Supporting Information

IDF model formalism:

Water diffusion rate in solutions over narrow temperature range can be described by Arrhenius formula, $D = A \cdot exp(-E_a/RT)$, (1,2) where A is molecule collision frequency factor and E_a is thermal activation energy, physically related to molecular concentration and hydrogen bonding strength, respectively. In general, both parameters tend to increase with macromolecule solute concentration with respect to values for pure water (1,2). In vivo, temperature is fixed at 37 °C allowing this relation to be expressed as $D = F \cdot D_W$, factoring out water diffusion rate components from those dependent on solute concentrations, $F \approx A/A_w \cdot \exp(-(E_a - E_w)/RT)$. Consequently, for nanoscale interaction between water and macromolecules (Figure1, left), the proposed model assumes two Arrhenius water pools with distinct diffusion constants (present both in intra- and extra-cellular space) that contribute to the apparent diffusion, Dc, for subcellular water fraction F_c,

$$D_c \approx F_{fc} \cdot D_f + IDF \cdot D_i; \quad F_c = F_{fc} + IDF \qquad \text{Eq.S1}$$

Here D_f is the free water diffusion rate (e.g., 3μs^2/ms at body temperature of 37^{o}C or 2.1μs^2/ms at typical scanner room temperature of 21 ^{o}C) of uncoordinated water fraction, F_{fc}; while D_i is an average collective diffusion rate of the impeded water coordinated around macromolecules and cell membranes (both inner and outer layers, Fig.1, left). Note that according to Arrhenius model, F_{fc} (D_c), is expected to scale inversely with the solute concentration (e.g., as observed for the polyvinylpyrrolidone (PVP) in Sup. Figure S2a, V3 & V5), due to generally steeper dependence on activation energy versus collision frequency. The D_i value is governed by the macromolecule diffusion and its ability to coordinate water, and is much slower than free water diffusion (D_i<0.03μm^2/ms << D_f) as observed by previous NMR and DLS studies (3,4). These two pools contribute to DWI signal from a (sub)cellular compartment as, F_cE_c; $E_c = exp(-bD_c)$, where b-value represent the strength of applied diffusion gradients. For typical clinical DWI scanning with b<2ms/μm^2, and D_i<<D_f, the following approximation is valid: $D_c \rightarrow F_{fc}D_f$;, $E_c \approx \exp(-bF_{fc}D_f)$.

On the microscale (Fig.1, middle), in addition to subcellular component, DWI signal includes contribution from bulk free diffusion fraction, F_f (e.g., in ductal lumen rich in ionic fluids and hydrophobic macromolecules that do not coordinate water diffusion), $F_fE_f = F_f\exp(-bD_f)$, and pseudo

diffusion, F_p, in randomly oriented capillaries, $F_pE_p = F_f\exp(-bD_p)$, such that $F_p + F_f + F_c = 1$. The pseudo-diffusion constant is assumed to be primarily determined by ~5mm-capillary pressure gradient (D_p>30μm^2/ms>>D_f) exceeding D_f, as measured e.g. by PET perfusion studies (5). Hence, pseudo-diffusion contribution to DWI signal becomes negligible ($E_p \rightarrow 0$) for b>.1 ms/μm^2.

On a typical clinical DWI voxel scale of 2x2x4 mm^3 (Fig.1, right) and over the range of SOC acquired *b*-values (between .1 and 2 ms/μm^2), with all contributing compartments at equilibrium in living tissue, and diffusion rates fixed to fundamental values for bulk free D_f, capillary pseudo-diffusion D_p and impeded D_i, the model is simplified to fit three partial volume fraction from voxel DWI signal, S_b, measured as a function of *b*-value:

$$\frac{S_b}{S_0} = \sum_j^{f,p,c} F_j E_j \approx F_f E_f + F_c E_c; E_{c,f} = \exp\left(-bD_{c,f}\right);\ F_c = 1 - F_p - F_f\ ;\ F_{fc} = \frac{D_c}{D_f} \qquad \text{Eq.S2}$$

Note that to satisfy the model constraints and properly quantify $IDF = 1 - F_p - F_f - F_{fc}$, it is essential that at least one DWI is acquired for 2>b>1ms/μm^2. The model fit for three-parameters (F_p, F_f, F_{fc}) in general involves computationally intensive non-linear error minimization with at least four b-value measurements on a voxel basis (e.g., Sup. Figure S1), when spatial mapping is desired.

