0.70$ between CD_{MRI} and CD_{histo}) correlated with their histological counterparts F_{histo}, vCS_{histo} and CD_{histo}. Table 1 also reports correlation coefficients for dMRI ADC and kurtosis K. Both ADC and K exhibit significant, moderate correlations with histological properties, i.e., negative/positive correlation of ADC/K with cell density CD_{histo} ($r = -0.47$ and 0.43 respectively). Significant correlations are also seen with F_{histo} ($r = 0.40$, $P = 0.048$ between K and F_{histo}).

Fig. S7 and Fig. S8 show Pearson's correlation coefficients for all possible pairs of metrics, in the form of correlation matrices. Correlations among dMRI metrics are seen, as a strong negative correlation between CD_{MRI} and vCS_{MRI} ($r = -0.84$ for model *Diff-in-exFast* fitted at high b-value). This finding, which indicates that tighter cell packings per unit volume are achieved with smaller cells, is mirrored by histological CD_{histo} and vCS_{histo} ($r = -0.88$ between CD_{histo} and vCS_{histo}). Other weak-to-moderate correlations are seen, e.g.: between K and vCS_{MRI} ($r = -0.47$) and K and CD_{MRI} ($r = 0.38$), which agree with the correlations observed between K and histological vCS_{histo} ($r = -0.31$) and CD_{histo} ($r = 0.43$). In general, metrics from dMRI models where $ADC_E > ADC_I$ show stronger correlations with their histological counterparts than models *Diff-in-exTD* and *Diff-in-ex*. We observe the strongest dMRI-histology correlations for model *Diff-in* fitted to high b-value images.

Metrics from the proposed dMRI model reveal intra-/inter-tumour characteristics

Fig. 3 shows maps from the proposed dMRI model *Diff-in* alongside histological metrics in 3 mouse livers, representative of the 3 phenotypes seen in our mouse data (*Control*, for normal liver structures; Pat_{inf1}, for small cell infiltration; Pat_{nec}, for necrosis). Visually, we observe excellent

co-localisation between MRI slices and histology sections. The histological details reveal higher cellularity in sample Pat_{infl} compared to *Control*, due to packing of small cells in between larger hepatocytes, or an alternation of areas with lower/ higher cell density in sample Pat_{nec}. These qualitative trends are confirmed in the histological maps F_{histo}, vCS_{histo} and CD_{histo}. Metrics values (Table S4 and Data set S1, e.g., F_{MRI} around 0.75 and vCS_{MRI} of the order of 20 μm), are physiologically plausible (39,40), although F_{MRI} slightly underestimates F_{histo}, while vCS_{MRI} slightly overestimates vCS_{histo}.

Maps F_{MRI}, vCS_{MRI} and CD_{MRI} replicate the contrasts seeing in their histological counterparts F_{histo}, vCS_{histo} and CD_{histo}. Fig. S9 shows standard DKI ADC and kurtosis K in the mouse livers. Visual trends highlight that the higher cell density of sample Pat_{infl} translated to remarkably reduced ADC and increased K compared to the *Control*. Lastly, Fig. S10 shows F_{MRI}, vCS_{MRI} and CD_{MRI} maps from model *Diff-in-exFast*, while Fig. S11 shows maps of intrinsic cytosol diffusivity $D_{0,I}$ and asymptotic apparent diffusivity $D_{E,\infty}$. The figures highlight that overall, spatial trends seen in maps from the selected model *Diff-in* agree with those seen in *Diff-in-exFast*, but metrics from the latter appear noisier. Metrics $D_{0,I}$ and $D_{E,\infty}$ show limited between-sample contrast, and are difficult to validate histologically.

Fig. 4 shows F_{MRI}, vCS_{MRI} and CD_{MRI} maps in patients, alongside HE biopsies. 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}, CD_{MRI} show contrasts that are plausible with such histopathological features. For example, in a breast cancer liver metastasis in Fig. 4, 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. S12 shows routine dMRI ADC and K in the same tumours in patients. Spatial trends are also compatible with the histology, e.g., high ADC and low K are seen in the necrotic core of the breast cancer tumour. Supplementary Fig. S13 shows F_{MRI}, vCS_{MRI} and CD_{MRI} from model *Diff-in-exFast*. Image contrasts match visually those seen in the same metrics from model *Diff-in* (the proposed approach), despite being noisier. Fig. S14 shows intra-cellular cytosol diffusivity $D_{0,I}$ and the asymptotic diffusivity $D_{E,\infty}$ in the same tumours. These, notoriously difficult to measure accurately *in vivo* (16,41), exhibit a speckled appearance.

Smaller baseline dMRI cell size is associated to faster cancer progression under immunotherapy, a finding confirmed by histology

Table 2 reports on the PFS stratification based on baseline dMRI metrics through multi-variate Cox regression models, which account for the potential confounding effect of age, sex and baseline tumour volume. There is a significant association between progression and baseline MRI estimates of cell size within liver tumours (Hazard Ratio (HR) for vCS_{MRI} of 0.45 (0.25; 0.81), P = 0.008), with smaller baseline cell size being associated to faster progression. Similarly, there is a significant association between DKI K and PFS from multi-variate Cox regression, with higher kurtosis values associated to faster progression (HR for K of 1.44 (1.02; 2.03), P = 0.036). Table

S5 reports results from Cox regression when only one *Diff-in* or DKI metric in turn was used as regressor, jointly with age, sex and baseline tumour volume. No statistically significant associations are observed in this case. However, the estimated HRs for vCS_{MRI} and K are in the same direction of those reported in Table 2, with $P < 0.100$ in both cases.

Fig. 5 reports on Spearman's correlation between PFS and baseline histological metrics (panels A to D), as well as corresponding HRs from univariate Cox regression models (panel E). The figure corroborates the associations between dMRI and clinical outcome reported above. There is a statistically significant, strong, positive Spearman's correlation between PFS and baseline tumour cell size vCS_{histo} ($r = 0.60$, $P = 0.039$; N = 12), which corresponds to a HR of 0.70 (0.47; 1.04), $P = 0.079$ in Cox regression, in the same direction of what was previously reported for vCS_{MRI}. We also observe a statistically significant, very strong, and negative Spearman's correlation between PFS and baseline Ki67 staining fraction F_{Ki67} ($r = -0.90$, $P = 0.006$; N = 7), which yields an HR of 1.55 (1.22; 1.98), $P = 0.00038$ in univariate Cox regression. Fig. 6 provides an example of the histological correlates of cancer progression under immunotherapy. The figure reports HE and Ki67 histological images of two patients with faster and slower cancer progression (patient 3: ovarian cancer, PFS of 11 days; patient 6: primary HCC, PFS of 110 days; patients also included in Fig. 4, S12, S13 and S14 with dMRI maps). In the case of faster progression, the biopsy shows intense Ki67 staining (i.e., fast cell proliferation, leading to high F_{Ki67}), while in the case of slower progression, lower Ki67 signal is detected. This contrast corresponds to differences in cell size and density on the HE-stained biopsied tissue (smaller cell size vCS_{histo} and higher cell density CD_{histo} in presence of higher Ki67 staining fraction F_{Ki67}), which agree with trends in dMRI vCS_{MRI} and CD_{MRI}.

Lastly, Table S6 reports the Spearman's correlation coefficient between each of vCS_{histo}, F_{histo} and CD_{histo} with F_{Ki67} (N = 7). vCS_{histo} is significantly, strongly, and negatively correlated with F_{Ki67} ($r = -0.667$, $P = 0.049$), while both F_{histo} and CD_{histo} are significantly, strongly, and positively correlated with F_{Ki67} ($r = 0.717$, $P = 0.029$ and $r = 0.900$, $P = 0.00094$).

Discussion

The latest liver dMRI models aim to disentangle intra-cellular and extra-cellular water contributions to the total signal (4,16,19,20). This powerful approach enables the estimation of innovative tissue property maps, but its clinical deployment is hampered by the high number of unknown tissue parameters to estimate, which requires impractically long acquisitions (16,42). With this challenge in mind, this paper delivers a practical implementation of a two-compartment dMRI signal model, tailored for liver imaging, and truly feasible in the clinic. Through histology-informed model selection, we design a compact framework consisting of fitting a one-pool model of restricted intra-cellular diffusion to highly DW images. This provides cell size and density estimates that correlate with histology, and which may provide additional information to standard-of-care imaging in applications such as immunotherapy.

To find the optimal dMRI signal implementation, we analysed co-localised dMRI and histology data (N = 25) from fixed mouse livers and from cancer patients. We compared 5 signal models,

each fitted according to two distinct strategies, and ranked them for their ability to estimate histological intra-cellular fraction and cell size. Rankings unequivocally suggest the highest radiological-histological agreement is obtained by *fitting a single-compartment model of restricted diffusion within spherical cells* – 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, confirmed by Monte Carlo computer simulations, 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 extra-cellular liver ADC (as high as 2.5 μm^2/ms (16)). Such a high ADC_E implies that the extra-cellular 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 = 1200 s/mm^2 = 1.2 ms^{-1} μm^2 and ADC_E = 2.5 μm^2 ms^{-1}), justifying the use of single-compartment model of intra-cellular diffusion (25,33,41).

Importantly, we conducted dMRI model selection using a variety of criteria, namely: TCS (overall correlation between dMRI and histology), HFC (accuracy in histological property estimation via dMRI), and BIC (dMRI model quality of fit, penalising model complexity). In general, all criteria point towards the same direction, with model *Diff-in* fitted to high b-value images being the top-ranking model. We would also like to emphasise that while we used 15 *in vivo* dMRI scans for both model design and for the clinical demonstration, the same liver dMRI model implementation would have been selected had we only looked at the *ex vivo* mouse data or at *in silico* signals (Tables S1 to S3). This gives confidence on the generalisability of our dMRI model design.

Our proposed dMRI modelling approach (e.g., fitting the one-compartment model of restricted diffusion *Diff-in* to highly DW data) 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. S10 and S13). 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 (41).

After selecting a practical dMRI model implementation, we investigated its utility in an exemplificatory clinical application, related to immunotherapy response assessment. Immunotherapy has shown promise in several cancers (13); nevertheless, only a small fraction of patients truly benefits from this class of drugs, and their identification prior to treatment is extremely challenging (14,43). With this in mind, we investigated whether it is possible to stratify the probability of progression under treatment given baseline dMRI metrics within liver tumours, using standard Cox regression models. To our knowledge, this is the first time that such advanced dMRI markers have been tested for patient stratification in immunotherapy *in vivo*. Results suggest that smaller values of the dMRI cell size index vCS_{MRI} at baseline are associated with faster

progression. We corroborated this finding by relating immunotherapy response to metrics derived from the biopsied tumours (vCS_{histo}, CD_{histo}, F_{histo} and F_{Ki67}), in those patients for whom a biopsy was available. Through correlation and Cox regression analyses, we found evidence that that smaller histological cell sizes at baseline, as well as higher levels of Ki67 staining (i.e., higher cell proliferation), are associated with faster progession. Despite the exploratory character of these findings, owing to the small sample size (N = 12 for HE-derived metrics, N = 7 for Ki67 data), it is notable that results from an independent imaging modality point exactly towards the same direction of what was obtained for dMRI: lower vCS_{histo} in liver malignancies at baseline is associated with faster progression, and so is lower vCS_{MRI}. This fact is in line with other studies focussing on different treatments, where higher tumour cellularity has been associated with higher cancer aggressiveness or worse prognosis (note that cell size is negatively correlated with cell density on both dMRI and histology – see Fig. S8). For example, in CRC liver metastasis resection, high cell density in resected metastases has been associated to shorter disease-free survival (44). Conversely, in breast cancer, lower tumour cellularity has been associated to pathologic complete response in chemo-free dual HER2 blockade treatment (for HER2-positive breast cancer) (45), as well as longer survivals in neoadjuvant chemotherapy (46). We speculate that the smaller cell sizes seen in tumour with faster progression may be a consequence of their fast proliferation, and hence, potentially higher malignancy (the histology-derived Ki67 staining fraction F_{Ki67} and the histology-derived cell size index vCS_{histo} are negatively correlated – see Supplementary Table S6). Overall, the association between clinical outcome and baseline dMRI metrics, and the fact that this agrees with the associations seen for histology, is highly encouraging, and motivates further dMRI research in immunotherapy.

We benchmarked our proposed dMRI approach against well-established DKI ADC and K (31), being these easy to compute with compact dMRI acquisitions, and sensitive to tissue microstructure. However, they are also known to be relatively unspecific, being surrogate metrics that conflate different histopathological characteristics into a single index (42). In our data, ADC and K correlate with cancer cellularity and cell size, a fact that agrees with known literature (42,47). Moreover, K showed a significant associations with PFS, with higher liver tumour K at baseline being associated with faster progression. This result, in line with the slower progression seen in presence of larger vCS_{MRI} (vCS_{MRI} and K are negatively correlated), demonstrates that dMRI techniques such as DKI can also offer value beyond tumour size assessment based on standard-of-care imaging. 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 limited 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 (42,48). In future, we warrant a detailed comparison of the inter-protocol robustness of our metrics, relating it to figures from DKI.

We would like to acknowledge the following potential limitations of our study. Firstly, our sample size (N = 25 for model development; N = 34 for the clinical application in immunotherapy) is relatively small, and findings were not validated, for example, in external, openly accessible cohorts. This was due to the fact that, to the best of our knowledge, no freely available data sets of

advanced, multi diffusion-time dMRI in immunotherapy exist. Our article aims to provides a solid methodological advancencement in cancer imaging, with a compelling exploratoty demonstration of the potential utility of the new approach. The demonstration is unique of its kind, since it reports heretofore undescribed associations between imaging and clinical outcomes in immunotherapy, which are supported by histological data. Nonetheless, while works proposing related dMRI techniques relied on similar (20,49), if not even smaller (16,17), sample sizes, we acknowledge that our preliminary report requires further confirmation in larger cohorts, as well as additional investigation linking dMRI metrics, immune cell infiltration, and immunotherapy response.

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 may imply that the biopsies are not fully representative of the tumour microenvironment in its entirety. Also, and most importantly, histology has its own limitations, since it provides cell property estimates that may not be, *per se*, fully accurate. 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), 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 (see Supplementary Methods). Nonetheless, our histology-derived estimates of cell properties are likely biased versions of the true figures.

We would also like to acknowledge that the proposed dMRI signal model neglects certain, potentially relevant microstructural properties, such as water exchange between intra-/extra-cellular spaces (18,50), distributions of cell sizes/cytosolic diffusivities (26), or intra-compartmental relaxation (49). On the one hand, ignoring these properties may have biased the estimation of F_{MRI} and vCS_{MRI} (18,50). On the other hand, properties such as exchange rates, overlooked here, may be relevant biomarkers *per se*. In future, we plan to incorporate these properties in our signal modelling.

To conclude, *this study delivers a practical liver dMRI signal model consisting of a single-compartment of restricted diffusion within spherical cells, which should be fitted to b-values higher than, approximately, 900 s/mm^2* in vivo. This model offers estimates of cell size and cell density that are correlated to the underlying histology, and which may provide complementary information to routine volumetric tumour burden assessment, for example for response assessment in immunotherapy. Striving to bringing precision imaging one step closer to the clinic, we release our approach as an easy-to-use, open-source Python implementation, now freely accessible online.

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 research has been 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), a CRIS Foundation Talent Award (TALENT19-05), the FERO Foundation through the XVIII Fero Fellowship for Oncological Research, the Instituto de Salud Carlos III-Investigación en Salud (PI18/01395 and PI21/01019), the Asociación Española Contra el Cancer (AECC) (PRYCO211023SERR) and the Generitat 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., A.V., and A.G.) 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).

Authors' 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. **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.

Authors' Disclosures

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.

Supplementary data

This article includes Supplementary Methods, Supplementary Figures S1 to S14, and Supplementary Tables S1 to S6.

References

1. Weiskopf N, Edwards LJ, Helms G, Mohammadi S, Kirilina E. Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. Nature Reviews Physics. Nature Publishing Group; 2021;3:570–88.

2. Jarrett AM, Kazerouni AS, Wu C, Virostko J, Sorace AG, DiCarlo JC, et al. Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting. Nat Protoc. 2021;16:5309–38.

3. Le Bihan D. Looking into the functional architecture of the brain with diffusion MRI. Nat Rev Neurosci. 2003;4:469–80.

4. Panagiotaki E, Walker-Samuel S, Siow B, Johnson SP, Rajkumar V, Pedley RB, et al. Noninvasive quantification of solid tumor microstructure using VERDICT MRI. Cancer Res. AACR; 2014;74:1902–12.

5. Baxter GC, Graves MJ, Gilbert FJ, Patterson AJ. A Meta-analysis of the Diagnostic Performance of Diffusion MRI for Breast Lesion Characterization. Radiology. pubs.rsna.org; 2019;291:632–41.

6. Stabile A, Giganti F, Rosenkrantz AB, Taneja SS, Villeirs G, Gill IS, et al. Multiparametric MRI for prostate cancer diagnosis: current status and future directions. Nat Rev Urol. 2020;17:41–61.

