A biomimetic tumor tissue phantom for validating diffusion-weighted MRI measurements
McHugh, Damien J; Zhou, Feng-Lei; Wimpenny, Ian; Poologasundarampillai, Gowsihan; Naish, Josephine H; Hubbard Cristinacce, Penny L; Parker, Geoffrey JM

DOI: 10.1002/mrm.27016
License: Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?).
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.
A Biomimetic Tumor Tissue Phantom for Validating Diffusion-Weighted MRI Measurements

Damien J. McHugh 1,2,* Feng-Lei Zhou 2,3† Ian Wimpenny,1,3 Gowsihan Poolgasundarampillai,3,4 Josephine H. Naish,1 Penny L. Hubbard Cristinacce,1 and Geoffrey J. M. Parker 1,2,5

Purpose: To develop a biomimetic tumor tissue phantom which more closely reflects water diffusion in biological tissue than previously used phantoms, and to evaluate the stability of the phantom and its potential as a tool for validating diffusion-weighted (DW) MRI measurements.

Methods: Coaxial-electrospraying was used to generate micron-sized hollow polymer spheres, which mimic cells. The bulk structure was immersed in water, providing a DW-MRI phantom whose apparent diffusion coefficient (ADC) and microstructural properties were evaluated over a period of 10 months. Independent characterization of the phantom's microstructure was performed using scanning electron microscopy (SEM). The repeatability of the construction process was investigated by generating a second phantom, which underwent high resolution synchrotron-CT as well as SEM and MR scans.

Results: ADC values were stable (coefficients of variation (CoVs) < 5%), and varied with diffusion time, with average values of 1.44 ± 0.03 μm²/ms (Δ = 12 ms) and 1.20 ± 0.05 μm²/ms (Δ = 45 ms). Microstructural parameters showed greater variability (CoVs up to 13%), with evidence of bias in sphere size estimates. Similar trends were observed in the second phantom.

Conclusion: A novel biomimetic phantom has been developed and shown to be stable over 10 months. It is envisaged that such phantoms will be used for further investigation of microstructural models relevant to characterizing tumor tissue, and may also find application in evaluating acquisition protocols and comparing DW-MRI-derived biomarkers obtained from different scanners at different sites. Magn Reson Med 000:000–000, 2017. © 2017 The Authors Magnetic Resonance in Medicine published by Wiley Periodicals, Inc. on behalf of International Society for Magnetic Resonance in Medicine. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Key words: tumor microstructure; diffusion MRI; biomimetic phantoms; hollow microspheres; coaxial electrospraying

INTRODUCTION

The use of diffusion-weighted (DW) MRI in oncology is motivated by the potential for inferring clinically useful information related to microstructural properties of tumors from the measured DW signal. Such information typically comes in the form of a biomarker, which may be used for a variety of applications including lesion detection, distinguishing between benign and malignant tissue, and predicting or evaluating response to treatment (1,2).

Depending on the tissue being imaged and the sequence parameters used for acquisition, a range of biomarkers can be derived from DW data, by modeling the signal in different ways. These models can broadly be split into two categories: phenomenological and biophysical (3). Examples of biomarkers from phenomenological models include the apparent diffusion coefficient (ADC) (4–10), diffusional kurtosis (11–16), and the stretched exponential (17,18). Biophysical models attempt to describe the DW signal in terms of specific microstructural tissue properties, potentially yielding biomarkers such as cell size, intracellular volume fraction, and compartment diffusivities. The greater specificity offered by biophysical models in comparison to phenomenological models has motivated extensive research for white matter applications (19–22), and recent applications to tumor tissue (23–26).

If such biomarkers are to become useful clinical tools, it is important that they are subjected to a process of validation to assess both their technical performance, for example their accuracy and precision, and their relationship to biological processes (27). To date, much of the validation of DW-MRI methods in oncology has focused on the technical validation of ADC using free-diffusion phantoms. For example, ice-water phantoms have been used to evaluate the repeatability and reproducibility of ADC values on clinical (28,29) and preclinical (30) scanners, as well as to investigate spatial variations in ADC due to gradient non-linearities (31). ADC stability has also been assessed using gels developed with a range of diffusivities and relaxation times (32).
While useful for ADC investigations, free-diffusion phantoms are not suitable for the validation of other DW-MRI-derived biomarkers, as they lack the cellular-level structure which underlies quantities such as diffusional kurtosis and microstructural parameters. As part of the validation of such biomarkers, it is therefore desirable to have physical phantoms which mimic the cellular structure of tissue, and whose microstructural properties can be controlled and/or characterized. A number of such phantoms have been used for studying diffusion in white matter, including solid fibers (33), hollow silica microcapillaries (34,35), plant tissue (36,37), and electrospun hollow fibers (38), but there is a notable absence of systems applicable to tumor tissue (1).

