Histology-informed liver diffusion MRI: biophysical model design and demonstration in cancer immunotherapy

Francesco Grussu[1]*, Kinga Bernatowicz[1], Marco Palombo[2,3], Irene Casanova-Salas[1], Ignasi Barba[1,4], Sara Simonetti[1], Garazi Serna[1], Athanasios Grigoriou[1,5], 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], 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) models enable *in vivo* mapping of biologically meaningful properties such as cell size, potential biomarkers in cancer. However, while cancers frequently spread to the liver, models tailored for liver applications and easy to deploy in the clinic are still sought. We tackle this unmet need by delivering a practical and clinically viable liver dMRI modelling framework. Through direct comparison of candidate dMRI approaches in mouse and cancer patients' data, we select a model of intra-cellular diffusion fitted to highly diffusion-weighted images, as it provides the strongest radiological-histological correlates. We demonstrate the potential application of the proposed model in cancer immunotherapy, stratifying the risk of progression based on baseline cell size and density measurements from dMRI. This result, heretofore unreported and not achievable with standard dMRI indices (e.g., apparent diffusion coefficient), suggests that our approach may become a useful tool for precision imaging in 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[1,3,4]. 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[5], such as the size of cells restricting water[6,7]. dMRI bridges the gap between macroscopic and microscopic imaging, and has found applications in brain[5], spinal cord[8], prostate[6], breast[9] imaging and beyond.

Innovative dMRI techniques are also urgently needed in abdominal imaging, as in liver MRI[10,11]. The liver is a frequent site for cancer metastasisation [12], and liver tumours are common targets for treatment response assessment in oncology. However, current response criteria such as RECIST[13] 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[14,15], and could better stratify patients eligible for treatments such as immunotherapy[16], given the challenge of predicting which patients can benefit from these innovative drugs[17]. This would be a major advancement in oncology, as it may allow for personalised treatment planning, reductions in sample sizes in clinical trials, and ultimately improve patient outcomes[18].

The most recent biophysical dMRI techniques describe the non-vascular liver tissue signal as the sum of contributions from intra-cellular and extra-cellular water[19–21]. While these models provide promising readouts[22], 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[23], beyond default examinations available in the scanner console. In this study we aim to tackle this unmet need 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 sequences. 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 approaches with unprecedented fidelity to cytoarchitecture[24,25], maximising biological specificity[26].

In this article, we aimed to identify a practical mathematicl model that maximises the agreement of dMRI estimates of metrics such as cell size, 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, corroborating results with computer simulations. Afterwards, we demonstrated the clinical utility of the designed approach in one exemplificatory task, assessing response to immunotherapy in patients suffering from advanced solid tumours – an urgent, unmet need, given the lack of robust predictors of treatment response for this class of drugs[17]. In summary, our study delivers 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, identified which patients progress faster from baseline dMRI scans. These results, while exploratory and requiring further confirmation, suggest that our dMRI framework could provide complementary information to standard-of-care imaging, and thus play a key role in oncology research and practice.

Results

Overview: data set

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*.

Preclinical data consists of pulsed gradient spin echo (PGSE) DW MRI scans of seven fixed mouse livers, performed *ex vivo* on a 9.4T Bruker system. It also includes whole-organ HE-stained sections, obtained at known radiographic position. We studied the livers of mice sacrificed as part of xenograft model development in prostate cancer. Six had been implanted with biopsies of prostate cancer patients, while one had not had any implantation. While the livers from the implanted mice did not grow any tumours, they feature a variety of pathologies, with three unique histopathological phenotypes (Fig. S1). The liver from the mouse with no implantation features normal liver structures, and we will refer to it as *Control*. Of the six implanted cases, two also show normal liver tissue, with normal representation of all hepatic structures. We will refer to these cases as Pat_{NA1} and Pat_{NA2} (patient biopsy implantation, but normal appearing). Another case exhibits generalised necrosis and diffuse acute and chronic inflammation surrounding necrotic areas, with presence of occluded thrombotic vessels. This specimen will be identified as Pat_{nec} (patient biopsy implantation, with necrosis). Finally, three specimens feature an immature, lymphoproliferative process, with various degrees of infiltration of small, lymphoid, atypical cells with abundant mitosis, which infiltrate portal vessels and sinusoidal capillaries, but without

producing tumours. These will be referred to as Pat_{inf1} to Pat_{inf3} (patient biopsy implantation, with lymphoid cell infiltration).

We obtained clinical data on cancer patients suffering from advanced solid tumours, participating in an ongoing imaging study at the Barcelona Vall d'Hebron Institute of Oncology (VHIO, Spain). The study involves the acquisition of MRI data, alongside clinical and biological information, in patients eligible for a phase I immunotherapy trial. We included data from 33 patients with liver malignancies (mean/std of age: 62.91/12.34 year; 16 male, 17 female). dMRI was based on diffusion-weighted (DW) echo planar imaging (EPI) scans performed on a 1.5T Siemens Avanto system (Twice-Refocussed Spin Echo (TRSE) DW-EPI) and on a 3T GE SIGNA Pioneer system (PGSE DW-EPI). We also obtained digitised HE-stained biopsies from one of the liver tumours. Biopsies were available for 18 patients, and were collected after dMRI, but before immunotherapy. Clinical outcome (progression-free survival (PFS)) was available for 30 patients, who effectively entered the immunotherapy trial.

We used dMRI-histology data to design the dMRI signal model (N = 25) and dMRI-PFS data for immunotherapy response assessment (N = 30).

Overview: dMRI signal models

We studied five dMRI biophysical models, accounting for restricted diffusion inside spherical cells and hindered diffusion in the extra-cellular space[6,19,22]. Models can be grouped into two families (Fig. 2.A; see Material and Methods).

The first family is more general, in that it does not make any assumption on which is higher between intra-/extra-cellular ADC (ADC_I and ADC_E from now on). It includes:

i. *Diff-in-exTD*: it accounts for restricted IC diffusion within spherical cells[6], and hindered diffusion in the extra-cellular space, with diffusion time dependence (TD)[27] in both intra-/extra-cellular spaces[28]. The diffusion time quantifies the time during which water molecules can sense cellular barriers, before the MR image is acquired.

ii. *Diff-in-ex*: as previous model *Diff-in-exTD*, but neglecting TD in the extra-cellular space. Popular techniques such as IMPULSED[19] or VERDICT[20] are essentially implementations of this model.

Conversely, the second family explicitly assumes that $ADC_E > ADC_I$, similarly to related dMRI techniques (e.g., Restriction Spectrum Imaging[29]; power-law axon radius mapping[7]). It includes:

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

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

iii. *Diff-in*: a model where it is hypothesised that due to fast, extra-cellular diffusion, the extra-cellular signal is negligible, and the measured signal is dominated by intra-cellular water.

Overview: dMRI metrics

All models enable the estimation of *volume-weighted mean cell size* (vCS_{MRI}, expressed in μm) and *intra-cellular signal fraction* (F_{MRI}, dimensionless), which can be combined into an apparent *cell density per unit volume* ($CD_{MRI} = F_{MRI}/vCS_{MRI}^3$, expressed in cell mm^{-3})[20]. For reference, we benchmarked these metrics against routine ADC (in μm^2 ms^{-1}) and apparent diffusion excess kurtosis K (dimensionless) from diffusion kurtosis imaging (DKI)[30], popular dMRI indices sensitive to cancer cellularity, easy to compute from short acquisitions[20,31].

We processed HE-stained histological data with automatic cell detection[32] to derive histological counterparts of vCS_{MRI} and F_{MRI} at known radiographic location. The histological metrics were: histological volume-weighted mean cell size (vCS_{histo}, in μm), intra-cellular area fraction (F_{histo}, dimensionless), and cell density per unit area (CD_{histo}, in cell mm^{-2}). We compared vCS_{histo} and F_{histo} to vCS_{MRI} and F_{MRI} (Fig. 2B).

A one-pool model of intra-cellular diffusion provides the most histologically meaningful metrics

Fig. 3A summarises the different dMRI models, while Fig. 3B reports values of the MRI-Histology *Total Correlation Score* (TCS) for all models. TCS measures the overall correlation between histological and radiological readouts of cell size and intra-cellular fraction, and is obtained by summing Pearson's correlation coefficients between vCS_{MRI} and vCS_{histo}, and between F_{MRI} and F_{histo} (see Materials and Methods). Higher TCS point towards stronger histological-radiological correlation. Negative correlations reduce TCS, so they are penalised.

The bar plot in Fig. 3B highlights that dMRI models where $ADC_E > ADC_I$, shown in violet shades, provide consistently higher TCS values than models that do not make such an assumption (orange shades). We observe the highest TCS for model *Diff-in*. Note that Fig. 3 refers to TCS values obtained by fitting dMRI models only to high b-value images, as this provided the highest TCS figures. Fig. S2 reports TCS for model fitting performed to the whole set of diffusion images. In this case, TCS is lower, but again, models where $ADC_E > ADC_I$ provide the highest TCS. In Fig. S2, *Diff-in-exFast* provides the highest TCS, although this is lower than *Diff-in* TCS in Fig. 3.

Fig. S3 reports rankings according to additional criteria, namely: the *Histology Fidelity Criterion* (HFC), measuring the sum of absolute errors in F and vCS estimation via dMRI, and the *Bayesian*

Information Criterion (BIC)[33]. BIC is commonly used in dMRI model development[34,35]; it quantifies the overall model fitting quality (penalizing model complexity), but without accounting for histological information. Lower HFC and lower BIC imply better model performance. Fig. S3 reports the number of times, in percentage terms, that a model provides the lowest HFC and BIC across our 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 do, according to HFC. The model *Diff-in* is the most selected model according to both BIC and HFC (fig. S3.B; fitting to high b-value images). Fig. S4 splits HFC and BIC rankings depending on the MRI scanner. In all cases, models with $ADC_E > ADC_I$ (*Diff-in*, *Diff-in-exFast*, *Diff-in-exFastTD*) are selected more frequently than models *Diff-in-ex* and *Diff-in-ex-TD*. When fitting is performed only on high b-value images, *Diff-in* is the most selected model according to both BIC and HFC.

Computer simulations confirm model selection from MRI measurements

We performed Monte Carlo computer simulations to corroborate the model selection performed on *ex vivo* and *in vivo* dMRI data. The simulations consisted in generating synthetic dMRI signals according to the three dMRI protocols used in this study. We synthesised signals for a substrate made of packed spherical cells (Fig. S5), a common body dMRI tissue model[6,19,22,23], and then performed model selection on the synthetic signals (see Materials and Methods). 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.

Our proposed approach: a one-compartment model of intra-cellular diffusion, fitted to high b-values

In view of all rankings reported above, our recommended modelling approach is the *fitting a one-compartment model of restricted, intra-cellular diffusion within spherical cells to high b-values images* ($\gtrapprox$ 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

We now report on the Pearson's correlation coefficient r between histology, *Diff-in* and DKI metrics. We consider correlations to be weak, moderate, and strong when $|r| < 0.4$, $|r| \geq 0.4$ but $|r| < 0.6$, and $|r| \geq 0.6$.

Table 1 shows that the correlation between *Diff-in* F_{MRI} (intra-cellular fraction), vCS_{MRI} (cell size index) and CD_{MRI} (cell density) with their counterparts F_{histo}, $vCS_{histo}$and CD_{histo} are respectively weak (r = 0.19 between F_{MRI} and F_{histo}), moderate (r = 0.44 between vCS_{MRI} and

vCS_{histo}) and strong (r = 0.70 between CD_{MRI} and CD_{histo}). The weak correlation between F_{MRI} and F_{histo} can be explained, at least in part, with the fact that F_{MRI} is a signal fraction, rather than an actual volume/area fraction (unlike F_{histo}), i.e., it is influenced by T2/T1 differences between intra-cellular and residual extra-cellular signals[36]. Moreover, F_{MRI} estimation can be biased by unaccounted exchange between intra-/extra-cellular water[21,37], which is not accounted for in our signal models. Conversely, the much higher correlations between vCS_{MRI} and vCS_{histo} and between CD_{MRI} and CD_{histo}, point towards the biological specificity of vCS_{MRI} and CD_{MRI}.

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) – a result entirely consistent with previous studies[38,39]. Significant correlations are also seen with F_{histo} (r = 0.40, p = 0.048 between K and F_{histo}). These findings are in line with known literature: ADC and K are sensitive to the underlying tissue microstructure, but they are also unspecific, being surrogate metrics that conflate different histopathological characteristics into a single number.

Fig. S6 and Fig. S7 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, appears biophysically plausible, being 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. 4 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_{infl}, 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}, CD_{histo}, with values in physiologically plausible ranges, as for example intra-cellular fractions around 0.75 and cell sizes

of the order of 20 μm[40,41]. Maps F_{MRI}, vCS_{MRI} and CD_{MRI} replicate the contrasts seeing in their histological counterparts F_{histo}, vCS_{histo} and CD_{histo}. Fig. S8 shows standard dMRI metrics, namely ADC and kurtosis excess K, in the same 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. S9 shows F_{MRI}, vCS_{MRI} and CD_{MRI} maps from another model (*Diff-in-exFast*). Fig. S10 instead shows maps of the other metrics provided by models *Diff-in-exFast* and *Diff-in* (intrinsic cytosol diffusivity $D_{0,I}$ and asymptotic ADC_E, $D_{E,\infty}$). 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.

Table S4 reports qualitative per-sample mean and standard deviation of all MRI and histology metrics in mice. F_{MRI} slightly underestimates F_{histo}, while vCS_{MRI} slightly overestimates vCS_{histo}. We speculate that the discrepancies may be due, at least in part, to unaccounted factors such as variability in intrinsic cell shape/cytosol diffusivity[42] or water exchange[21], and by the difficulty of relating accurately 2D histology to 3D MRI[43].

Fig. 5 shows F_{MRI}, vCS_{MRI} and CD_{MRI} maps in patients, alongside 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. 5, we observe a core of low intra-cellular fraction F_{MRI} and low cell density CD_{MRI}, compatible with necrosis. In a HCC case instead, we see areas of high F_{MRI} and high CD_{MRI}, sourrounded by lower F_{MRI} and lower CD_{MRI}, potentially indicating the alternation of high cell densities with fibrotic tissue. Fig. S11 shows routine dMRI ADC and K in the same tumours. 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. S12 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), giving confidence of the overall robustness of the biophysical modelling framework. Nevertheless, maps appear noisier. Fig. S13 shows intra-cellular cytosol diffusivity $D_{0,I}$ asymptotic ADC_E ($D_{E,\infty}$) in the same tumours. Their speckled appearance suggests that these metrics are difficult to measure accurately *in vivo*[19,44].

Metrics from the proposed dMRI model stratify the risk of cancer progression in immunotherapy

Finally, we demonstrate the potential utility of the proposed liver dMRI model in an exemplificatory response assessment task. Fig. 6 reports on the PFS stratification based on *Diff-in* metrics. Panels on the left report results from Kaplan-Meier analyses, log-rank tests and Cox regressions performed after binarising dMRI metrics (higher/lower than the median of the cohort). Panels on the right report results from Cox regressions assessing the continuous dependence of PFS on *Diff-in* metrics. We detect a statistically significant dependence of the risk of progression on baseline vCS_{MRI} (log-rank test: $p = 0.047$, Fig. 6C) and CD_{MRI} (log-rank test: $p = 0.035$, Fig. 6E). These differences correspond to statistically significant Hazard Ratios (HRs) from Cox regression (HR = 0.47, $p = 0.050$ for binarised vCS_{MRI}; HR = 2.36, $p = 0.043$ for binarised CD_{MRI}). The risk of progression is about twice as high in patients whose baseline cell size vCS_{MRI} is smaller than the median vCS_{MRI}, or whose cell density CD_{MRI} is higher than the median CD_{MRI}. Importantly, we obtained similar results when vCS_{MRI} and CD_{MRI} were not binarised, but rather used as continuous predictors in Cox regressions (right panels in Fig. 6). In this latter case, the HR was statistically significant for vCS_{MRI} (HR = 0.65, $p = 0.034$, Fig. 6D), and it approached statistical significance for CD_{MRI} (HR = 1.40, $p = 0.055$, Fig. 6F). These association are not confounded by age, sex or baseline tumour volume (Supplementary Table S5; HR = 0.59, $p = 0.02$ for vCS_{MRI}; HR = 1.65, $p = 0.01$ for CD_{MRI}).