7. Donato H, França M, Candelária I, Caseiro-Alves F. Liver MRI: From basic protocol to advanced techniques. Eur J Radiol. 2017;93:30–9.

8. Tapper EB, Lok AS-F. Use of Liver Imaging and Biopsy in Clinical Practice. N Engl J Med. 2017;377:756–68.

9. Tsilimigras DI, Brodt P, Clavien P-A, Muschel RJ, D'Angelica MI, Endo I, et al. Liver metastases. Nat Rev Dis Primers. 2021;7:27.

10. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228–47.

648 11. Buikhuisen JY, Torang A, Medema JP. Exploring and modelling colon cancer inter-tumour
649 heterogeneity: opportunities and challenges. Oncogenesis. 2020;9:66.

650 12. Vitale I, Shema E, Loi S, Galluzzi L. Intratumoral heterogeneity in cancer progression and response
651 to immunotherapy. Nat Med. 2021;27:212–24.

652 13. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science
653 to clinical practice. Nat Rev Immunol. 2020;20:651–68.

654 14. Pilard C, Ancion M, Delvenne P, Jerusalem G, Hubert P, Herfs M. Cancer immunotherapy: it's time
655 to better predict patients' response. Br J Cancer. 2021;125:927–38.

656 15. O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, et al. Imaging
657 biomarker roadmap for cancer studies. Nat Rev Clin Oncol. 2017;14:169–86.

658 16. Jiang X, Xu J, Gore JC. Mapping hepatocyte size in vivo using temporal diffusion spectroscopy
659 MRI. Magn Reson Med. Wiley; 2020;84:2671–83.

660 17. Panagiotaki E, Chan RW, Dikaios N, Ahmed HU, O'Callaghan J, Freeman A, et al. Microstructural
661 Characterization of Normal and Malignant Human Prostate Tissue With Vascular, Extracellular, and
662 Restricted Diffusion for Cytometry in Tumours Magnetic Resonance Imaging. Invest Radiol.
663 journals.lww.com; 2015;50:218.

664 18. Gardier R, Villarreal Haro JL, Canales-Rodríguez EJ, Jelescu IO, Girard G, Rafael-Patiño J, et al.
665 Cellular Exchange Imaging (CEXI): Evaluation of a diffusion model including water exchange in
666 cells using numerical phantoms of permeable spheres. Magn Reson Med [Internet]. 2023; Available
667 from: http://dx.doi.org/10.1002/mrm.29720

668 19. Hoffmann E, Gerwing M, Niland S, Niehoff R, Masthoff M, Geyer C, et al. Profiling specific cell
669 populations within the inflammatory tumor microenvironment by oscillating-gradient diffusion-
670 weighted MRI. J Immunother Cancer. BMJ Specialist Journals; 2023;11:e006092.

671 20. Jiang X, Dudzinski S, Beckermann KE, Young K, McKinley E, J McIntyre O, et al. MRI of tumor T
672 cell infiltration in response to checkpoint inhibitor therapy. J Immunother Cancer [Internet]. 2020;8.
673 Available from: http://dx.doi.org/10.1136/jitc-2019-000328

674 21. Ye Z, Price RL, Liu X, Lin J, Yang Q, Sun P, et al. Diffusion Histology Imaging Combining
675 Diffusion Basis Spectrum Imaging (DBSI) and Machine Learning Improves Detection and
676 Classification of Glioblastoma Pathology. Clin Cancer Res. 2020;26:5388–99.

677 22. Howard AF, Mollink J, Kleinnijenhuis M, Pallebage-Gamarallage M, Bastiani M, Cottaar M, et al.
678 Joint modelling of diffusion MRI and microscopy. Neuroimage. Elsevier; 2019;201:116014.

679 23. Xu J, Jiang X, Li H, Arlinghaus LR, McKinley ET, Devan SP, et al. Magnetic resonance imaging of
680 mean cell size in human breast tumors. Magn Reson Med. 2020;83:2002–14.

681 24. Reynaud O, Winters KV, Hoang DM, Wadghiri YZ, Novikov DS, Kim SG. Pulsed and oscillating
682 gradient MRI for assessment of cell size and extracellular space (POMACE) in mouse gliomas.
683 NMR Biomed. 2016;29:1350–63.

25. Reynaud O. Time-Dependent Diffusion MRI in Cancer: Tissue Modeling and Applications. Frontiers in Physics [Internet]. 2017;5. Available from: https://www.frontiersin.org/articles/10.3389/fphy.2017.00058

26. Grussu F, Bernatowicz K, Casanova-Salas I, Castro N, Nuciforo P, Mateo J, et al. Diffusion MRI signal cumulants and hepatocyte microstructure at fixed diffusion time: Insights from simulations, 9.4T imaging, and histology. Magn Reson Med [Internet]. Wiley; 2022; Available from: https://onlinelibrary.wiley.com/doi/10.1002/mrm.29174

27. Xu J, Xie J, Semmineh NB, Devan SP, Jiang X, Gore JC. Diffusion time dependency of extracellular diffusion. Magn Reson Med. 2023;89:2432–40.

28. Conlin CC, Feng CH, Digma LA, Rodríguez-Soto AE, Kuperman JM, Rakow-Penner R, 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. 2023;5:e210115.

29. Veraart J, Novikov DS, Christiaens D, Ades-Aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. Neuroimage. 2016;142:394–406.

30. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn Reson Med. 2016;76:1574–81.

31. Jensen JH, Helpern JA, Ramani A, Lu H. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. in Medicine: An … [Internet]. Wiley Online Library; 2005; Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.20508

32. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7:16878.

33. Veraart J, Nunes D, Rudrapatna U, Fieremans E, Jones DK, Novikov DS, et al. Noninvasive quantification of axon radii using diffusion MRI. Elife. eLife Sciences Publications, Ltd; 2020;9:e49855.

34. Ourselin S, Roche A, Subsol G, Pennec X, Ayache N. Reconstructing a 3D structure from serial histological sections. Image Vis Comput. 2001;19:25–31.

35. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage. 2003;20:870–88.

36. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal. 2008;12:26–41.

37. Schwarz G. Estimating the Dimension of a Model. aos. Institute of Mathematical Statistics; 1978;6:461–4.

38. Rafael-Patino J, Romascano D, Ramirez-Manzanares A, Canales-Rodríguez EJ, Girard G, Thiran J-P. Robust Monte-Carlo Simulations in Diffusion-MRI: Effect of the Substrate Complexity and Parameter Choice on the Reproducibility of Results. Front Neuroinform. 2020;14:8.

39. Martin NC, McCullough CT, Bush PG, Sharp L, Hall AC, Harrison DJ. Functional analysis of mouse hepatocytes differing in DNA content: volume, receptor expression, and effect of IFNgamma. J Cell Physiol. 2002;191:138–44.

40. Zhou Z, Xu M-J, Gao B. Hepatocytes: a key cell type for innate immunity. Cell Mol Immunol. nature.com; 2016;13:301–15.

41. Jelescu IO, Veraart J, Fieremans E, Novikov DS. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. NMR Biomed. 2016;29:33–47.

42. Fokkinga E, Hernandez-Tamames JA, Ianus A, Nilsson M, Tax CMW, Perez-Lopez R, et al. Advanced Diffusion-Weighted MRI for Cancer Microstructure Assessment in Body Imaging, and Its Relationship With Histology. J Magn Reson Imaging. 2024;60:1278–304.

43. Usset J, Rosendahl Huber A, Andrianova MA, Batlle E, Carles J, Cuppen E, et al. Five latent factors underlie response to immunotherapy. Nature Genetics. Nature Publishing Group; 2024;1–9.

44. Serrablo A, Paliogiannis P, Pulighe F, Moro SS-M, Borrego-Estella V, Attene F, et al. Impact of novel histopathological factors on the outcomes of liver surgery for colorectal cancer metastases. Eur J Surg Oncol. 2016;42:1268–77.

45. Nuciforo P, Pascual T, Cortés J, Llombart-Cussac A, Fasani R, Paré L, et al. A predictive model of pathologic response based on tumor cellularity and tumor-infiltrating lymphocytes (CelTIL) in HER2-positive breast cancer treated with chemo-free dual HER2 blockade. Ann Oncol. 2018;29:170–7.

46. Gentile D, Sagona A, De Carlo C, Fernandes B, Barbieri E, Di Maria Grimaldi S, et al. Pathologic response and residual tumor cellularity after neo-adjuvant chemotherapy predict prognosis in breast cancer patients. Breast. 2023;69:323–9.

47. Tang W-J, Yao W, Jin Z, Kong Q-C, Hu W-K, Liang Y-S, 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. 2022;56:1912–23.

48. Lee H-H, Papaioannou A, Novikov DS, Fieremans E. In vivo observation and biophysical interpretation of time-dependent diffusion in human cortical gray matter. Neuroimage. 2020;222:117054.

49. Palombo M, Valindria V, Singh S, Chiou E, Giganti F, Pye H, et al. Joint estimation of relaxation and diffusion tissue parameters for prostate cancer with relaxation-VERDICT MRI. Sci Rep. 2023;13:2999.

50. Jiang X, Devan SP, Xie J, Gore JC, Xu J. Improving MR cell size imaging by inclusion of transcytolemmal water exchange. NMR Biomed. 2022;35:e4799.

Figures

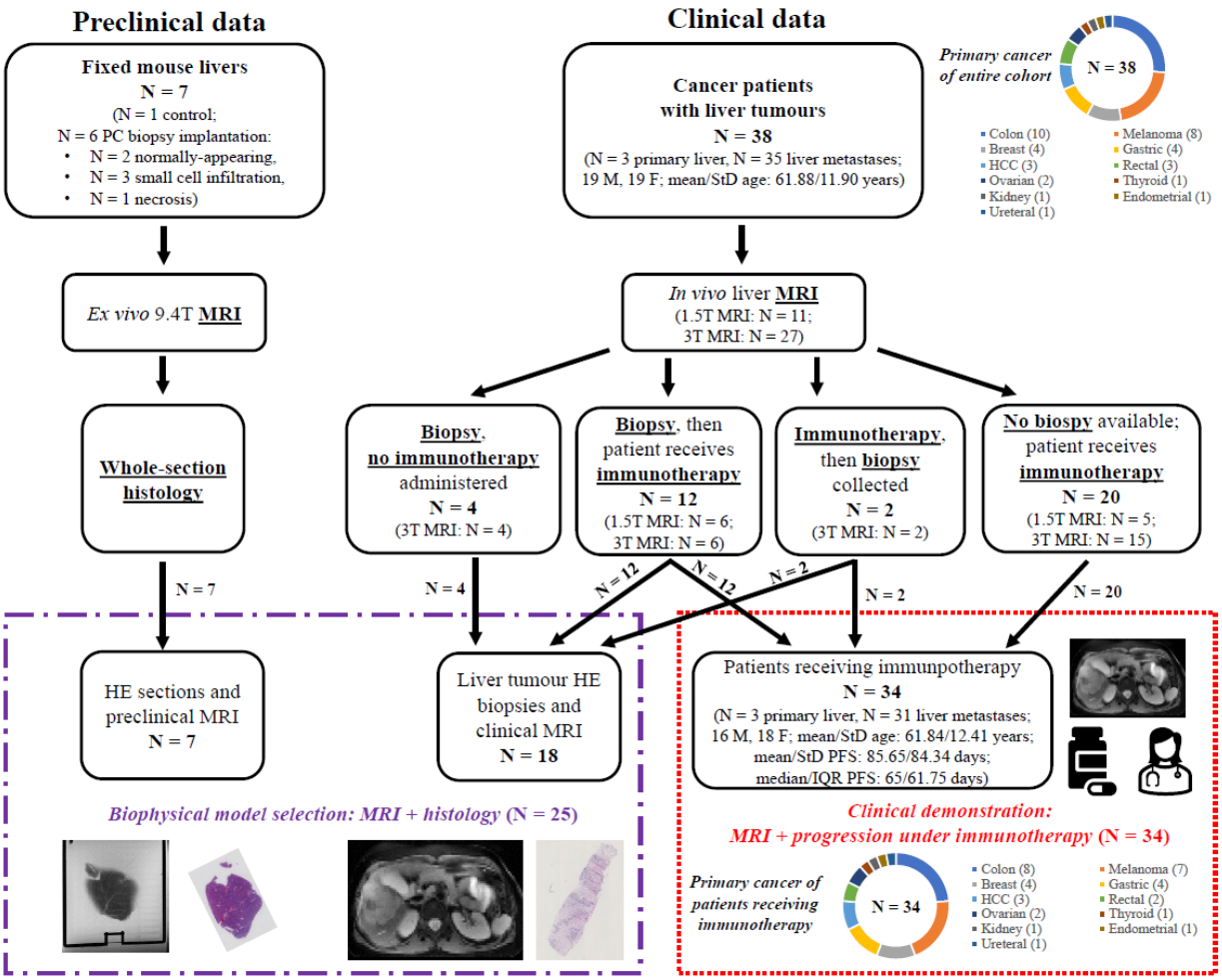

Figure 1. Illustration of the liver MRI and histology data used in the study. Our data set consisted of preclinical and clinical data. The preclinical data encompasses dMRI scans of seven fixed livers from mice (six implanted with tissue from biopsies of patients suffering from prostate cancer; one without 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. For 18 patients, histological material from a biopsy taken from one of the imaged liver tumours was also available. 34 out of 38 patients received immunotherapy as part of a phase I immunotherapy trial. For them, a clinical outcome was available, in terms of Progression Free Survival (PFS). Of the 18 patients for whom a biopsy was available, 4 did not receive immunotherapy, while 14 receive it (the biopsy was collected before starting therapy in 12 cases, while after starting receiving the treatment in 2 cases). We used a total sample size of N = 25 for dMRI model design and of N = 34 for response assessment. In the figure, PC stands for prostate cancer and HCC for hepatocellular carcinoma.

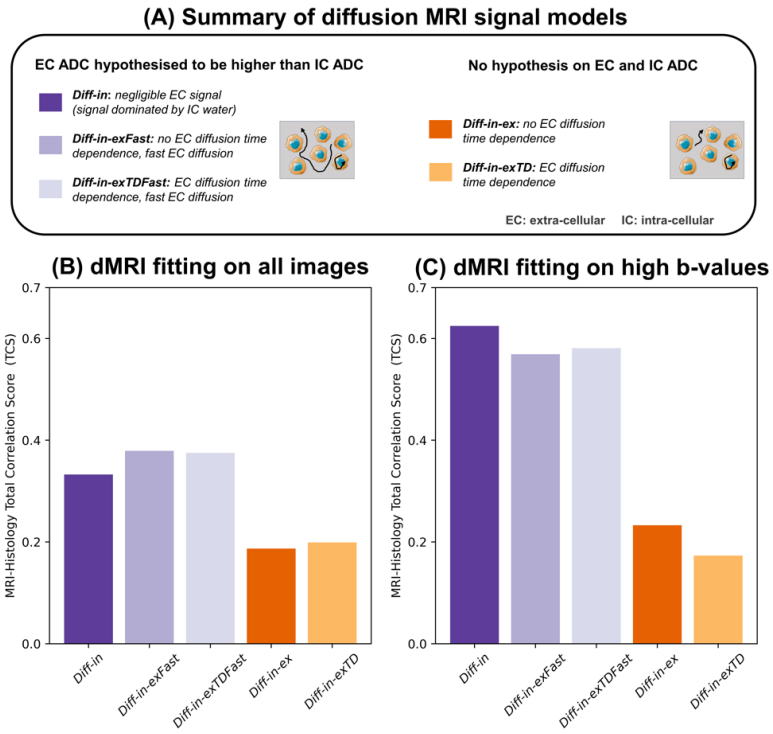

Figure 2. Biophysical dMRI signal model selection based on the MRI-histology Total Correlation Score (TCS). (**A**): panel summarising the salient differences between the biophysical dMRI models compared in this study. Models can be divided in two families, i.e.: i) models where it is hypothesised that 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*). In this figure, pictures from Servier Medical Art have been used. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (https://creativecommons.org/licenses/by/3.0).

