

MAGORINO: Magnitude-only fat fraction and R_2^* estimation with Rician noise modeling

Timothy J. P. Bray^{1,2}  | Alan Bainbridge^{1,3} | Emma Lim⁴  | Margaret A. Hall-Craggs^{1,3}  | Hui Zhang⁵ 

¹Centre for Medical Imaging, University College London, London, United Kingdom

²Department of Imaging, University College London Hospital, London, United Kingdom

³Department of Medical Physics, University College London Hospitals, London, United Kingdom

⁴Department of Imaging, Imperial College Healthcare NHS Trust, London, United Kingdom

⁵Department of Computer Science and Centre for Medical Image Computing, University College London, London, United Kingdom

Correspondence

Timothy J.P. Bray, Centre for Medical Imaging, University College London, 2nd Floor Charles Bell House, 43-45 Foley Street, W1W 7TS, London, United Kingdom.

Email: t.bray@ucl.ac.uk

Funding information

Research Trainees Coordinating Centre, Grant/Award Number: CL-2019-18-001; National Institute for Health Research; NIHR; UCLH Biomedical Research Centre

Purpose: Magnitude-based fitting of chemical shift-encoded data enables proton density fat fraction (PDFF) and R_2^* estimation where complex-based methods fail or when phase data are inaccessible or unreliable. However, traditional magnitude-based fitting algorithms do not account for Rician noise, creating a source of bias. To address these issues, we propose an algorithm for magnitude-only PDFF and R_2^* estimation with Rician noise modeling (MAGORINO).

Methods: Simulations of multi-echo gradient-echo signal intensities are used to investigate the performance and behavior of MAGORINO over the space of clinically plausible PDFF, R_2^* , and SNR values. Fitting performance is assessed through detailed simulation, including likelihood function visualization, and in a multisite, multivendor, and multi-field-strength phantom data set and in vivo.

Results: Simulations show that Rician noise-based magnitude fitting outperforms existing Gaussian noise-based fitting and reveals two key mechanisms underpinning the observed improvement. First, the likelihood functions exhibit two local optima; Rician noise modeling increases the chance that the global optimum corresponds to the ground truth. Second, when the global optimum corresponds to ground truth for both noise models, the optimum from Rician noise modeling is closer to ground truth. Multisite phantom experiments show good agreement of MAGORINO PDFF with reference values, and in vivo experiments replicate the performance benefits observed in simulation.

Conclusion: The MAGORINO algorithm reduces Rician noise-related bias in PDFF and R_2^* estimation, thus addressing a key limitation of existing magnitude-only fitting methods. Our results offer insight into the importance of the noise model for selecting the correct optimum when multiple plausible optima exist.

KEYWORDS

computer-assisted imaging processing, magnetic resonance imaging, radiology

1 | INTRODUCTION

In recent years, chemical shift-encoded MRI (CSE-MRI) has emerged as the leading method for quantifying proton density fat fraction (PDFF), an accurate and “confounder-corrected” biomarker of tissue fat content.^{1–4} PDFF measurements are now established for the assessment of hepatic steatosis^{5–7} and are used increasingly for other applications including the pancreas, muscle, and bone marrow.^{8–13} Chemical shift-encoded MRI typically makes use of gradient echo-based acquisitions, which have the advantage that R_2^* measurements can be extracted simultaneously, enabling quantification of iron or calcium.^{12,14–16} Gradient echo-based CSE-MRI is therefore a flexible way to quantify a variety of biologic and pathological processes.

Chemical shift-encoded MRI relies on the fact that fat and water have different resonant frequencies, and therefore develop time-dependent phase differences in an MRI experiment. These differences in phase can be exploited using complex-based fitting methods, which utilize the phase of the MRI signal to separate the fat and water signals.^{1–3,17–21} Complex fitting also offers an accurate method to measure R_2^* ,¹⁴ and avoids the need for procedures such as baseline fitting, in which an additional parameter is introduced into the model to capture the noise floor,^{22,23} or truncation, whereby data from longer TEs are discarded,^{24,25} which have otherwise been used to avoid noise-related bias at high R_2^* values. However, complex signal-based fitting methods may suffer from inaccuracies in field-map estimation and phase errors, may fail in areas of large B_0 inhomogeneity, and require phase data to be accessible and reliable. Although this may be realistic in a research setting, it can be challenging in multicenter studies and in standard care, where sites may not all have access to expensive research agreements or dedicated software packages (this may particularly apply in countries or regions with less funding). This limits the feasibility and increases the cost of using these measurements in clinical trials and standard care.

An alternative, potentially simpler approach is magnitude-based fitting,^{26–28} whereby the phase is discarded. However, existing signal magnitude-based fitting fails to consider the Rician distribution of noise, leading to inaccuracies in the estimation of both PDFF and R_2^* .

To address the vulnerability of magnitude-based fitting to Rician noise, we propose MAGORINO (magnitude-only fat fraction and R_2^* estimation with Rician noise modeling), a new fitting algorithm combining (1) Rician noise-based likelihood optimization based on the signal magnitude and (2) two-point search of the likelihood function to ensure that both global and local optima are explored.

2 | THEORY

2.1 | Tissue and noise models

With a gradient echo-based CSE-MRI acquisition, assuming that the fat and water signals have equal phase at $t = 0$ (a reasonable assumption for multi-echo gradient-echo sequences), the noise-free complex signal S acquired at TE t can be modeled as

$$S(t | \rho_W, \rho_f, R_2^*, f_B) = \left(\rho_W + \rho_f \sum_{m=1}^M r_m \exp(i2\pi f_{F,m} t) \right) \exp(i2\pi f_B t) \exp(-tR_2^*), \quad (1)$$

where ρ_W and ρ_f are the amplitudes of water and fat components; $f_{F,m}$ is the frequency of each spectral fat component; r_m is the relative amplitude of each spectral fat component; M is the total number of fat components; f_B is the frequency offset due to B_0 inhomogeneity; and $R_2^* = 1/T_2^*$ (s^{-1}) is an unknown relaxation constant. The standard approach is to assume that the relative amplitudes and frequency offsets of each fat component are known a priori; therefore, the unknown parameters are ρ_W , ρ_f , f_B , and R_2^* .

With the addition of complex Gaussian noise, the measured signal S' is modeled as

$$S'(t | \rho_W, \rho_f, R_2^*, f_B, \sigma^2) = S(t | \rho_W, \rho_f, R_2^*, f_B) + N(0, \sigma^2) + iN(0, \sigma^2), \quad (2)$$

where $N(0, \sigma^2)$ is Gaussian noise present in both real and imaginary channels, and σ^2 is the noise variance.

For a single measurement, the log likelihood for the measured signal is given by

$$\log L(S', S | \sigma^2) = -2 \log \left(\sqrt{2\pi\sigma^2} \right) - \frac{|S' - S|^2}{2\sigma^2}. \quad (3)$$

For a set of measured signals, the log likelihood becomes

$$\log L(\{S'_i\}, \{S_i\} | \sigma^2) = -2n \log \left(\sqrt{2\pi\sigma^2} \right) - \sum_{i=1}^n \frac{|S'_i - S_i|^2}{2\sigma^2}, \quad (4)$$

where $\{S'_i\}$ is the set of measured signals; $\{S_i\}$ is the corresponding set of predicted signals based on the parameter estimates; and n is the number of measurements (double the number of TEs for complex data, or the number of TEs for magnitude data).

The second term is the sum of squared errors (SSE) divided by $2\sigma^2$. The maximum value for Equation (4)

therefore corresponds to the minimum SSE value, meaning that the maximum likelihood estimate can be obtained minimizing the sum of squared errors (SSE), which is the widely used nonlinear least-squares approach.