In addition, free-diffusion phantoms do not capture potentially important ways in which ADC can vary with sequence parameters. For example, free-diffusion phantoms do not exhibit a dependence of ADC on diffusion time, which is a general phenomenon in biological tissue (39) and has been observed in tumor tissue (40,41). As such, although ice-water ADC has been shown to be reproducible across acquisitions with different scan parameters (30), this will not necessarily be the case for tumor ADC.

These considerations motivate the current work, which describes the construction and characterization of a phantom designed as a simple mimic of tumor cellular structure, building on a preliminary report of an earlier phantom design (42). Results from DW-MRI experiments performed to investigate the phantom’s temporal stability are presented, allowing assessment of its potential use as a long-term test object in multi-center studies. Experiments designed to evaluate its potential as a tool for validating microstructural measurements and comparing acquisition protocols are also presented, with DW-MRI characterization compared with independent microstructural measurements (43).

**METHODS**

**Phantom Construction and Characterization**

The phantom consists of a collection of approximately spherical, micron-scale hollow polymer spheres, which mimic cells. The spheres were produced by coaxial electrospraying (44), extending the approach described previously for generating solid spheres (45); a related technique, coaxial electrospinning, has been used to generate hollow fibers for mimicking white matter (38). Coaxial electrospraying was performed using polyethylene glycol (PEG) and poly(β,γ-lactic-co-glycolic acid) (PLGA) solutions for the core and shell of the microspheres, respectively. PEG and PLGA solutions were injected into the inner and outer needles of a coaxial spinneret, with resulting hollow spheres collected on a copper wire. The spheres were collected on a copper wire connected to the ground electrode, generating a bulk sample in approximately 1 h (Fig. 1a). The wire was then removed, leaving the bulk phantom structured as a hollow cylinder approximately 4 cm long, with inner and outer diameters of approximately 1.8 mm and 3 mm.

The phantom was then split into two sections, with one used for MR experiments and the other for characterization with scanning electron microscopy (SEM). The MR sample was placed in a 5 mm NMR tube which was then filled with deionized water (Fig. 1b). At the same time, the SEM sample was also immered in deionized water in a separate NMR tube. In order to assess the effect of prolonged immersion on the phantom’s microstructure, sections of the SEM sample were scanned over a 6-month period (see Supporting Information), quantifying the outer radius of the spheres, $R_o$ (42). The term ‘outer radius’ is used because the SEM measurements reflect the exterior size of the spheres, and do not quantify the non-zero wall thickness.

The repeatability of the construction process was investigated by generating a second phantom, keeping all electrospraying parameters the same. Similar room temperature and relative humidity conditions were used in both cases (22.6°C, 30% and 23.5°C, 28%), as these variables are known to influence the properties of electrospayed fibers and particles (47). Analogous to the methods described above for the first phantom, sections of the second phantom were used for MR experiments and SEM analysis. In addition, high resolution synchrotron-CT (sCT) scans were performed on two sections of the second phantom, with one section immersed in water and one kept dry; full details of the sCT acquisition and analysis is given in the Supporting Information. Briefly, these scans were used to characterize the sphere wall thickness (from manual measurements, Supporting Fig. S1) and the sphere volume fraction (from segmenting the images into ‘sphere’ and ‘non-sphere’ regions, Supporting Fig. S2). The first and second phantoms will be referred to as phantoms A and B, respectively.
MR Acquisition

MR experiments with phantom A were carried out over a period of approximately 10 months, at the following post-immersion time points: ~2, 6, 24, 72 h, 1, 2, 3, 4, 9, 16, 20, 26, and 42 weeks. All time points except the first and last had corresponding SEM analysis. Scans with phantom B were carried out over 1 month, at ~7, 27 h, 1 and 4 weeks post-immersion. All scans were performed on a 7 T horizontal bore magnet (Magnex Scientific Ltd., Abingdon, UK) interfaced to a Bruker Avance III console (Bruker BioSpin, Ettlingen, Germany), with the phantom(s) and a control NMR tube containing only deionized water placed inside a transmit/receive volume coil. Each scan session included pulsed gradient spin-echo (PGSE) acquisitions for evaluating ADC and microstructural parameters. Room temperature was monitored and varied by a maximum of 0.7°C within a given scan session, with a mean ± standard deviation (SD) of 24 ± 1°C over all time points.