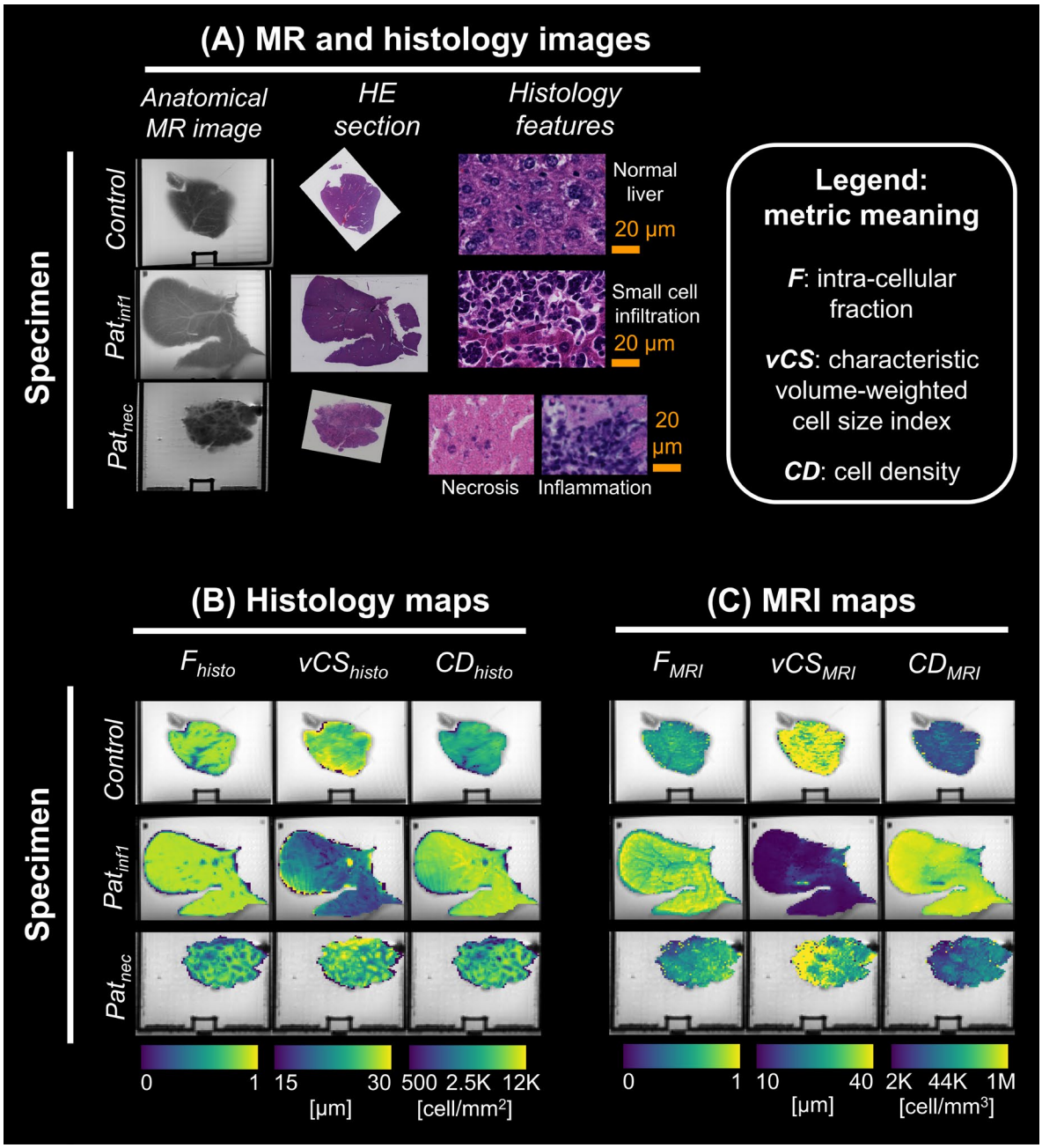

Figure 3. Maps from the selected dMRI model *Diff-in* with their histological counterparts in the fixed mouse livers scanned at 9.4T *ex vivo*. The figure reports MRI and histology data for 3 specimens, representative of the 3 microstructural phenotypes observed in our *ex vivo* data set, namely: normal 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).

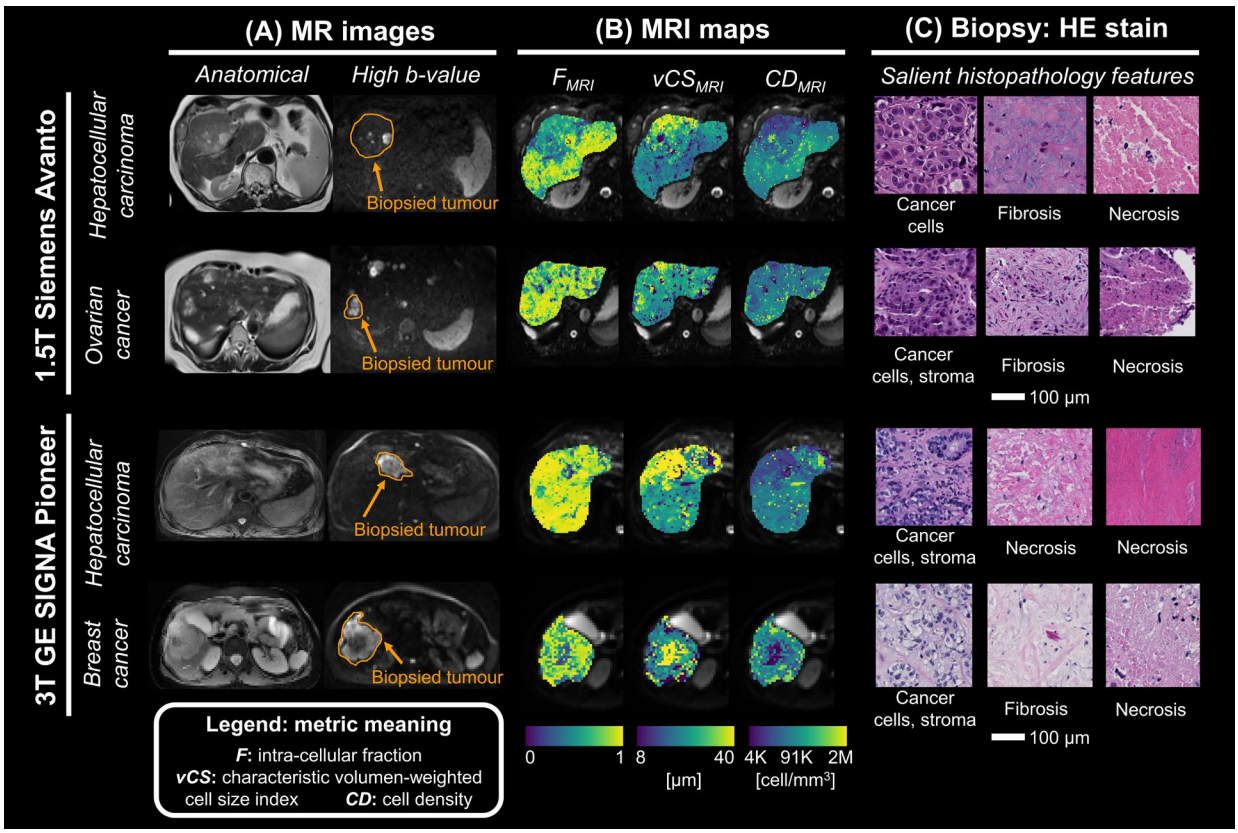

Figure 4. Examples of maps from the proposed dMRI model *Diff-in* in liver tumours of patients scanned at 1.5T and 3T *in vivo*, with co-localised biopsies. MRI maps are shown in a biopsied liver tumour in two patients for each MRI scanner, arranged along rows. **(A)**: examples of slices from the high-resolution anatomical T2-w image and from a high b-value image, with biopsied tumour outlined. **(B)**: 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 a primary hepatocellular carcinoma (HCC) case, and a case with liver metastases from ovarian cancer. For the 3T GE scanner (third and fourth rows from top) we report a primary HCC case and a case with liver metastases from breast cancer.

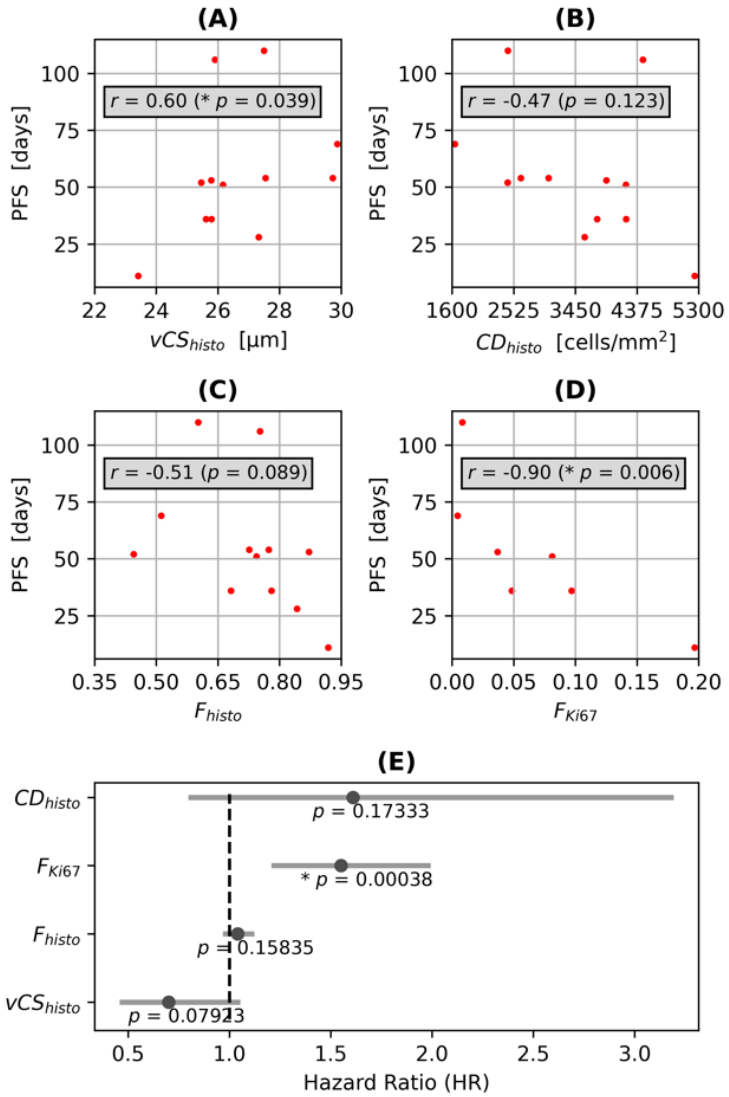

Figure 5. Relationship between immunotherapy clinical outcome and histological properties of the biopsied liver tumours. (**A**) to (**D**): Spearman's correlation coefficient between Progression Free Survival (PFS) and cellular characteristics evaluated on digitised liver tumour biopsies collected at baseline, before immunotherapy administration (**A**: histological cell size index vCS_{histo}; **B**: cell density per unit area CD_{histo}; **C**: intra-cellular fraction F_{histo}; **D**: Ki-67 staining fraction F_{Ki67}). (**E**): vCS_{histo}, CD_{histo}, F_{histo} and F_{Ki67} Hazard Ratios (HR), with P values and confidence intervals, as obtained through univariate Cox regression models that assess the association between PFS and vCS_{histo}, CD_{histo}, F_{histo} and F_{Ki67}.

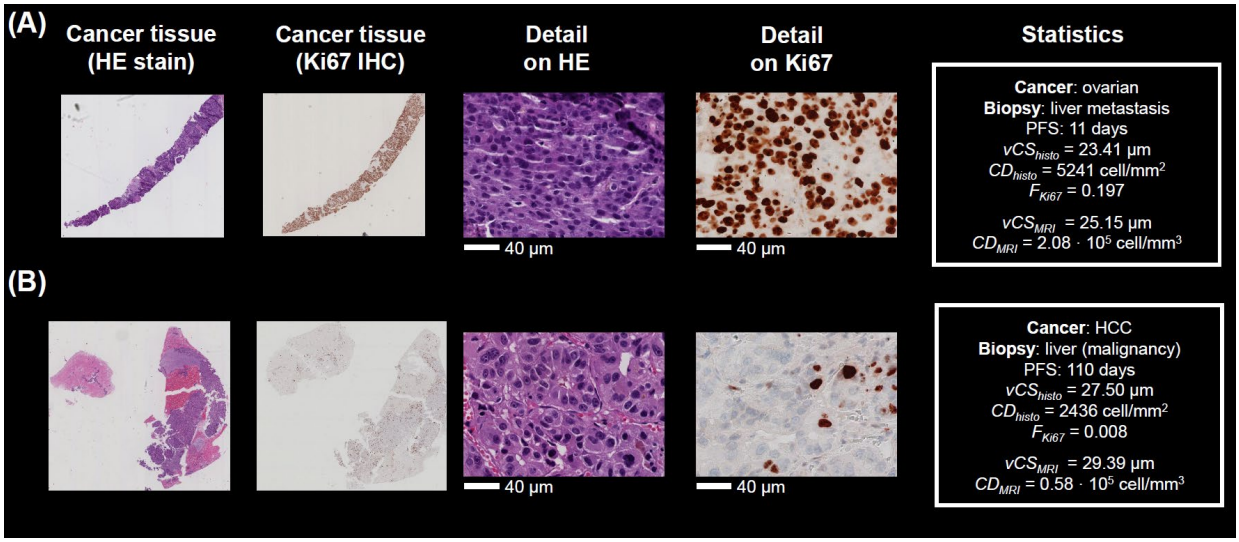

Figure 6. Example of histological correlates of cancer progression under immunotherapy. (**A**): HE image and Ki67 immunohistochemistry (IHC) with high-resolution details from a liver biopsy collected from a patient with fast cancer progression (patient 3: PFS of 11 days, liver metastasis from ovarian cancer). The figure also includes statistics of cell size, density and replication derived from the quantitative analysis of the biopsy, as well as the MRI counterparts of such metrics (cell size index vCS_{histo}/vCS_{MRI} from the biopsy/dMRI scan; cell density CD_{histo}/CD_{MRI} from the biopsy/dMRI scan; Ki67 staining fraction F_{Ki67} from the biopsy). (**B**): same information, but for a patient with slower progression (patient 6: PFS of 110 days, primary hepatocellular carcinoma (HCC)). MRI maps from these two patients are reported in Fig. 4 (first and second rows; both patients were scanned on the 1.5T).

852 **Tables**

853 **Table 1. Correlation between dMRI metrics and histological metrics.** The table reports
854 Pearson's correlation coefficients r and corresponding p-values p of dMRI metrics F_{MRI} (intra-
855 cellular fraction), vCS_{MRI} (volume-weighted cell size index) and CD_{MRI} (cell density per unit
856 volume) with their histological pairs (F_{histo}, vCS_{histo} and CD_{histo} respectively) for the selected
857 dMRI model (*Diff-in*, fitted to high b-value images). The table also reports correlation coefficients
858 between routine ADC and K from DKI and each of F_{histo}, vCS_{histo} and CD_{histo}. The sample size
859 was N = 25, so that $P < 0.05$ if $| r | > 0.3961$. When $P < 0.05$, grey shadowing is used.

dMRI technique	Histology F_{histo}	Histology vCS_{histo}	Histology CD_{histo}
Biophysical model *Diff-in*	With F_{MRI}: $r = 0.19$; $P = 0.37$	With vCS_{MRI}: $r = 0.44$; $P = 0.029$	With CD_{MRI}: $r = 0.70$; $P = 0.0001$
Routine DKI	With ADC: $r = -0.28$; $P = 0.18$	With ADC: $r = 0.49$; $P = 0.014$	With ADC: $r = -0.47$; $P = 0.017$
	With K: $r = 0.40$; $P = 0.048$	With K: $r = -0.31$; $P = 0.13$	With K: $r = 0.43$; $P = 0.033$

860

861

862

863

864

865

866

867

868

869

870

871

872

Table 2. Immunotherapy response assessment based on baseline MRI metric values. The table reports the cancer progression Hazard Ratio (HR, with 95% confidence interval and *P* value) for DKI metrics ADC and K and for the volume-weighted Cell Size (vCS_{MRI}) and Cell Density (CD_{MRI}) metrics, as obtained through fitting the selected biophysical liver dMRI model *Diff-in* to high b-value images. HRs were estimated by fitting a multi-variable proportional hazard Cox model. Events were represented by cancer progression under immunotherapy as recorded by the Progression Free Survival (PFS) metric. Age, sex and baseline liver tumour volume were taken into account as potential confounding factors, and included as regressors in the Cox model. Regressors whose HR features $P < 0.05$ are flagged by grey shadowing.