Fig. 7 and supplementary Fig. S14, Fig. S15 and Fig. S16 report on response assessment based on on all other dMRI metrics (routine ADC and K in Fig. 7; cytosol diffusivity $D_{0,I}$ for model *Diff-in* in Fig. S14; vascular fraction f_V in Fig. S15; all metrics from model *Diff-in-exFast* in Fig. S16). While the estimated HRs for metrics vCS_{MRI}, F_{MRI} and CD_{MRI} from dMRI model *Diff-in-exFast* are in the same direction as those from model *Diff-in*, their association with the probability of progression is weaker (HR = 1.53, $p = 0.05$ for CD_{MRI}; Table S5). We do not detect any association between baseline ADC and K and the probability of progression (Fig. 7 and Table S5).

Discussion

The latest liver dMRI models aim to disentangle intra-cellular and extra-cellular water contributions to the total signal[6,19,22,23]. 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 dMRI acquisitions[19,45,46]. 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 dMRI framework consisting of fitting a one-pool model of restricted intra-cellular diffusion to highly DW images. The framework provides cell size and density estimates that correlate with histology, and which enable the stratification of the risk of cancer progression under 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 intra-cellular fraction and cell size, as seen on histology. 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 fact that simpler models of diffusion can provide the highest fidelity to the underlying histology, if deployed in appropriate measurement regimes. In practice, our approach suggests that focussing on measurement regimes where the signal is dominated by intra-cellular diffusion, may enable the deployment of simpler models, which still suffice to capture salient microstructural features[7,47–49]. Notably, our model selection results are consistent with the fact recent estimates of the extra-cellular liver ADC, as high as 2.5 μm^2/ms[19]. 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[7,44,50].

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[35] (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 (i.e., fitting model *Diff-in* on high b-value images) would have been selected had we only looked at the *ex vivo* mouse data or at *in silico* signals (Fig. S4, Tables S1 to S3). This fact gives confidence on the robustness and generalisability of our model design. Another important observation is that both TCS and HFC, which are histology-informed, suggest that models constraining the extra-cellular ADC to be higher than the intra-cellular ADC, outperform models without such a constraint. Such a better performance does not stand out as clearly when looking at BIC rankings instead (e.g., Fig. S3.A, Fig. S4.A, Fig. S4.B, Fig. S4.C). This minor discrepancy can be understood considering that a good fitting quality may not necessary imply accurate parameter estimation in presence of noise[26,44]. It also stresses the importance of informing dMRI modelling with histology, for improved biological specificity.

After selecting a practical dMRI model implementation, we investigated its utility in an exemplificatory response assessment task. Immunotherapy has shown promise in several cancers[16]. However, only a small fraction of patients truly benefits from this class of drugs, and their identification prior to treatment is extremely challenging[17]. With this in mind, we investigated

whether it is possible to stratify the probability of progression under immunotherapy given baseline dMRI metrics within liver tumours (N = 30). Several, independent statistical tests consistently point towards the fact that smaller baseline dMRI cell sizes and higher dMRI cell densities are associated to faster cancer progression (shorter PFS). To our knowledge, this is the first time that such advanced dMRI markers have been tested for patient stratification in immunotherapy *in vivo*. The deleterious impact of higher cell density on PFS is in line with other studies focussing on different treatments, where higher tumour cellularity has been associated with higher cancer aggressiveness or worse prognosis. In CRC liver metastasis resection, for example, high cell density in resected metastases has been associated to shorter disease-free survival[51]. 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)[52], as well as longer survivals in neoadjuvant chemotherapy[53]. The significant association between PFS and dMRI cell size/density reported here is promising, and motivates future studies to confirm it.

We benchmarked the proposed dMRI approach against well-established DKI apparent diffusion and kurtosis coefficients (ADC and K)[30]. ADC and K are easy to compute with compact dMRI acquisitions, and are sensitive to cancer cellularity[46], a fact confirmed by our data. However, in our cohort, neither of ADC and K showed statistically significant associations with clinical PFS. This may due, at least partly, to the fact that ADC and K are semi-quantitative, protocol-dependent metrics, whose value can change as function of factors as the diffusion time[46,54]. Here, we did not perform inter-scanner harmonisation deliberately, to stress the quantitative nature of our cell size/density mapping approach, which inherently accounts for inter-scanner protocol differences. However, it is possible that better performances for semi-quantitative ADC and K may be obtained by adopting inter-scanner harmonisation[55].

We would like to acknowledge the following potential limitations. Firstly, our sample size (N = 25 for model development; N = 30 for the clinical demonstration) is relatively small. This paper provides a first demonstration of the potential utility of the proposed *Diff-in* approach. The demonstration is unique of its kind, since it reports heretofore undescribed dMRI-based stratification in immunotherapy. Nonetheless, while works proposing related dMRI techniques relied on similar[23,36], if not even smaller[19,20], sample sizes, we acknowledge that our exploratory findings require further confirmation in larger cohorts.

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[56]), and the automatic processing of large fields-of-view requires trading off between sensitivity and specificity. We took steps to mitigate these issues, e.g., by accounting for biases due to tissue shrinkage. Nonetheless, our histology-derived estimates of cell properties are likely biased versions of the true figures.

We would also like to acknowledge that the proposed dMRI approach neglects other potentially relevant microstructural properties, such as water exchange between intra-/extra-cellular spaces[21,37], presence of cell size/cytosolic diffusivity distributions[42,57], or intra-compartmental T2 or T1[36]. On the one hand, ignoring these properties may have biased the estimation of F_{MRI} and vCS_{MRI}[21,37]. On the other hand, properties such as exchange rates, overlooked in our model, may be relevant markers of cellular stress *per se*. In future, we plan to incorporate these properties in our models, while ensuring the clinical feasibility of the dMRI protocols required to fit them.

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/mm²* 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 by stratifying the risk of cancer progression 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, which will be freely accessible online.

Methods

dMRI models

Common biophysical body dMRI signal models[6,19,20,50,58,59] 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[60]. *In vivo*, the IVIM vascular ADC ranges approximately[61] [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[19]

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

A common model for a_I in Eq. 2 is that of restricted diffusion within spheres of diameter L[19,20]:

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

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[62]. 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[7,42], 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[19,27,59]:

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

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[59] 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[20,36].

Fitting

We fitted the 5 models using custom-written Python routines, based on objective function minimisation initialised by a grid search. The objective function was $f_{obj} = -ln(\lambda)$, where λ is the offset-Gaussian likelihood[34]. 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[20]

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

Preclinical data

Animals

We obtained data from 7 fixed livers of NOD.Cg-Prkdc^{scid} IL2rg^{tm1WjI}/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. S17A; 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[63] (kernel: 7×7×3), ii) mitigated Gibbs ringing (MrTrix3 local sub-voxel shift method[64]),

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[65] (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.

Histology

After MRI, samples underwent histology. We cut two 4 μm-thick histological sections at known position, stained them with HE, and digitised them (Hamamatsu C9600-12 slide scanner; 0.227 μm resolution). An experienced pathologist (S.S.) inspected images qualitatively. We then processed them with the automatic cell detection tool of QuPath[32], 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}[42]. vCS_{histo}, defined as

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

is a more accurate counterpart of dMRI cell size than the arithmetic mean[7,42] $aCS_{histo} = < l >$. 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[56]). 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%[56]). Lastly, we co-registered histological maps to MRI[42] using DiPy[66].

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, eligible for a phase I immunotherapy trial at VHIO (Barcelona, Spain), provided informed written consent to participate in the imaging study. We included 33 patients with liver malignancies (mean/std of age: 62.91/12.34 year; 16 male, 17 female), of which 3 suffered from primary HCC, while 30 had liver metastases from different primary cancers (10 colon, 8 melanoma, 3 rectal, 2 ovarian, 2 gastric, 2 breast, 1 renal, 1 endometrial, 1 ureteral). We obtained baseline dMRI scans (i.e., acquired immediately before starting immunotherapy), and digitised HE-stained biopsies from one of the imaged liver tumours. We obtained biopsies from 18 patients (6 scanned at 1.5T, 12 at 3T), collected after baseline dMRI. 30 out of 33 patients finally entered the immunotherapy trial after screening; for them, a clinical outcome in the form of PFS was available. PFS represents the lag between therapy starting date and progression or death (whichever occurs first), with progression determined via RECIST[13], or in case of established clinical worsening.

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. S17B) 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. S17B) 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. S10A) EPI (dMRI scan time: 16 minutes). It featured: resolution: 2.4 × 2.4 × 6 mm^3; 32 slices; TR = 6000 ms; bandwidth 1953 Hz/pixel; averaging of 3 orthogonal diffusion directions × 2 signal averages (effective NEX = 6); ASSET factor of 2. The dMRI protocol consisted of b = {0, 50, 100, 400, 900, 1200, 1500} s/mm^2, each for TE = {75, 90, 105} ms. The gradient timings (Fig. S17A) were: gradient duration δ = {0.0, 3.9, 5.2, 9.2, 15.0, 18.2, 21.0} ms for TE = 75 ms, δ = {0.0, 3.9, 5.2, 9.2, 13.0, 15.8, 18.5} ms for TE = 90 ms and 105 ms; gradient separation Δ = {0.0, 27.8, 29.0, 33.0, 28.7, 31.8, 34.7} ms for TE = 75 ms and Δ = {0.0, 27.8, 29.0, 33.0, 37.0, 39.6, 42.3} ms for TE = 90 ms and TE = 105 ms.

dMRI post-processing consisted of slice-wise Python MP-PCA denoising (kernel: 5 × 5)[63]; MRTrix3 Gibbs unringing[64]; motion correction via affine co-registration[67]; FSL distortion correction[68] (1.5T data only). An experienced radiologists (R.P.L.) segmented tumours on the T2-w scan, enabling per-patient tumour volume computation. Afterwards, we warped the tumour mask to dMRI using ANTs[69] non-linear co-registration, and fitted the 5 dMRI models, fixing again f_V and $T2_T$ to previously computed values[65] (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 $\Delta_{1,2}$ + δ_2 in place of Δ in Eq. 6 (Fig. S10B), 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[70].

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, HE staining 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. Finally, we rescaled vCS_{histo} and CD_{histo}, by multiplying vCS_{histo} from image processing 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[56]). 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%[56].

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

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

where $r(vCS_{MRI}, vCS_{histo})$ and $r(F_{MRI}, F_{histo})$ are the correlation coefficients of vCS_{MRI} and F_{MRI} with histological vCS_{histo} and F_{histo}. The correlation between CD_{MRI} and CD_{histo} was not included in Eq. 13 since CD_{MRI} is determined analytically from vCS_{MRI} and F_{MRI}. For TCS computation, we pooled together mouse and human data (N = 25).

We also performed model selection using a *Histology Fidelity Criterion* (HFC), and popular *Bayesian Information Criterion* (BIC)[33,34]. 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.

Simulated dMRI model selection

We synthesised signals via Monte Carlo diffusion random walks for each of the three dMRI protocols of this study, using the MCDC simulator[70]. We seeded walkers in a substrate made of spherical cells of identical diameter[6,19,22,23] (Fig. S5), 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 to assess whether these could stratify the probability of progression under immunotherapy, measured by PFS (N = 30). Firstly, we performed the stratification after binarising all MRI metrics as lower/higher than the median of the cohort, creating two groups. We evaluated group-wise survival curves with the Kaplan-Meier estimator, and compared them with a log-rank test. We also fitted a proportional hazard Cox regression, where the binarised MRI metric was the only regressor. Secondly, we stratified PFS without binarising MRI metrics. We fitted a proportional hazard Cox model using each continous dMRI metric as the only regressor (in the form of a z-score), and then fitted the statistical model again, including age, sex and total baseline tumour volume as confounding factors. We performed all analyses in Python, using *lifelines*.

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).

Author information

Francesco Grussu and Raquel Perez-Lopez are joint corresponding authors.

Authors and affiliations

Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain. Francesco Grussu, Kinga Bernatowicz, Irene Casanova-Salas, Ignasi Barba, Sara Simonetti, Garazi Serna, Athanasios Grigoriou, Anna Voronova, Rodrigo Toledo, Paolo Nuciforo & Raquel Perez-Lopez.

Cardiff University, Cardiff, United Kingdom. Marco Palombo.

University of Vic - Central University of Catalonia (UVic-UCC), Vic, Spain. Ignasi Barba.

Department of Biomedicine, Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain. Athanasios Grigoriou & Anna Voronova.

PET/MR Unit, CETIR-ASCIRES, Barcelona, Spain. Valezka Garay.

Department of Radiology, Hospital Universitari Vall d'Hebron, Barcelona, Spain. Juan Francisco Corral, Xavier Merino, Richard Mast, Núria Roson & Manuel Escobar.

Institut de Diagnòstic per la Imatge (IDI), Barcelona, Spain. Juan Francisco Corral, Xavier Merino, Richard Mast, Núria Roson & Manuel Escobar.

Siemens Healthineers, Madrid, Spain. Marta Vidorreta.

GE HealthCare, Madrid, Spain. Pablo García-Polo García.

Medical Oncology Service, Vall d'Hebron Barcelona Hospital Campus, Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain. Maria Vieito, Joaquin Mateo & Elena Garralda

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. Investigation: F.G., R.P.L., K.B., I.C.S., I.B., S.S., G.S., A.G., 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. Data curation: F.G., R.P.L., K.B., A.V., G.S., I.C.S., A.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

Corresponding authors

Correspondence to Francesco Grussu or Raquel Perez-Lopez.

Ethics declarations

Competing interests

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

Ethics

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.

Supplementary Information

This article includes Supplementary Figure S1 to S17 and Supplementary Tables S1 to S5.

References

1. Weiskopf, N., Edwards, L. J., Helms, G., Mohammadi, S. & Kirilina, E. Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. *Nature Reviews Physics* **3**, 570–588 (2021).
2. Jarrett, A. M. *et al.* Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting. *Nat. Protoc.* **16**, 5309–5338 (2021).
3. Kiselev, V. G. Fundamentals of diffusion MRI physics. *NMR Biomed.* **30**, (2017).
4. Novikov, D. S., Fieremans, E., Jespersen, S. N. & Kiselev, V. G. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR Biomed.* **32**, e3998 (2019).
5. Le Bihan, D. Looking into the functional architecture of the brain with diffusion MRI. *Nat. Rev. Neurosci.* **4**, 469–480 (2003).
6. Panagiotaki, E. *et al.* Noninvasive quantification of solid tumor microstructure using VERDICT MRI. *Cancer Res.* **74**, 1902–1912 (2014).