For ADC calculations, DW data were acquired with $b = 0, 150, 500, 1000$ s/mm$^2$, $\delta = 4$ ms, $\Delta = 12$ ms ($G = 0, 117.6, 214.7, 303.7$ mT/m) and 45 ms ($G = 0, 58.3, 106.4, 150.5$ mT/m), with TE = 21.3, 54.3 ms, respectively, and TR = 2500 ms. In addition, DW data were also acquired with $G = 0, 70, 140, 210$ mT/m, $\delta = 4$ ms, $\Delta = 12, 23, 45$ ms, with TE = 21.3, 32.3, 54.3 ms, respectively, and TR = 2500 ms: $b$-values were 53.1, 212.5, 487.1, 107.5, 430.0, 967.6, 216.3, 865.1, 1946.5 s/mm$^2$. In a subset of the experiments, the $\Delta = 12$ ms ADC acquisition was repeated at the end of the scan session, to assess short-term (~2 h) repeatability. All imaging data were acquired with a 30 mm × 30 mm field of view, 128 × 128 matrix, and 10 axial slices of 1 mm thickness.

MR Analysis

ADC maps for $\Delta = 12$ and 45 ms were generated using maximum likelihood (ML) fitting (48), with the noise, $\sigma$, estimated from a region of interest (ROI) drawn in the background: $\sigma = S_{bg}\sqrt{2/\pi}$, where $S_{bg}$ is the mean signal intensity in the background ROI (49). Using a single Rician probability density function (PDF) in the objective function was appropriate here as the signals used for ADC fitting were not averaged (50). The phantom material region was obtained using a semi-automated method which first separates the NMR tubes from the background, then thresholds the $b=0$ s/mm$^2$ images to remove high signal voxels corresponding to free water; minor manual adjustment then provided the phantom ROI. Free water ADC values were obtained from the control NMR tube. The signal-to-noise ratio (SNR) was calculated as $S_{bg}/\sigma$, where $S_{bg}$ is the mean $b=0$ s/mm$^2$ signal in the phantom.

For the multi-$G$, multi-$\Delta$ dataset, signals for each $\Delta$ acquisition were normalized to their $G = 0$ mT/m scan, and these normalized signals were analyzed by fitting a two-compartment microstructural model combining restricted diffusion inside a sphere with hindered extraspHERE (analogous to extra-cellular) diffusion, yielding three model parameters: sphere radius, $R$, intra-sphere (analogous to intra-cellular) volume fraction, $f_i$, and free diffusivity, $D$. As the same fluid is inside and outside of the spheres, a single diffusivity was used in the model. The normalized DW-MRI signal, $S/S_0$, is given by

$$S/S_0 = f_i S_i + (1 - f_i) S_v,$$

where

$$S_i = \text{exp} \left( -2\gamma^2 G^2 \sum_{m=1}^{\infty} \frac{1}{\alpha_m^2} \left( \frac{2\alpha_m D}{\alpha_m^2} - 2 \right) + \frac{2e^{-\alpha_m^2 D G^2} - 2e^{-\alpha_m^2 D G^2} - 2e^{-\alpha_m^2 D G^2} - 2}{\alpha_m^4 D^2} \right),$$

$$S_v = \text{exp} \left( -2\gamma^2 G^2 (\Delta - \delta/3) - \frac{D}{1 + f_i^2/2} \right).$$

Equation [2] is the PGSE signal for diffusion restricted within an impermeable sphere, assuming a Gaussian phase distribution: $x_m$ is obtained from the $m$th root of $\alpha_m R^2 \gamma^2 (\sigma_m R) - 1 = J_{1/2}(\sigma_m R) = 0$, where $J_{1/2}$ is the Bessel function of the first kind, order 3/2 (51,52). Equation [3] gives the signal for hindered extracellular diffusion with the diffusivity reduced by a tortuosity factor, $1 + f_i/2$ (53).