	Hazard Ratio (HR)	***P* value**
Cox model with *Diff-in* metrics		
Age	0.86 (0.63; 1.18)	0.363
Male sex	0.81 (0.39; 1.69)	0.579
Liver tumour volume	0.89 (0.75; 1.05)	0.152
vCS_{MRI}	0.45 (0.25; 0.81)	0.008
CD_{MRI}	0.55 (0.22; 1.36)	0.196
Cox model with DKI metrics		
Age	0.96 (0.67; 1.37)	0.806
Male sex	0.73 (0.34; 1.57)	0.418
Liver tumour volume	0.87 (0.75; 1.02)	0.078
ADC	1.17 (0.84; 1.62)	0.354
K	1.44 (1.02; 2.03)	0.036

Supplementary Figures

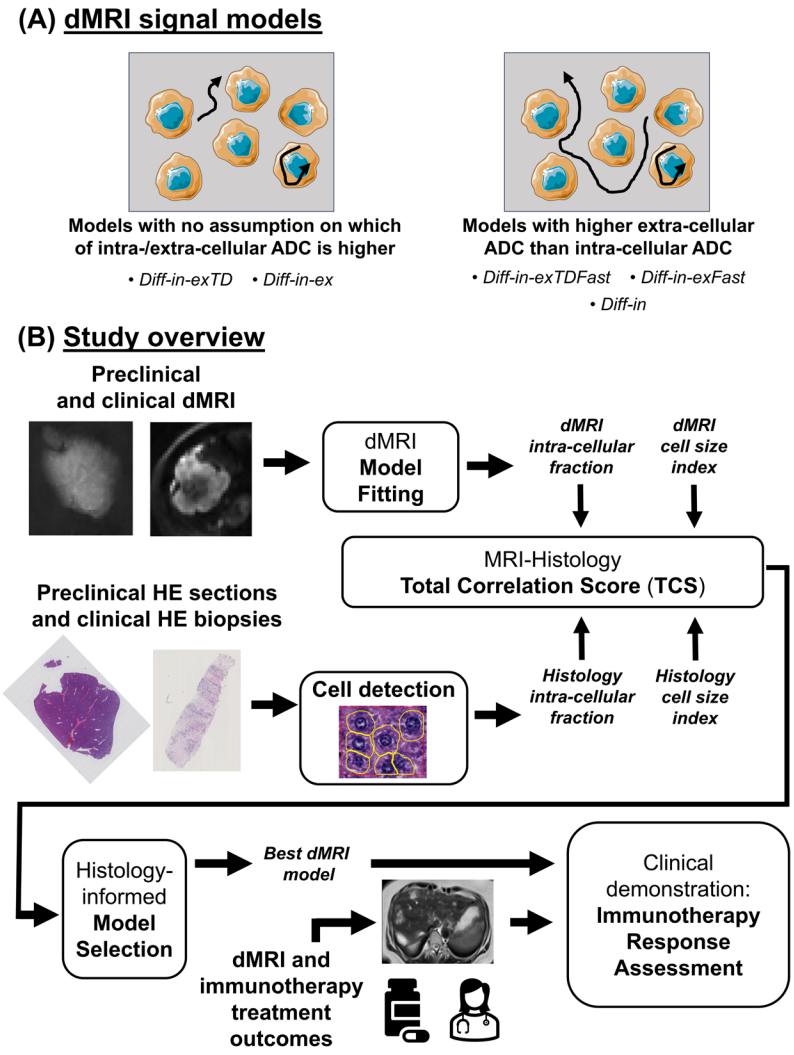

Figure S1. Description of the dMRI signal models and study overview. (**A**), top: cartoon illustrating the two families of dMRI models considered in this study, consisting of 1) models with no assumption of which of intra/extra-cellular ADC is higher, and 2) models where the extra-cellular ADC is hypothesised to be higher than the intra-cellular ADC. (**B**), bottom: study overview. We analysed dMRI data from fixed mouse livers (preclinical data) and from cancer patients imaged in vivo (clinical data) 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. The utility of the model was then demonstrated in immunotherapy response assessment in vivo. In this figure, pictures from Servier Medical Art have been used. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (https://creativecommons.org/licenses/by/3.0).

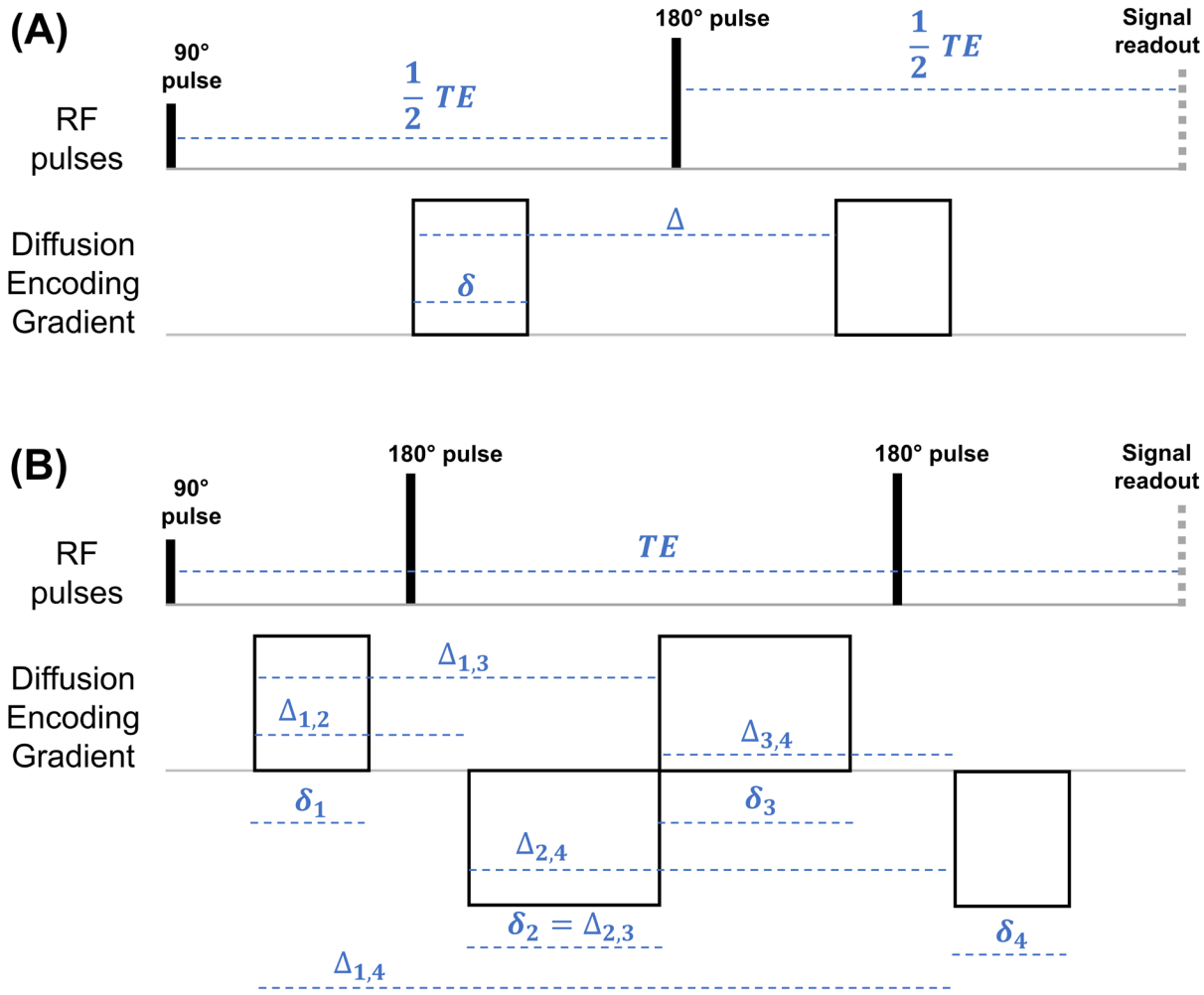

Figure S2. Schematic of the dMRI sequences used in this study.
(**A**): pulsed gradient spin echo (PGSE sequence, also known as Stejskal-Tanner sequence, pulsed-field gradient (PFG), or single linear diffusion encoding) used to acquire data on the 9.4T Bruker system on fixed mouse livers *ex vivo* and on the 3T GE system on patients *in vivo*. δ and Δ respectively indicate the diffusion gradient duration and separation, while TE is the echo time. (**B**): twice-refocussed diffusion-weighted spin echo sequence used to acquire data on the 1.5T Siemens system on patients *in vivo*. δ_n and $\Delta_{n,m}$ respectively indicate the duration of the *n*-th gradient lobe and the separation time between the *n*-th and *m*-th gradient lobes, for *n,m* = 1, …, 4. TE is again the echo time. In both panels, "Signal readout" corresponds to sampling the center of the k-space (i.e., zero spatial frequency).

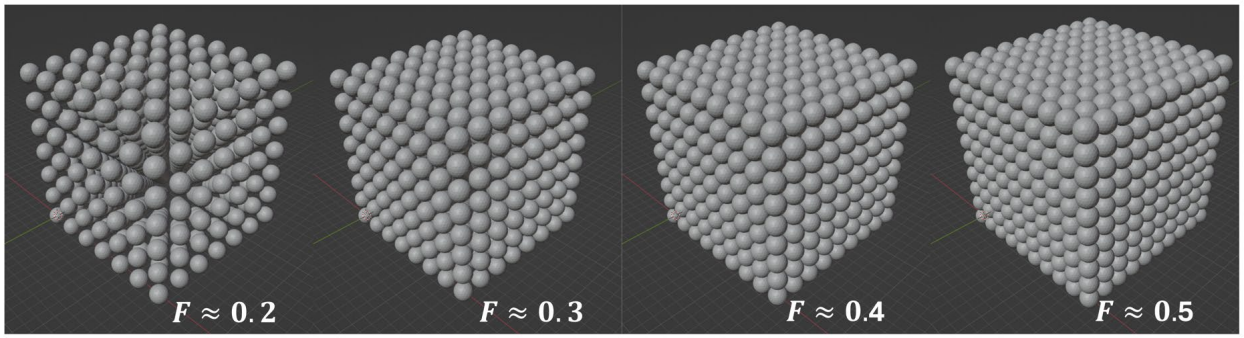

Figure S3. Environments used to generate synthetic dMRI signals in computer simulations. The synthetic environment consisted of meshed spheres of fixed diameter, representing cells, as this is a common biophysical model used in several dMRI techniques (e.g., VERDICT, IMPULSED). We used the synthetic environment to generate dMRI signals via Monte Carlo simulations for each of the 3 dMRI protocols considered in this study (the PGSE protocol used on the *ex vivo* mouse livers at 9.4T; the PGSE protocol used in patients *in vivo* at 3T; the DW TRSE protocol used in patients *in vivo* at 1.5T). Afterwards, we performed performed dMRI model selection on the synthetic signals, following the same procedures implemented for actual *ex vivo* and *in vivo* dMRI data. We controlled the intra-sphere fraction F by adding gaps of increasing size in-between abutting spheres packed in an ideal cubic lattice. We probed 4 different values of F (approximately equal to 0.197, 0.323, 0.406, 0.523; notice that the maximum theoretical value of F for cubic lattice packing is equal to 0.5236). For each value of F, we varied the cell diameter (8, 16, 22 and 30 μm), intra-sphere diffusivity (10 linearly-spaced values in the ranges [0.8; 2.6] $\mu m^2\ ms^{-1}$ and [0.8; 3.0] $\mu m^2\ ms^{-1}$ for the *ex vivo* and *in vivo* protocols respectively) and extra-sphere intrinsic diffusivity (again, 10 linearly-spaced values in the ranges [0.8; 2.6] $\mu m^2\ ms^{-1}$ and [0.8; 3.0] $\mu m^2\ ms^{-1}$ for the *ex vivo* and *in vivo* protocols respectively), generating a total of 1600 synthetic voxels. Before dMRI signal model fitting, we corrupted synthetic signal with Rician noise at a signal-to-noise ratio (SNR) of 30 on the b = 0 signal $s(b=0)$ (SNR = $s(b=0)/\sigma$, where σ is the noise standard deviation).

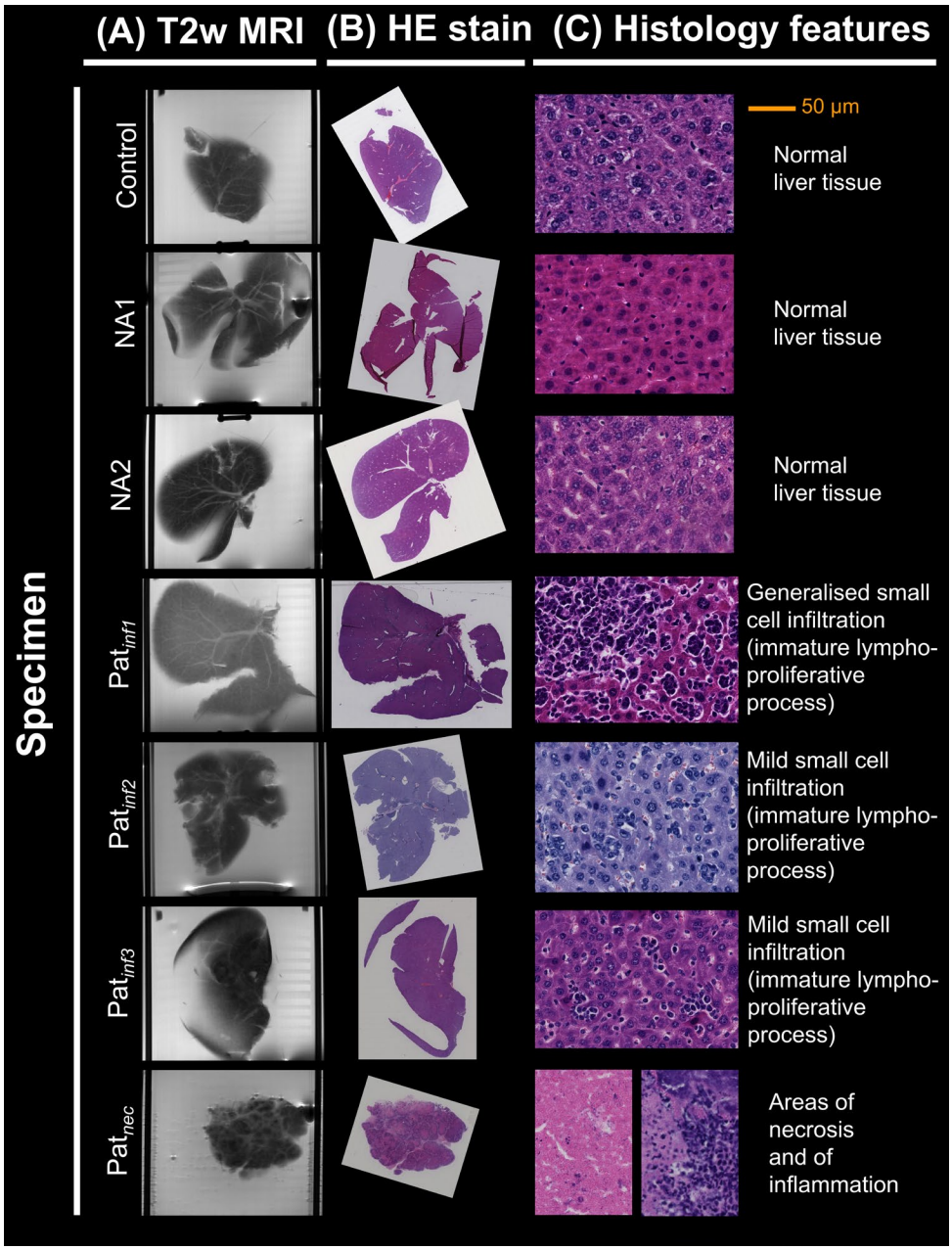

Figure S4. Radiological-histological co-localisation of the *ex vivo* mouse liver data.

Illustration of the radiological-histological co-localisation on the 7 fixed mouse livers obtained from mice implanted with a biopsy from a prostate cancer patient. **(A)**, left: illustrative slice of the high-resolution anatomical T2-weighted fast spin echo. **(B)**, centre: hematoxylin and eosin (HE)-stained section, taken from the MRI slice shown to the left. **(C)**, right: detail of the microstructure characterising each specimen, as assessed by an experienced pathologist (SS). Different specimens are arranged along different rows. From top to bottom: *Control*, normal liver structures (no biopsy implantation); Pat_{NA1} and Pat_{NA2}, normal appearing normal liver structures after prostate cancer biopsy implantation; Pat_{inf1}, Pat_{inf2} and Pat_{inf3}: pathology following implantation, consisting of an immature, lympho-proliferative process (infiltration of small cells in sinusoidal spaces); Pat_{nec}, pathology following implantation, consisting of necrosis and inflammation.