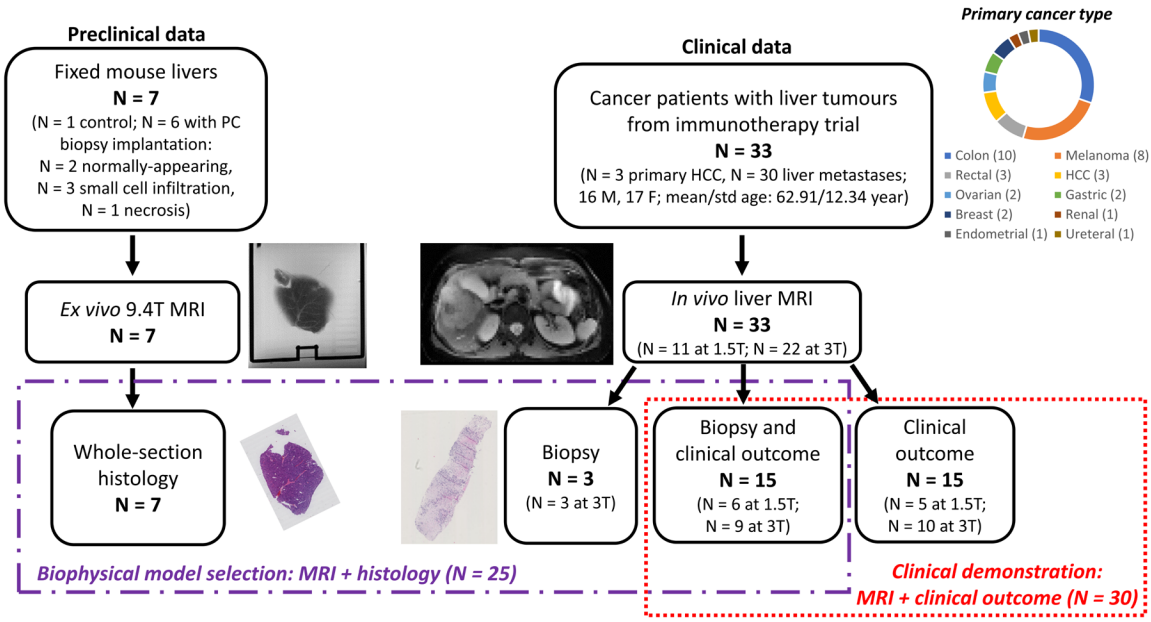
7. Veraart, J. *et al.* Noninvasive quantification of axon radii using diffusion MRI. *Elife* **9**, e49855 (2020).
8. Cohen-Adad, J. *et al.* Generic acquisition protocol for quantitative MRI of the spinal cord. *Nat. Protoc.* **16**, 4611–4632 (2021).
9. Baxter, G. C., Graves, M. J., Gilbert, F. J. & Patterson, A. J. A Meta-analysis of the Diagnostic Performance of Diffusion MRI for Breast Lesion Characterization. *Radiology* **291**, 632–641 (2019).
10. Donato, H., França, M., Candelária, I. & Caseiro-Alves, F. Liver MRI: From basic protocol to advanced techniques. *Eur. J. Radiol.* **93**, 30–39 (2017).
11. Tapper, E. B. & Lok, A. S.-F. Use of Liver Imaging and Biopsy in Clinical Practice. *N. Engl. J. Med.* **377**, 756–768 (2017).
12. Tsilimigras, D. I. *et al.* Liver metastases. *Nat Rev Dis Primers* **7**, 27 (2021).
13. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
14. Buikhuisen, J. Y., Torang, A. & Medema, J. P. Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges. *Oncogenesis* **9**, 66 (2020).
15. Vitale, I., Shema, E., Loi, S. & Galluzzi, L. Intratumoral heterogeneity in cancer progression and response to immunotherapy. *Nat. Med.* **27**, 212–224 (2021).
16. Waldman, A. D., Fritz, J. M. & Lenardo, M. J. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat. Rev. Immunol.* **20**, 651–668 (2020).
17. Pilard, C. *et al.* Cancer immunotherapy: it's time to better predict patients' response. *Br. J. Cancer* **125**, 927–938 (2021).
18. O'Connor, J. P. B. *et al.* Imaging biomarker roadmap for cancer studies. *Nat. Rev. Clin. Oncol.* **14**, 169–186 (2017).
19. Jiang, X., Xu, J. & Gore, J. C. Mapping hepatocyte size in vivo using temporal diffusion spectroscopy MRI. *Magn. Reson. Med.* **84**, 2671–2683 (2020).
20. 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).
21. Gardier, R. *et al.* Cellular Exchange Imaging (CEXI): Evaluation of a diffusion model including water exchange in cells using numerical phantoms of permeable spheres. *Magn. Reson. Med.* (2023) doi:10.1002/mrm.29720.
22. 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).
23. Jiang, X. *et al.* MRI of tumor T cell infiltration in response to checkpoint inhibitor therapy. *J Immunother Cancer* **8**, (2020).
24. Ye, Z. *et al.* Diffusion Histology Imaging Combining Diffusion Basis Spectrum Imaging (DBSI) and Machine Learning Improves Detection and Classification of Glioblastoma Pathology. *Clin. Cancer Res.* **26**, 5388–5399 (2020).

25. Howard, A. F. *et al.* Joint modelling of diffusion MRI and microscopy. *Neuroimage* **201**, 116014 (2019).
26. Novikov, D. S., Kiselev, V. G. & Jespersen, S. N. On modeling. *Magn. Reson. Med.* **79**, 3172–3193 (2018).
27. Xu, J. *et al.* Diffusion time dependency of extracellular diffusion. *Magn. Reson. Med.* **89**, 2432–2440 (2023).
28. Neuman, C. H. Spin echo of spins diffusing in a bounded medium. *J. Chem. Phys.* **60**, 4508–4511 (1974).
29. Conlin, C. C. *et al.* Improved Characterization of Diffusion in Normal and Cancerous Prostate Tissue Through Optimization of Multicompartmental Signal Models. *J. Magn. Reson. Imaging* **53**, 628–639 (2021).
30. Jensen, J. H., Helpern, J. A., Ramani, A. & Lu, H. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *in Medicine: An …* (2005).
31. Hectors, S. J. *et al.* Advanced Diffusion-weighted Imaging Modeling for Prostate Cancer Characterization: Correlation with Quantitative Histopathologic Tumor Tissue Composition-A Hypothesis-generating Study. *Radiology* **286**, 918–928 (2018).
32. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
33. Schwarz, G. Estimating the Dimension of a Model. *aos* **6**, 461–464 (1978).
34. Panagiotaki, E. *et al.* Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *Neuroimage* **59**, 2241–2254 (2012).
35. Ferizi, U. *et al.* A ranking of diffusion MRI compartment models with in vivo human brain data. *Magn. Reson. Med.* **72**, 1785–1792 (2014).
36. Palombo, M. *et al.* Joint estimation of relaxation and diffusion tissue parameters for prostate cancer with relaxation-VERDICT MRI. *Sci. Rep.* **13**, 2999 (2023).
37. Jiang, X., Devan, S. P., Xie, J., Gore, J. C. & Xu, J. Improving MR cell size imaging by inclusion of transcytolemmal water exchange. *NMR Biomed.* **35**, e4799 (2022).
38. Rosenkrantz, A. B. *et al.* Assessment of hepatocellular carcinoma using apparent diffusion coefficient and diffusion kurtosis indices: preliminary experience in fresh liver explants. *Magn. Reson. Imaging* **30**, 1534–1540 (2012).
39. Tang, W.-J. *et al.* Evaluation of the Effects of Anti-PD-1 Therapy on Triple-Negative Breast Cancer in Mice by Diffusion Kurtosis Imaging and Dynamic Contrast-Enhanced Imaging. *J. Magn. Reson. Imaging* **56**, 1912–1923 (2022).
40. Martin, N. C. *et al.* Functional analysis of mouse hepatocytes differing in DNA content: volume, receptor expression, and effect of IFNgamma. *J. Cell. Physiol.* **191**, 138–144 (2002).
41. Zhou, Z., Xu, M.-J. & Gao, B. Hepatocytes: a key cell type for innate immunity. *Cell. Mol. Immunol.* **13**, 301–315 (2016).

42. 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.
43. Morawski, M. *et al.* Developing 3D microscopy with CLARITY on human brain tissue: Towards a tool for informing and validating MRI-based histology. *Neuroimage* **182**, 417–428 (2018).
44. Jelescu, I. O., Veraart, J., Fieremans, E. & Novikov, D. S. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR Biomed.* **29**, 33–47 (2016).
45. Nilsson, M. *et al.* Mapping prostatic microscopic anisotropy using linear and spherical b-tensor encoding: A preliminary study. *Magn. Reson. Med.* **86**, 2025–2033 (2021).
46. Fokkinga, E. *et al.* Advanced Diffusion-Weighted MRI for Cancer Microstructure Assessment in Body Imaging, and Its Relationship With Histology. *J. Magn. Reson. Imaging* (2023) doi:10.1002/jmri.29144.
47. Jensen, J. H., Russell Glenn, G. & Helpern, J. A. Fiber ball imaging. *Neuroimage* **124**, 824–833 (2016).
48. McKinnon, E. T., Helpern, J. A. & Jensen, J. H. Modeling white matter microstructure with fiber ball imaging. *Neuroimage* **176**, 11–21 (2018).
49. Warner, W. *et al.* Temporal Diffusion Ratio (TDR) for imaging restricted diffusion: Optimisation and pre-clinical demonstration. *Neuroimage* **269**, 119930 (2023).
50. Reynaud, O. Time-Dependent Diffusion MRI in Cancer: Tissue Modeling and Applications. *Frontiers in Physics* **5**, (2017).
51. Serrablo, A. *et al.* Impact of novel histopathological factors on the outcomes of liver surgery for colorectal cancer metastases. *Eur. J. Surg. Oncol.* **42**, 1268–1277 (2016).
52. Nuciforo, P. *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.* **29**, 170–177 (2018).
53. Gentile, D. *et al.* Pathologic response and residual tumor cellularity after neo-adjuvant chemotherapy predict prognosis in breast cancer patients. *Breast* **69**, 323–329 (2023).
54. Lee, H.-H., Papaioannou, A., Novikov, D. S. & Fieremans, E. In vivo observation and biophysical interpretation of time-dependent diffusion in human cortical gray matter. *Neuroimage* **222**, 117054 (2020).
55. Fortin, J.-P. *et al.* Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* **161**, 149–170 (2017).
56. 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).
57. Xu, J. *et al.* MRI-cytometry: Mapping nonparametric cell size distributions using diffusion MRI. *Magn. Reson. Med.* **85**, 748–761 (2021).
58. Xu, J. *et al.* Magnetic resonance imaging of mean cell size in human breast tumors. *Magn. Reson. Med.* **83**, 2002–2014 (2020).

59. 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).
60. Le Bihan, D. *et al.* MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* **161**, 401–407 (1986).
61. 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).
62. 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).
63. Veraart, J. *et al.* Denoising of diffusion MRI using random matrix theory. *Neuroimage* **142**, 394–406 (2016).
64. 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).
65. 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).
66. Garyfallidis, E. *et al.* Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.* **8**, 8 (2014).
67. 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).
68. 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).
69. 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).
70. 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).

970 # Figures



971

972 **Fig. 1. Illustration of the liver MRI and histology data used in the study.** Our data set consisted
973 of preclinical and clinical data. The preclinical data encompasses dMRI scans of seven fixed livers
974 from mice (six implanted with tissue from biopsies of patients suffering from prostate cancer; one
975 without any implantation). We scanned the livers *ex vivo* on a 9.4T system, and obtained HE
976 histological sections at known position. The clinical data includes *in vivo* liver dMRI scans
977 performed on 33 patients suffering from advanced liver tumours. Scans were performed on clinical
978 1.5T and 3T MRI systems. For 18 patients, HE-stained material from a biopsy taken from one of
979 the imaged liver tumours was also available. 30 out of 33 patients effectively participated in a
980 phase I immunotherapy trial, and clinical outcome was available as PFS. We used a total sample
981 size of N = 25 for dMRI model design and of N = 30 for response assessment. In the figure, PC
982 standard fort prostate cancer and HCC for hepatocellular carcinoma.

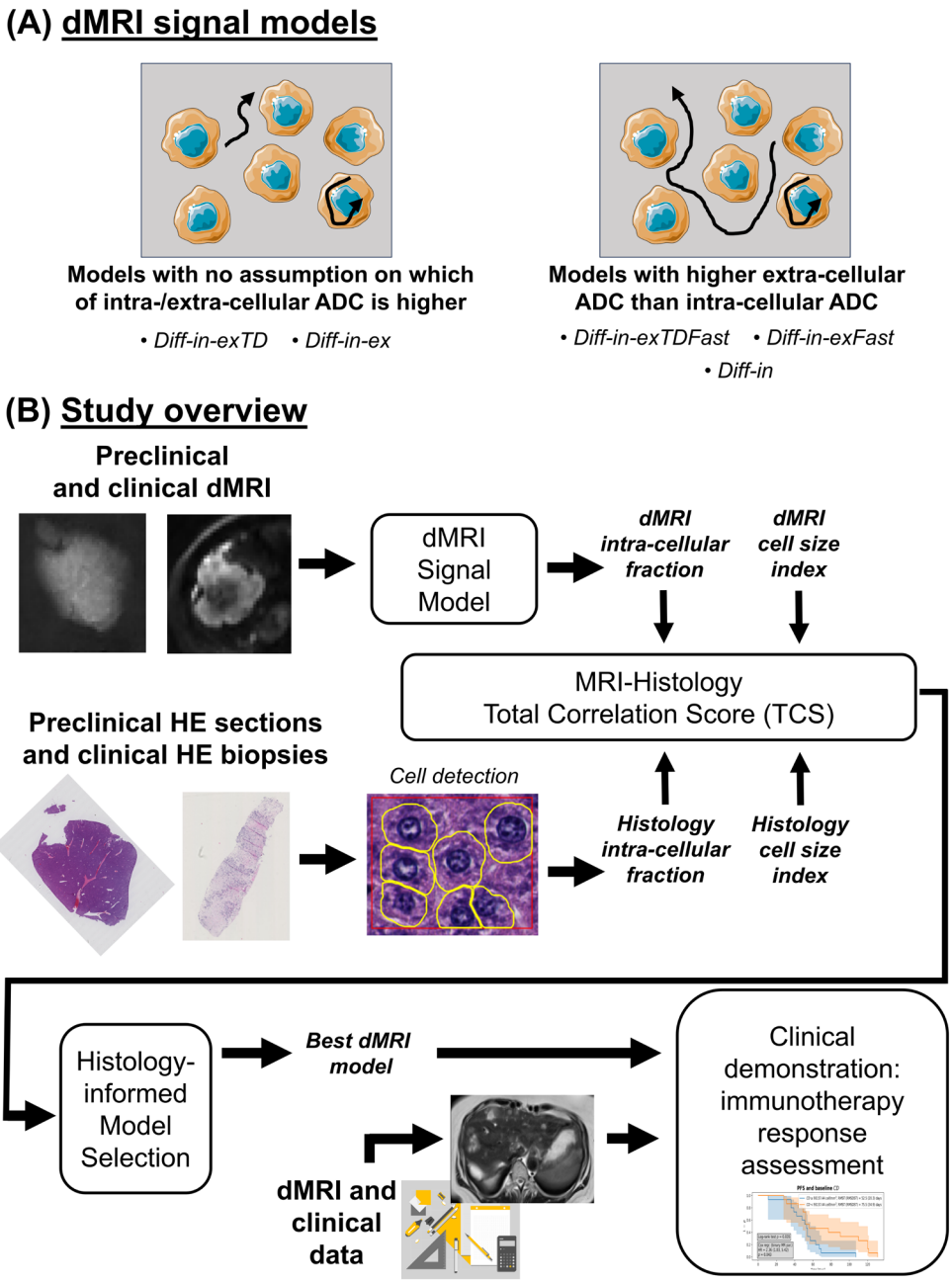


Fig. 2. Description of the dMRI signal models and study overview. (**A**), top: cartoon illustrating the two families of dMRI models considered in this study, consisting of 1) models with no assumption of which of intra/extra-cellular ADC is higher, and 2) models where the extra-cellular ADC is hypothesised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

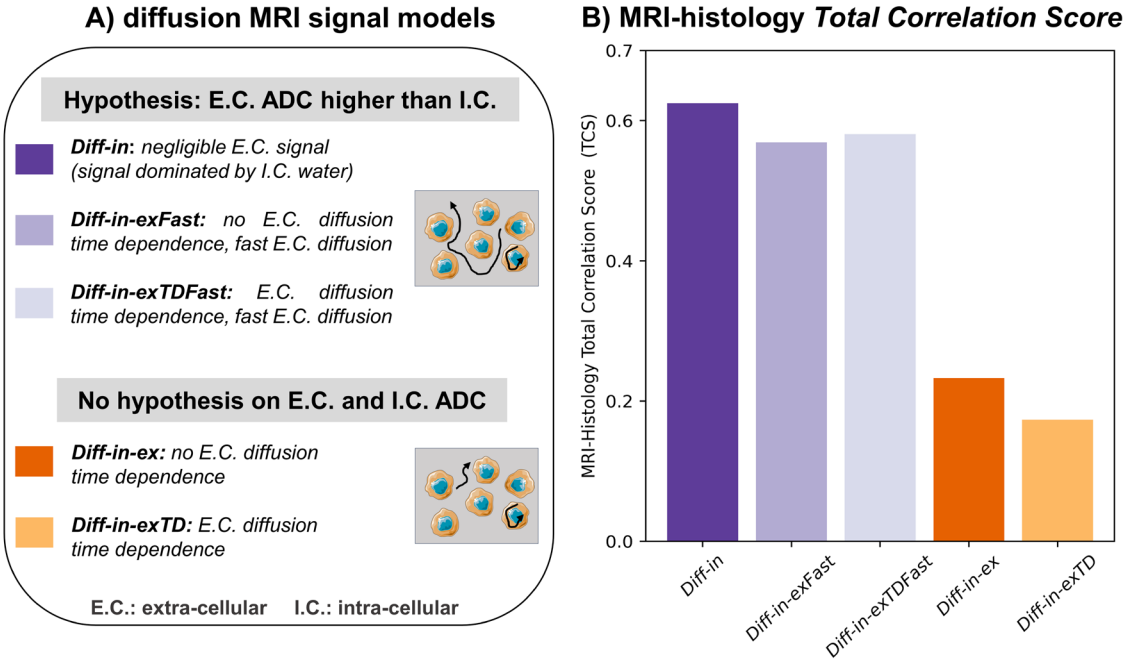


Fig. 3. Biophysical dMRI signal model selection based on the MRI-histology Total Correlation Score (TCS). (**A**): panel summarising the salient differences between the biophysical dMRI models compared in this study. Models can be divided in two families, i.e.: i) models where it is hypothesised that the extra-cellular ADC is higher than the intra-cellular ADC, and ii) models with no hypothesis on which, between intra-/extra-cellular ADC is higher. Violet shades are used for the first family (models *Diff-in*, *Diff-in-exFast* and *Diff-in-exTDFast*), while orange shades for the second family (models *Diff-in-ex* and *Diff-in-exTD*). (**B**): values of TCS for all models, as obtained by fitting them on high b-value images (b > 1800 s/mm^2 in the fixed mouse livers; b > 900 s/mm^2 *in vivo*). We performed model selection using a sample size of N = 25 (see Fig. 1). In Fig. 3, 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).