Having estimated ρ_F and ρ_W , the PDFF is calculated as

$$PDFF = \frac{\rho_F}{\rho_F + \rho_W}. \quad (5)$$

For the signal magnitude, the noise-free signal in Equation (1) becomes

$$M(t | \rho_W, \rho_f, R_2^*) = \left| \rho_W + \rho_f \sum_{m=1}^M r_m \exp(i2\pi f_{F,m} t) \right| \exp(-tR_2^*), \quad (6)$$

with only three unknown parameters (ρ_W , ρ_f , and R_2^*).

The corresponding noisy signal becomes

$$M'(t | \rho_W, \rho_f, R_2^*, \sigma^2) = M(t | \rho_W, \rho_f, R_2^*) + N_R(0, \sigma^2), \quad (7)$$

where $N_R(0, \sigma^2)$ denotes Rician noise, and 0 and σ^2 denote the mean and variance of the underlying complex Gaussian distribution. Importantly, this distribution has a nonzero mean that depends on σ^2 .

Under Rician noise, the assumption of nonlinear least squares that the minimum SSE corresponds to the maximum likelihood no longer holds, and parameter estimates obtained in this way are subject to a noise-related bias arising from the nonzero mean of Rician noise.

To address this problem, parameter estimates can be obtained directly by maximizing the log likelihood:

$$\log L(\{M'_i\}, \{M_i\} | \sigma^2) = \sum_{i=1}^n \left[\log \frac{M'_i}{\sigma^2} - \frac{M_i'^2 + M_i^2}{2\sigma^2} \right. \\ \left. \times \log I_0\left(\frac{M_i' + M_i}{\sigma^2}\right) \right], \quad (8)$$

where $\{M'_i\}$ is the set of measured magnitude signals at different TEs; and $\{M_i\}$ is the corresponding set of predicted magnitude signals, where I_0 is the zeroth-order modified Bessel function of the first kind.

Note that this optimization needs to be performed with sigma estimated a priori;^{29–35} although it is possible to estimate sigma as a floating parameter in a “one-step” fitting process, this leads to underestimation in low SNR and/or high R_2^* voxels. This underestimation arises due to overfitting, as the optimizer can increase the likelihood by simultaneously reducing sigma and minimizing SSE. Underestimation of sigma effectively increases the SNR estimate toward the Gaussian regime and reduces the effect of the Rician noise model.

A priori sigma estimation can be achieved in multiple ways, including by acquiring multiple realizations of the same image or from the signal intensities within a homogenous region of interest (ROI).^{29–35} The latter approach has the advantage that an additional acquisition is not required but has the disadvantage that it can be susceptible to spatial inhomogeneity in the ROI (also note that the previous approach of estimating sigma from background is not commonly used now, as it can be inaccurate with parallel imaging).

To address this susceptibility to inhomogeneity in the ROI, we propose an extension to the ROI-based method, in which fitting with floating sigma is performed in the individual voxels within an ROI with good SNR (ie, in the Gaussian regime, where sigma can be estimated accurately), before performing fitting with fixed sigma for the whole image. In vivo, muscle provides an appropriate tissue, providing good SNR, low fat fraction and low R_2^* , enabling sigma estimation. Simulations show that this approach provides similar sigma estimates to the conventional ROI-based approach but is less vulnerable to spatial inhomogeneity (Supporting Information Figure S1).

2.2 | Dual optima problem and two-point search method

With magnitude-based fitting, the likelihood function has two optima: one “true” solution corresponding closely to the ground truth and one incorrect “swapped” solution with a PDFF value at the opposite end of the range (Figure 1). To ensure that both optima are explored, Triay Bagur et al developed a two-point search method, in which the fitting is initialized twice: once assuming a pure-water voxel and once assuming a pure-fat voxel, where the solution with the lower error is taken as the output for the fit.²⁸ However, this approach can fail in the presence of noise, because the “true” solution is not always globally optimal (Figure 2). Here, to increase the chance that the true solution does correspond to the global optimum, we combine this two-point search with optimization of the Rician likelihood function in Equation (8). It should be noted that the two-point search method refers to exploration of the likelihood function and is distinct from the traditional use of “two-point Dixon,” which refers to the number of TEs.

2.3 | Effect of parallel imaging on noise distribution

The choice of coil combination method can affect the noise distribution. If SENSE reconstruction is used, the noise

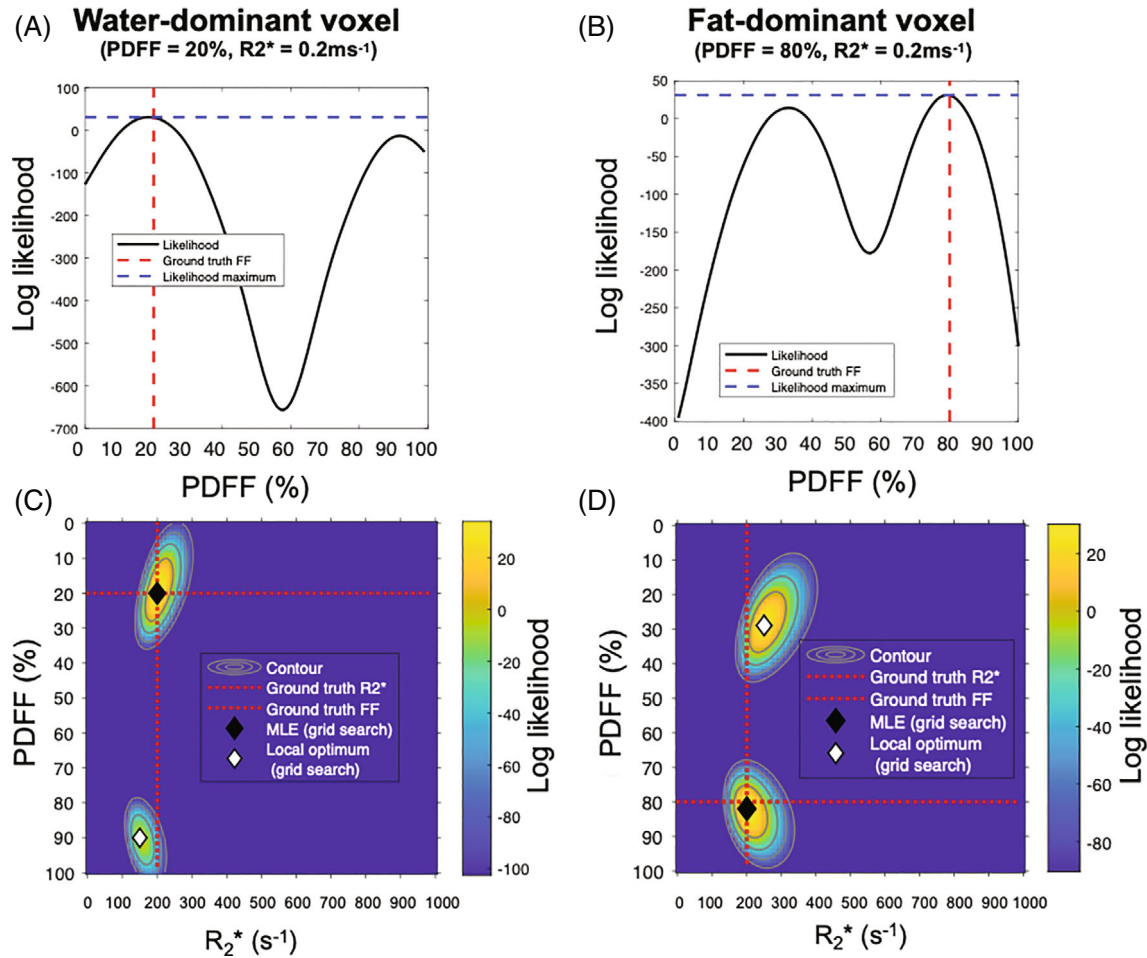


FIGURE 1 Conceptual illustration of magnitude-only resolution of fat–water ambiguity. The Gaussian likelihood of proton density fat fraction (PDFF) estimates are shown for water-dominant voxels (PDFF = 20%) (A, C) and fat-dominant voxels (PDFF = 80%) (B, D). A, B, One-dimensional likelihood plots. C, D, Two-dimensional likelihood plots. The 2D plot shows the color-coded likelihood over the clinically feasible space of possible PDFF and R_2^* values for a given pair of ground-truth parameter measurements. On the one-dimensional (1D) plots, ground-truth estimates and the maximum likelihood are shown as red dotted lines and blue dotted lines, respectively. On the 2D plots, the ground truth, maximum likelihood estimates from 2D PDFF/ R_2^* grid search (MLE grid search), and local likelihood optima (ie, the likelihood peak that is not globally optimum) are shown as black and white diamonds, respectively. For both water-dominant and fat-dominant voxels, there are two likelihood maxima occurring at low PDFF and high PDFF; in each case, the true (nonswapped) maximum can be identified on the basis that it has higher likelihood than the swapped solution, and thus the chosen solutions from the fitting correspond closely to the ground truth