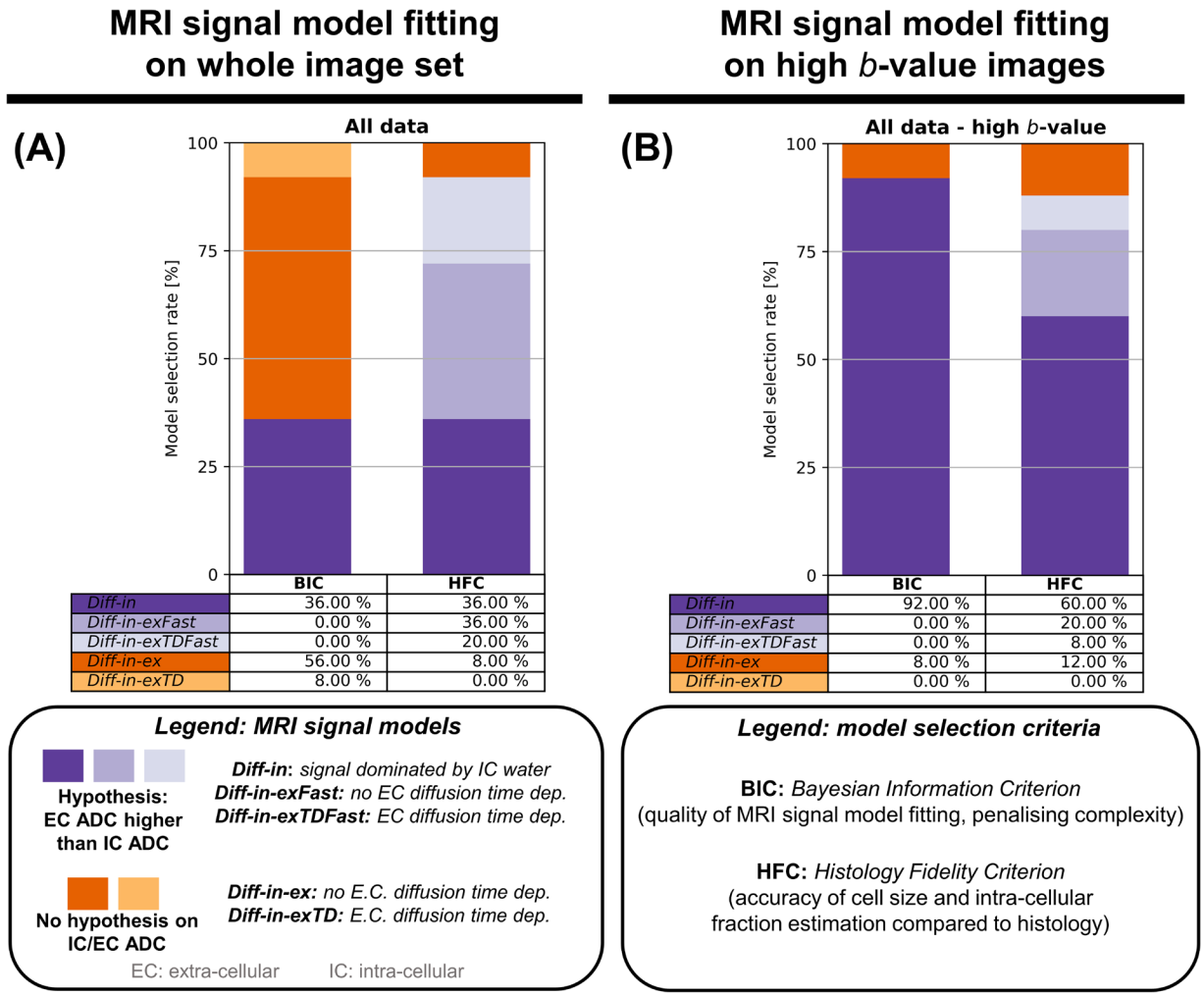

Figure S5. Biophysical dMRI signal model selection based on BIC and HFC criteria.
Frequency of model selection based on the *Bayesian Information Criterion* (BIC, quantifying how well a model fits the dMRI signals, penalising model complexity) and on the *Histology-fidelity Criterion* (HFC, quantifying how accurately a dMRI models estimates the intra-cellular fraction and the volume-weighted cell size as seen on histology). (**A**) reports results when models are fitted to the entire set of measurements with negligible vascular constributions (b > 1000 s/mm^2 for suppression of PBS fluid within vessels in the fixed mouse livers; b > 100 s/mm^2 for IVIM signal suppression *in vivo* on clinical systems), while (**B**) reporting results obtained when fitting models only on high b-value images (b > 1800 s/mm^2 in fixed mouse livers; b > 900 s/mm^2 *in vivo*).

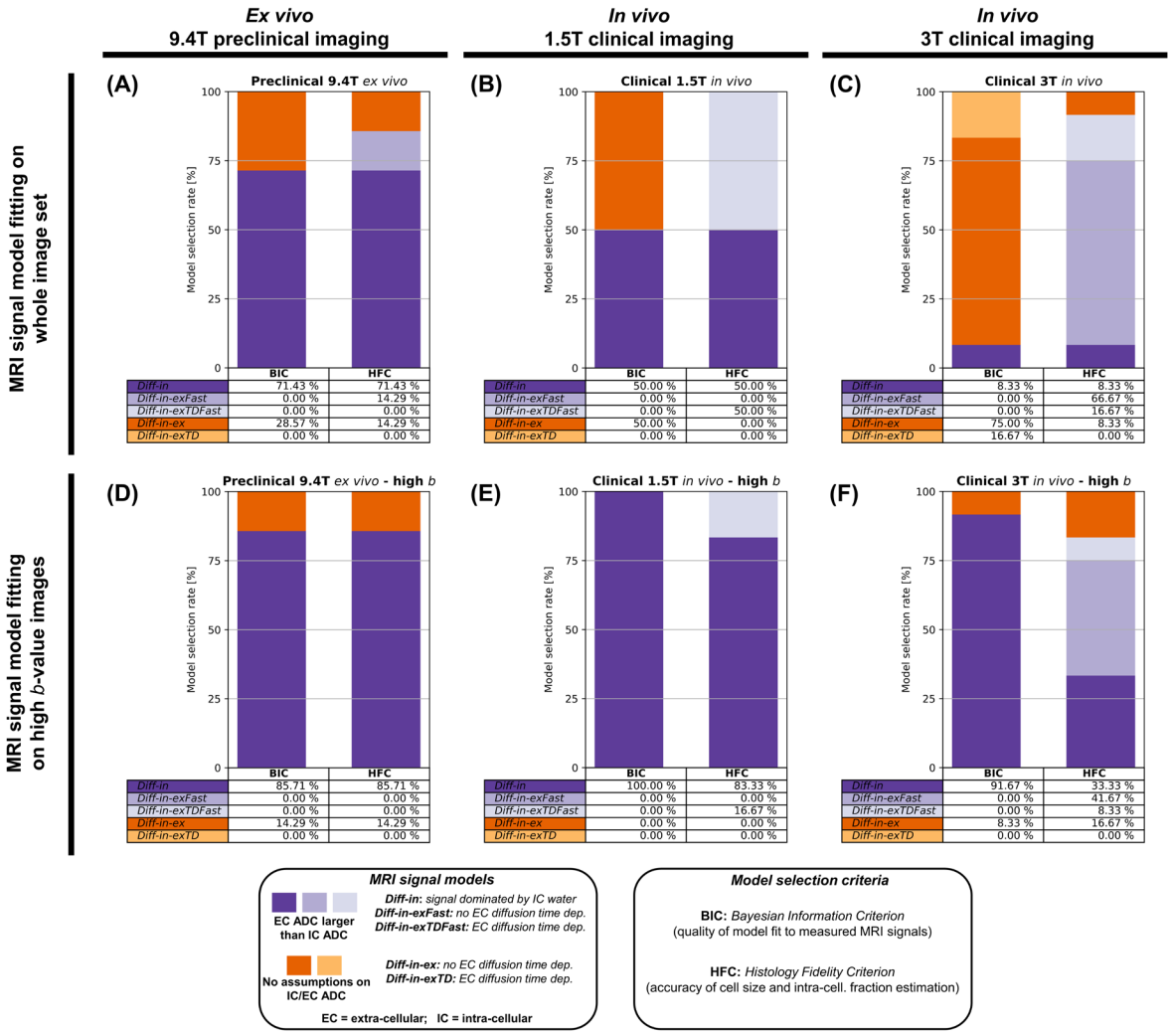

Figure S6. Biophysical model selection with BIC and HFC criteria across different MRI scanners and data subsets.

Frequency of selection of each of 5 biophysical dMRI models on 3 MRI-histology data subsets. First column: selection on 7 fixed mouse livers scanned *ex vivo* on a preclinical 9.4T Bruker system (**A** and **D**, left); Second column: selection on 6 liver tumours imaged *in vivo* on a clinical 1.5T Siemens system (**B** and **E**, middle); Third column: selection on 12 liver tumours imaged *in vivo* on a clinical 3T GE system (**C** and **F**, right). Plots on top (**A** to **C**) refer to dMRI model fitting performed on images where the vascular signal was suppressed ("whole image set fitting", b > 1000 s/mm^2 for suppression of PBS fluid within vessels on the 9.4T; b > 100 s/mm^2 for IVIM signal suppression on clinical systems). Plots to the bottom (**D** to **F**) refer to dMRI model fitting performed on images where both vascular and extra-cellular, extra-vascular signals were suppressed ("high b-value fitting", b > 1800 s/mm^2 on the 9.4T; b > 900 s/mm^2 on clinical systems). Violet: models where extra-cellular ADC is larger than intra-cellular ADC; orange: models with no constraints on which is larger between intra-/extra-cellular ADC. The *Bayesian Information Criterion* (BIC) selects a model depending on the goodness of model fitting. The *Histology Fidelity Criterion* (HFC) selects a model depending on the agreement between MRI volume-weighted Cell Size (*vCS*) and intra-cellular fraction (*F*) estimation with histology.

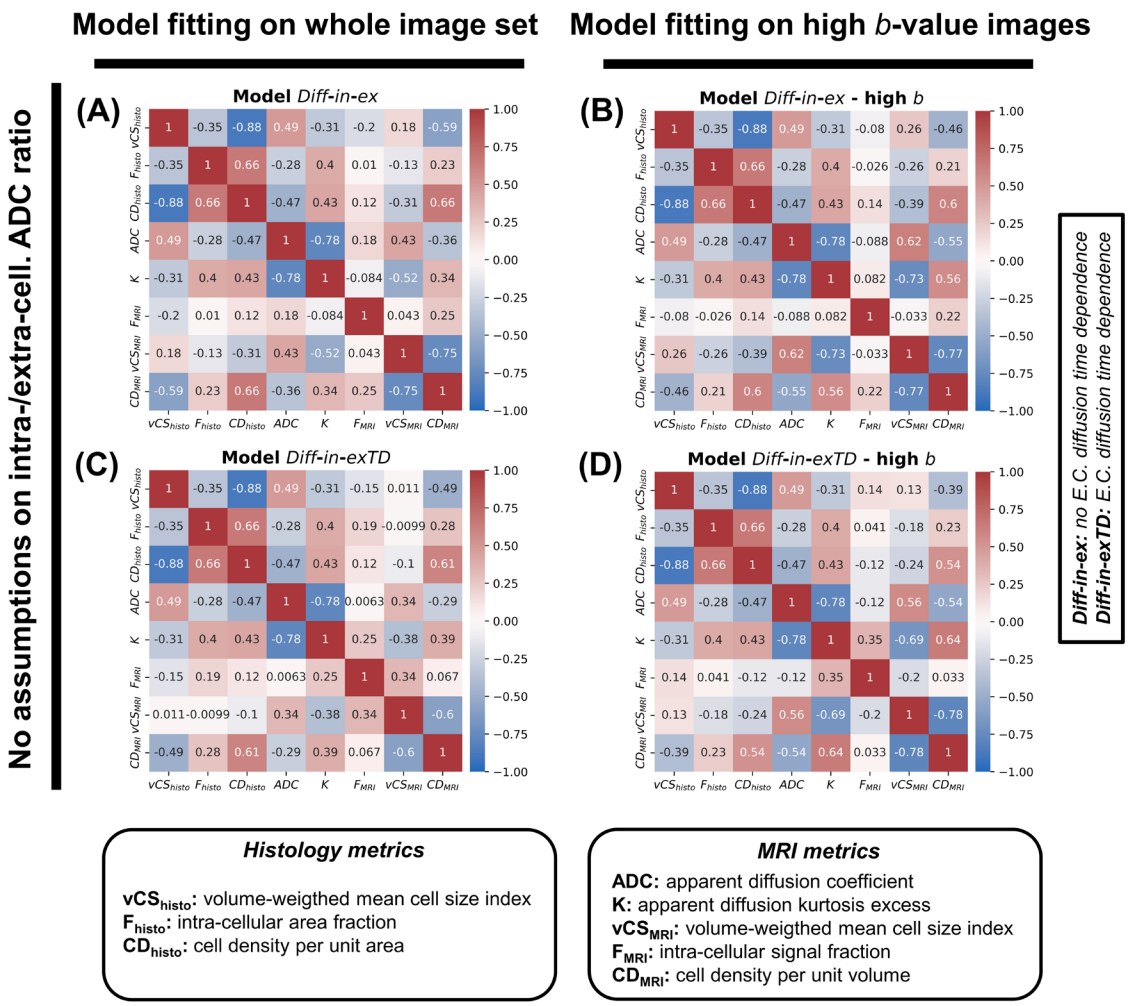

Figure S7. MRI-histology correlations for models with no assumptions on which is larger between intra-cellular and extra-cellular ADC. Matrices illustrating Pearson's correlation coefficients among all possible pairs of MRI and histology metrics. Histological metrics are: intra-cellular area fraction F_{histo}; volume-weighted mean cell size index vCS_{histo}; cell density per unit area CD_{histo}. MRI metrics are: apparent diffusion coefficient ADC; apparent diffusion excess kurtosis K; intra-cellular area fraction F_{MRI}; volume-weighted mean cell size index vCS_{MRI}; cell density per unit area CD_{MRI}. Metrics F_{MRI}, vCS_{MRI} and CD_{MRI} were obtained by fitting models with no assumptions on which is larger between intra-cellular and extra-cellular ADC (*Diff-in-ex* and *Diff-in-exTD*). The 4 panels refer to models *Diff-in-ex* and *Diff-in-exTD* fitted according to 2 different strategies. Panel (**A**): model *Diff-in-ex* fitted on the whole set of measurements with vascular signal suppression (b > 100 s/mm^2 *in vivo*, b > 1000 s/mm^2 *ex vivo*); panel (**B**): model *Diff-in-ex* fitted on high b-value measurements (b > 900 s/mm^2 *in vivo*, b > 1800 s/mm^2 *ex vivo*); panel (**C**): model *Diff-in-exTD* fitted on the whole set of measurements with vascular signal suppression; panel (**D**): model *Diff-in-exTD* fitted on high b-value measurements. We calculated correlation coefficients using a sample size of N = 25.

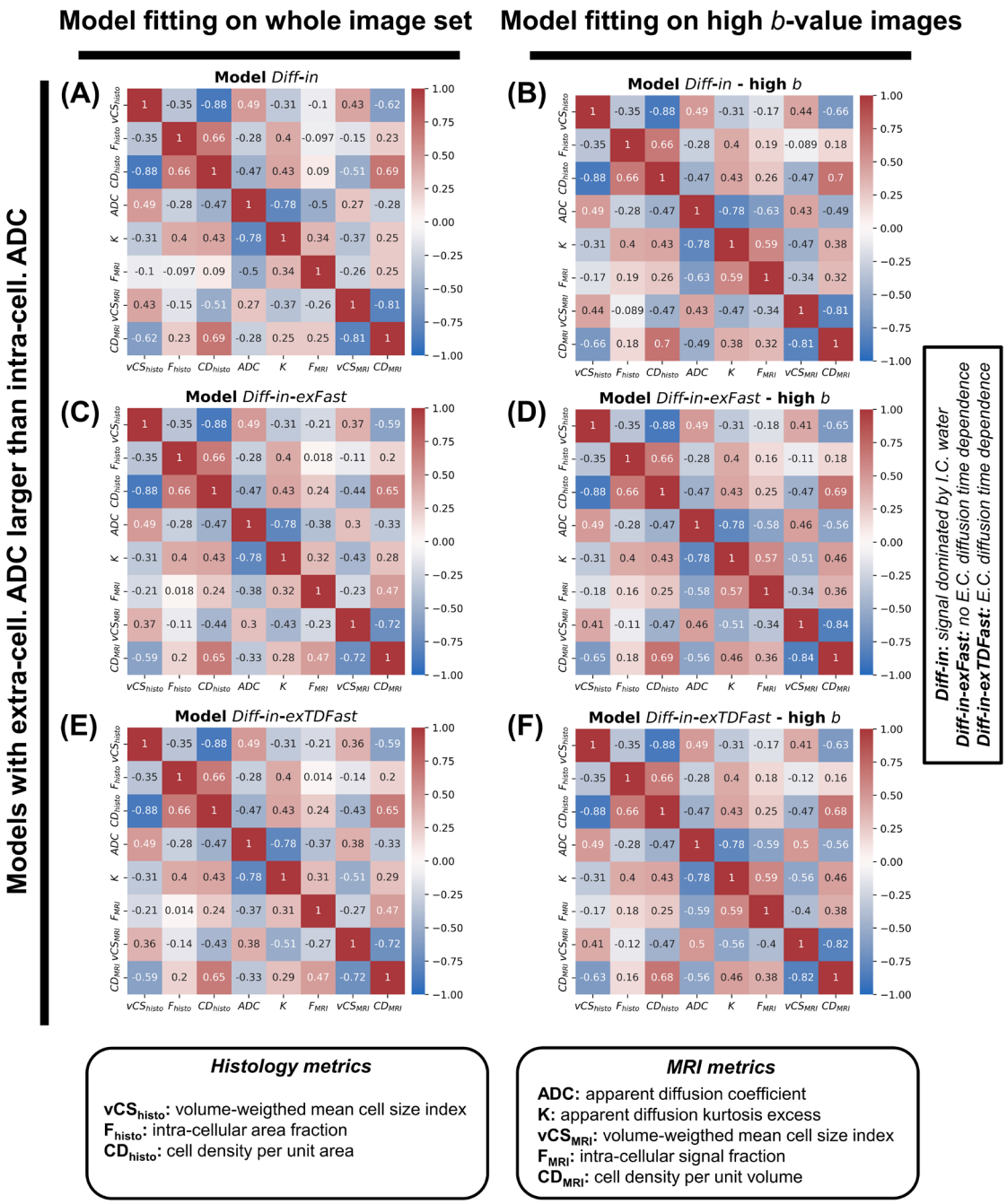

Figure S8. MRI-histology correlations for models where the extra-cellular ADC is constrained to be larger than the intra-cellular ADC. Matrices illustrating Pearson's correlation coefficients among all possible pairs of MRI and histology metrics. Histological metrics are: intra-cellular area fraction F_{histo}; volume-weighted mean cell size index vCS_{histo}; cell density per unit area CD_{histo}. MRI metrics are: apparent diffusion coefficient ADC; apparent diffusion excess kurtosis K; intra-cellular area fraction F_{MRI}; volume-weighted mean cell size index vCS_{MRI}; cell density per unit area CD_{MRI}. Metrics F_{MRI}, vCS_{MRI} and CD_{MRI} were obtained by fitting models that assume that the extra-cellular ADC is always larger than the intra-cellular ADC (*Diff-in*, *Diff-in-exFast* and *Diff-in-exTDFast*). The 6 panels refer to models *Diff-in*, *Diff-in-exFast* and *Diff-in-*

exTDFast fitted according to 2 different strategies. Panel (**A**): model *Diff-in* fitted on the whole set of measurements with vascular signal suppression (b > 100 s/mm^2 *in vivo*, b > 1000 s/mm^2 *ex vivo*); panel (**B**): model *Diff-in* fitted on high b-value measurements (b > 900 s/mm^2 *in vivo*, b > 1800 s/mm^2 *ex vivo*); panel (**C**): model *Diff-in-exFast* fitted on the whole set of measurements with vascular signal suppression; panel (**D**): model *Diff-in-exFast* fitted on high b-value measurements; panel (**E**): model *Diff-in-exTDFast* fitted on the whole set of measurements with vascular signal suppression; panel (**F**): model *Diff-in-exTDFast* fitted on high b-value measurements. We calculated correlation coefficients using a sample size of N = 25.