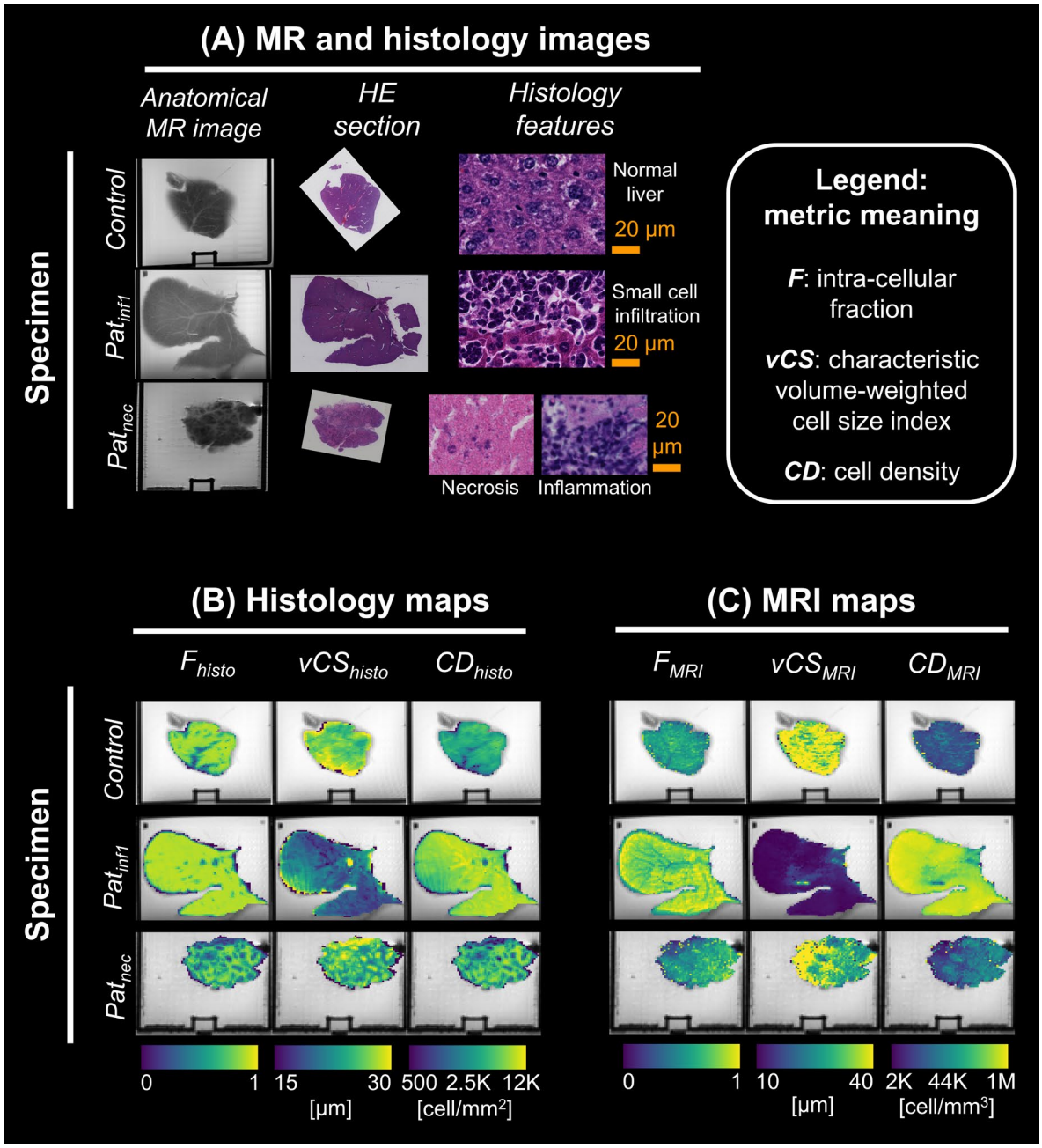


Fig. 4. Maps from the selected dMRI model *Diff-in* with their histological counterparts in the fixed mouse livers scanned at 9.4T *ex vivo*. The figure reports MRI and histology data for 3 specimens, representative of the 3 microstructural phenotypes observed in our *ex vivo* data set, namely: normal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

and CD_{MRI} from the selected dMRI signal model (model *Diff-in*, fitted to high *b*-value images, i.e., $b > 1800$ s/mm^2).

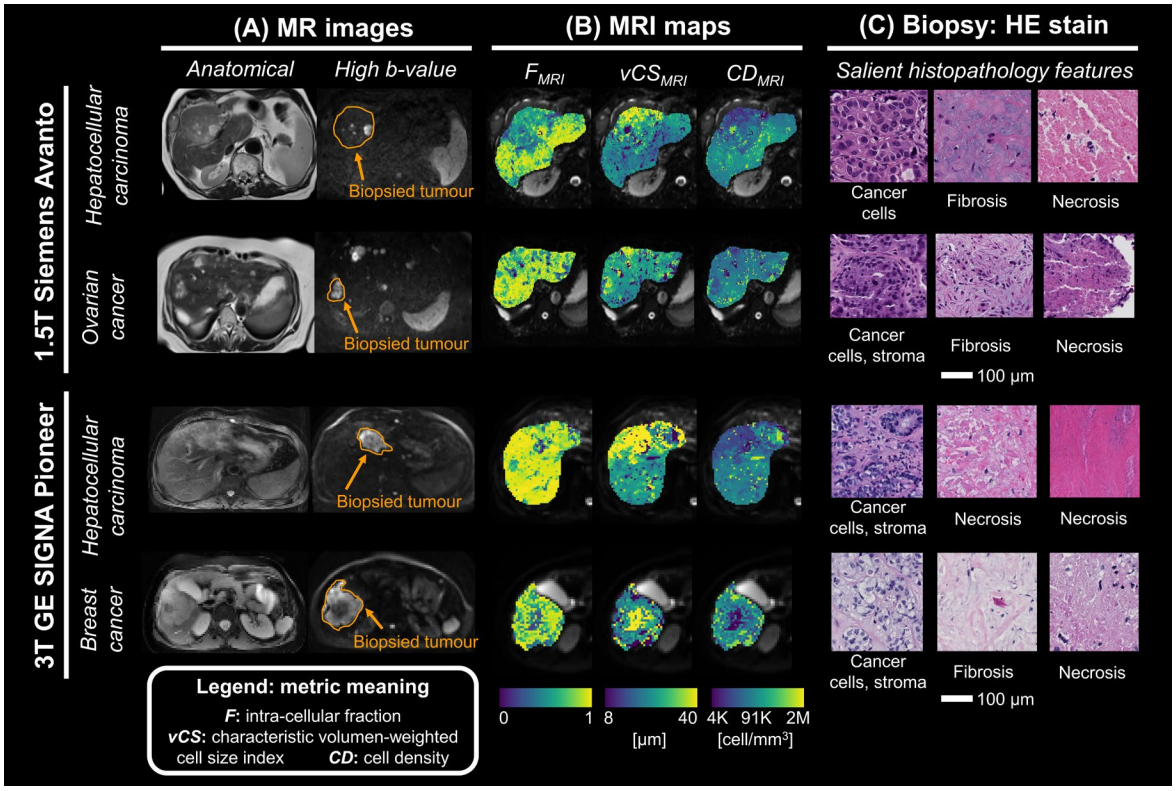


Fig. 5. Examples of maps from the proposed dMRI model *Diff-in* in liver tumours of patients scanned at 1.5T and 3T *in vivo*, with co-localised biopsies. MRI maps are shown in a biopsied liver tumour in two patients for each MRI scanner, arranged along rows. **(A)**: examples of slices from the high-resolution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

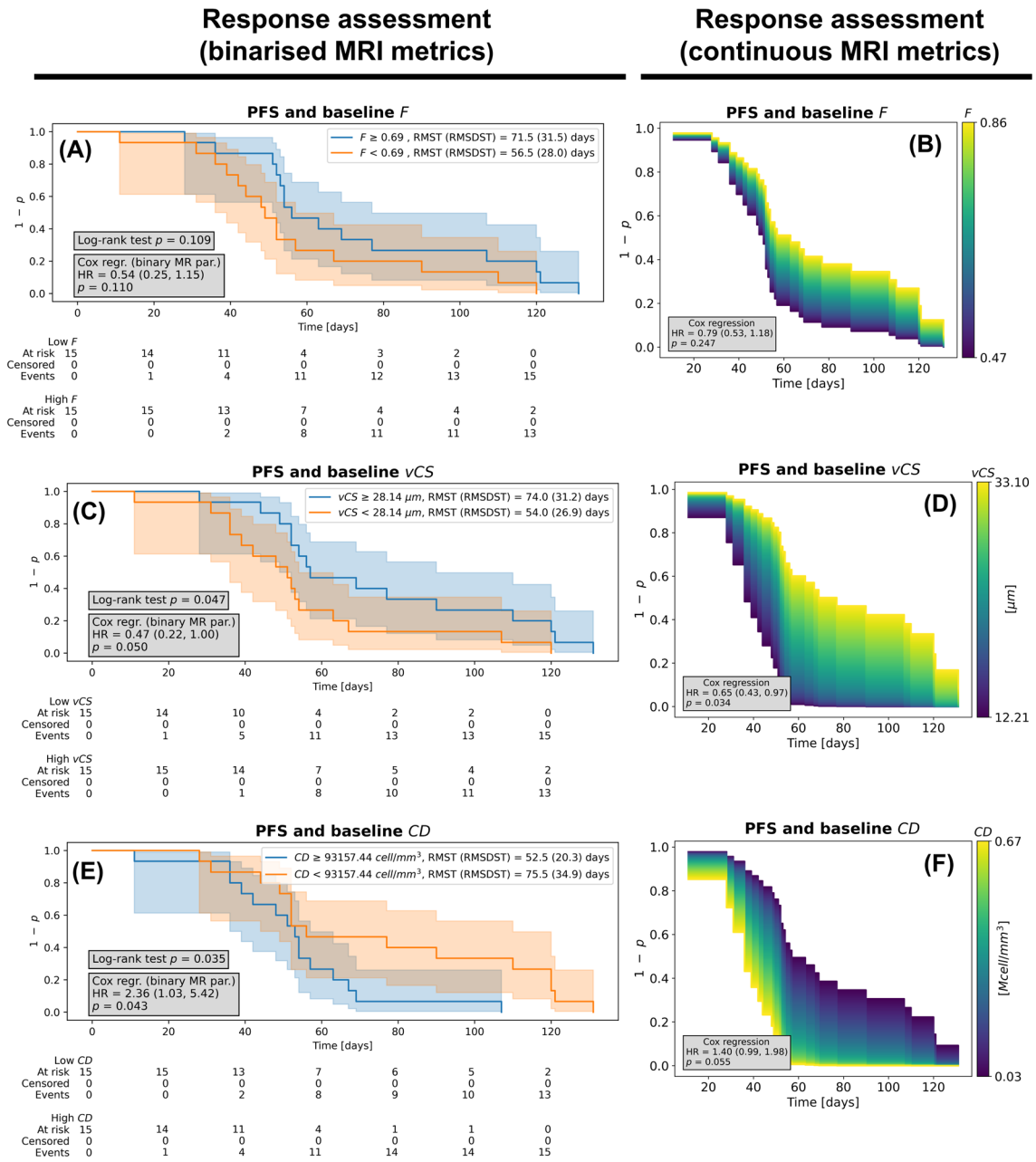


Fig. 6. Immunotherapy response assessment based on metrics from the proposed model *Diff-in* within liver tumours at baseline. This figure reports on the dependence of patients' progression-free survival (PFS) on the average value of F, vCS and CD within liver tumours at baseline (i.e., before starting immunotherapy), as obtained by fitting model *Diff-in* at high b-value. *Left*: Kaplan-Meier (KM) survival curves of two groups obtained by splitting patients based on baseline F (panel A), vCS (panel C) and CD (panel D) (lower/higher than the sample median). The grey panel reports the p-values of a log-rank sum test comparing the KM curves, and of a Cox regression based on the binarised MRI metric (with the corresponding hazard ratio (HR) estimate

and 95% confidence interval). The legend reports the Restricted Mean Survival Time (RMST) and Restricted Standard Deviation of Survival Time (RSDST) for each KM curve. *Right*: results from univariate Cox regression where the baseline F (panel B), vCS (panel D) and CD (panel F) is a continuous predictor of the survival. The panel shows how changes in baseline F, vCS and CD modulate the survival curve, given the HR estimated for each metric. In the grey box, the p-value and HR (with 95% CI) corresponding to the baseline MRI metric are reported. In all panels, the y-axis shows $1 - p$, with p being the probability of progression, while the x-axis shows the time to progression (in days). We performed the response assessment using a sample size of N = 30 (Fig. 1).

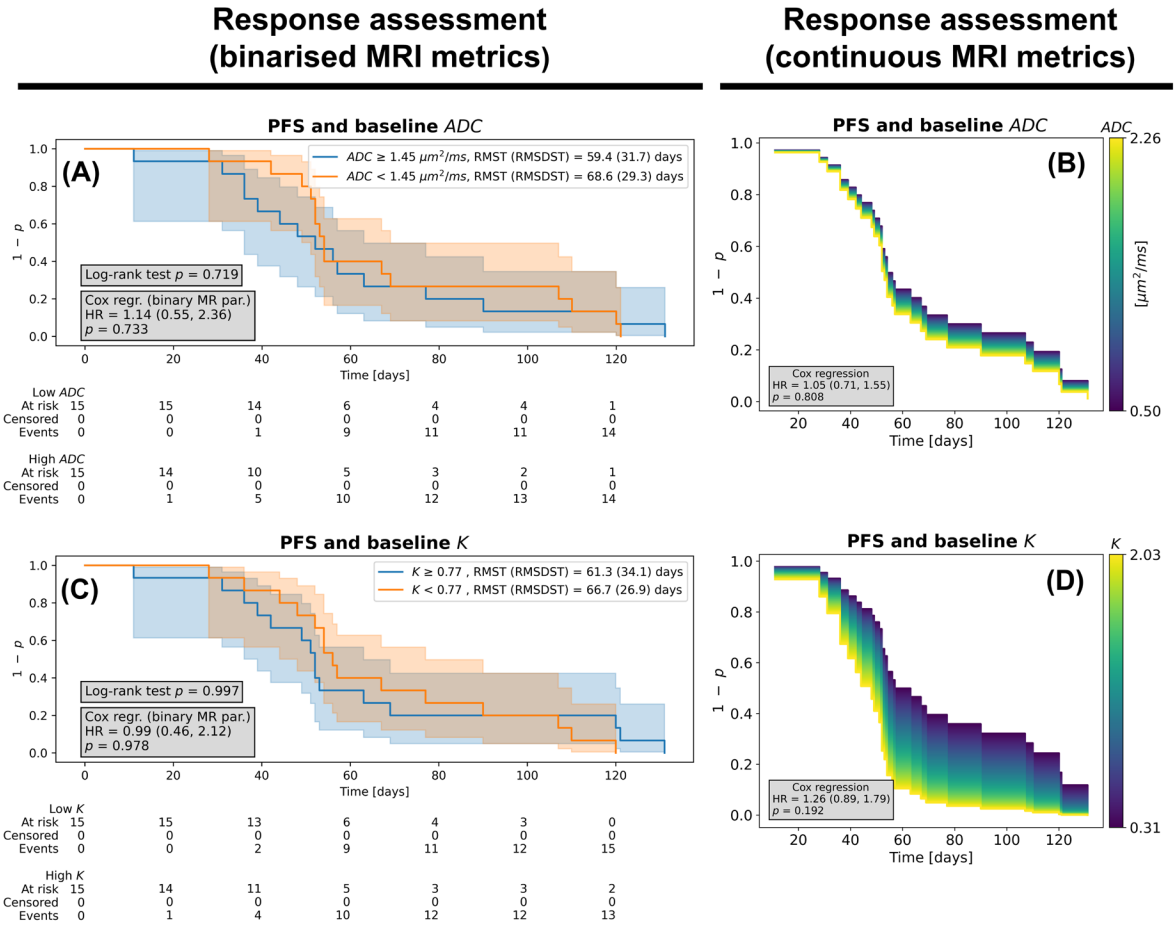


Fig. 7. Immunotherapy response assessment based on mean *ADC* and kurtosis *K* within liver tumours at baseline. This figure reports on the dependence of patients' progression-free survival (PFS) on the average value of *ADC* and *K* within liver tumours at baseline (i.e., before starting immunotherapy). *Left*: Kaplan-Meier (KM) survival curves of two groups obtained by splitting patients based on baseline *ADC* (panel A) and *K* (panel C) (lower/higher than the sample median).