Two fitting procedures were performed: first, all model parameters were estimated in the fitting; second, $D$ was fixed to the median ADC (at $\Delta = 12$ ms) measured in the water-only NMR tube, which serves as a ground truth for $D$. For both procedures, fitting was performed both on a voxel-wise basis and using whole-ROI averaged signals (fitting to the mean signal from the entire phantom ROI). In each case, fitting was performed for 100 starting values, taking the final result as the fit with the lowest value of the objective function. As the microstructural model fitting involves averaging and/or normalizing signals, a single Rician PDF no longer characterizes the distribution of the signals and cannot be used in the objective function, making the ML fitting method described above no longer appropriate. Instead, least squares fitting was used, with potential bias mitigated by discarding signals lower than 25$S_{\text{noise}}$, where $S_{\text{noise}}$ is the mean signal in a noise ROI (54). Fitting was performed using a Nelder-Mead simplex algorithm, with parameters constrained to be within plausible biological limits: $0.1 \leq R \leq 25$ μm, $0.1 \leq D \leq 3$ μm$^2$/ms, $0.01 \leq f_i \leq 1$. When fitting to whole-ROI averaged signals, the precision of the model parameters was assessed by bootstrapping the residuals. Due to the SNR differences for different $\Delta$ acquisitions, bootstrapped datasets were generated for a given $\Delta$ using the residuals for that $\Delta$; for example, a residual for a $\Delta = 45$ ms data point would not be added to a $\Delta = 12$ ms data point. One thousand bootstrap samples were generated, and 95% confidence interval (CI) limits were taken as the 2.5% and 97.5% quantiles of the bootstrap distribution (55). The bootstrapping results were also used to investigate correlations between the model parameters. The effect of acquisition protocol on microstructural estimates was assessed by fitting the model to whole-ROI averaged signals using only the data from the $\Delta = 12$ ms and 23 ms acquisitions, that is, excluding all data from the longest diffusion time.

Coefficients of variation (CoVs) were calculated to assess measurement repeatability, and two-sample $t$-tests were used for statistical analyses, with $P < 0.05$ taken to indicate significant differences. All analyses were carried out with MATLAB 2014a (The MathWorks, Inc., Natick, MA).
RESULTS

SEM Characterization of Phantom Microstructure and Stability

SEM images (Fig. 2a, phantom A) show that the spheres tend to group together, indicating that the phantom’s microstructure is not simply a packing of discrete idealized spheres, but consists of extended clumps of spheres. The mean ± SD of the baseline radii was 5.7 ± 0.7 μm. The CoV of the mean post-immersion values was 4.2%, and the maximum difference in means between the first post-immersion time point and subsequent points was 0.46 μm. Averaging over the mean values at each post-immersion time point gave $R_o = 5.2 ± 0.2$ μm, which was taken as the ground truth outer sphere radius for phantom A (Fig. 2b). For phantom B, measurements over 1 month gave a CoV of 4.7% and $R_o = 5.6 ± 0.3$ μm. Comparing $R_o$ values for the two phantoms at equivalent post-immersion time points gave a mean difference of 0.4 μm.

Stability and Time-Dependence of Phantom ADC

Figure 3a shows example DW images and ADC maps for phantom A at the 6-h time point. The lower signal annular region corresponds to the phantom, with the water in the center filling the space left by the wire used to collect the spheres during production. Mean ± SD SNR in phantom A, averaged over all slices and time points, was 27 ± 3 and 17 ± 2 for the first $Δ = 12$ ms and the $Δ = 45$ ms scans, respectively. ADC was consistently higher at the shorter diffusion time, with a mean ± SD of ROI median values over each time point of $1.44 ± 0.03 \, \mu m^2/ ms$ and $1.20 ± 0.05 \, \mu m^2/ ms$ for the first $Δ = 12$ ms and the $Δ = 45$ ms scans, respectively (Fig. 3b). Median ADC values at the two diffusion times were significantly different ($P < 0.001$), with a mean percentage difference of 16%. Such a dependence was not observed in the free water, where the mean ± SD of ROI median values over each time point was $2.01 ± 0.04 \, \mu m^2/ ms$ and $2.01 ± 0.03 \, \mu m^2/ ms$ for the first $Δ = 12$ ms and the $Δ = 45$ ms scans, respectively ($P = 0.62$).

Figure 3b also demonstrates the stability of ADC values over the 10-month period, with CoVs of 2.4% and 4.3% for $Δ = 12$ ms and $45$ ms, respectively. In the seven scan sessions where the $Δ = 12$ ms acquisition was repeated, the mean absolute percentage difference in median ADC values in the phantom was 1.1%, and no significant difference was found ($P = 0.18$). While weeks 9, 16, and 26 showed a trend for a lower ADC at the end of the experiment, both in the phantom and free water, this was not observed consistently.

Phantom B also exhibited stable ADC values, with CoVs of 1.0% and 2.1% for $Δ = 12$ ms and $45$ ms, respectively. ADC values for both diffusion times were lower than in phantom A, with $1.32 ± 0.01 \, \mu m^2/ ms$ and $1.02 ± 0.02 \, \mu m^2/ ms$ for the first $Δ = 12$ ms and the $Δ = 45$ ms scans, respectively.