properties are equivalent to those of a spatial matched filter reconstruction (which is the optimal coil combination with a maximized SNR of the resulting image), and a Rician distribution is expected in the magnitude data.³⁶ If a sum-of-squares reconstruction is used, the distribution is noncentral chi.³⁶ In this case, the assumption of Rician noise is violated but remains a more appropriate approximation than the assumption of Gaussian noise. We suggest that any fitting method should be implemented with knowledge of the reconstruction method. However, an appropriate spatial matched filter-based or SENSE reconstruction should be available on most modern scanners.

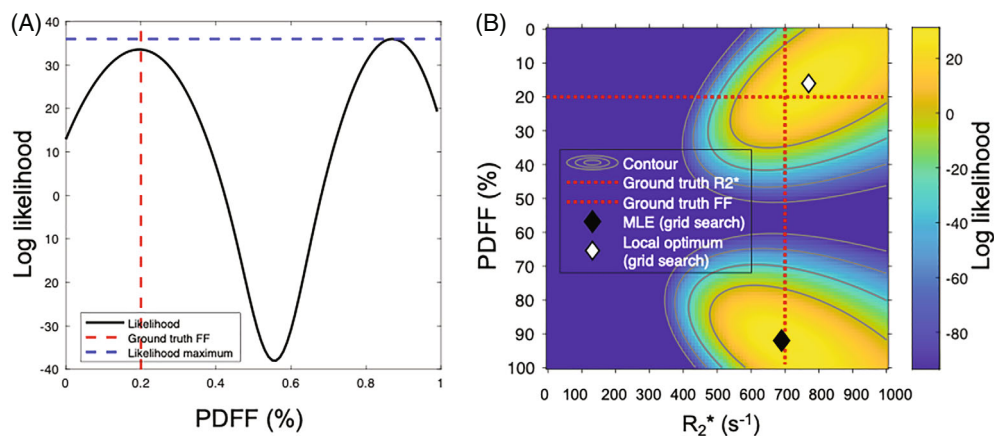
3 | METHODS

3.1 | Simulation experiments

3.1.1 | Fitting implementation

We implement and compare three fitting algorithms: (1) magnitude fitting with Gaussian noise model, which is an equivalent implementation of the “MAGO” algorithm described by Triay Bagur et al²⁸; (2) magnitude fitting with Rician noise model; and (3) complex fitting implementation including estimation of f_B . Each method is implemented twice using two different start points (ie, using

FIGURE 2 Conceptual illustration of failure in magnitude-only resolution of fat-water ambiguity in the presence of noise. A, B, The Gaussian likelihood of fat-fraction estimates for a truly water-dominant voxel (PDFF = 20%) is shown in 1D (A) and in 2D (B). In this case, the swapped (incorrect) PDFF maximum has a higher likelihood than the true (nonswapped) solution, leading the algorithm to select the wrong solution. In the 1D plot, this manifests as a reversal in the size of the two likelihood peaks, whereas in the 2D plot it manifests as a swap in the positions of the MLE and local optimum (ie, the black and white diamonds have swapped position compared with Figure 1C)



a “two-point search” method); fitting is therefore run 6 times in total for each voxel. The initial values of ρ_w (for water-dominant initialization) and ρ_f (for fat-dominant initialization) are set to the maximum signal magnitude from the multi-echo data, $\max_t |S_t|$, multiplied by a constant C . The constant C compensates for reduction in the signal magnitude due to R_2^* decay and chemical shift, and avoids the need for empirical manual adjustment of initial values depending on scanner gain, as performed in Triay Bagur et al.²⁸ Here, we use $C = \exp(t_{Smax} R_{2init}^*)$, where t_{Smax} is the TE corresponding to the maximum signal magnitude, and R_{2init}^* is the initialization R_2^* value. Specifically, for water-dominant initialization, the initial values are $\{\rho_w, \rho_f, R_2^*\} = \{C * \max_t |S_t|, 0.001, 0.1 \text{ms}^{-1}\}$; and for fat-dominant initialization, the initial values are $\{\rho_w, \rho_f, R_2^*\} = \{0.001, C * \max_t |S_t|, 0.1 \text{ms}^{-1}\}$. Each of these parameters is assigned a lower bound of 0, and R_2^* is assigned an upper bound of 2ms^{-1} . For complex fitting during implementation, f_B is correctly initialized at $f_B = 0$ and is not constrained with either upper or lower bounds. All fitting is performed by maximization of likelihood functions (equivalent to minimization of error functions under Gaussian noise); this approach ensures consistency across noise models (the error function is not defined for the Rician case). For each of the three methods, the solution providing the highest likelihood is chosen as the fit output.

The frequency offsets and relative amplitudes for the multippeak fat spectrum are matched to those used in Triay Bagur et al.²⁸ and Hernando et al.³⁷ (ie, frequency shifts relative to the water peaks of -3.90 , -3.50 , -2.70 , -2.04 , -0.49 , and $+0.50$ ppm and relative amplitudes of 0.087 , 0.694 , 0.128 , 0.004 , 0.039 , and 0.048).

To mimic the proposed extension for sigma estimation (detailed in section 2.1), we first implement Rician fitting with σ included as floating parameter in the model (maximization of Equation [8]) in 100 simulated pure-water voxels with low R_2^* (which have good SNR for each TE, meaning that sigma can be estimated accurately) before fitting the full set of simulated voxels with the sigma estimated from the first step fixed. The choice of 100 voxels approximately matches the size of a typical ROI in vivo, and the use of pure-water, low R_2^* voxels mimics the properties of muscle, which we use to estimate sigma in vivo. To compensate for slight overfitting (and underestimation of sigma), we use a simulation-derived correction factor $k = 1.163$. Details on the derivation of k are found in Supporting Information Figure S1. Subsequent fitting of the whole image is performed with sigma fixed to the estimate from the first step.

All fitting is performed in *MATLAB* 2020a (MathWorks, Natick, MA) using the *fmincon* minimizer with an interior point algorithm on an Apple iMac with 3.8-GHz 8-Core Intel i7 processor. The processing time for both

MAGORINO and MAGO was approximately 0.01 s per voxel, and approximately 20 min for a 320×320 image slice.

3.1.2 | Experiment design

To determine the effect of varying PDFF and R_2^* on parameter estimation, simulations were performed across a dense grid of PDFF and R_2^* combinations, with PDFF values between 0% and 100% (at 2% intervals) and R_2^* values between 0 and 1 ms^{-1} (at 0.1 ms^{-1} intervals). For each PDFF/ R_2^* combination, 1000 noise-free signals were simulated using Equation (1) and sampled at TEs corresponding to a typical in vivo protocol at 3 T using the shortest available TEs ($TE_1 = 1.1 \text{ ms}$ and $\Delta TE = 1.1 \text{ ms}$).²⁸ Gaussian noise was added to the noise-free signals in both the real and imaginary channels according to the SNR. The simulations were performed at “typical SNR” for a 3T protocol in vivo (SNR = 60)²⁸ and at “low SNR” (SNR = 20). Two-point Gaussian magnitude fitting (MAGO), Rician magnitude fitting (MAGORINO), and complex fitting were then applied to the noisy signals to obtain PDFF and R_2^* estimates.