The grey panel reports the p-values of a log-rank sum test comparing the KM curves, and of a Cox regression based on the binarised MRI metric (with the corresponding hazard ratio (HR) estimate and 95% confidence interval). The legend reports the Restricted Mean Survival Time (RMST) and Restricted Standard Deviation of Survival Time (RSDST) for each KM curve. *Right*: results from univariate Cox regression where the baseline ADC (panel B) and K (panel D) is a continuous predictor of the survival. The panel shows how changes in baseline ADC and K modulate the survival curve, given the HR estimated for each metric. In the grey box, the p-value and HR (with 95% CI) corresponding to the baseline MRI metric are reported. In all panels, the y-axis shows 1 – p, with p being the probability of progression, while the x-axis shows the time to progression (in days). We performed the response assessment using a sample size of N = 30 (Fig. 1).

Tables

Table 1. Correlation between dMRI metrics and histological metrics. The table reports Pearson's correlation coefficients r and corresponding p-values p of dMRI metrics F_{MRI} (intra-cellular fraction), vCS_{MRI} (volume-weighted cell size index) and CD_{MRI} (cell density per unit volume) with their histological pairs (F_{histo}, vCS_{histo} and CD_{histo} respectively) for the selected dMRI model (*Diff-in*, fitted to high b-value images). The table also reports correlation coefficients between routine ADC and K from DKI and each of F_{histo}, vCS_{histo} and CD_{histo}. The sample size was N = 25, 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}
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: $r = 0.43$; $p = 0.033$

Supplementary Information for

Histology-informed liver diffusion MRI: biophysical model design and demonstration in cancer immunotherapy

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

*Corresponding authors: Francesco Grussu, email: fgrussu@vhio.net, Raquel Perez-Lopez, email: rperez@vhio.net

Table of contents

Supplementary Table S2: results of the model selection based on the Histology Fidelity Criterion (HFC) as obtained on simulated dMRI signals.
Supplementary Table S3: results of the model selection based on the Bayesian Information Criterion (BIC) as obtained on simulated dMRI signals.
Supplementary Table S4: descriptive statistics of histology and MRI metrics in the fixed mouse livers.
Supplementary Table S5: hazard ratios obtained from Cox regression models controlling for sex, age, and baseline tumour volume.

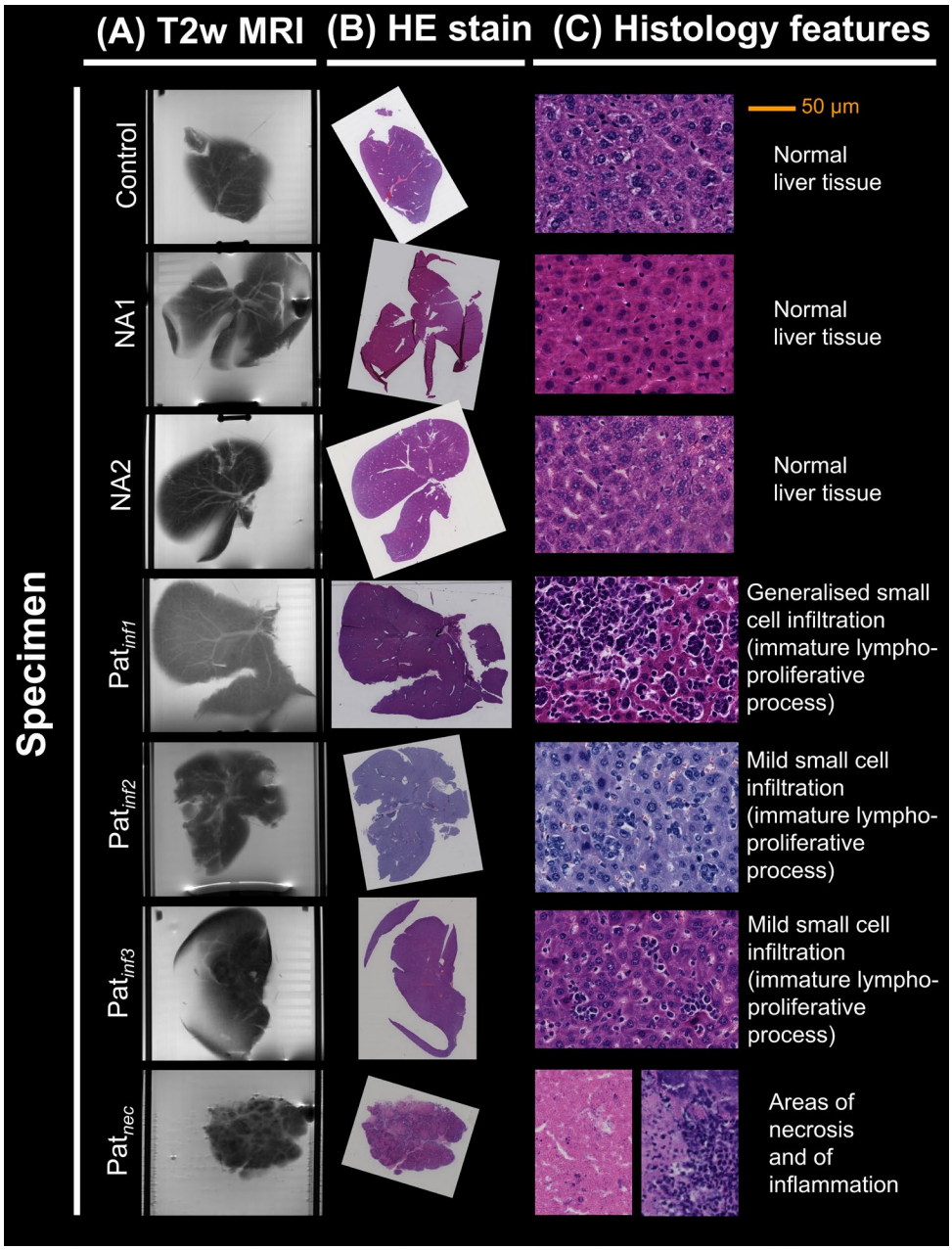


Fig. S1: radiological-histological co-localisation of the *ex vivo* mouse liver data.

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_{infl}, 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.

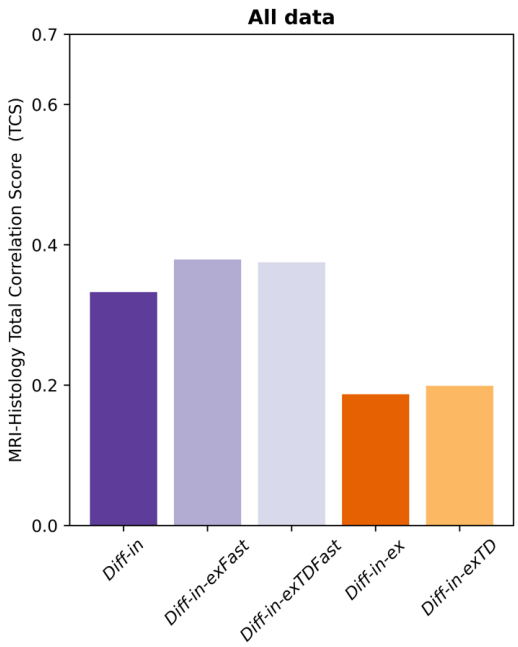


Fig. S2: MRI-histology Total Correlation Score (TCS) for biophysical dMRI model selection, as obtained when fitting dMRI signal models on the whole image set.

Values of TCS for all models, as obtained by fitting models on the whole image set (set of images with negligible vascular signal contributions, i.e., $b > 1000$ s/mm^2 in the fixed mouse livers and $b > 100$ s/mm^2 *in vivo*). We evaluated TCS for histology-informed model selection using a sample size of N = 25.

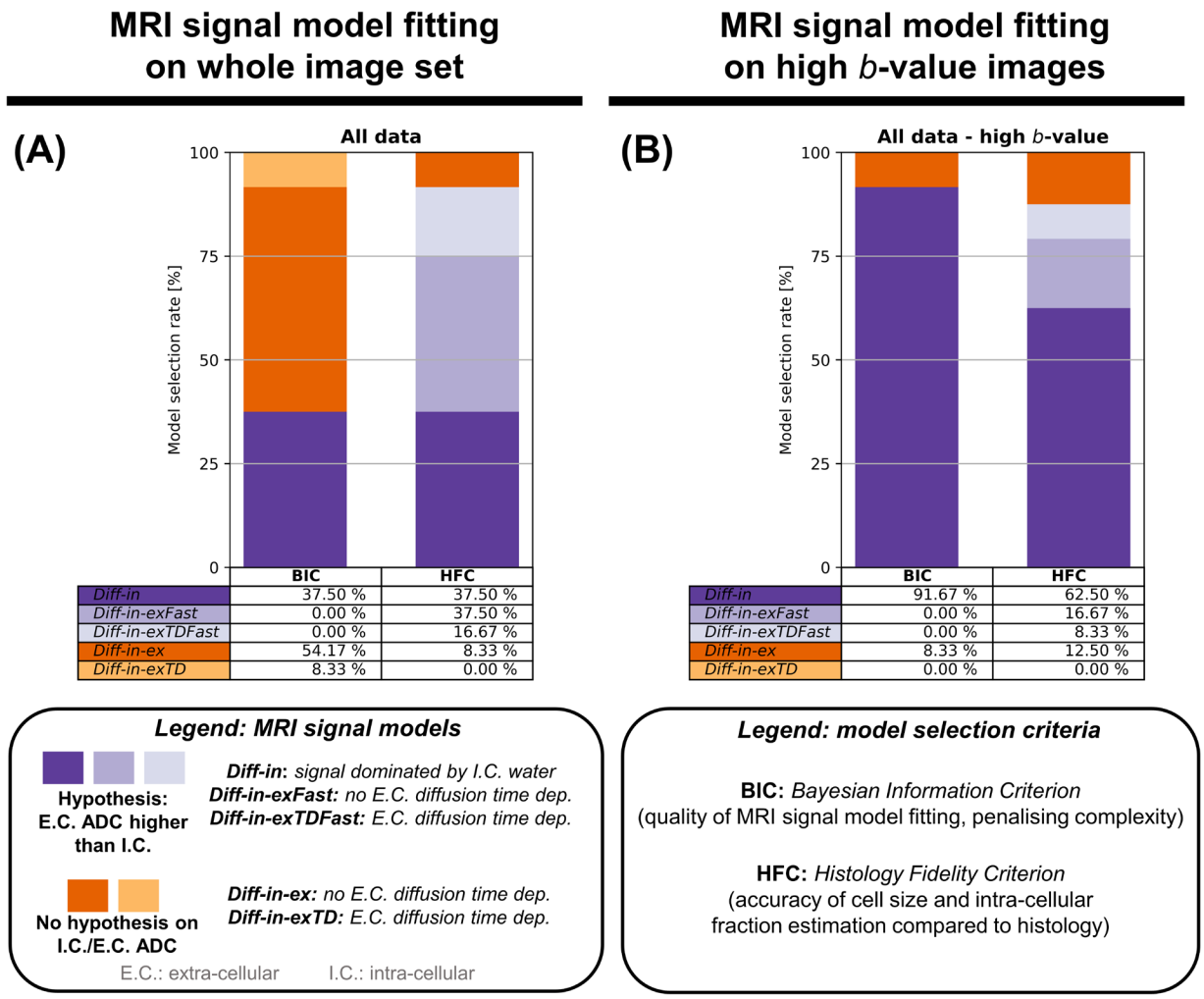


Fig. S3: biophysical dMRI signal model selection based on BIC and HFC criteria.

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 signal 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 the fixed mouse livers; $b > 900$ s/mm^2 *in vivo*).

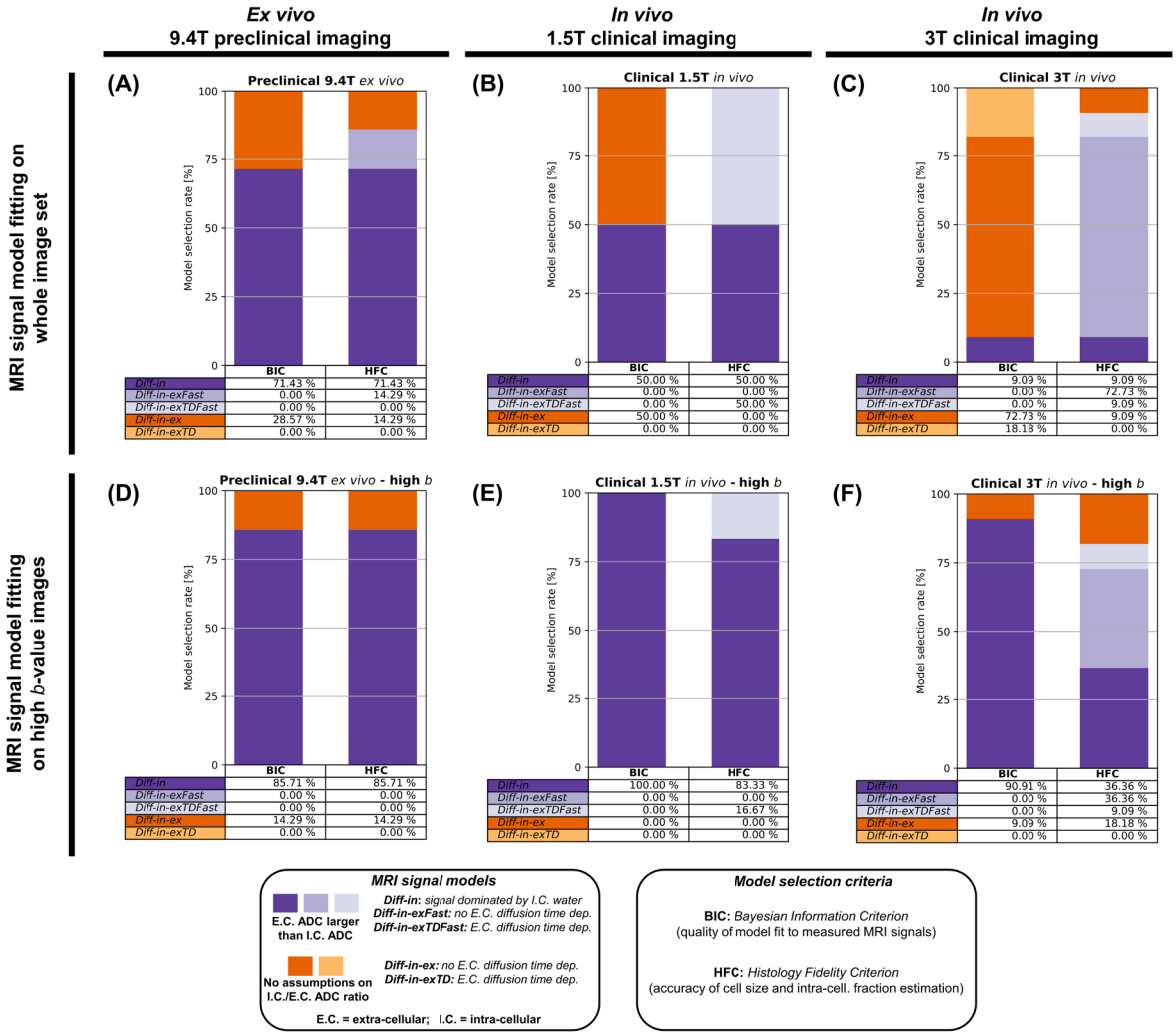


Fig. S4: biophysical model selection across different MRI scanners and data subsets.

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 MRI signal fitting. The *Histology Fidelity Criterion* (HFC) selects a model depending on the overall agreement between MRI volume-weighted Cell Size (*vCS*) and intra-cellular fraction (*F*) with their histology counterparts.