Application of Phantom for Microstructural Model Evaluation

Figure 4 shows example maps and histograms of $R$, $D$, and $f_o$ at the 1-week time point for phantom A. Large variations in parameter values were observed, indicating that voxel-wise estimates had poor precision. In particular, fits in a number of voxels resulted in values at or near the fit constraints. For example, at the 1-week time point the percentage of voxels with values within 1% of the constraints was 21%, 8%, and 14% for $R$, $D$, and $f_o$, respectively.
respectively. Fixing $D$ had little impact on these percentages (for $R$ and $f_i$), showing that fixing the diffusivity did not improve the precision of parameter estimates (Supporting Fig. S3). The relatively low spread in $R_o$ values from SEM suggests that the variation in $R$ stems from imprecision, rather than reflecting genuine
heterogeneity in the phantom. In general, parameter estimates suffered from poor precision with voxel-wise fitting, an observation which was consistent across time points.

Figure 5a shows example fits using whole-ROI averaged signals, when fitting all parameters; model parameters and the coefficient of determination, $R^2$, are shown in each case. The model fits the data well, with a mean $R^2$ value (over all time points) of 0.9998 for both fitting procedures. At 6 and 72 h, and 42 weeks, the 2SN$_{\text{noise}}$ threshold applied to remove low SNR data points resulted in the highest-$G$, highest-$\Delta$ signal being excluded from the fitting, while it was included at all other time points. Compared with fits where this data point was explicitly excluded, including it for weeks 1–20 tended to slightly increase $R$ estimates and slightly decrease $D$ estimates (mean percentage differences of 5% and −2%, respectively), with negligible effect on $f_i$; at week 26 it had negligible effect on any parameter. For consistency across time points, subsequent analyses focus on fits where the highest-$G$, highest-$\Delta$ signal was always excluded. Mean $R^2$ values for these fits were 0.9998 (fitting all parameters) and 0.9996 (fixing $D$).

Example parameter correlations, obtained from bootstrapping, are shown in Figure 5b, where the bootstrapped parameter values for the 1-week time point are plotted as bivariate histograms. Here, the strongest correlation was between $R$ and $f_i$ ($\rho = 0.70$, Pearson’s correlation coefficient), reflecting the fact that larger cells with higher volume fractions and smaller cells with lower volume fractions can give rise to similar signals. Trends in correlations over time were broadly consistent with those in Figure 5b, except at 6 h and 9 weeks, where positive correlations between $R$ and $D$ were observed ($\rho = 0.14$, 0.11), along with stronger correlations between $D$ and $f_i$ ($\rho = 0.74$, 0.81).

Phantom A’s microstructural estimates are shown in Figure 6. $R$ was consistently overestimated compared with SEM measurements of $R_0$. When fitting all parameters, the mean $\pm$ SD over all time points for $R$ was $8.3 \pm 0.4 \, \mu$m, compared with $5.2 \pm 0.2 \, \mu$m from SEM. Mean $\pm$ SD over all time points for $D$ was $1.91 \pm 0.05 \, \mu$m$^2$/ms, showing that $D$ was consistently underestimated compared with the free water ADC of $2.01 \pm 0.04 \, \mu$m$^2$/ms ($P < 0.001$, comparing median free water ADCs from the first $\Delta = 12$ ms acquisition and $D$ values from whole-ROI averaged fitting). Fixing $D$ resulted in slightly lower $R$ and higher $f_i$ estimates, trends expected based on the correlations shown in Figure 5b. Note that as $D$ was fixed to the free water ADC measured at each time point, the blue data points in Figure 6’s central plot reflect the variation in free water ADC (CoV = 1.8%).

For phantom B, $R = 7.9 \pm 0.3 \, \mu$m, again overestimated compared with $R_0 = 5.6 \pm 0.3 \, \mu$m from SEM. In contrast to phantom A, $D$ for phantom B was consistent with free water ADC; fixing $D$ therefore had less impact on $R$ and $f_i$ than for phantom A. Figure 7 plots microstructural estimates for both phantoms, averaged over 10 months (A) and 1 month (B), for both fitting procedures. In terms of percentage differences, the intracellular volume fraction differs most between the two phantoms, with $f_i$ values significantly higher for B ($P < 0.01$ for both fitting methods).

For both fitting procedures, $f_i$ showed the greatest variability, with the other parameters yielding CoVs of less than 5% (Table 1). Fixing $D$ had the greatest effect on the $f_i$ CoV for phantom A, with a reduction by a factor of 1.3 compared with fitting all parameters, but in general fixing $D$ had little impact on the stability of $R$ or $f_i$.