Algorithm performance was assessed in three domains: (1) parameter error, specifically the mean error on PDFF, R_2^* , and S_0 estimates, where $S_0 = \rho_f + \rho_w$; (2) parameter SD for PDFF, R_2^* , and S_0 ; and (3) fitting error, assessed in terms of (i) the sum-of-squared error (SSE), (ii) the sum-of-squared error relative to the noiseless signal generated directly from the ground truth parameter values in the simulation, referred to here as the “true SSE”, and (iii) the estimated SSE of the noise compared with the true noise SSE. Note that (ii) and (iii) inform on the degree of overfitting, which results in an increase in true SSE and a reduction in the estimated noise. Note also that high performance in PDFF estimation should produce both low parameter error and low SD, and that in some cases consistently poor performance (consistent fat–water swaps) can produce low SD.

3.1.3 | Interrogation of specific PDFF/ R_2^* combinations

To gain further insights into the differences in behavior between Gaussian and Rician fitting, we selected specific PDFF/ R_2^* combinations showing larger error for more detailed analysis. Specifically, for the chosen PDFF/ R_2^* combinations, the parameter estimates from fat-dominant and water-dominant initializations were displayed for each simulation instantiation using (1) fit success histograms and (2) likelihood difference plots. The fit success histograms show the frequency of parameter estimates

over all simulation instantiations, displayed on a histogram relative to the ground truth. The likelihood difference plots are scatterplots in which the parameter estimates are plotted against the difference in likelihood for the fat-dominant and water-dominant initializations. Assuming that the initialization functions correctly (and both fat-dominant and water-dominant optima are obtained from the fits), there are two main possibilities that can be captured by the likelihood difference plot: (1) The water-dominant solution is more likely than the fat-dominant solution, resulting in a positive likelihood difference for a low PDFF estimate; or (2) the fat-dominant solution is more likely than the fat-dominant solution, resulting in a negative likelihood difference for a high PDFF estimate. Additionally, there are several further possibilities that can occur if the optimization unexpectedly finds the opposite optimum to its initialization (ie, a fat-dominant optimization finds a water-dominant solution, or vice versa). These are (3) both initializations find the same local optimum, resulting a likelihood difference of 0; or (4) both initializations find the wrong local optimum, resulting in a “reversed” likelihood difference such that a water-dominant solution has negative likelihood difference or a fat-dominant solution has a positive likelihood difference.

We also generated PDFF/ R_2^* scatterplots to investigate the distribution of R_2^* values arising from the true and swapped likelihood optima.

3.1.4 | Likelihood function visualization

To gain deeper insights into the behavior observed using fit success histograms and likelihood difference plots, we computed and visualized the likelihood functions for the chosen PDFF/ R_2^* combinations. First, noise-free data were simulated based on the PDFF/ R_2^* values chosen for interrogation, and Gaussian noise was added in real and imaginary channels. A grid of “candidate” PDFF/ R_2^* values was generated (PDFF 0%–100% and R_2^* 0– 1000 s^{-1}), and the likelihood at each point on the grid was computed based on Equations (3) and (8). This 2D likelihood plot was displayed using a colormap, enabling identification of “true” optima (corresponding closely to the ground truth) and swapped optima (typically with a PDFF in the opposite half of the range to the ground truth).

Having generated the 2D likelihood plot, the noisy complex signal was passed to the fitting algorithm as done previously. The positions of the two candidate solutions (arising from fat-dominant and water-dominant initializations) were recorded and displayed on the likelihood function, with the chosen solution highlighted. The paths taken by the optimizer for both initializations

were also displayed. A further estimate of the global optimum was obtained using a search over the generated 2D grid of likelihood values. Note that the values from this search will generally be close to but not exactly match the values obtained from the fitting, because the 2D nature of the search means that the value of S_0 is fixed; this provides a useful simplification that reduces the degrees of freedom and thus reduces the potential for overfitting.

3.1.5 | Sigma uncertainty

To assess the impact of inaccuracies in sigma on PDFF/ R_2^* estimation, the simulations described in section 3.1.2 were repeated with inaccurate sigma assumptions. Specifically, the simulations were performed with a 30% overestimate and a 30% underestimate and compared against the simulations performed with accurate sigma estimates. Note that these are relatively large sigma errors that are greater than those expected based on simulation data (see Supporting Information S1), thus providing a relatively conservative measure of the robustness of the method.

3.2 | Phantom experiments

To ensure the feasibility of MAGORINO-based fitting on a range of scanners, MAGORINO was evaluated in a publicly available multisite, multivendor, and multi-field-strength phantom data set.³⁷ Full details of the phantom data set are given in Hernando et al.³⁷ Briefly, the data set consists of fat-water mixtures with known varying fat fraction, scanned at 1.5 T and 3 T at six centers (two centers each for GE, Philips, and Siemens). Data acquisition was performed using each site's version of a multi-echo 3D spoiled gradient-echo CSE sequence, with two different protocols (each performed at 1.5 T and 3 T) used to test the reproducibility across different acquisition parameters. Protocols 1 and 2 were performed at 1.5 T, and protocols 2 and 4 were performed at 3 T. Protocols 1 and 3 generated approximately in-phase and opposed-phased echoes, whereas protocols 2 and 4 used the shortest echoes achievable. To obtain PDFF measurements, circular ROIs were manually placed on each of the vials by a radiologist (EL). The MAGORINO algorithm was implemented as described previously, and sigma was estimated a priori using fitting from the 0% PDFF vial, again using the simulation-derived correction factor $k = 1.163$ to compensate for slight overfitting. The relationship between PDFF estimates and references values was assessed using linear regression for both MAGORINO and MAGO.

3.3 | In vivo imaging

To evaluate the feasibility of MAGORINO in vivo, we imaged the pelvis and lower legs of 2 healthy volunteers. These scans were performed with institutional review board approval (Queen Square Research Ethics Committee, London, REC 15/LO/1475), and both subjects provided written informed consent. The data were acquired on a 3T Philips Ingenia system using a multi-echo 3D spoiled gradient-echo sequence with monopolar readouts and flyback gradients ($TE_1 = 1.2$ ms, $\Delta TE = 1.6$ ms, flip angle = 5° , TR = 25 ms, matrix size = 320×320 , and pixel spacing = 1.8×1.8 mm). Coil combination was performed using SENSE (factor 1). The MAGORINO algorithm was implemented as described previously, and sigma was estimated from homogenous regions of muscle close to the tissues of interest. To provide a qualitative performance assessment, the images were assessed and compared by a consultant radiologist with 9 years of MRI research experience. To provide a quantitative performance assessment, particularly focusing on high R_2^* tissues (where performance is expected to diverge from MAGO), we evaluated differences between parameter estimates in bone marrow and cortex, which are healthy tissues that have high R_2^* . To enable a direct comparison between the in vivo data and the predictions of simulation, simulations were rerun using the same PDFF and R_2^* values observed in the in vivo data. Bias was determined for Gaussian fitting relative to Rician fitting for both in vivo and simulation data.

4 | RESULTS

4.1 | Simulations

Results of the simulation experiments are shown in Figure 3 and Supporting Information Figures S2–S4. Figure 3 shows the parameter error; Supporting Information Figure S2 shows the parameter SD; and Supporting Information Figure S3 shows the fitting error on PDFF, R_2^* , and S_0 at good SNR. Supporting Information Figure S4 shows the parameter error for low SNR. The subsequent analysis for specific “interrogated” PDFF/ R_2^* combinations, using fit success histograms, likelihood difference plots, and likelihood function visualization, is shown in Figures 4–7.