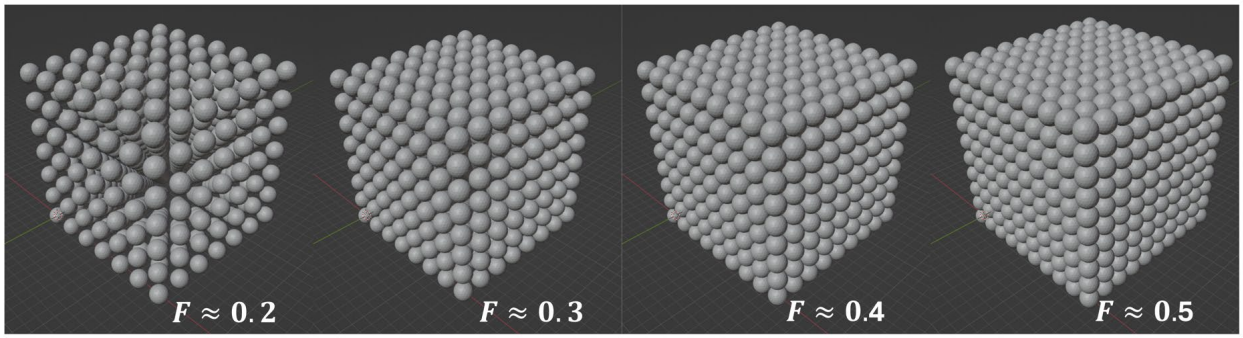


Fig. S5: environments used to generate synthetic dMRI signals in computer simulations.
The synthetic environment consisted of meshed spheres of fixed diameter, representing cells, as this is a common biophysical model used in several dMRI techniques (e.g., VERDICT, IMPULSED). We used the synthetic environment 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).

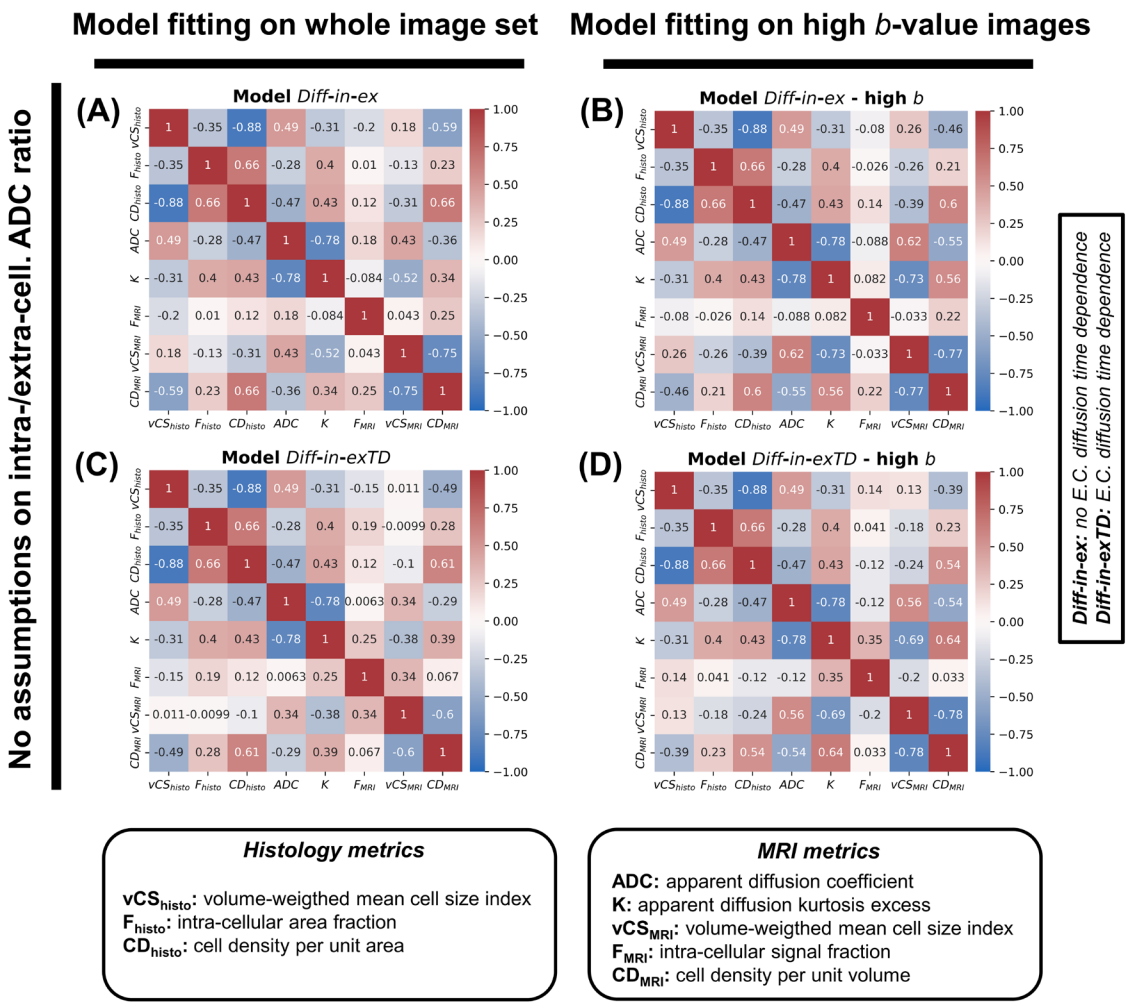


Fig. S6: MRI-histology correlations for models with no assumptions on which is larger between intra-cellular and extra-cellular ADC. Matrices illustrating Pearson's correlation coefficients among all possible pairs of MRI and histology metrics. Histological metrics are: intra-cellular area fraction F_{histo}; 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 (Fig. 1).

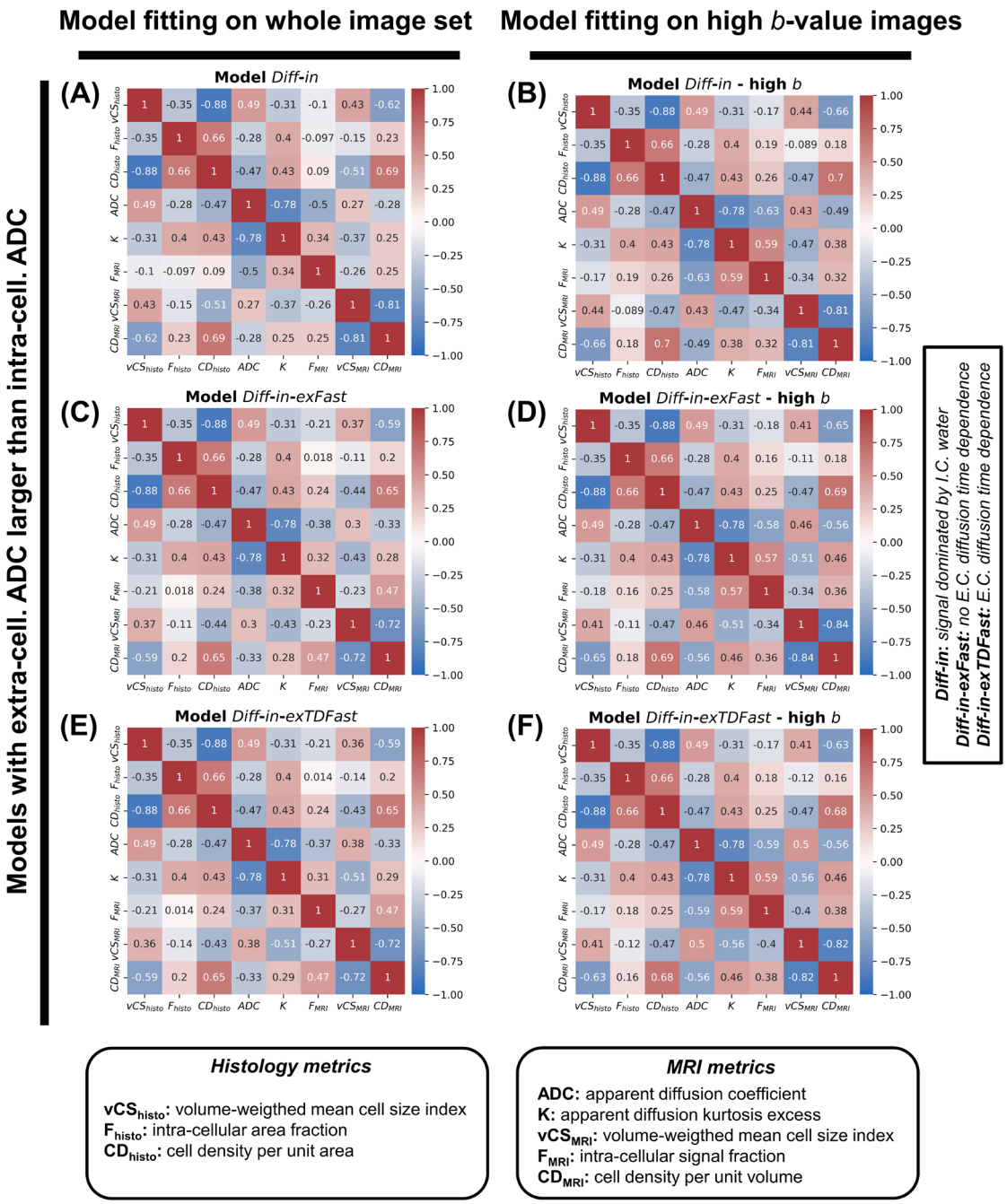


Fig. S7: MRI-histology correlations for models where the extra-cellular ADC is constrained to be larger than the intra-cellular ADC. Matrices illustrating Pearson's correlation coefficients among all possible pairs of MRI and histology metrics. Histological metrics are: intra-cellular area fraction F_{histo}; 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 (Fig. 1).

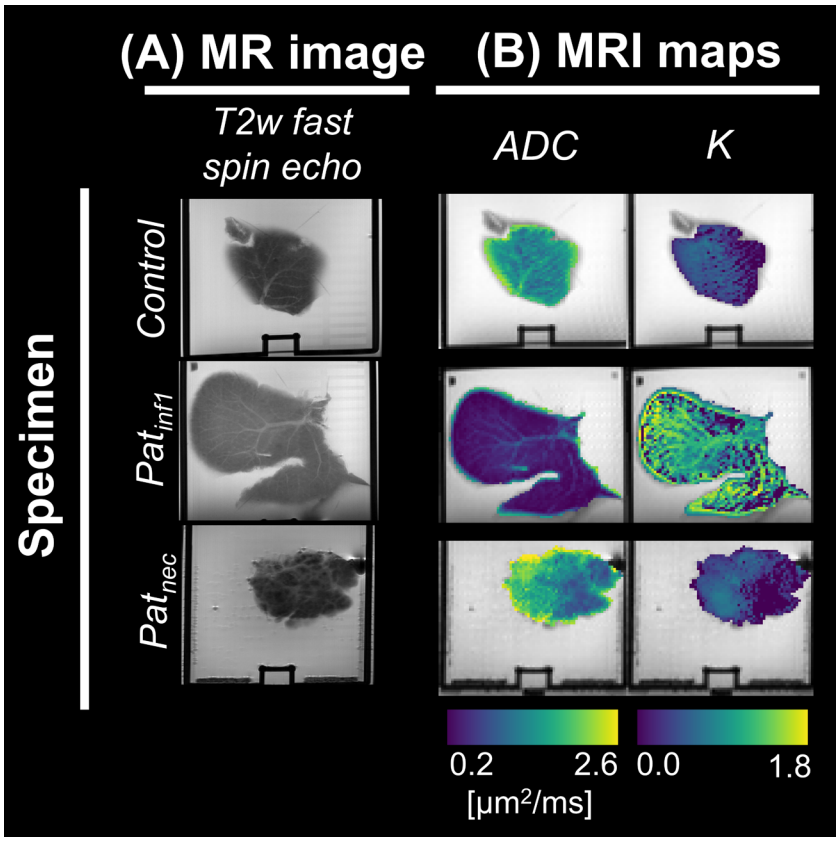


Fig. S8: standard diffusion MRI metrics in fixed mouse livers *ex vivo*.

(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}).

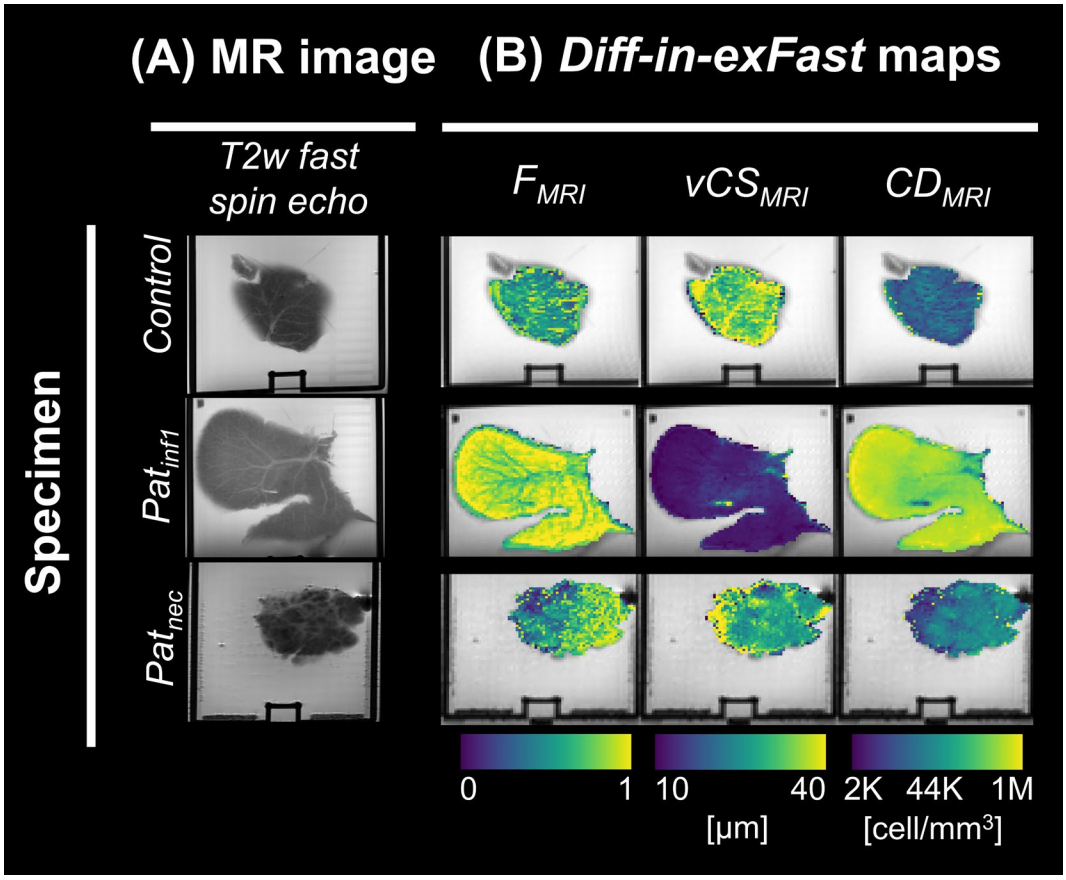


Fig. S9: key parametric maps of the *Diff-in-exFast* model on *ex vivo* mouse livers.

(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}).

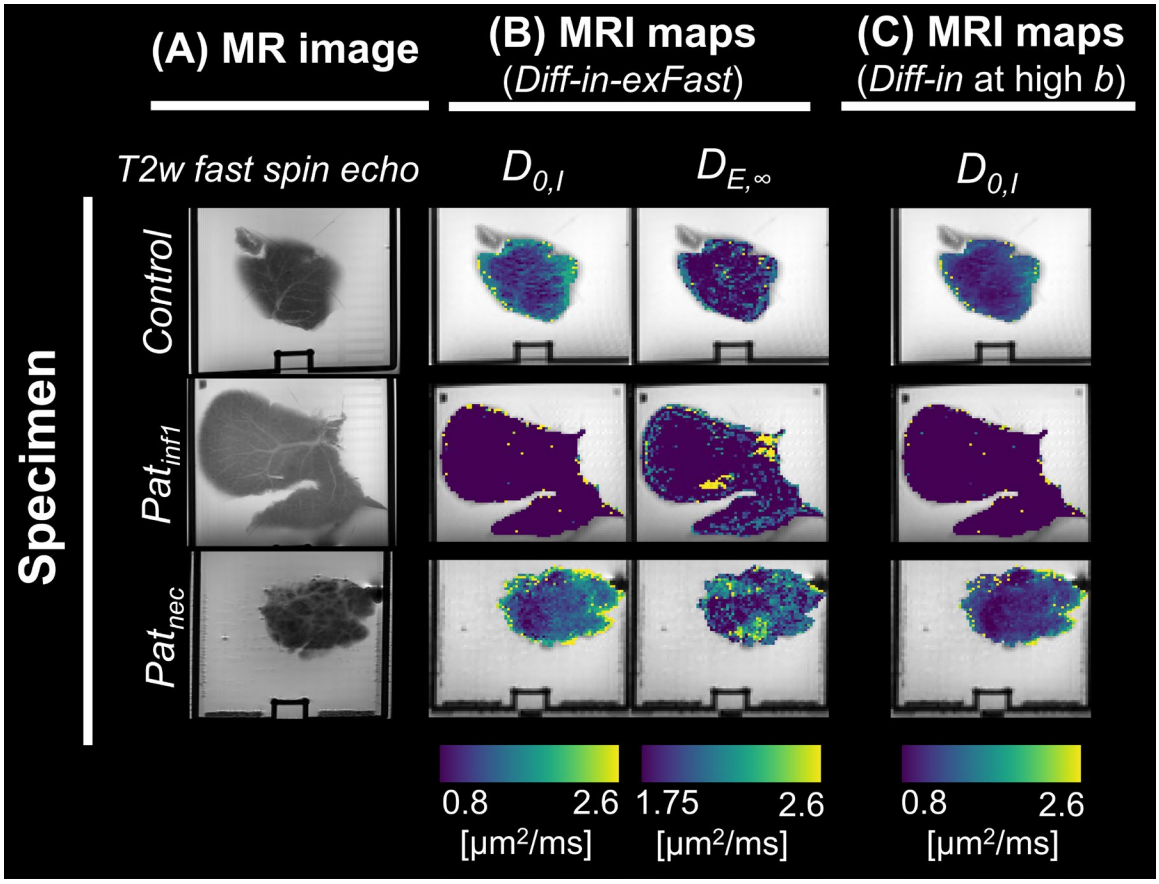


Fig. S10: diffusivity metrics from biophysical MRI models in fixed *ex vivo* mouse livers.