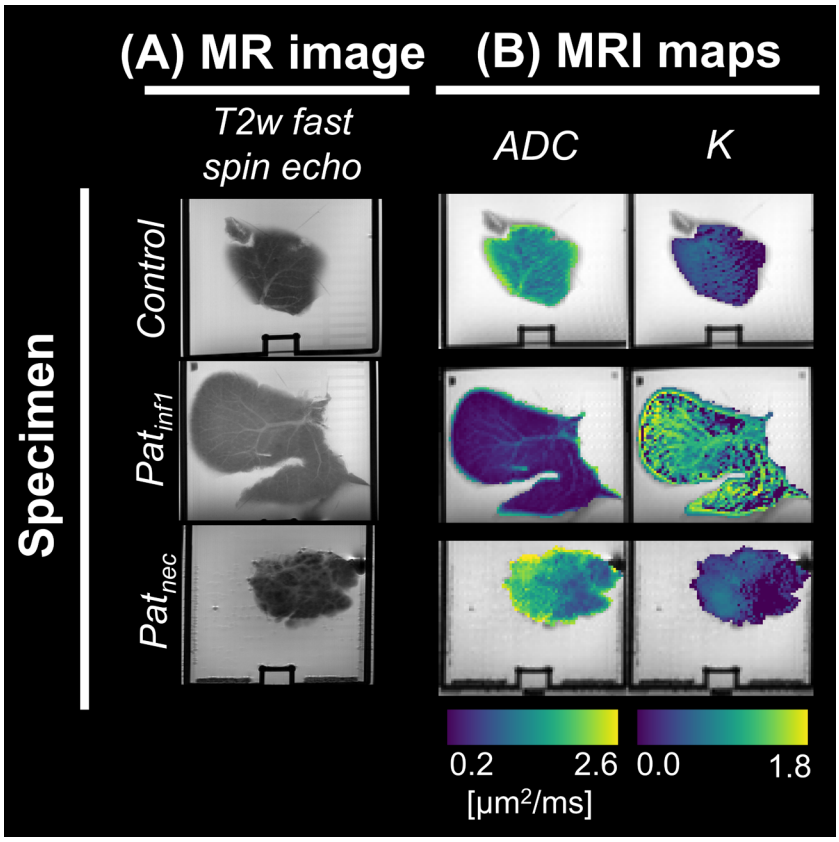

Figure S9. Standard diffusion MRI metrics in fixed mouse livers *ex vivo*.

(A): high-resolution fast spin echo scan acquired in fixed mouse livers scanned *ex vivo* on the 9.4T Bruker system. **(B)**: standard diffusion metrics, namely *ADC* (apparent diffusion coefficient) and *K* (apparent diffusion kurtosis excess). These metrics were obtained by fitting the standard diffusion kurtosis signal representation $s = s_0 \exp(- b\ ADC + K\ (b\ ADC)^2/6)$ to the set of measurements at fixed TE = 45 ms and Δ = 30 ms. From top to bottom, the figure reports maps from 3 specimens, representative of the 3 different microstructural phenotypes seen in our mouse liver data. These are: normal liver structures (illustrated by the *Control* case, e.g., mouse with no biopsy implantation); pathology following biopsy implantation, consisting of an immature, lympho-proliferative process (infiltration of small cells in sinusoidal spaces, illustrated by case Pat_{inf}); pathology following biopsy implantation, consisting of necrosis and inflammation (illustrated by case Pat_{nec}).

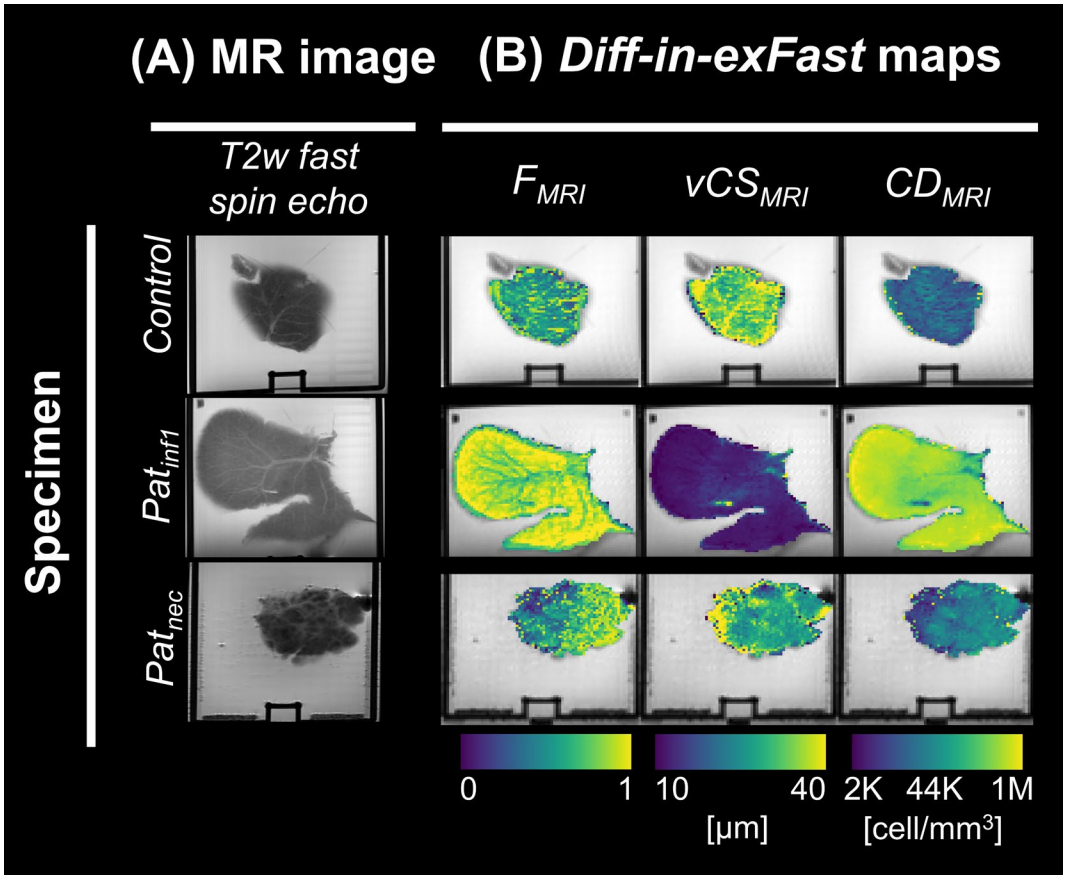

Figure S10. Key parametric maps of the *Diff-in-exFast* model on *ex vivo* mouse livers.

(A): high-resolution fast spin echo scan acquired in fixed mouse livers scanned *ex vivo* on the 9.4T Bruker system. **(B)**: metrics from the *Diff-in-exFast* model fitted to the whole DW image set (b-values with negligible vascular signal contributions, i.e., $b > 1000$ s/mm^2 on fixed *ex vivo* tissue, to suppress signal from PBS-filled vessels). From left to right: intra-cellular signal fraction F_{MRI}; volume-weighted cell size index vCS_{MRI} index; cell density per unit volume CD_{MRI}. Maps from 3 specimens are reported along different rows. The specimens are representative of the 3 different microstructural phenotypes seen in our mouse liver data. From top to bottom, these are: normal liver structures (illustrated by the *Control* case, e.g., mouse with no biopsy implantation); pathology following biopsy implantation, consisting of an immature, lympho-proliferative process (infiltration of small cells in sinusoidal spaces, illustrated by case Pat_{inf}); pathology following biopsy implantation, consisting of extended necrosis and inflammation (illustrated by case Pat_{nec}).

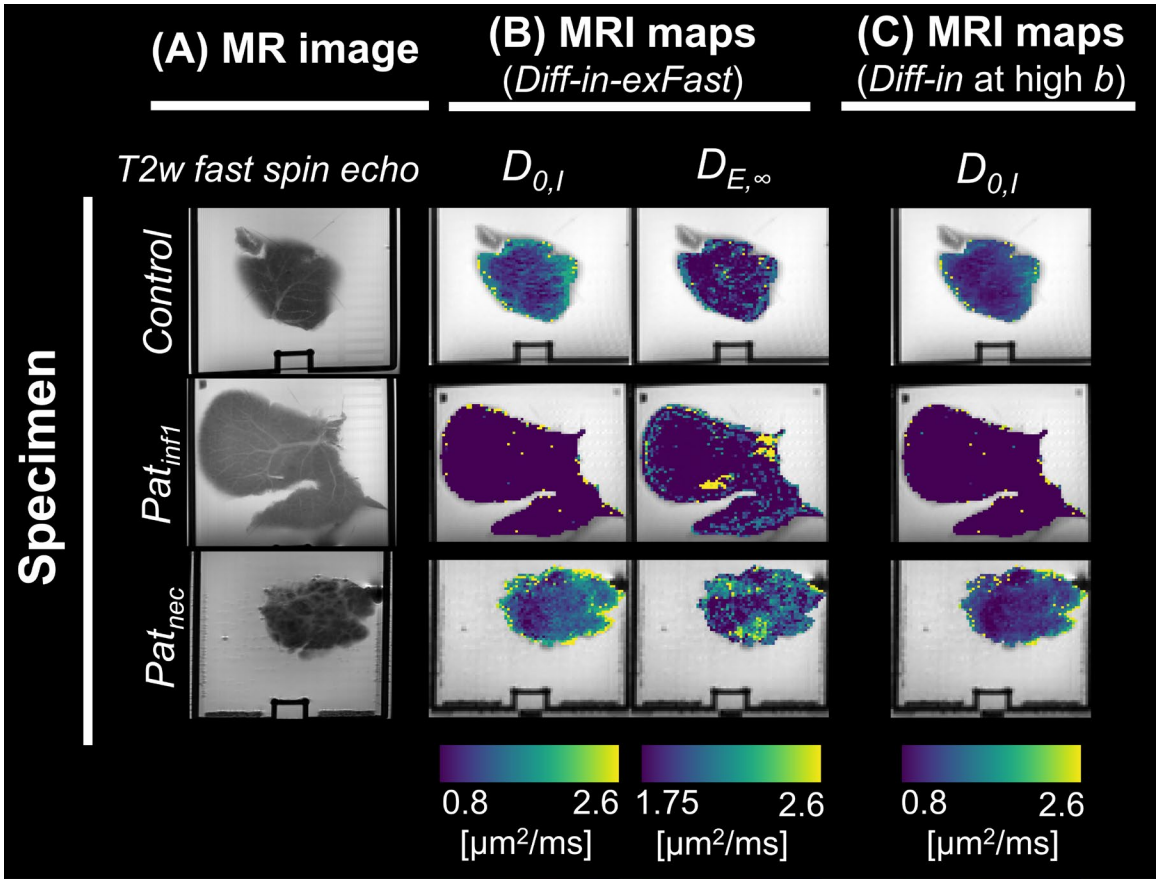

Figure S11. Diffusivity metrics from biophysical MRI models in fixed *ex vivo* mouse livers. **(A)**: high-resolution fast spin echo scan acquired in fixed mouse livers scanned *ex vivo* on the 9.4T Bruker system. **(B)**: diffusivity metrics from biophysical model *Diff-in-exFast*, namely: intrinsic intra-cellular cytosolic diffusivity $D_{0,I}$; asymptotic extra-cellular diffusion coefficient $D_{E,\infty}$. **(C)**: intrinsic intra-cellular cytosolic diffusivity $D_{0,I}$ from model *Diff-in* fitted to high b-value images (b > 1800 s/mm^2). Maps from 3 specimens are reported along different rows. The specimens are representative of the 3 different microstructural phenotypes seen in our mouse liver data. From top to bottom, these are: normal liver structures (illustrated by the *Control* case, e.g., mouse with no biopsy implantation); pathology following biopsy implantation, consisting of an immature, lympho-proliferative process (infiltration of small cells in sinusoidal spaces, illustrated by case Pat_{inf}); pathology following biopsy implantation, consisting of extended necrosis and inflammation (illustrated by case Pat_{nec}).

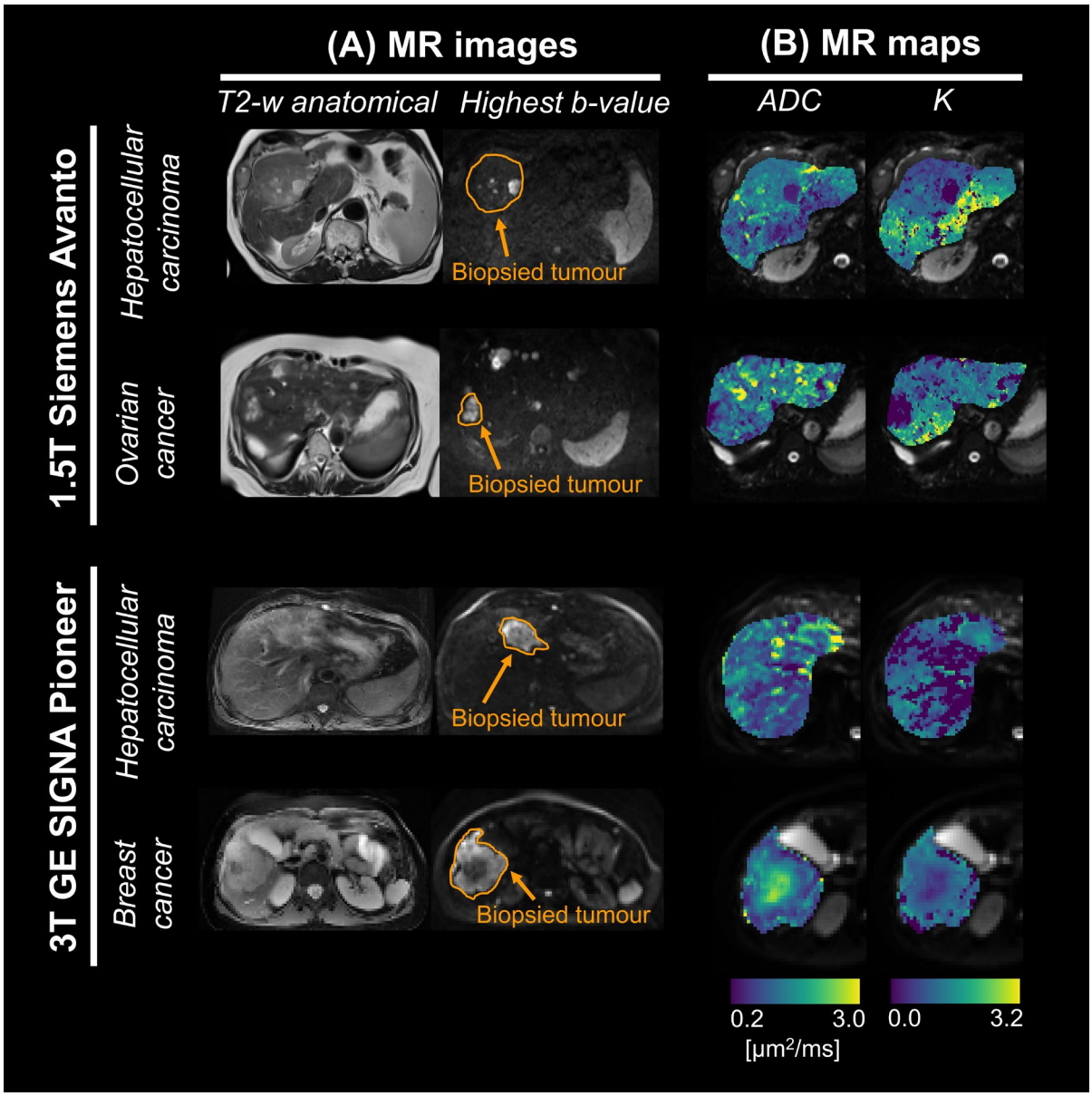

Figure S12. Standard diffusion MRI metrics in patients *in vivo*.