Linear IDF approximation:

Further model simplification is possible for 2>b >1 ms/μm^2, where bulk free water contribution is negligible, $E_f \rightarrow 0$ (independent of organ), allowing linear IDF model fit with just three acquired *b*-values. Thus, for enhanced contrast-to-noise of dense-tumor IDF ($e.g.$, D_c <1.5μm^2/ms, F_{fc}<0.5), the model predicts optimal acquisition *b*-range between 1<b<2ms/μm^2 that minimizes free and capillary water contribution (E_f, $E_p \rightarrow 0$). For these conditions, IDF model (Eq.S2) can be "linearized" on log-signal scale to derive linear approximation, IDF_L as:

$$\log\left(\frac{S_b}{S_0}\right) = C_1 + C_2 \cdot b;\ \ F_{fc} = -C_2/D_f;\ \ F_c = \exp(C_1);\ IDF_L = \exp(C_1) + C_2/D_f. \qquad \text{Eq.S3}$$

This practical approximation allows for computationally efficient vectorization of whole-volume IDF_L fit. However, similar to mono-exponential DWI ADC model, finite bias is expected for linear IDF_L model fit near the noise floor, as well as depending on utilized (low) b-values. IDF_L fit bias evaluation can be performed empirically by comparing to non-linear fit IDF values of Eq.S2. Additionally, when some clinical protocols do not acquire b=0, but use 0.01<b_0<0.05ms/mm^2 (e.g., to alleviate blood flow pulsatility artifacts), the linear model fit still can be performed using effective *b*-values b_{eff}=b-b_0.

***PCa DWI acquisition protocol bias correction*:**

The derived fractions will generally depend on acquisition TR and TE, as (6):

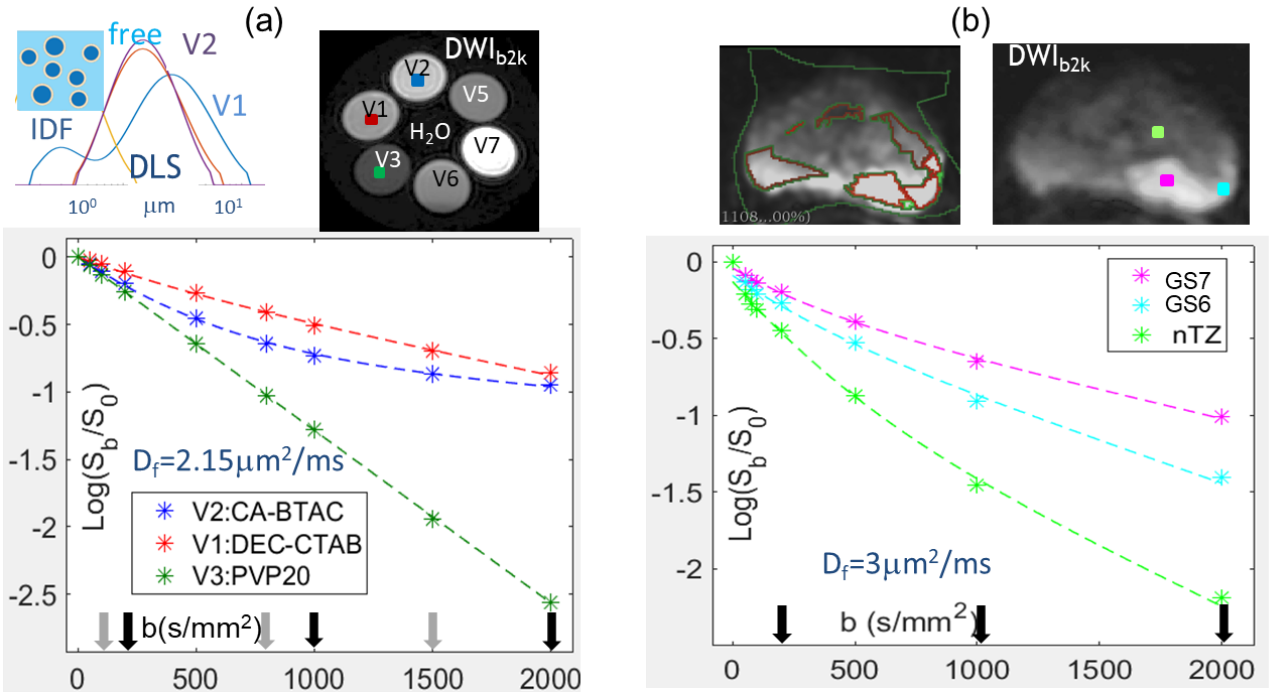
$$F = \frac{S_{0F}(1-E1_F)E2_F}{\sum_F^{f,p,c} S_{0F}(1-E1_F)E2_F}; E1 = \exp\left(-\frac{TR}{T1}\right); E2 = \exp\left(-\frac{TE}{T2}\right) \quad \text{Eq.S4}$$

Where S_{0F} is full (unbiased) proton density MRI signal for the corresponding compartment fraction F= *f, p, c*. The acquisition-induced absolute bias in compartment fractions can be evaluated using compartment-specific magnetization relaxation times T1 and T2 (7,8), and applied acquisition parameters, TR and TE. Consistent measurements (fixed bias) across acquisition protocols may be insured by using standardized TE and TR values. Alternatively, to generalize diagnostic essay thresholds, e.g., for retrospective analysis between two acquisition protocols with different TR and TE, one could adjust for the difference between protocol-dependent IDF (e.g., derived from Eq.S4 and subtracted from Eq.S2 values). Note that bias evaluation is required for fit F_f and F_p, since F_c=1-F_f-F_p.