(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}).

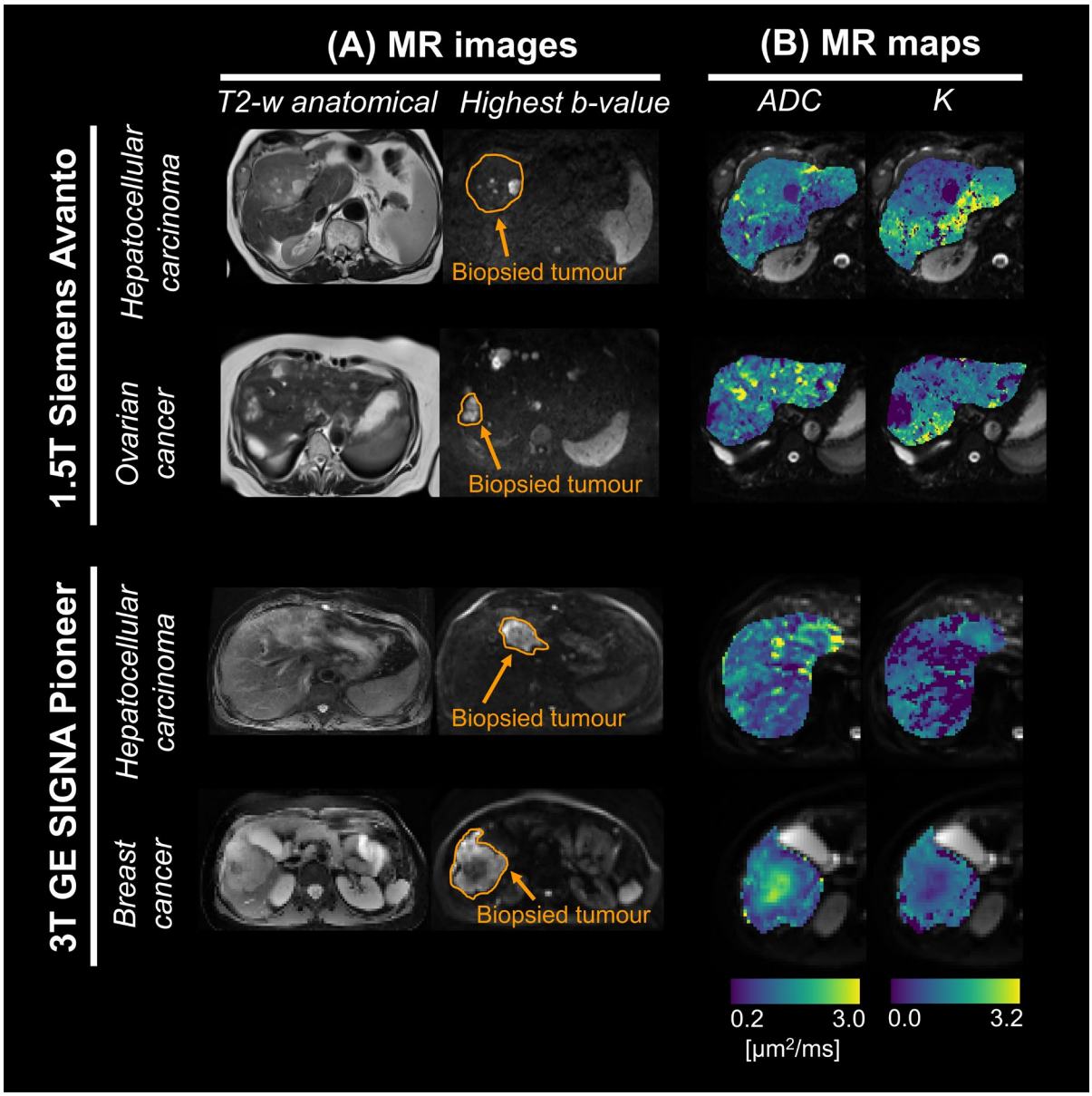


Fig. S11: standard diffusion MRI metrics in patients *in vivo*.

(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): patient 6 (primary hepatocellular carcinoma) and patient 3 (liver metastases from ovarian cancer). For the 3T GE scanner (third and fourth rows from top): patient 24 (primary hepatocellular carcinoma) and patient 30 (liver metastases from breast cancer).

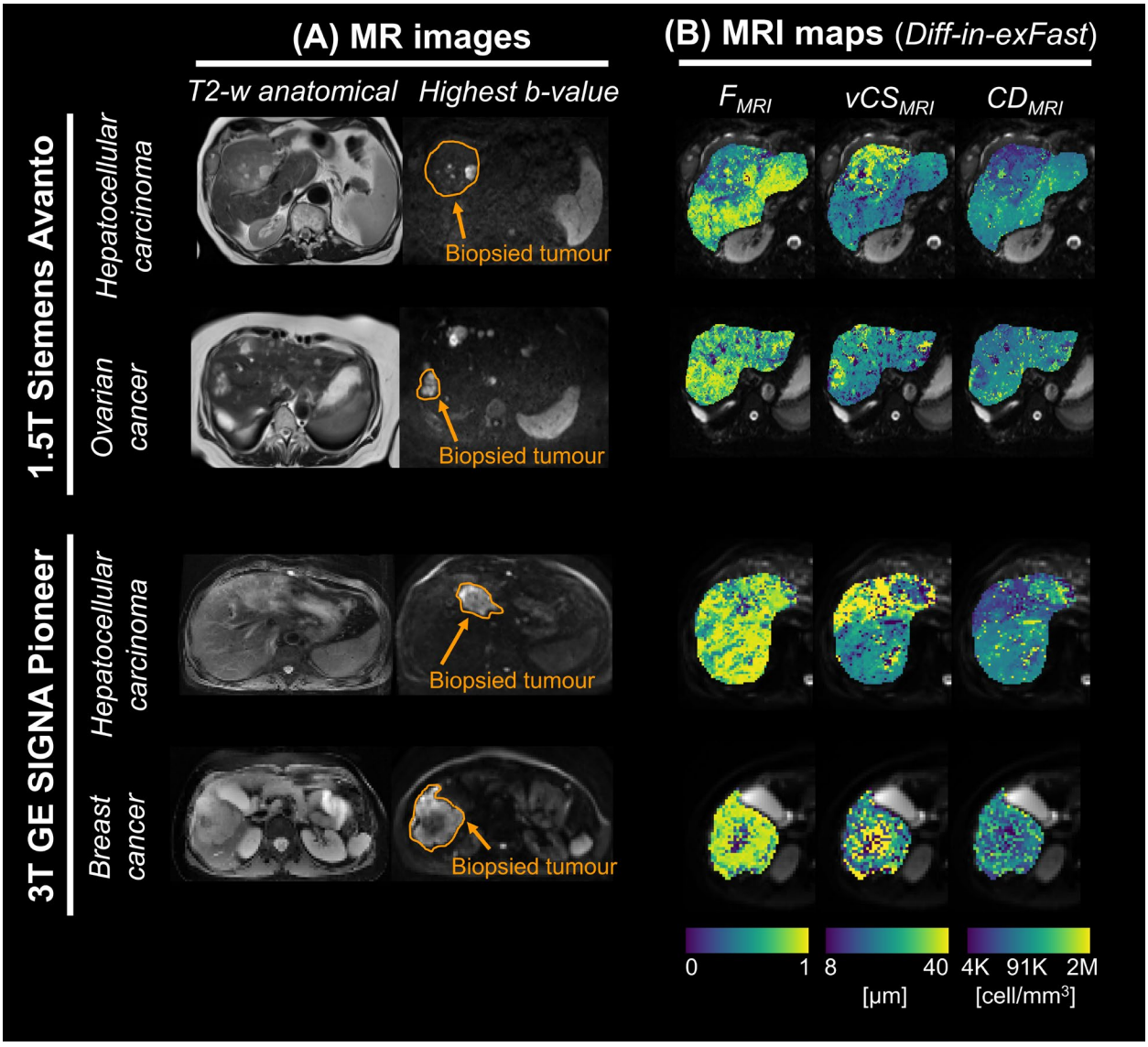


Fig. S12: key parametric maps of the *Diff-in-exFast* model in patients *in vivo*.

(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): patient 6 (primary hepatocellular carcinoma) and patient 3 (liver metastases from ovarian cancer). For the 3T GE scanner (first and second rows from bottom): patient 24 (primary hepatocellular carcinoma) and patient 30 (liver metastases from breast cancer).

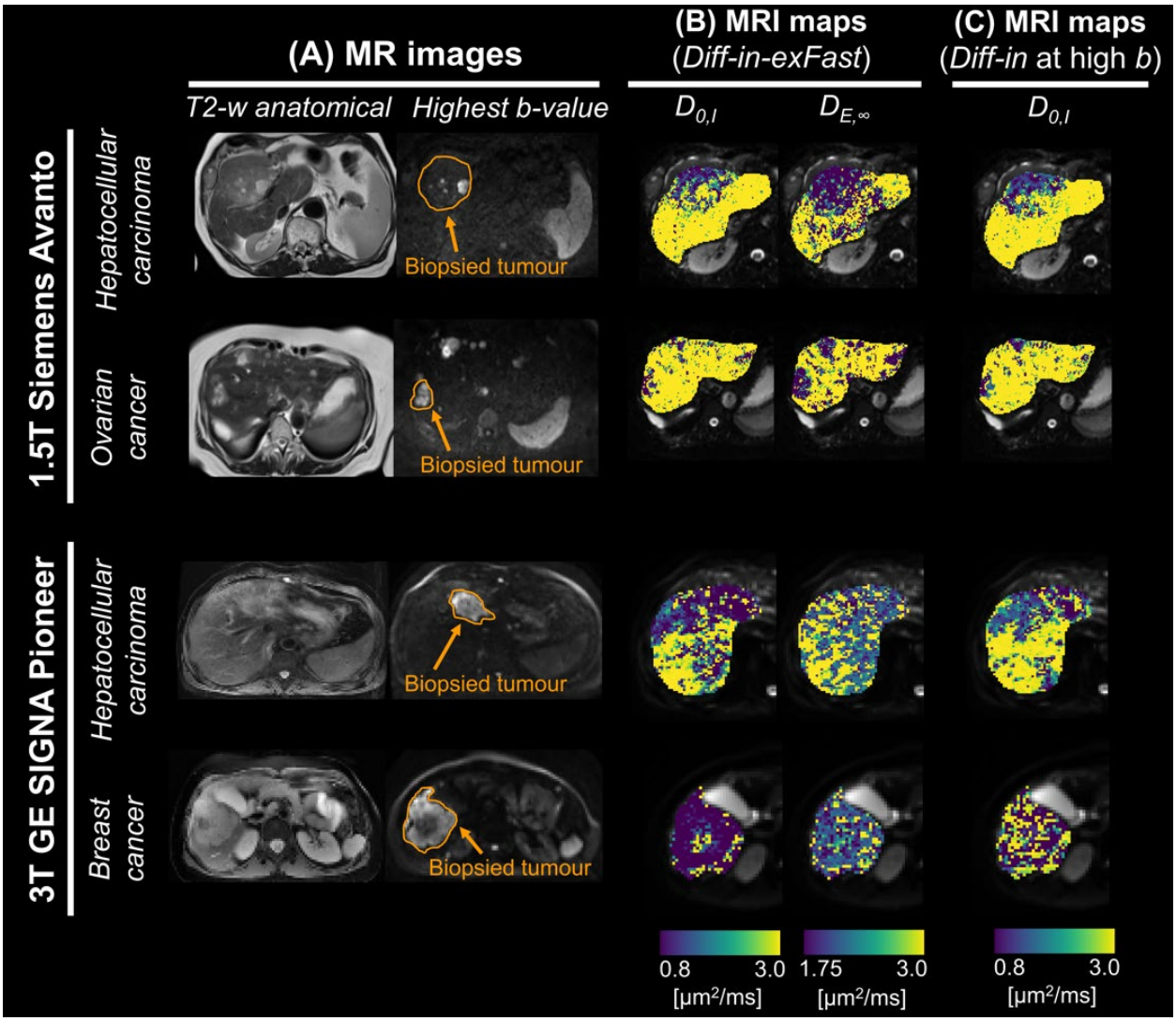


Fig. S13: diffusivity metrics from biophysical MRI models in patients *in vivo*.

(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): patient 6 (primary hepatocellular carcinoma) and patient 3 (liver metastases from ovarian cancer). For the 3T GE scanner (first and second rows from bottom): patient 24 (primary hepatocellular carcinoma) and patient 30 (liver metastases from breast cancer).

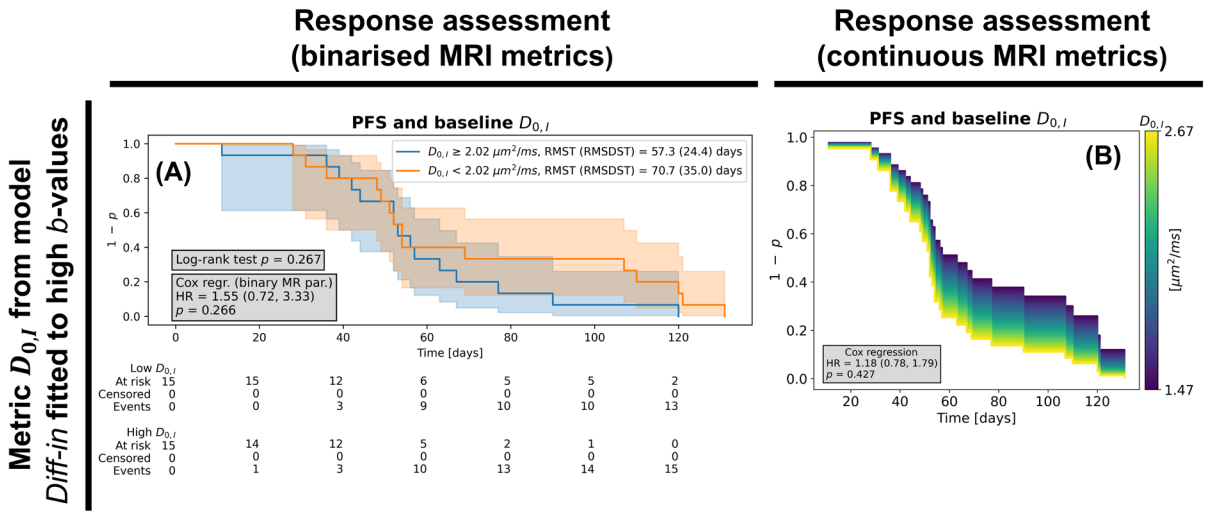


Fig. S14: immunotherapy response assessment based on *Diff-in* cytosol diffusivity estimates.