(A): high-resolution fast spin echo scan as well as a high b-value diffusion image, with biopsied tumour outlined. **(B)**: standard diffusion metrics in the biopsied tumour. Metrics are: apparent diffusion coefficient (*ADC*) and apparent diffusion kurtosis excess (*K*). These were obtained by fitting the standard diffusion kurtosis signal representation $s = s_0 \exp(-b\,ADC + K\,(b\,ADC)^2/6)$ to the set of measurements at fixed, minimum TE and $b > 100$ s/mm^2. Maps are shown in four representative patients (two patients for each MRI scanner), along different rows. For the 1.5T Siemens scanner (first and second rows from top): a primary hepatocellular carcinoma case and a patient with liver metastases from ovarian cancer. For the 3T GE scanner (third and fourth rows from top): a primary hepatocellular carcinoma case and a patient with liver metastases from breast cancer.

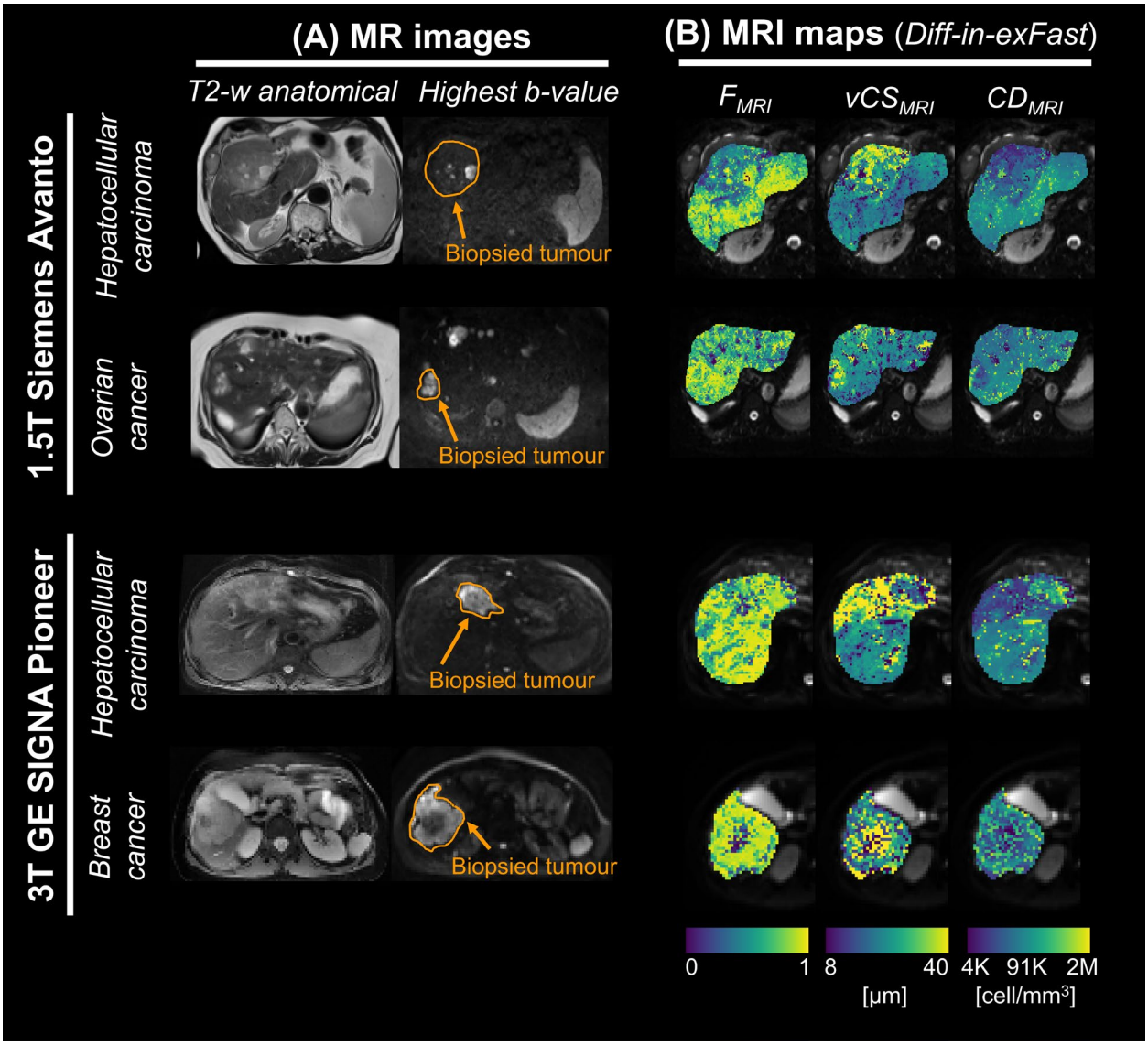

Figure S13. Key parametric maps of the *Diff-in-exFast* model in patients *in vivo*.

(A): high-resolution fast spin echo scan as well as a high b-value diffusion image, with biopsied tumour outlined. **(B)**: salient metrics of the *Diff-in-exFast* model fitted to the whole set of images with negligible vascular signal contributions (b > 100 s/mm^2). Metrics are shown in the biopsied tumour. From left to right: intra-cellular signal fraction F_{MRI}; volume-weighted cell size index vCS_{MRI} index; cell density per unit volume CD_{MRI}. Metrics are shown in four representative patients (two patients for each MRI scanner), along different rows. For the 1.5T Siemens scanner (first and second rows from top): a primary hepatocellular carcinoma case and a patient with liver metastases from ovarian cancer. For the 3T GE scanner (third and fourth rows from top): a primary hepatocellular carcinoma case and a patient with liver metastases from breast cancer.

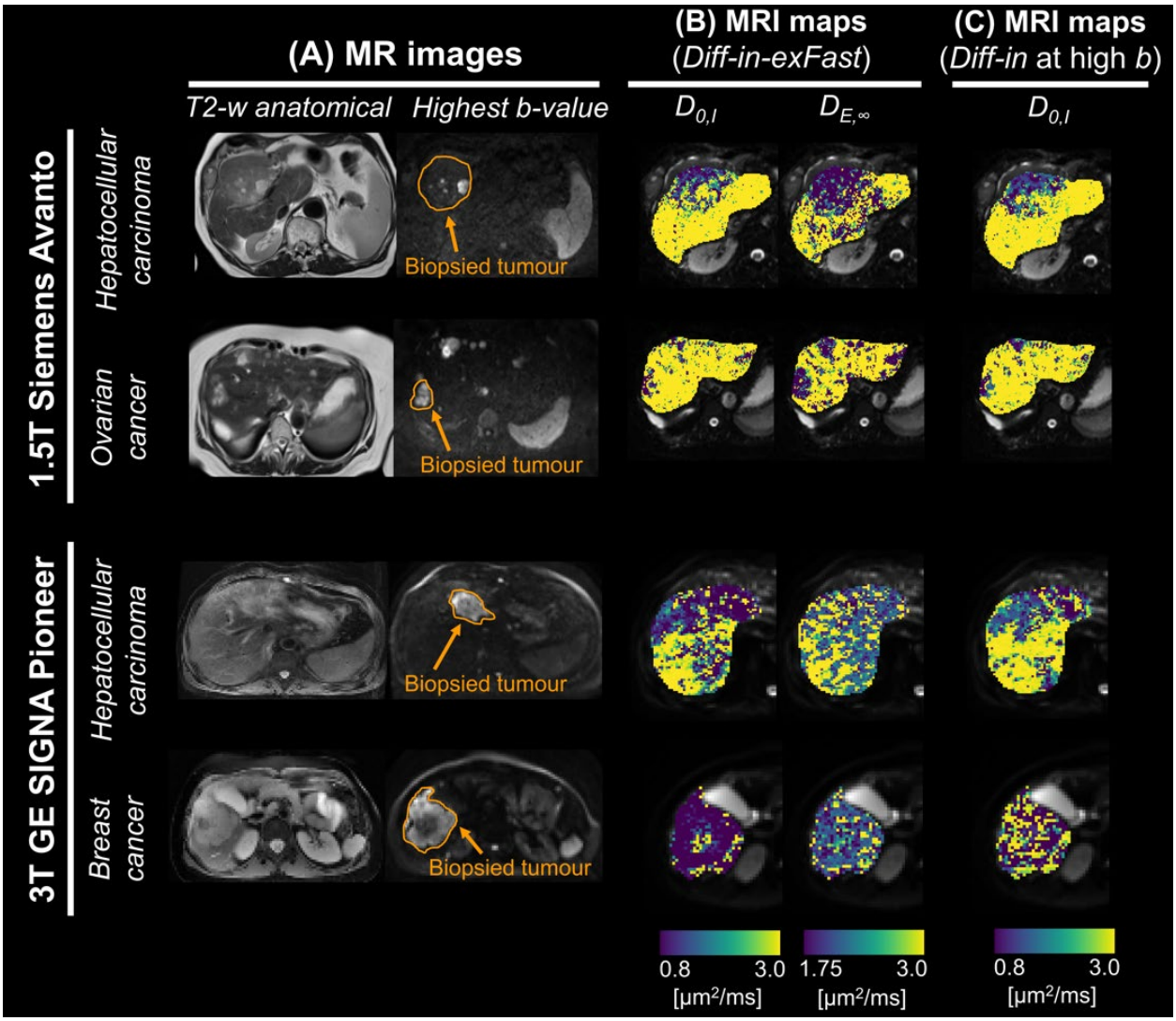

Figure S14. Diffusivity metrics from biophysical MRI models in patients *in vivo*.

(A): high-resolution fast spin echo scan as well as a high b-value diffusion image, with biopsied tumour outlined. **(B)**: diffusivity maps from biophysical model *Diff-in-exFast* in the biopsied tumour. Metrics are: intra-cellular cytosolic diffusivity $D_{0,I}$ and asymptotic extra-cellular diffusion coefficient $D_{E,\infty}$. **(C)**: intra-cellular cytosolic diffusivity $D_{0,I}$ for biophysical model *Diff-in* fitted being fitted only to high b-value images ($b > 900$ s/mm^2). Metrics are shown in four representative patients (two patients for each MRI scanner), along different rows. For the 1.5T Siemens scanner (first and second rows from top): a primary hepatocellular carcinoma case and a patient with liver metastases from ovarian cancer. For the 3T GE scanner (third and fourth rows from top): a primary hepatocellular carcinoma case and a patient with liver metastases from breast cancer.

Supplementary tables

Table S1: results of the model selection based on the Total Correlation Score (TCS) as obtained on simulated dMRI signals.

We performed model selection on synthetic signals simulated for all the dMRI protocols considered in this study (*ex vivo* PGSE, used on fixed mouse livers; *in vivo* PGSE and DW TRSE, used in patients *in vivo*; see Methods for a full description of the protocols). We fitted the models on protocol subsets obtained with the same b-value thresholds used when analysing actual MRI signals ("Regular fit": fitting on all b-values with negligible vascular contributions; "High *b* only fit": fitting on b-values minimising extra-cellular signal contributions). The Table reports the value of $TCS = r(vCS_{est}, vCS_{gt}) + r(F_{est}, F_{gt})$, where vCS is the cell size, F the intra-cellular fraction, $r(x,y)$ the Pearson's correlation between variables x and y computed pooling together all synthetic voxels, and where subscripts *est* and *gt* respectively indicate estimated and ground truth values. Higher values of *TCS* point towards better model performance. For each protocol and fitting strategy, the model with the highest *TCS* is flagged by gray shadowing and bold font.

	Protocol: *ex vivo*		**Protocol: *in vivo* PGSE**		**Protocol: *in vivo* TRSE**	
	Regular fit	High *b* only fit	Regular fit	High *b* only fit	Regular fit	High *b* only fit
Model						
Diff-in-exTD	0.217	–0.111	0.312	0.202	0.604	0.483
Diff-in-ex	0.406	0.089	0.472	0.335	**0.700**	0.550
Diff-in-exTDFast	0.948	0.827	0.536	0.336	0.618	0.544
Diff-in-exFast	0.952	0.850	0.563	0.349	0.626	0.547
Diff-in	**1.222**	**0.977**	**0.773**	**0.462**	0.543	**0.630**

Table S2. Results of the model selection based on the Histology Fidelity Criterion (HFC) as obtained on simulated dMRI signals.

We performed model selection on synthetic signals simulated for all the dMRI protocols considered in this study (*ex vivo* PGSE, used on fixed mouse livers; *in vivo* PGSE and DW TRSE, used in patients *in vivo*; see Methods for a full description of the protocols). We fitted the models on protocol subsets obtained with the same b-value thresholds used when analysing actual MRI signals ("Regular fit": fitting on all b-values with negligible vascular contributions; "High *b* only fit": fitting on b-values minimising extra-cellular signal contributions). For each model, the table reports the percentage of synthetic voxels where $HFC = |vCS_{est} - vCS_{gt}|/vCS_{gt} + |F_{est} - F_{gt}|/F_{gt}$ was the lowest across all models. Above, vCS is the cell size, F the intra-cellular fraction, and subscripts *est* and *gt* respectively indicate estimated and ground truth values. Higher percentages indicate smaller estimation errors, and therefore point towards better model performance. For each protocol and fitting strategy, the model with the highest proportion of synthetic voxels with minimum HFC is flagged by gray shadowing and bold font.

	Protocol: *ex vivo*		**Protocol: *in vivo* PGSE**		**Protocol: *in vivo* TRSE**	
	Regular fit	High *b* only fit	Regular fit	High *b* only fit	Regular fit	High *b* only fit
Model						
Diff-in-exTD	12.62%	13.56%	21.69%	22.10%	26.69%	20.69%
Diff-in-ex	21.19%	16.62%	**25.94%**	20.35%	**28.94%**	24.31%
Diff-in-exTDFast	15.50%	14.94%	10.75%	12.27%	19.69%	14.69%
Diff-in-exFast	21.81%	10.50%	22.25%	11.33%	19.12%	12.12%
Diff-in	**28.88%**	**44.38%**	19.38%	**33.94%**	5.56%	**28.19%**

Table S3. Results of the model selection based on the Bayesian Information Criterion (BIC) as obtained on simulated dMRI signals.

We performed model selection on synthetic signals simulated for all the dMRI protocols considered in this study (*ex vivo* PGSE, used on fixed mouse livers; *in vivo* PGSE and DW TRSE, used in patients *in vivo*; see Methods for a full description of the protocols). We fitted the models on protocol subsets obtained with the same b-value thresholds used when analysing actual MRI signals ("Regular fit": fitting on all b-values with negligible vascular contributions; "High *b* only fit": fitting on b-values minimising extra-cellular signal contributions). For each model, the table reports the percentage of synthetic voxels where the Bayesian Information Criterion (BIC, a standard metric of model fitting quality that penalises model complexity) was the lowest across all models. Higher percentages indicate smaller BIC values across synthetic voxels, and therefore point towards better model fitting quality. For each protocol and fitting strategy, the model with the highest proportion of synthetic voxels with minimum HFC is flagged by gray shadowing and bold font.

	Protocol: *ex vivo*		**Protocol: *in vivo* PGSE**		**Protocol: *in vivo* TRSE**	
	Regular fit	High *b* only fit	Regular fit	High *b* only fit	Regular fit	High *b* only fit
Model						
Diff-in-exTD	0.12%	0.00%	0.00%	0.19%	0.00%	0.00%
Diff-in-ex	31.25%	15.88%	17.44%	4.69%	**40.69%**	13.69%
Diff-in-exTDFast	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Diff-in-exFast	3.44%	0.06%	6.69%	0.00%	21.38%	1.06%
Diff-in	**65.19%**	**84.06%**	**75.88%**	**95.12%**	38.00%	**85.25%**

Table S4. Descriptive statistics of histology and MRI metrics in the fixed mouse livers.