In this work, IDF protocol bias due to different T2 weighting was calculated by Eq.S4 for the subjects acquired with TE=69ms versus TE=99ms. The true input fractions (F_f, F_p) = (0.1; 0.05); (0.15, 0.07) and (0.2, 0.07) provided the range of observed F_f and F_p. The compartmental T1f=4s, T2f=2s, T1p=1.7s, T2p=0.15s (based on literature values for free water and blood, respectively) were used for analytical simulation according to Eq.S4. The cellular compartment F_c bias (T1c=1.3s; T2c=0.07s) was equal to the sum of F_f and F_p bias. The difference F_f bias between two protocols followed a linear function of 0.154*F_f+0.036, with observed median of positive 7%. The relative F_p bias was 1%, and IDF bias -0.122*IDF+0.144 with median of negative 8%. The corresponding relative bias correction was applied to the F_f, F_p and IDF of three subject scans with TE=69ms. Application of median or linearly dependent bias correction lowered the threshold of maximum GS7 sensitivity to IDF>0.51 not effecting GS6 FPR=0/6. (The corresponding threshold before bias correction was IDF>0.57, FPR=0/6).

References:

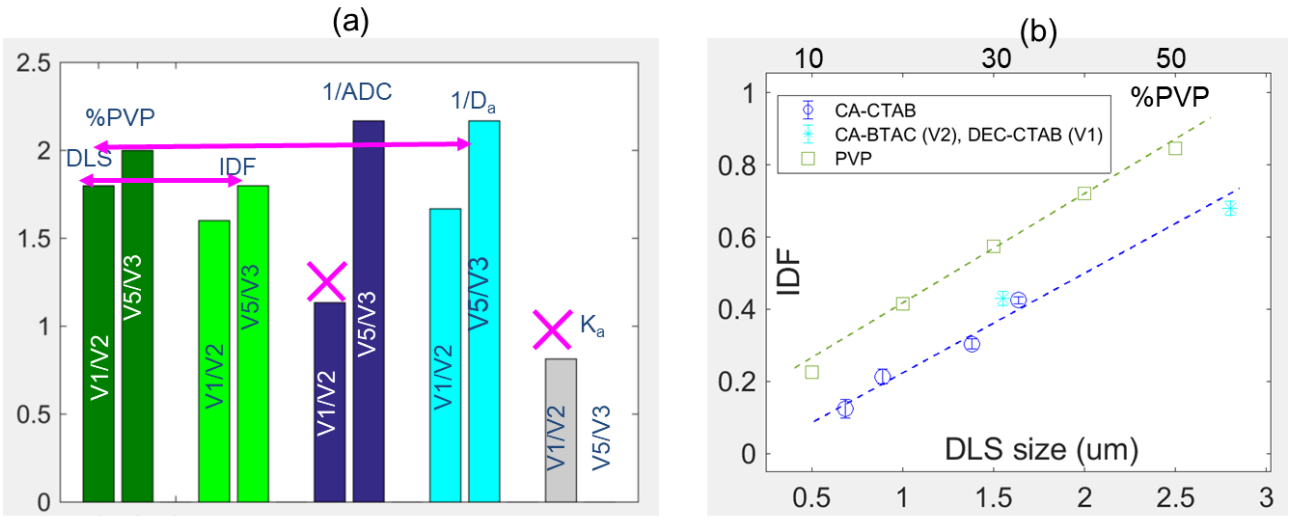
1. Holz M, Heil SR, Sacco A. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate H-1 NMR PFG measurements. Phys Chem Chem Phys. 2000;2(20):4740-2. PubMed PMID: WOS:000089772500041.
2. Pullens P, Bladt P, Sijbers J, Maas AR, Parizel PM. Technical Note: A safe, cheap, and easy-to-use isotropic diffusion MRI phantom for clinical and multicenter studies. Med. Phys. 2017, 44:1063-1070.
3. Polnaszek CF, Bryant RG. Self-Diffusion of Water at the Protein Surface: A Measurement J. Am. Chem. Soc. 1984, 106: 429-30
4. Trovato F, Tozzini V. Diffusion within the Cytoplasm: A Mesoscale Model of Interacting Macromolecules. Biophysical Journal. 2014;107(11):2579-91. doi: 10.1016/j.bpj.2014.09.043. PubMed PMID: WOS:000345859500018
5. Inaba T. Quantitative Measurements of Prostatic Blood Flow and Blood Volume by Positron Emission Tomography. J Urol. 1992;148:1457-60.
6. Gilani N, Malcolm P, Johnson G. A model describing diffusion in prostate cancer. Magnetic Resonance in Medicine. 2017;78(1):316-26. doi: 10.1002/mrm.26340. PubMed PMID: WOS:000403803900032
7. Stanisz GJ, Odrobina EE, Pun J, Escaravage M, Graham SJ, Bronskill MJ, Henkelman RM. T1, T2 relaxation and magnetization transfer in tissue at 3T. Magn Reson Med. 2005;54(3):507-12. Epub 2005/08/09. doi: 10.1002/mrm.20605. PubMed PMID: 16086319
8. Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3 T? Magn Reson Imaging. 2017;35:69-80. Epub 2016/10/28. doi: 10.1016/j.mri.2016.08.021. PubMed PMID: 27594531
9. Hurrell SL, McGarry SD, Kaczmarowski A, Iczkowski KA, Jacobsohn K, Hohenwalter MD, Hall WA, See WA, Banerjee A, Charles DK, Nevalainen MT, Mackinnon AC, LaViolette PS. Optimized b-value selection for the discrimination of prostate cancer grades, including the cribriform pattern, using diffusion weighted imaging. J Med Imaging (Bellingham). 2018;5(1):011004. Epub 2017/11/04. doi: 10.1117/1.JMI.5.1.011004. PubMed PMID: 29098169; PMCID: PMC5658575