Repeating the fitting with whole-ROI averaged signals from only the $\Delta = 12$ ms and 23 ms acquisitions resulted
in lower mean $R$ and $f_i$ values for both phantoms when fitting all parameters (decreases $\sim 15$–20\% for $R$ and $f_i$ relative to fitting to the full dataset).

sCT images obtained from sections of phantom B showed a clear difference in contrast between the immersed and dry conditions, with the hollow structure of the spheres evident in the immersed state (Fig. 8). This difference in contrast is hypothesized to be related to the influence of water on the core polymer, PEG, which is water-soluble, and as such is expected to dissolve when

FIG. 6. Microstructural estimates for phantom A (excluding the 24-h and 4-week time points, due to misalignment between acquisitions and severe image artefacts, respectively). Data points are the values obtained from whole-ROI averaged fitting, and error bars represent 95\% CI limits obtained from bootstrapping. Results from the two fitting procedures are shown in different colors; when $D$ is fixed (blue points), $D$ values have no error associated with them from bootstrapping. Solid lines show mean parameter estimates over all time points, and dashed lines show 95\% CI. Note that time points are plotted evenly spaced, as opposed to on an absolute time scale.

FIG. 7. Microstructural estimates for phantoms A and B, for both fitting procedures. Mean $\pm$ SD values are plotted, averaged over 10 months and 1 month for A and B, respectively.
the spheres are immersed. The wall thickness, estimated from manual measurements of 50 spheres in the immersed state (Supporting Fig. S1), was 2.1 ± 0.3 µm. Note that such measurements will tend to overestimate the true wall thickness, as the slices cut through spheres at different angles (56). Although the manual method used for the thickness measurements does not satisfy all criteria for using stereological corrections, applying the \( \frac{\pi}{4} \) factor stated in (56) brings the estimate down to 1.6 µm, which is consistent with the lower end of the measured values (see histogram in Fig. 8), and is likely more representative of the wall thickness. Figure 8 also shows an example of the segmentation obtained from one slice of the dry spheres (also see Supporting Fig. S2). From area fraction measurements from 150 ROIs, the sphere volume fraction was estimated as 0.22 ± 0.05 (Fig. 8).

### DISCUSSION

The low CoV of mean \( R_o \) values in phantom A, and the fact that the maximum difference post-immersion was less than half a micron, suggest that the phantom microstructure shows little variation over 6 months. Although phantom A’s baseline \( R_o \) was higher than all post-immersion time points, a trend not observed in phantom B, the similarity in post-immersion \( R_o \) values for the two phantoms provides evidence of the repeatability of the construction process, in that samples with comparable sphere sizes can be generated. SEM characterization demonstrates the stability of both phantoms, and also shows that the sphere size is appropriate for mimicking tumor cells (25,57).

The observed tendency for the spheres to group together has both negative and positive implications. On the one hand, the aggregation is not ideal in that such a structure is not generally considered in the types of microstructural model the phantom is designed to validate, where tissue is modeled as a collection of individual spherical cells. But on the other hand, the aggregation may better reflect biological tissue, where cell adhesion plays an important role in forming and maintaining tissue structure (58). It is hypothesized that the extent of the aggregation in the bulk phantom

<table>
<thead>
<tr>
<th></th>
<th>( R ) CoV (%)</th>
<th>( D ) CoV (%)</th>
<th>( f ) CoV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit all parameters</td>
<td>4.7 4.2 3.9</td>
<td>2.7 1.9 2.5</td>
<td>13 8.1 13</td>
</tr>
<tr>
<td>Fix ( D )</td>
<td>4.6 4.2 4.2</td>
<td>1.9 1.6 1.6</td>
<td>9.9 7.3 9.9</td>
</tr>
</tbody>
</table>

**FIG. 8.** sCT images of phantom B. Example images are shown for immersed and dry spheres, along with an example segmentation of a dry sphere image. The wall thickness histogram (from measurements of 50 spheres) and area fraction estimates (from measurements on 50 slices, with 3 ROIs per slice) are shown on the bottom row.
depends on the electrospaying parameters, and further work is required to assess the degree to which it can be controlled. While the size and aggregation of the spheres show similarity to tumor tissue, it should be noted that the current phantom clearly oversimplifies the tumor microenvironment. For example, structures such as cell nuclei, collagen fibers, and blood vessels are not mimicked, nor are different cell populations, such as tumor and immune cells. While future work could increase the complexity of the phantom, the current phantom still has utility in investigating microstructural models which are themselves simplifications and do not attempt to capture all aspects of tumor microstructure. By controlling the phantom’s ground truth microstructure and investigating the resulting DW-MRI parameters, insight may be gained into the sensitivity of DW-MRI to different aspects of tissue structure, with the phantoms potentially aiding the development of microstructural models.