4.1.1 | Parameter error

Figure 3 shows the mean error on PDFF, R_2^* , and S_0 relative to the ground-truth values. Note that, for fat fraction estimates (top row), areas of positive error at low fat fraction

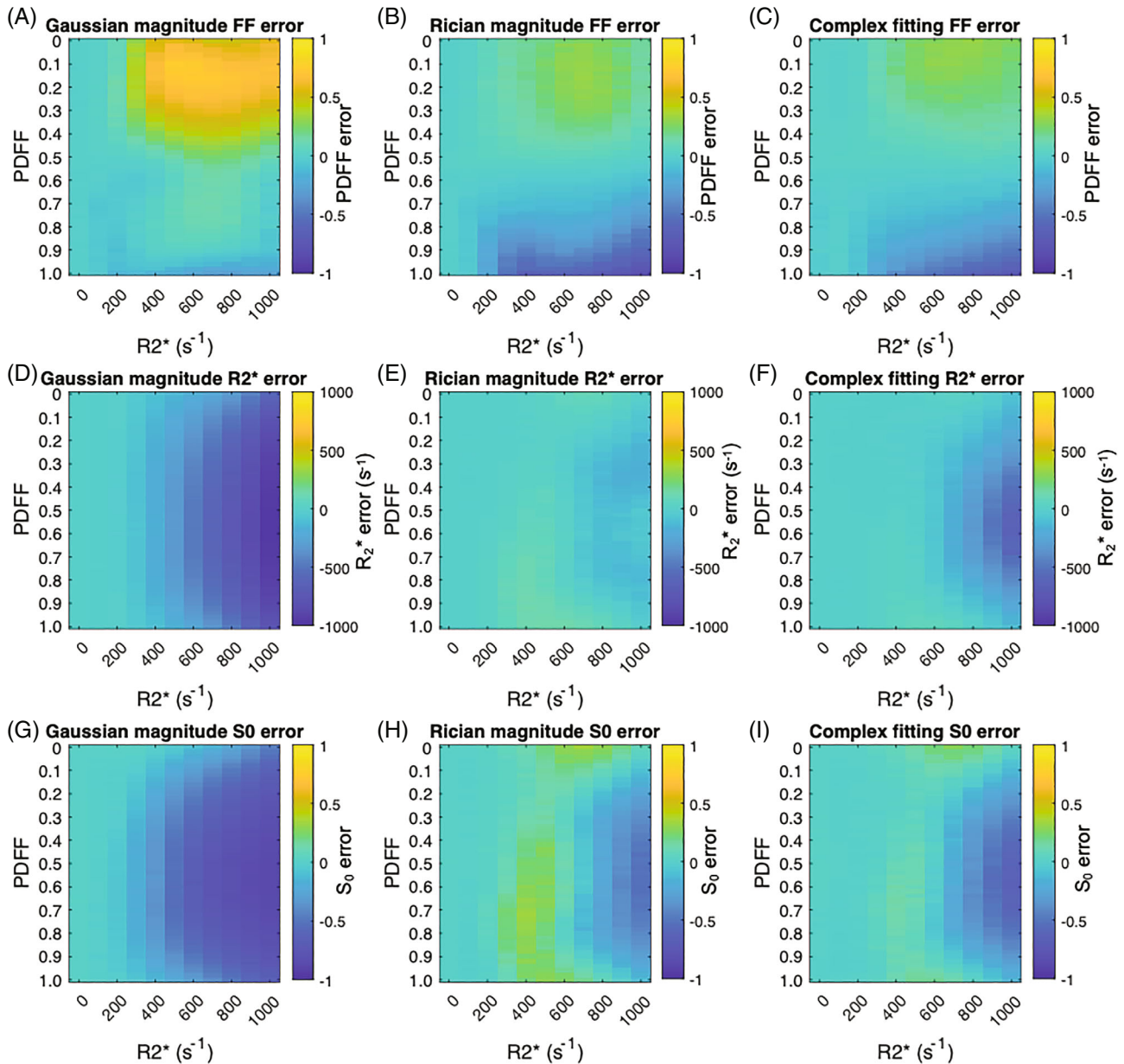


FIGURE 3 Parameter error is lower for Rician than for Gaussian fitting. The plots show the color-coded error in PDF (A–C), R_2^* (D–F), and S_0 (G–I) estimates for each combination of PDF and R_2^* values over all simulations, with SNR = 60. For fat-fraction measurements (top row), Gaussian fitting suffers from fat-water swaps as R_2^* increases, particularly in the low PDF range (producing the bright yellow area at the top of [A]). This problem is substantially mitigated using Rician fitting (B), which approaches the performance of complex fitting (C), although with a tradeoff of some increase in error at high PDF and high R_2^* . At low R_2^* , both Gaussian and Rician magnitude fitting (A, B) show lower error in PDF than complex fitting (C); the figures below show that this is because complex fitting does not reach the true (nonswapped) likelihood maximum in every case, resulting in a small positive bias and increase in parameter SD. D–F, For R_2^* measurements (second row), Rician fitting (E) substantially reduces the negative bias occurring at high R_2^* values for Gaussian fitting (D), with similar performance to complex fitting (F)

and negative error at high fat fraction arise predominantly from fat–water swaps.

For PDF measurements (top row) at low R_2^* (left edge of plots), all three algorithms show minimal bias over the full range of PDF values. For higher R_2^* values, Gaussian fitting shows a substantial positive bias in PDF

measurements for PDF values $< 50\%$ with $R_2^* > 400 s^{-1}$ (see top-right quadrant of Figure 3A), indicating frequent fat–water swaps. Rician and complex fitting both substantially reduce this bias (upper-right quadrant of Figure 3B). A caveat is that, for Rician fitting, there is a small increase in bias at high fat fraction and high R_2^* (bottom-right