Supporting Information Figure S1: Summary of DWI IDF analysis workflow for physical phantom (a), and prostate cancer (PCa, b). Top left figure in (a) illustrates dynamic light scattering (DLS) measurements for V1 and V2 phantom samples with insert showing schematic of their vesicular suspension comprised of free (light blue) and impeded (dark blue) water compartments. Top right image in (a) shows phantom DWI (b=2ms/μm^2) with material vial labels and color-coded voxel locations corresponding to the IDF fit data in the lower pane legend. The chemical sample abbreviations in the legend denote CA: cetearyl alcohol, DEC: decyl alcohol, CTAB: cetyltrimethylammonium bromide, BTAC: behentriammonium chloride). Top images in (b) illustrate an example of PCa lesion annotations from whole-mount histopathology [Hurrel, et al. JMI (2017)], registered to b=0 DWI and traced on DWI (b=2ms/μm^2). Voxel samples for Gleason (GS) 6 and 7 and normal transition zone (nTZ), defined away from lesion boundaries are color-coded for the IDF fit data in the lower pane legend. Dashed curves in lower panes show model-based single-voxel fit for acquired log-DWI data (asterisks) as a function of b-value (fit fractions summarized in Sup. Table 1). Arrows mark b>0 subsets tested for the SOC protocols (versus all-b>0.1ms/μm^2 fit): two for phantom (gray and black arrows) and one used for PCa analysis (black arrows).

Supporting Information Table S1: IDF model nonlinear fit fractions (±2%) for single-voxel DWI of phantom V1-V3 vials, and PCa normal transition zone (nTZ) and Gleason GS6, GS7 lesions in Supporting Information Figure S1.

(%)\Sample	V1	V2	V3	GS7	GS6	nTZ	
F_p	.01	.01	.02	4.0	8.5	12.1	
F_f	15.1	51.2	1.4	19.9	20.3	40.6	
F_{fc}	16.7	6.1	59.4	12.5	18.4	25.1	
IDF	68.2	42.7	39.2	63.6	52.8	22.3	



Supporting Information Figure S2: (a) Bar-plot of relative characteristics for V1/V2 and V5/V3 phantom samples is compared for DLS-derived vesicle sizes and PVP concentrations (dark green) versus different diffusion model parameters: IDF (light green), inverse ADC (dark blue), inverse apparent diffusion (cyan) and apparent kurtosis (gray, no K_a values for mono-exponential V5/V3, K_a~0). Horizontal arrows connect the diffusion parameters with the highest correlation to physical phantom properties. Magenta crosses mark the trends opposite to physical property ratios. (b) Illustrates apparent linear relation between IDF and %PVP (10% to 50%, top axis, green) and DLS vesicle size (bottom axis) for V1: DEC-CTAB, V2: CA-BTAC (cyan) and CA-CTAB (3:1 to 6:1, blue) samples. Dashed lines mark observed linear trends.