The observation of a higher ADC at a lower diffusion time provides evidence of hindered and/or restricted diffusion in the phantom, consistent with the dependence of ADC on diffusion time that has previously been observed in biological tissue in general and in tumor tissue specifically. Also, the absence of such a dependence in freely diffusing water is due to the lack of structures to impede diffusion, as is the case with the free-diffusion phantoms previously used for DW-MRI validation (e.g., ice-water and gels). These findings, along with the observation that the multi-G, multi-Δ data are well described by a microstructural model, demonstrate that this phantom more closely reflects diffusion in tissue than previous phantoms, and emphasize the need to consider diffusion times, and not only b-values, when comparing ADC values across studies. For example, if different studies use different diffusion times but the same b-values, in vivo ADC values may differ while free-diffusion phantoms may give comparable values. As such, it should be emphasized that b-values alone do not fully characterize a scan when measuring diffusion in systems such as biological tissue (59), and this should be considered when discussing standardizing acquisitions (1,2). Using the biomimetic phantom in multi-center studies could therefore enable a more comprehensive validation of DW-MRI, enabling comparison of ADC time-dependence, and microstructural estimates, between scanners. This would be another way of addressing the recently noted need for ‘more data on interplatform reproducibility’ (2), in addition to studies utilizing free-diffusion phantoms and healthy volunteers (60,61). The fact that this is a water-based phantom is advantageous for multi-center studies, as the phantoms could be distributed dry and immersed in water on-site. This is more practical than using phantoms immersed in organic solvents (38,42), which require the use of protective clothing and fume cupboards. Although the structural robustness of the phantoms has not been fully evaluated, initial experience suggests that they can be successfully transported between sites without obvious degradation. The small size of the bulk phantom is a limitation in terms of its use on clinical scanners, making the current phantom better suited to preclinical scanners. Developing larger phantoms suitable for clinical scanners is a focus of ongoing work.

The CoVs of less than 5% indicate that ADC repeatability in the phantom is very good, and provide evidence that the phantom remains stable over 10 months. As PLGA is known to degrade over time, with many factors influencing the degradation process (62,63), it is expected that the phantom will eventually become unstable, potentially limiting the extent to which it can be used as a long-term test object. This study suggests that the phantom can be used for at least 10 months, with further longitudinal analysis required to track longer-term stability.

While SEM characterization showed that phantoms A and B had comparable sphere sizes, providing evidence of the repeatability of the construction process, the ADC differences suggest that the underlying microstructure does vary between the two phantoms. The microstructural modelling suggests that phantom B has a higher fi than phantom A, which is consistent with the observed lower ADC in phantom B, as ADC is expected to decrease as intracellular volume fraction increases (64,65). This suggests that phantom A’s sphere volume fraction is lower than 0.22, the sCT-derived volume fraction for phantom B, although the lack of a ground truth volume fraction for phantom A precludes a full validation of this finding. Note that as the sphere volume fraction obtained from the segmented sCT images includes both the sphere wall and the hollow interior, it is expected to be larger than the MR-derived fi, which is taken to reflect only the hollow interior. As such, there is a clear overestimate from the microstructural model, which yields fi = 0.37 ± 0.03 for phantom B.

While the similarity of MR-estimated radii for the two phantoms is consistent with the similarity observed on SEM, the difference in absolute values between the two modalities indicates the microstructural model also overestimates the sphere size. As R is expected to reflect the inner radius, as opposed to the outer radius seen with SEM, the overestimation is even greater when considering the relatively thick walls observed with sCT. Microstructural modelling in white matter has revealed a tendency to overestimate axonal radii (21), with recent work suggesting that this trend may be driven by the unmodeled influence of time-dependent extracellular diffusion (66). However, the time-dependence of extra-cellular diffusion is expected to be weaker in three-dimensional geometries, such as the sphere packings considered in the present work, than in the two-dimensional geometries relevant to axonal packings (25). As such, neglecting such time-dependence may have less of an impact on compartment size estimates in three-dimensions, and the present work’s use of a time-independent extracellular diffusivity is consistent with that used in previous approaches to tumor microstructural modelling (23,25).