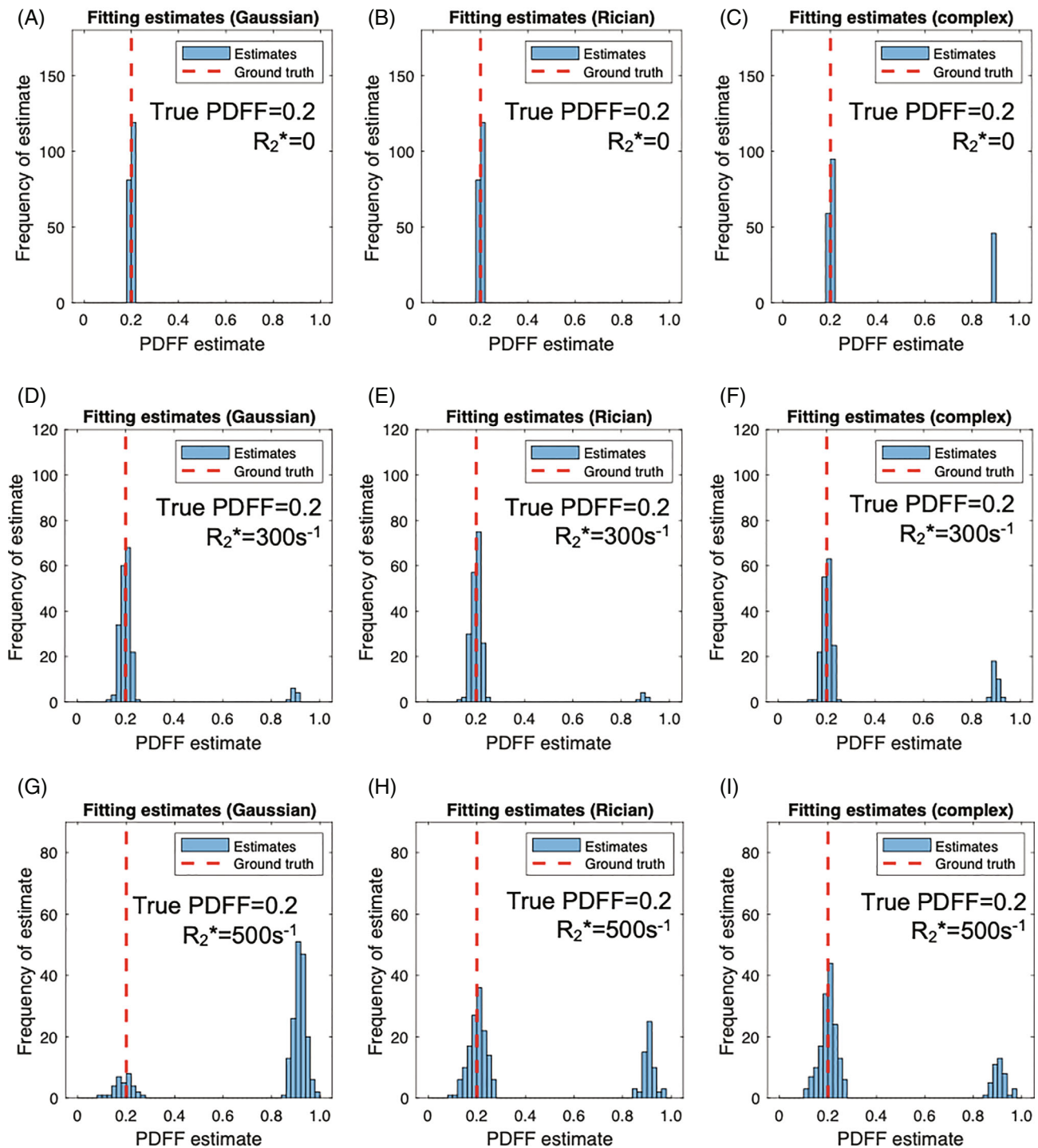


FIGURE 4 PDFF error arises from incorrect selection of the correct optimum-fit success histograms. Each plot shows the frequency of fat-fraction estimates relative to the ground-truth value. All plots were generated with PDFF = 20%, while three different R_2^* values (0.0, 0.3, and 0.5 ms^{-1}) were used to demonstrate the effect of varying R_2^* , with each R_2^* value on a separate row. Plots have been generated for Gaussian fitting (A, D, G), Rician fitting (B, E, H), and complex fitting (C, F, I). For Gaussian fitting, as R_2^* increases, the likelihood of the swapped solution arising increases, but this increase is mitigated by the use of Rician or complex fitting. For Gaussian fitting, at high R_2^* , most of the fitted solutions are incorrect (swapped) (E), whereas most of the solutions are correct (nonswapped) for Rician fitting (H) and complex fitting (I)

quadrant of Figure 3B). Complex fitting with f_B fixed (Figure 3D) almost eliminates the bias observed with the first three methods except for R_2^* values close to 1000 s^{-1} .

For R_2^* measurements (middle row), Gaussian fitting carries a substantial negative bias in R_2^* measurements, which is most severe at high R_2^* and in the intermediate

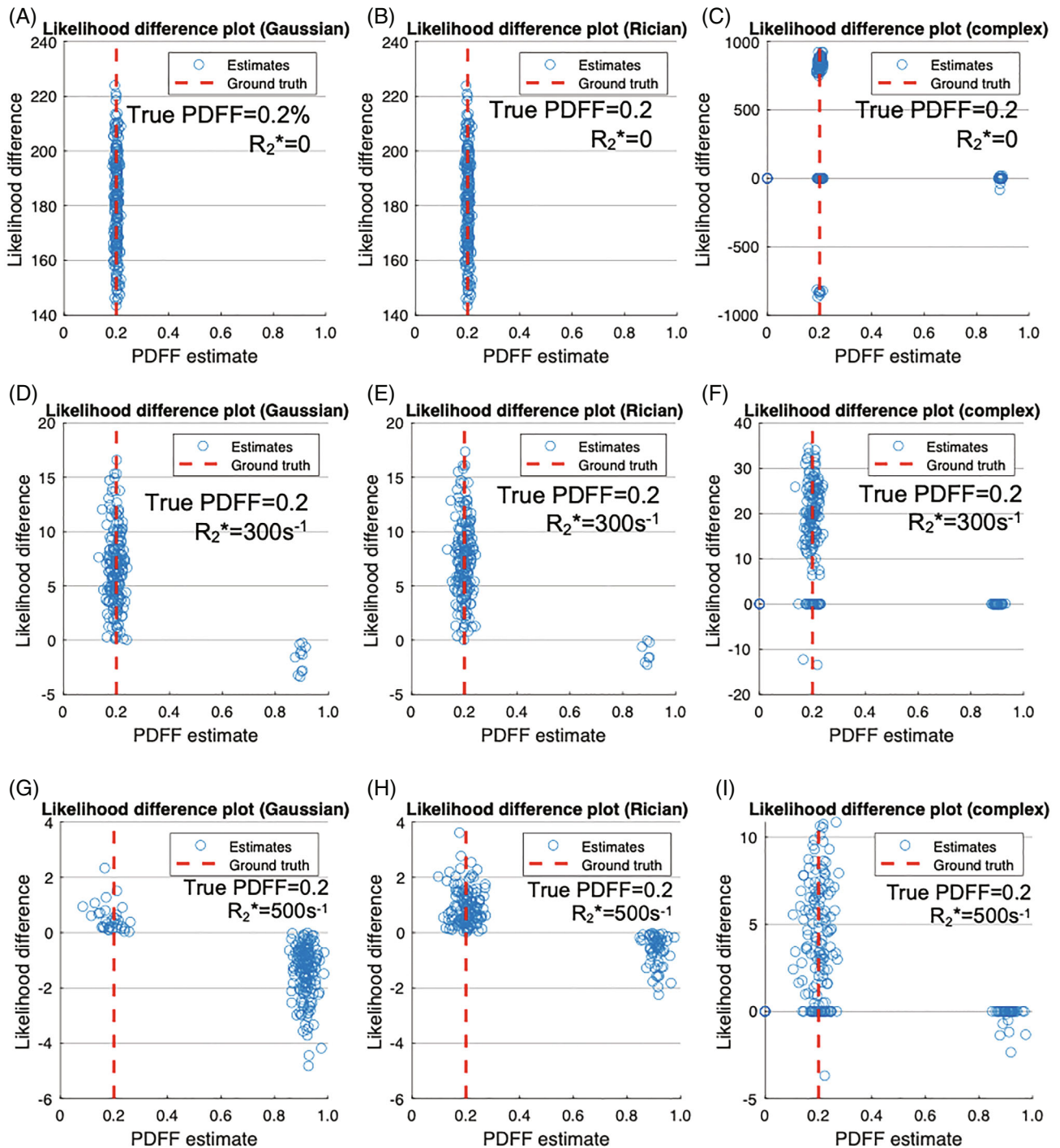


FIGURE 5 PDFF error arises from incorrect selection of the correct optimum-likelihood difference plots. For each plot, the y-axis shows the difference in likelihood between water-dominant and fat-dominant solutions. A positive likelihood difference indicates that the water-dominant (low PDFF) solution is more likely, whereas a negative likelihood indicates that the fat-dominant (high PDFF) solution is more likely. All plots were generated with PDFF = 20%, while three different R_2^* values (0, 300, and 500 s⁻¹) were used to demonstrate the effect of varying R_2^* , with each R_2^* value on a separate row. Plots have been generated for Gaussian fitting (A, D, G), Rician fitting (B, E, H), and complex fitting (C, F, I). For Gaussian fitting, the incorrect (swapped) solution shows greater likelihood in most simulations at high R_2^* (G), whereas the true (nonswapped) solution shows greater likelihood in most simulations for Rician fitting (H) and complex fitting (I). Note that in complex fitting, we can sometimes find the same minimum from both fat-dominant and water-dominant initializations, resulting in a likelihood difference of zero, or find the wrong minima for both initializations, resulting in “reversed” likelihood values (negative at low PDFF or positive at high PDFF)