This figure reports on the dependence of patients' progression-free survival (PFS) on the average value of the intrinsic intra-cellular cytosol diffusivity $D_{0,I}$ within liver tumours at baseline (i.e., before starting immunotherapy), as obtained by fitting model *Diff-in* at high b-value. Left (panel **A**): Kaplan-Meier (KM) survival curves of two groups obtained by splitting patients based on baseline $D_{0,I}$ (lower or higher than the sample median). The grey panel reports the p-values of a log-rank sum test comparing the KM curves, and of a Cox regression based on the binarised MRI metric (with the corresponding hazard ratio (HR) estimate and 95% confidence interval). The legend reports the Restricted Mean Survival Time (RMST) and Restricted Standard Deviation of Survival Time (RSDST) for each KM curve. Right (panel **B**): results from univariate Cox regression where the baseline $D_{0,I}$ is a continuous predictor of the survival. The panel shows how changes in baseline $D_{0,I}$ modulate the survival curve, given the HR estimated for each metric. In the grey box, the p-value and HR (with 95% CI) corresponding to the baseline MRI metric are reported. In all panels, the y-axis shows $1 - p$, with p being the probability of progression, while the x-axis shows the time to progression (in days).

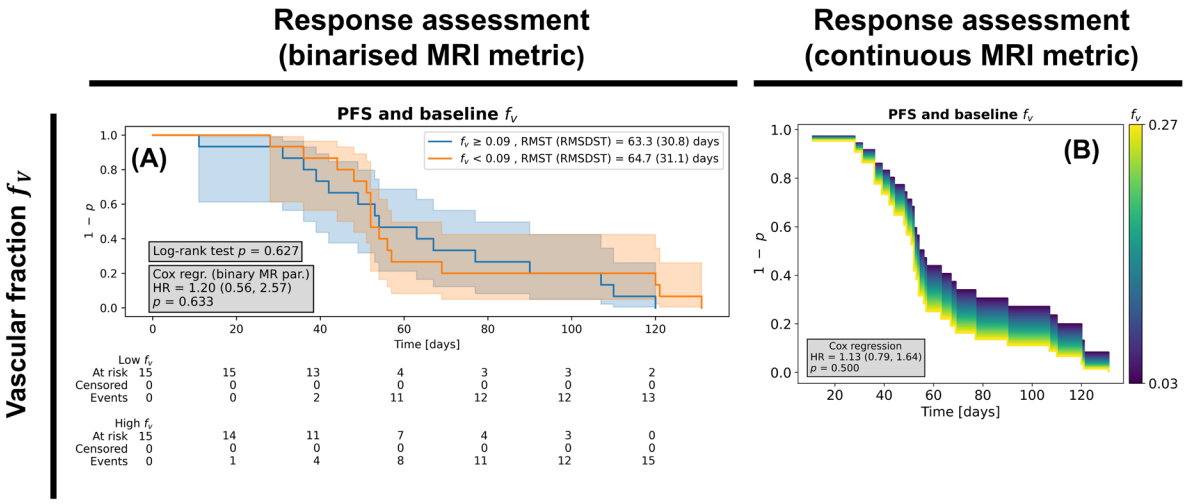


Fig. S15: immunotherapy response assessment based on vascular fraction estimates.

This figure reports on the dependence of patients' progression-free survival (PFS) on the average value of the vascular signal fraction f_v within liver tumours at baseline (i.e., before starting immunotherapy). The vascular signal fraction f_v was computed in the initial fitting step, which disentangles the vascular from the non-vascular (tissue) signal, before the latter is split into intra-/extra-cellular contributions in the biophysical model fitting step. The same representation layout as in Fig. S6 was used. Left (**A**): Kaplan-Meier (KM) analysis, log-rank sum test and Cox regression based on the binarised f_v (higher/lower than the sample median). Right (**B**): Cox regression modelling the probability of survival as a continuous function of baseline f_v. In all panels, the y-axis shows $1 - p$, with p being the probability of progression, while the x-axis shows the time to progression (in days).

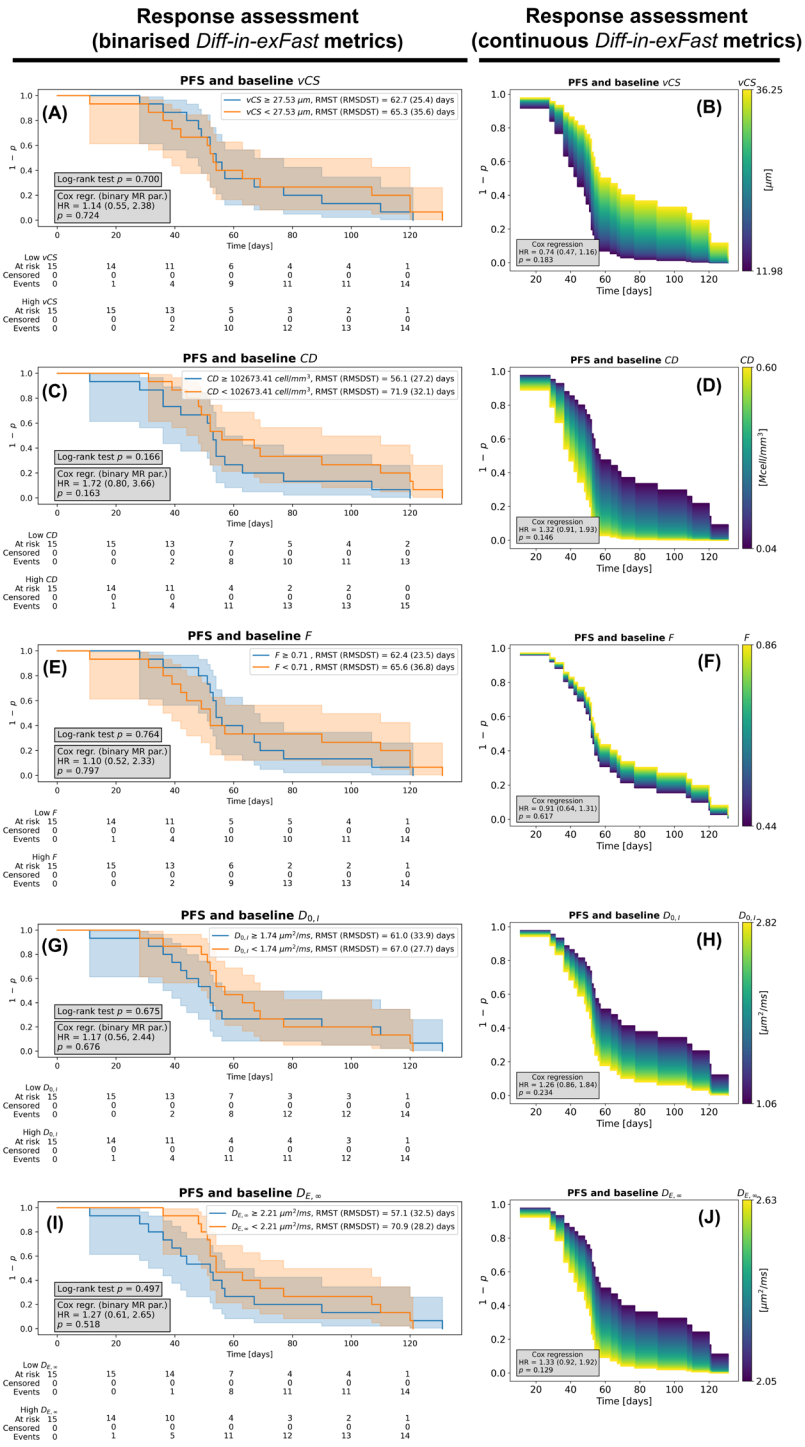


Fig. S16: immunotherapy response assessment based on *Diff-in-exFast* MRI metrics.

This figure reports on the dependence of patients' progression-free survival (PFS) on the average value of all *Diff-in-exFast* metrics within liver tumours at baseline (i.e., before starting immunotherapy). In each row, from top to bottom: PFS based on baseline volume-weighted Cell Size vCS (panels **A** and **B**), Cell Density CD (panels **C** and **D**), intra-cellular fraction F (panels **E**

and **F**), intrinsic intra-cellular cytosol diffusivity $D_{0,I}$ (panels **G** and **H**), asymptotic extra-cellular diffusion coefficient $D_{E,\infty}$ (panels **I** and **J**). The same representation layout as in Fig. S6 was used. Left (**A**, **C**, **E**, **G**, **I**): Kaplan-Meier (KM) analysis, log-rank sum test and Cox regression based on the binarised MRI metrics (higher/lower than the sample median). Right (**B**, **D**, **F**, **H**, **J**): Cox regression modelling the probability of survival as a continuous function of baseline MRI metrics. In all panels, the y-axis shows $1 - p$, with p being the probability of progression, while the x-axis shows the time to progression (in days).

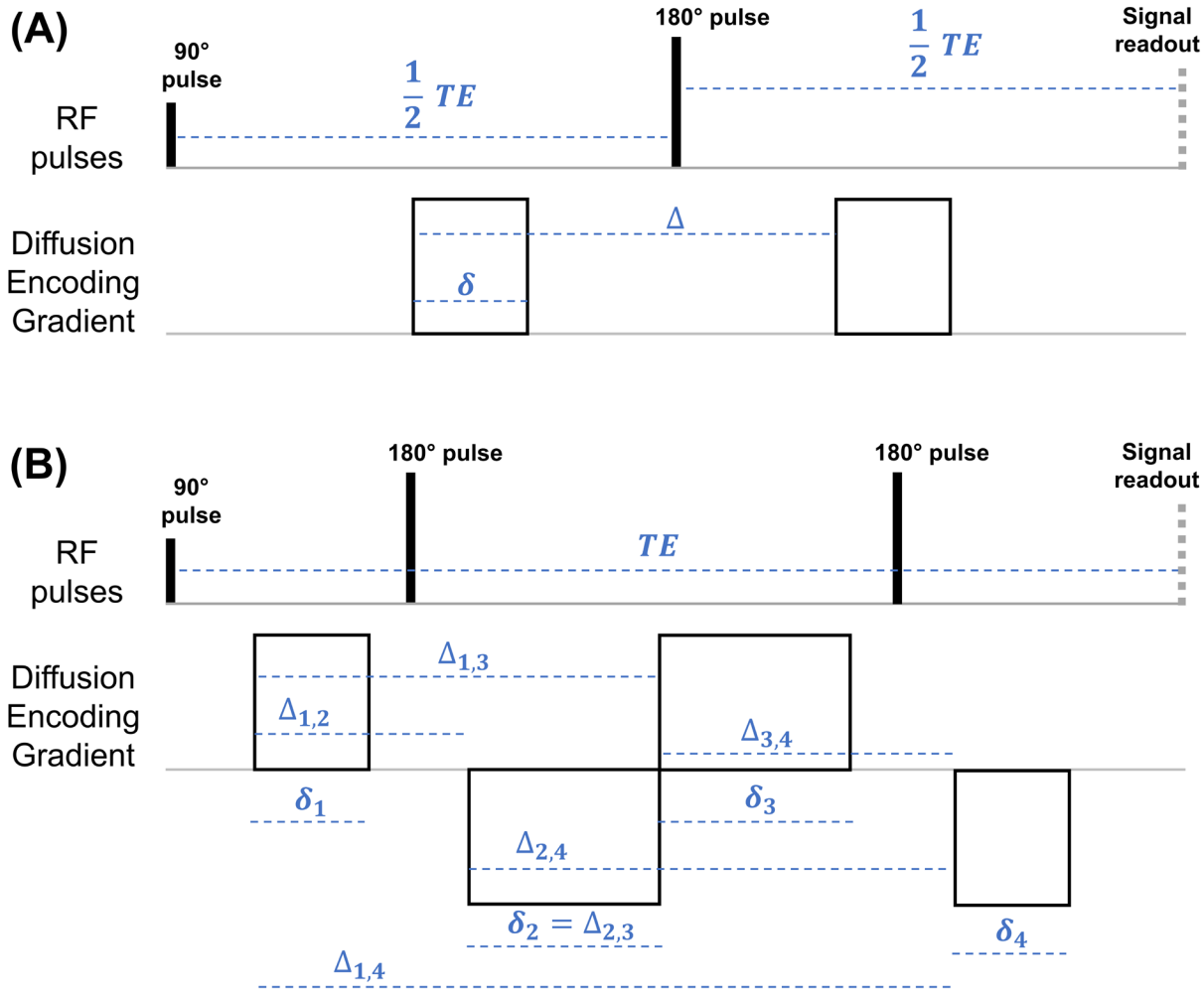


Fig. S17: schematic of the dMRI sequences used in this study.
(**A**): pulsed gradient spin echo (PGSE sequence, also known as Stejskal-Tanner sequence, pulsed-field gradient (PFG), or single linear diffusion encoding) used to acquire data on the 9.4T Bruker system on fixed mouse livers *ex vivo* and 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 $n,m = 1, \ldots, 4$. 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).

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. Hazard ratios obtained from Cox regression models controlling for sex, age, and baseline tumour volume.

The table reports the Hazard Ratios (HR) for different MRI metrics, with relative 95% confidence interval and p-value, estimated through Cox proportional hazard regressions. The models assessed the dependence of the probability of progression on the baseline mean value of MRI metrics within liver tumours, accounting for sex, age and tumour volume. Results are shown for standard diffusion metrics (apparent diffusion and excess kurtosis coefficients, ADC and K), for the vascular signal fraction f_v, for metrics from model *Diff-in-exFast* (intra-cellular fraction F, volume-weighted cell size vCS, cell density per unit volume CD, intrinsic intra-cellular cytosolic diffusivity $D_{0,I}$, extra-cellular asymptotic diffusion coefficient $D_{E,\infty}$) and for metrics from model *Diff-in* (intra-cellular fraction F, volume-weighted cell size vCS, cell density per unit volume CD, intrinsic intra-cellular cytosolic diffusivity $D_{0,I}$) fitted at high b-value (b > 900 s/mm^2). Grey shadowing highlights HRs whose p-value is ≤ 0.05.

	HR of MRI metric	**HR of male sex**	**HR of age**	**HR of tumour volume**
Standard diffusion MRI metrics				
ADC	1.00 (0.67; 1.49); p = 1.00	0.71 (0.32; 1.55); p = 0.39	0.97 (0.94; 0.99); p = 0.02	0.32 (0.06; 1.70); p = 0.18
K	1.36 (0.94; 1.97); p = 0.11	0.64 (0.29; 1.41); p = 0.27	0.97 (0.94; 1.00); p = 0.02	0.35 (0.08; 1.54); p = 0.17
Other metrics (from vascular vs non-vascular fitting initialisation step)				
f_v	1.16 (0.79; 1.71); p = 0.44	0.77 (0.35; 1.68); p = 0.51	0.96 (0.93; 0.99); p = 0.01	0.34 (0.17; 2.56); p = 0.17
Metrics from MRI model *Diff-in-exFast*				
F	0.87 (0.61; 1.26); p = 0.46	0.78 (0.35; 1.72); p = 0.54	0.97 (0.94; 1.00); p = 0.02	0.31 (0.06; 1.58); p = 0.16
vCS	0.70 (0.44; 1.12); p = 0.14	0.78 (0.36; 1.68); p = 0.53	0.97 (0.94; 0.99); p = 0.02	0.35 (0.08; 1.46); p = 0.15
CD	1.53 (1.00; 2.34); p = 0.05	0.63 (0.29; 1.38); p = 0.25	0.96 (0.93; 0.99); p = 0.01	0.32 (0.06; 1.81); p = 0.20
$D_{0,I}$	1.45 (0.98; 2.14); p = 0.06	0.68 (0.31; 1.47); p = 0.33	0.96 (0.93; 0.99); p = 0.01	0.36 (0.09; 1.35); p = 0.13
$D_{E,\infty}$	1.41 (0.97; 2.05); p = 0.07	0.69 (0.32; 1.48); p = 0.34	0.96 (0.94; 0.99); p = 0.01	0.40 (0.13; 1.26); p = 0.12
Metrics from MRI model *Diff-in* fitted to high *b*-value images				
F	0.79 (0.54; 1.17); p = 0.24	0.80 (0.36; 1.74); p = 0.57	0.97 (0.94; 0.99); p = 0.02	0.33 (0.07; 1.49); p = 0.15
vCS	0.59 (0.37; 0.93); p = 0.02	0.66 (0.30; 1.43); p = 0.29	0.96 (0.94; 0.99); p = 0.01	0.40 (0.12; 1.37); p = 0.14
CD	1.65 (1.12; 2.44); p = 0.01	0.61 (0.28; 1.35); p = 0.22	0.96 (0.93; 0.99); p = 0.01	0.36 (0.09; 1.52); p = 0.17
$D_{0,I}$	1.28 (0.82; 1.98); p = 0.28	0.67 (0.31; 1.45); p = 0.31	0.96 (0.94; 0.99); p = 0.01	0.37 (0.09; 1.48); p = 0.16