Exchange of water across the sphere wall is another unmodeled effect that may contribute to bias in R and fi (67). As longer diffusion times are expected to increase sensitivity to exchange, it may be hypothesized that, depending on the rate of exchange, excluding the longest diffusion time data from the fitting would reduce such
sensitivity and therefore improve the accuracy of \( R \). Excluding the \( \Delta = 45 \) ms data did reduce \( R \), bringing the estimate closer to the SEM \( R_o \), though a bias still remains, suggesting that sensitivity to exchange may still be present at \( \Delta = 23 \) ms. As faster exchange is expected to lead to a greater underestimate in \( f_i \) if not accounted for (67), it may be hypothesized that \( f_i \) estimates would increase when using an acquisition less sensitive to exchange, such as excluding longer diffusion times; however, the opposite trend was observed here, suggesting that exchange is not the dominant effect on \( f_i \) estimates. Moreover, the sCT data suggests that \( f_i \) is overestimated, not underestimated, from the microstructural model. Again, further work, such as a filter exchange imaging (FEXI) experiment (68), is needed to characterize the permeability of the spheres in order to assess these effects, and, more generally, the dependence of microstructural estimates on acquisition parameters warrants further investigation. Another factor which may influence the microstructural parameters is the presence of large pores in the extracellular space, due to the extended clumping of the spheres. Depending on the size of these regions in relation to the diffusivity and diffusion time, water here may appear restricted at longer diffusion times, therefore decreasing when excluding the longest diffusion time, as diffusion in these regions may appear free, again, depending on their size.

Compartmental differences in \( T_2 \) may also contribute to the observed bias in \( R \) and \( f_i \), with simulations (not shown) indicating that if \( T_2 \) is higher inside than outside the spheres, all model parameters are overestimated, with the magnitude of the bias increasing as the \( T_2 \) difference increases. Such an effect is consistent with the experimentally observed overestimation of \( R \) and \( f_i \), although an overestimation of \( D \) was not seen. A combination of compartmental differences in \( T_2 \) and permeability may influence the model parameters, and further work is needed to understand potential surface effects as water interacts with the spheres’ inner and outer shells.

A limitation of the comparisons between modalities is that different sections of the bulk phantom have been used for sCT, SEM and MR, and further work is required to characterize the homogeneity of the microstructure within a given sample as well as between different samples. Such developments would enhance the utility of the phantoms as tools for validating DW-MRI microstructural measurements.

**CONCLUSIONS**

A novel biomimetic tumor tissue phantom has been developed and shown to be stable over a period of 10 months. The phantom exhibits time-dependent diffusion and signals are well described by a microstructural model, indicating that the phantom more closely reflects diffusion in tissue than previously used free-diffusion phantoms. Microstructural estimates were found to be more variable than ADC measurements, with evidence of bias in \( R \) and \( f_i \). It is envisaged that such phantoms will be used for further investigation of microstructural models relevant to characterizing tumor tissue, and may also find application in evaluating acquisition protocols and comparing DW-MRI-derived biomarkers obtained from different scanners at different sites.

**ACKNOWLEDGMENTS**

This is a contribution from the Cancer Imaging Centre in Cambridge & Manchester, which is funded by the EPSRC and Cancer Research UK. The authors thank the Diamond Light Source and Diamond-Manchester Collaboration for beamtime MT15507 on the I13 Diamond-Manchester Branchline. We also acknowledge Dr Shashidhara Maratha, Miss Jekaterina Maksimcuka, Mr Cian Vyasar and Dr Gil Costa Machado for assistance with data collection. This work was made possible by the facilities and support provided by the Diamond-Manchester Collaboration and the Research Complex at Harwell. The authors thank Dr Ben Dickie for many useful discussions. Data and analysis code will be made available via www.qbi-lab.org/software.

**REFERENCES**


SUPPORTING INFORMATION
Additional Supporting Information may be found in the online version of this article.

Fig S1. Wall thickness measurement for phantom B. Four measurements were made for each sphere, with positions guided by perpendicular lines through the sphere centre, as well as by comparing overlays on original (left) and mean filtered (right) images.

Fig S2. sCT segmentation for estimating phantom B’s sphere volume fraction. On a single cropped slice, manual labelling of ‘sphere’ (green) and ‘non-sphere’ (red) regions was performed, and used to train a classifier to segment the entire image. The classifier was then applied to 49 other slices, providing segmentations from which regional area fractions, \(a_f\), were obtained.

Fig S3. Voxel-wise microstructural estimates, for phantom A. (a) Example maps of each model parameter are shown (columns) for two fitting procedures (rows). Images are a representative slice from the 1 week time point. (b) Histograms of each model parameter (columns) for two fitting procedures (rows). Quoted values are the median ± IQR, for all phantom voxels (black) and, where applicable, excluding voxels where at least one parameter value was within 1% of the fit constraints (red). The first and third row form Figure 4 in the main text.