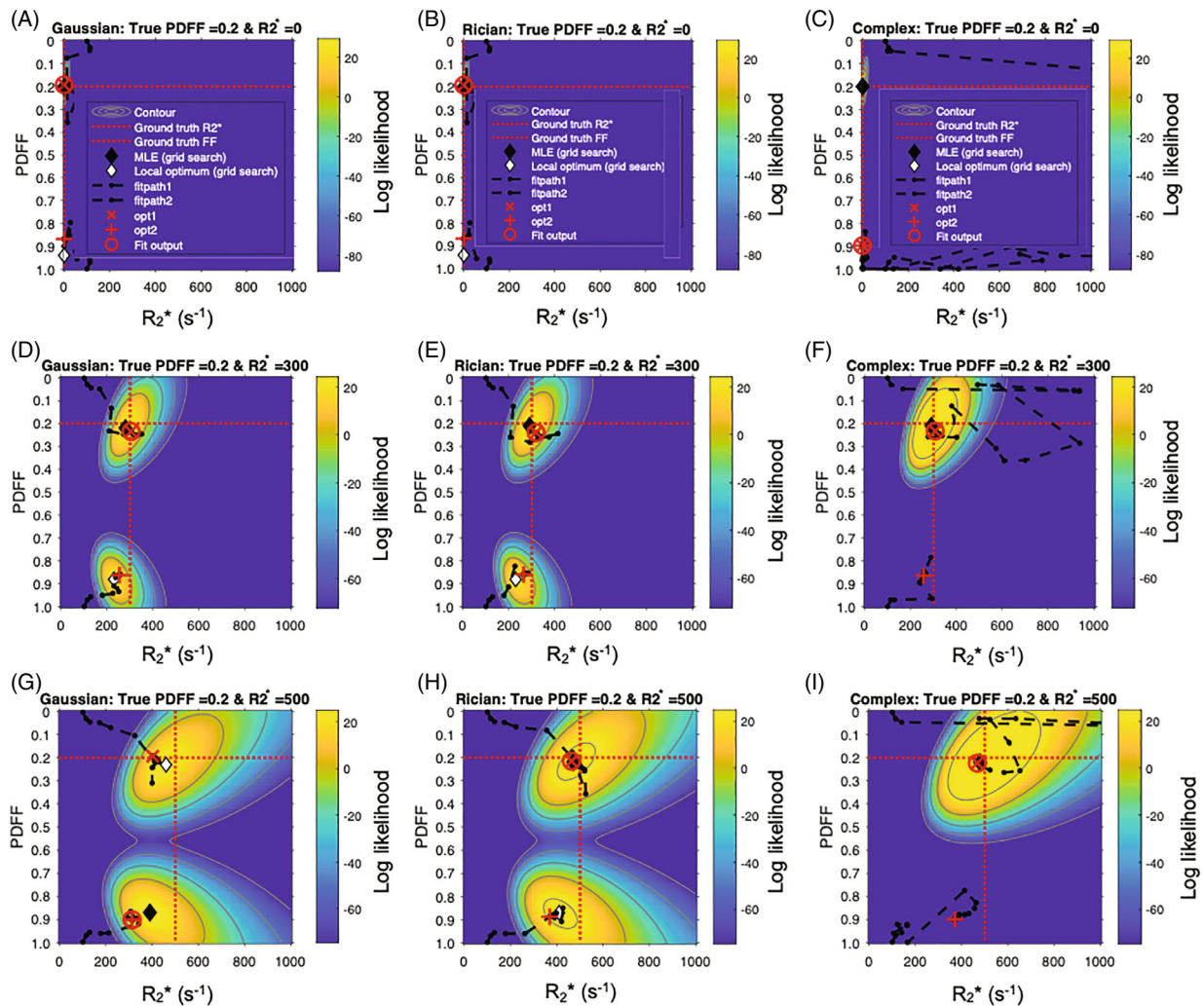


FIGURE 6 PDF error arises from incorrect selection of the correct optimum-likelihood function visualization. To gain further insight into the behaviors observed in Figures 3–5, likelihood functions (and fitted solutions) were visualized for three R_2^* values (0, 300, and 500 s^{-1}) with a fixed PDFF = 20%. Each plot shows the color-coded likelihood over the clinically feasible space of possible PDFF and R_2^* values for a given pair of ground-truth parameter measurements. Plots have been generated for Gaussian fitting (A, D, G), Rician fitting (B, E, H), and complex fitting (C, F, I). Each plot labels the ground-truth fat fraction and R_2^* , maximum likelihood estimate from 2D PDFF/ R_2^* grid search (MLE grid search), local optimum from grid search, likelihood optima from water-dominant and fat-dominant initializations (opt1 and opt2, with the chosen solution circled as the fit output), and paths on the objective function (path1 and path2 for opt1 and opt2, respectively). Note that all three methods arrive at the true (nonswapped solution) for $R_2^* = 0$ and $R_2^* = 300$, but at $R_2^* = 500$ only Rician and complex fitting correctly resolve the fat-water ambiguity. Note that there is a small discrepancy between the position of the MLE from the grid search and the fitting outputs; this arises because the grid search was performed in two dimensions (over PDFF and R_2^* values, to match the dimensions of the likelihood plot), whereas the fitting includes S_w and S_f as separate parameters, which is more realistic but leads to a greater degree of overfitting than the idealized grid search. For complex fitting, note that only the true (nonswapped) optimum is visible on the plots; this is because the swapped optimum occurs at a different value for f_B and is therefore not observed in this 2D grid of likelihood values that effectively has fixed f_B

PDFF range. This bias is substantially reduced for by Rician and complex fitting, with similar performance for both algorithms.

For S_0 measurements (bottom row), the results largely mirror those of R_2^* error: Gaussian fitting carries a negative bias at high R_2^* measurements, and the bias is reduced for Rician and complex fitting.

The benefits of Rician fitting over Gaussian fitting in terms of reduced bias are even more pronounced

at low SNR, with substantial reductions in bias for PDFF and R_2^* estimation (Supporting Information Figure S4).

4.1.2 | Parameter SD

Supporting Information Figure S2 shows the SD of PDFF, R_2^* , and S_0 estimates.