The table reports mean and standard deviation (within brackets) of histology and dMRI metrics in the 7 fixed mouse livers that were scanned on a 9.4T Bruker system. Histological maps were computed within patches matching the in-plane MRI resolution and then warped non-linearly to dMRI space. The histological maps are: per-patch intra-cellular area fraction F_{histo}, per-patch arithmetic mean cell size aCS_{histo}; per-patch volume-weighted mean cell size vCS_{histo}, cell density per unit patch area CD_{histo}. dMRI metrics are: apparent diffusion coefficient ADC, apparent diffusion kurtosis excess K, intra-cellular signal fraction F_{MRI}, volume-weighted cell size index vCS_{MRI}, apparent cell density per unit volume CD_{MRI}. Metrics F_{MRI}, vCS_{MRI} and CD_{MRI} are reported for both models *Diff-in-exFast* and model *Diff-in*, with *Diff-in* fitted only to high b-value images (b > 1800 s/mm^2). In model *Diff-in-exFast*, the extra-cellular ADC does not feature diffusion time dependence and is constrained to be larger than the intra-cellular ADC. In model *Diff-in*, the extra-cellular signal is modelled as negligible compared to the intra-cellular one (i.e., total signal dominated by intra-cellular water). Specimens are: *Control* (normal liver structures); *NA1* and *NA2* (normal appearing cases, i.e., normal liver structures despite sub-cutaneous biopsy implantation); $Pat_{inf1\text{-}3}$ (cases developing liver pathology following sub-cutaneous biopsy implantation, consisting of small cell infiltration in sinusoidal spaces, in between larger hepatocytes); Pat_{nec} (case developing liver pathology following sub-cutaneous biopsy implantation, consisting of necrosis and inflammation). aCS_{histo}, always considerably lower than vCS_{histo}, was included to highlight the impact of the largest cells in the computation of statistics based on weighting by cell volume (vCS_{histo}).

	Histology				Standard diffusion metrics		*Diff-in-exFast* model			*Diff-in* model at high *b*		
Specimen	F_{histo}	aCS_{histo} [μm]	vCS_{histo} [μm]	$CD_{histo}/10^2$ [cell/mm^2]	ADC [μm^2/ms]	K	F_{MRI}	vCS_{MRI} [μm]	$CD_{MRI}/10^5$ [cell/mm^3]	F_{MRI}	vCS_{MRI} [μm]	$CD_{MRI}/10^5$ [cell/mm^3]
Control	0.75 (0.18)	21.63 (1.85)	27.1 (1.6)	3.2 (1.1)	1.50 (0.32)	0.33 (0.20)	0.60 (0.18)	33.2 (5.7)	0.38 (1.69)	0.56 (0.11)	36.5 (5.1)	0.28 (1.44)
NA1	0.59 (0.25)	23.97 (3.05)	27.8 (3.2)	2.1 (1.1)	1.61 (0.37)	0.10 (0.15)	0.66 (0.25)	30.5 (8.4)	0.69 (2.18)	0.54 (0.16)	32.2 (7.5)	0.45 (1.68)
NA2	0.76 (0.14)	23.46 (1.41)	29.4 (1.2)	2.7 (0.7)	1.43 (0.44)	0.51 (0.29)	0.80 (0.16)	22.2 (4.8)	1.01 (1.35)	0.71 (0.12)	22.8 (5.0)	0.83 (1.15)
Pat_{inf1}	0.80 (0.20)	15.73 (2.18)	20.8 (2.7)	6.7 (2.2)	0.58 (0.41)	0.98 (0.49)	0.83 (0.13)	13.4 (3.1)	4.20 (2.22)	0.77 (0.14)	12.4 (3.0)	4.86 (2.27)
Pat_{inf2}	0.79 (0.20)	21.41 (2.29)	26.8 (2.7)	3.5 (1.2)	1.67 (0.39)	0.17 (0.17)	0.37 (0.22)	31.2 (7.5)	0.44 (2.06)	0.41 (0.14)	37.1 (6.0)	0.30 (1.57)
Pat_{inf3}	0.70 (0.27)	20.95 (1.75)	27.6 (1.7)	3.1 (1.3)	1.57 (0.62)	0.43 (0.31)	0.63 (0.23)	23.6 (8.8)	1.27 (2.62)	0.59 (0.20)	24.2 (8.0)	0.97 (2.01)
Pat_{nec}	0.52 (0.25)	19.10 (3.33)	25.9 (2.8)	3.4 (2.3)	1.49 (0.44)	0.31 (0.22)	0.61 (0.23)	28.2 (6.6)	0.58 (1.93)	0.54 (0.16)	31.6 (6.7)	0.42 (1.66)

Table S5. Immunotherapy response assessment based on regressing MRI metrics in turn. The table reports the cancer progression Hazard Ratio (HR, with 95% confidence interval and *P* value) for DKI metrics ADC and K and for all metrics computed through Diff-in model analysis, fitted on high b-value data (volume-weighted Cell Size vCS_{MRI}; Cell Density CD_{MRI}; intra-cellular fraction F_{MRI}; intrinsic intra-cellular, cytosolic diffusivity $D_{0,I}$; vascular signal fraction f_V). HRs were estimated by fitting a multi-variable proportional hazard Cox model where all metrics were included as regressors in turn (one MRI metric per regression). Additionally, all Cox models included age, sex and baseline liver tumour volume as potential confounding factors. Events were represented by cancer progression under immunotherapy as recorded by the Progression Free Survival (PFS) metric.

	Hazard Ratio (HR)	***P* value**
Cox models with *Diff-in* metrics included in turn		
vCS_{MRI}	0.72 (0.49; 1.05)	0.09
CD_{MRI}	1.11 (0.63; 1.95)	0.72
F_{MRI}	0.79 (0.54; 1.15)	0.22
$D_{0,I}$	1.0 (0.69; 1.43)	0.98
f_V	1.05 (0.71; 1.57)	0.79
Cox models with DKI metrics included in turn		
ADC	0.98 (0.65; 1.46)	0.91
K	1.34 (0.98; 1.85)	0.07

Table S6. Spearman's correlation coefficient between each of the HE-derived histological properties and Ki67 staining fraction F_{Ki67}. The table reports the Spearman's correlation coefficient *r* and the corresponding *P* value between each of the histological metrics derived from HE-stained biopsies (volume-weighted Cell Size index vCS_{histo}, intra-cellular fraction F_{histo} and cell density per unit area CD_{histo}) and the Ki67 staining fraction from IHC (F_{Ki67}). Stasticially significant correlations ($P < 0.05$) are flagged by grey shadowing.

HE-derived histological metric	Spearman's correlation coefficient
vCS_{histo}	$r = -0.667$; $P = 0.04987$
F_{histo}	$r = 0.717$; $P = 0.02981$
CD_{histo}	$r = 0.900$; $P = 0.00094$

Supplementary methods

This document provides additional details for each section of the Methods in the main manuscript.

All experimental protocols in animals 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. The imaging study in cancer patients was approved by the Vall d'Hebron University Hospital Ethics committee (PR(AG)29/2020), Barcelona, Spain. Patients provided informed written consent to participate in the study.

dMRI models

Common biophysical body dMRI signal models[1–6] 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[7]. *In vivo*, the IVIM vascular ADC ranges approximately[8] in [15; 60] μm^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[3]

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

A common model for a_I in Eq. 2 is that of restricted diffusion within spheres of diameter L[2,3]:

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

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[9]. Above, α_m is the m-th root of $\alpha_m R \, J'_{3/2}(\alpha_m R) - 0.5 \, J_{3/2}(\alpha_m R) = 0$, $J_{3/2}(x)$ is the Bessel function of the first kind and order 3/2, and $J'_{3/2}(x)$ its first-order derivative. 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[10,11], 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[3,5,12]:

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

with

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

In Eq. 6, $D_{E,\infty}$ is the asymptotic[5] 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[2,13].

<u>*Fitting*</u>

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[14]. Fitting provides estimates of vCS and voxel intra-cellular signal fraction

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

We also combined vCS and F into a cell density per unit volume[2]

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

Preclinical data

Animals

We obtained data from 7 fixed livers of NOD.Cg-Prkdc^{scid} IL2rg^{tm1Wjl}/SzJ mice. 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. 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. 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 anesthesia when tumour volume exceeded 2000 mm^3. We collected the livers, fixed them overnight in formalin, and transferred them to phosphate-buffered saline (PBS) solution.

MRI

We scanned livers on a 9.4T Bruker Avance system at room temperature. Livers were tightened with sewing thread to a histology cassette and placed into a Falcon® tube, filled with PBS solution. A 1-channel birdcage coil was used (excitation/reception). The protocol included a T2-weighted fast spin echo sequence (resolution: 144 µm × 144 µm × 2.216 mm) and PGSE dMRI (Fig. S2; TR = 2700 ms; resolution: 386 µm × 386 µm; matrix size: 86 × 86; 4 slices, 2.216 mm-thick, NEX = 1). The protocol featured: δ = 10 ms, Δ = {15, 30} ms, 10 linearly spaced b-values for each Δ (minimum/maximum nominal *b*: 0/2800 s/mm^2). DW images corresponding to Δ = 15 ms were acquired at each of TE = {31, 45, 65} ms, and to Δ = 30 ms at each of TE = {45, 65} ms. We i) denoised dMRI scans with Marchenko-Pastur Principal Component Analysis (MP-PCA) Python denoising[15] (kernel: 7×7×3), ii) mitigated Gibbs ringing (MrTrix3 local sub-voxel shift method[16]), and iii) corrected temporal signal drifts by assessing signal changes in a PBS solution region, accounting for TE (PBS T2: 500 ms).

Finally, we fitted the *Diff-in-exTD*, *Diff-in-exTDFast*, *Diff-in-ex*, *Diff-in-exFast* and *Diff-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; 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] 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*). For fitting, we fixed f_V and $T2_T$ to values obtained through a a two-pool vascular-tissue model[17] (fitting bounds: [0; 1] for f_V; [5; 80] ms for $T2_T$). Fitting was performed i) on all images with b > 1000 s/mm^2 (suppressing vascular signals, referred to as *fitting 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.

<u>*Histology*</u>

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[18], 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}[11]. 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[10,11] $aCS_{histo} = < l >$. 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[19]). 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%[19]). Lastly, we co-registered histological maps to MRI[11] using DiPy[20].

Clinical data

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 Ethics committee (PR(AG)29/2020). Patients provided informed written consent to participate in the imaging study. We included 38 patients with liver malignancies (mean/std of age: 61.88/11.90 years; 19 male, 19 female), of which 3 suffered from primary 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). 34 out of 38 patients received immunotherapy after dMRI, as part of an ongoing phase I clinical trial at the Vall d'Hebron Institute of Oncology of Barcelona. Immunotherapy included immune-checkpoint inhibitors (anti-PD1, anti-PDL1, anti-CTLA4, LAG3) in monotherapy or in combination with other drugs, and other cancer immunotherapies as for example bispecific antibodies, T-cell engagers and others (approvals PR(AG)252/2016, PR(AG)321/2018 and PR(AG)70/2018). For patients receiving immunotherapy, clinical outcome was available in terms of Progression Free Survival (PFS). PFS represents the lag between therapy starting date and progression or death (whichever occurs first), with progression determined via RECIST[21], or in case of established clinical worsening. Biopsies from one of the imaged liver tumours were also available in 18 patients. Of these 18 patients, 4 did not receive immunotherapy; 12 receieved immontherapy after the collection of the biopsy; and 2 received immunotherapy before the collection of the biopsy. In those two cases, additional dMRI scans were acquired immediately before biopsy collection, and hence after immunotherapy administration had started.

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. The protocol included a T2-weighted fast spin echo scan (resolution: 1.4 × 1.4 × 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. S2) EPI (dMRI scan time: 16 minutes). It featured: resolution: 1.9 × 1.9 × 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. S2) 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

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

We scanned 22 more 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 protocol 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. 2) 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. S2) 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)[15]; MRTrix3 Gibbs unringing[16]; motion correction via affine co-registration[22]; FSL distortion correction[23] (1.5T data only). An experienced radiologist (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[24] non-linear co-registration, and fitted the 5 dMRI models, fixing again f_V and $T2_T$ to previously computed values[17] (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. S2), 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[25].

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). The biological material underwent standard processing, followed by HE staining, Ki67 immunohistochemical labelling demonstrating cell replication, and digitalisation (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 the tumours. 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. Raw vCS_{histo} and CD_{histo} values were rescaled by multiplying vCS_{histo} by 1.503 (1.503 = 1.1806×1.2732, where 1.2732 accounts for biases from 2D sectioning, and 1.1806 accounts for a plausible tissue shrinkage of 15.3% following fixation, dehydration, clearing and paraffin embedding[19]). 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%[19]. Ki67 images were instead processed with custom-written macros within QuPath, estimating the fraction of histological material within the tumour labelled by Ki67 staining (metric F_{Ki67} from now on).

Statistical 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

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

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).

We also performed model selection using a *Histology Fidelity Criterion* (HFC), and popular *Bayesian Information Criterion* (BIC)[14,26]. 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.

<u>*Simulated dMRI model selection*</u>

We synthesised signals via Monte Carlo diffusion random walks for each of the three dMRI protocols of this study, using the MCDC simulator[25]. We seeded walkers in a substrate made of spherical cells of identical diameter[1,3,27,28] (Fig. S3), 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 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. We corrupted synthetic signals with Rician noise (b = 0 signal-to-noise ratio: 30), and performed model selection according to TCS, HFC and BIC.

dMRI-histology correlation analysis

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. In doing so, ADC was normalised to the ADC of the PBS solutions in mice and to the free water diffusivity in patients (3.0 $\mu m^2 ms^{-1}$ at 37 °C), given the difference in temperature.

Response assessment in immunotherapy

We studied mean values of *in vivo* baseline dMRI metrics within liver tumours (in the form of z-scores) to stratify the risk of progression under immunotherapy (N = 34). We fitted multi-variate proportional hazard Cox regression models accounting for age, sex and total baseline liver tumour volume (TV) of the form

$$\lambda(t) = \lambda_0(t)\, exp(\beta_1 m + \beta_2 Age + \beta_3 Sex + \beta_4 TV). \quad (16)$$

In Eq. 16, $\lambda(t)$ is the hazard function, $\lambda_0(t)$ is the baseline hazard (modelled through the Breslow's estimator[29]), and m the generic dMRI metric z-score. We also fitted two additional multi-variate models similar to the above, but including i) both DKI metrics ADC and K, i.e.,

$$\lambda(t) = \lambda_0(t)\, exp(\beta_1 ADC + \beta_2 K + \beta_3 Age + \beta_4 Sex + \beta_5 TV), \quad (17)$$

or ii) including the two metrics from the selected dMRI model with the highest MRI-histology correlation (i.e., vCS_{MRI} and CD_{MRI}), i.e.,

$$\lambda(t) = \lambda_0(t)\, exp(\beta_1 vCS_{MRI} + \beta_2 CD_{MRI} + \beta_3 Age + \beta_4 Sex + \beta_5 TV). \quad (18)$$

Finally, we also studied the relationship between histological properties and PFS, thus corroborating dMRI results. We calculated the Spearman's correlation coefficient between PFS and baseline vCS_{histo}, CD_{histo}, F_{histo} and F_{Ki67}, and fitted univariate Cox regression models (F_{histo} and F_{Ki67} expressed in percetange terms). Spearman's correlation coefficients between F_{Ki67} and each of vCS_{histo}, CD_{histo} and F_{histo} were also evaluated.

We performed all statistical analyses in Python, using the *SciPy* and *lifelines* packages.

Supplementary references

1. Panagiotaki, E. *et al.* Noninvasive quantification of solid tumor microstructure using VERDICT MRI. *Cancer Res.* **74**, 1902–1912 (2014).
2. 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).
3. Jiang, X., Xu, J. & Gore, J. C. Mapping hepatocyte size in vivo using temporal diffusion spectroscopy MRI. *Magn. Reson. Med.* **84**, 2671–2683 (2020).
4. Xu, J. *et al.* Magnetic resonance imaging of mean cell size in human breast tumors. *Magn. Reson. Med.* **83**, 2002–2014 (2020).
5. 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).
6. Reynaud, O. Time-Dependent Diffusion MRI in Cancer: Tissue Modeling and Applications. *Frontiers in Physics* **5**, (2017).
7. Le Bihan, D. *et al.* MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* **161**, 401–407 (1986).
8. 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).
9. 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).

10. Veraart, J. *et al.* Noninvasive quantification of axon radii using diffusion MRI. *Elife* **9**, e49855 (2020).

11. 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.* (2022) doi:10.1002/mrm.29174.

12. Xu, J. *et al.* Diffusion time dependency of extracellular diffusion. *Magn. Reson. Med.* **89**, 2432–2440 (2023).

13. Palombo, M. *et al.* Joint estimation of relaxation and diffusion tissue parameters for prostate cancer with relaxation-VERDICT MRI. *Sci. Rep.* **13**, 2999 (2023).

14. Panagiotaki, E. *et al.* Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *Neuroimage* **59**, 2241–2254 (2012).

15. Veraart, J. *et al.* Denoising of diffusion MRI using random matrix theory. *Neuroimage* **142**, 394–406 (2016).

16. 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).

17. 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).

18. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).

19. 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).

20. Garyfallidis, E. *et al.* Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.* **8**, 8 (2014).

21. 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).

22. 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).

23. 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).

24. 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).

25. 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).

26. Schwarz, G. Estimating the Dimension of a Model. *aos* **6**, 461–464 (1978).

27. Jiang, X. *et al.* MRI of tumor T cell infiltration in response to checkpoint inhibitor therapy. *J Immunother Cancer* **8**, (2020).

28. 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).

29. Lin, D. Y. On the Breslow estimator. *Lifetime Data Anal.* **13**, 471–480 (2007).