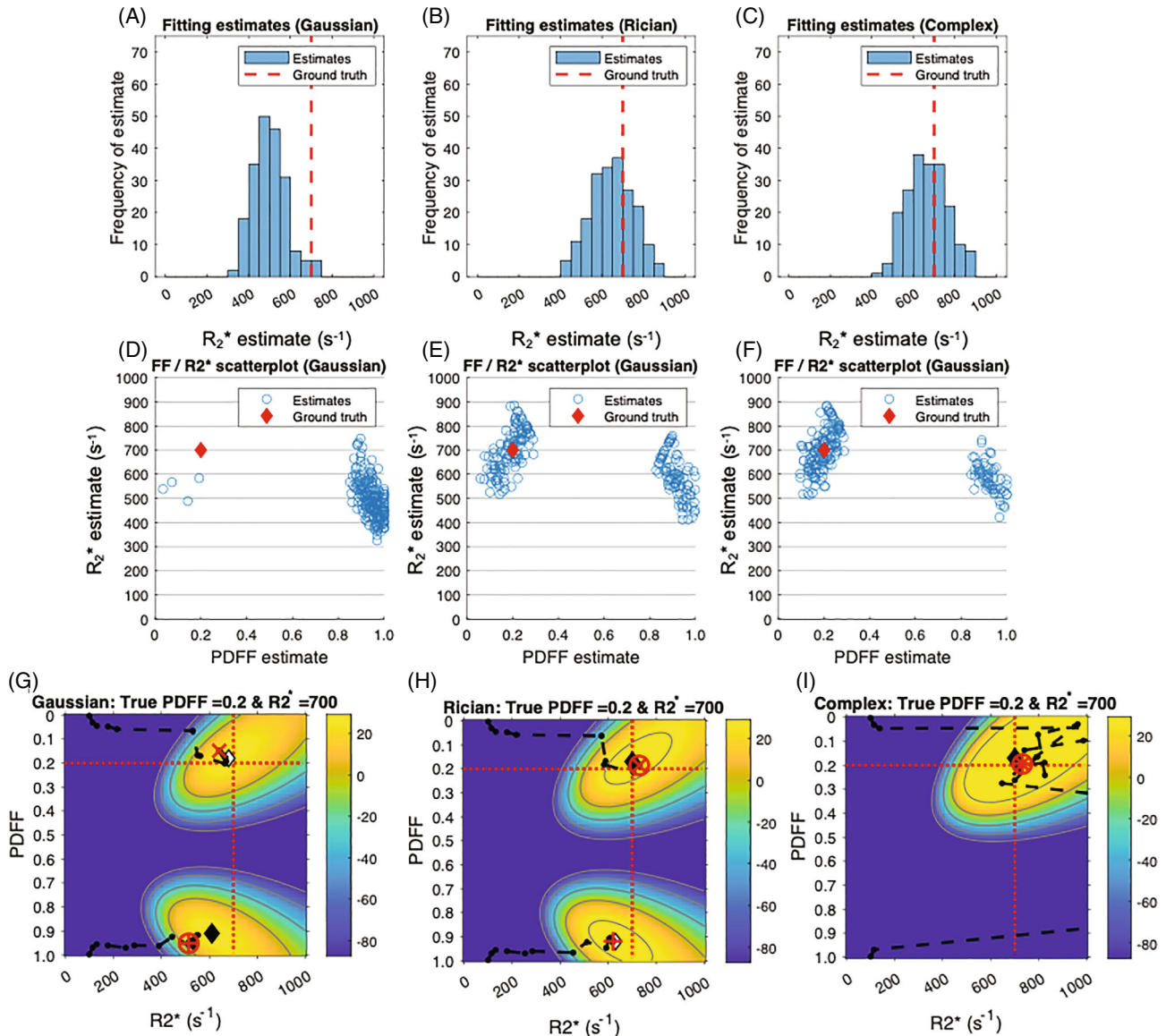


FIGURE 7 The R_2^* error arises from shifts in the position of the optima and incorrect selection of the correct optimum: fit success histograms (A–C), PDFF/ R_2^* scatterplots (D–F), and likelihood plots (G–I). For (A)–(C), the y-axis shows the frequency of R_2^* estimates relative to the ground-truth value for Gaussian (A), Rician (B), and complex (C) fitting. For (D)–(F), the scatterplots show PDFF and R_2^* parameter estimates. All plots were generated with PDFF = 20%, $R_2^* = 700 \text{ s}^{-1}$, and SNR = 60. The histograms show a substantial downward shift away from the ground-truth R_2^* value for Gaussian fitting; for Rician and complex fitting this bias is substantially reduced, and most estimates cluster around the ground-truth value. The scatterplots show that there are two separate reasons for the R_2^* bias observed with Gaussian fitting: (i) fat–water swaps, with the swapped solution having lower R_2^* than the true solution (this can be considered as an “ R_2^* swap”); and (ii) a further negative bias for both true and swapped solutions relative to the ground truth and relative to solutions obtained from Rician and complex fitting. G–I, Likelihood plots for Gaussian magnitude, Rician magnitude, and complex fitting, respectively. For both Rician and complex fitting, the fit results are closer to the ground-truth solution than for Gaussian fitting. The distance between the MLE and opt1 is reduced for Rician fitting compared with Gaussian fitting, indicating reduced overfitting. Abbreviation: FF, fat fraction

For PDFF measurements (top row), as R_2^* increases, all three algorithms show an increase in PDFF SD. Note that Gaussian fitting shows areas of low-PDFF SD in areas of frequent swapping (ie, areas of high bias in Figure 3A), whereas there are no corresponding areas of low-PDFF SD

associated with high bias for either Rician or complex fitting. The PDFF SD is markedly reduced for complex fitting with fixed f_B .

For both R_2^* and S_0 , parameter variance increases with increasing R_2^* and is broadly similar between algorithms.

4.1.3 | Fitting error

Supporting Information Figure S3 shows the SSE (top row), “true SSE” (ie, SSE relative to ground-truth parameter estimates) (middle row), and estimated noise SD relative to the true noise SD (bottom row).

For Gaussian fitting, there is a substantial increase in the true SSE (Supporting Information Figure 3E) at high R_2^* , which is not seen in the standard SSE value (Supporting Information Figure 3A) and is accompanied by a reduction in the noise estimate relative to the true noise (Supporting Information Figure 3I), indicating overfitting. For Rician and complex fitting, this overfitting is markedly reduced.

For complex fitting, there are areas of increased SSE, true SSE, and overestimation of the noise at low R_2^* . This could be eliminated by fixing f_B , suggesting incorrect estimation of f_B (despite the correct initialization) as a potential cause. The following likelihood function analyses give further insight into this behavior.

4.1.4 | Interrogation of specific PDFFF/ R_2^* combinations

Figures 4, 5, and 6 provide further insight into the behavior of the algorithms for three PDFFF/ R_2^* combinations corresponding to the top half of the PDFFF error plots shown in Figure 3A–C. All three combinations consisted of a PDFFF of 20%; the R_2^* values were 0, 300, and 500 s^{-1} (corresponding to the top, middle, and bottom rows, respectively, for each Figure), which were chosen to illustrate the behavior observed in Figure 3.

Figure 4 shows fit success histograms illustrating the frequency at which the two-point Gaussian, Rician, and complex fitting algorithms find the correct likelihood maximum (optimum) at different R_2^* values. As R_2^* increases, the frequency with which the swapped solution is selected also increases, but this increase is mitigated by the use of Rician and complex fitting. For Gaussian fitting at high R_2^* , most of the fitted solutions are incorrect (swapped) (Figure 4G), whereas most of the solutions are correct (nonswapped) for Rician and complex fitting (Figure 4H,I).

Similarly, Figure 5 plots the difference in likelihood for the true and swapped solutions from the fitting algorithms against the chosen (higher likelihood) estimate. Again, as R_2^* increases, the likelihood of the swapped solution arising increases, but this increase is mitigated by Rician and complex fitting. For Gaussian fitting, the incorrect (swapped) solution shows greater likelihood in most simulations at high R_2^* (Figure 5G), whereas the true (nonswapped) solution shows greater likelihood in most simulations for Rician and complex fitting (Figure 5H,I).

Figure 6 shows the likelihood functions obtained for a single noise instantiation for the three chosen PDFFF/ R_2^* combinations. At low R_2^* (top and middle rows), all three methods could identify the true (nonswapped) solution. However, at higher R_2^* (third row) for Gaussian fitting, the swapped solution assumes a higher likelihood and is chosen as the fit output by the MAGO algorithm. For Rician and complex fitting, the true solution has a higher likelihood and is chosen as the fit output.

Figure 7 provides insight into the origin of R_2^* error observed at high R_2^* values, and includes fit success histograms (top row) and PDFFF/ R_2^* scatterplots (bottom row). Figure 7A,D shows that Gaussian R_2^* estimates are negatively shifted relative to the ground-truth value. Figure 7D shows that this arises as a result of selection of the swapped optimum and due to a downward shift in the position of both the true and swapped optimum relative to the ground truth. For Rician and complex fitting, both the number of swaps and the negative shift in the positions of the optima are reduced, contributing to a reduction in bias. Figure 7G–I shows example likelihood functions for Gaussian, Rician, and complex fitting at high R_2^* , and offers further insight into the R_2^* bias observed in Figure 3E–H and Figure 7A–F. For Gaussian fitting, there are two sources of negative bias in R_2^* . First, the position of the true optimum (closest to the ground truth) is negatively shifted relative to the ground truth, as evidenced by the position of opt1 and the maximum likelihood estimation from grid search. Second, the fit has chosen the swapped optimum due to its higher likelihood, resulting in a further downward bias (this can be considered as an “ R_2^* swap”). For Rician and complex fitting, the local optimum is closer to the ground truth value and has also been chosen correctly as the fit output, mitigating bias and accounting for the behavior in Figure 7A–F.

4.1.5 | Sigma uncertainty

The accuracy of sigma estimation using three different methods is shown in Supporting Information Figure S1. All three methods provide accurate sigma estimates when sigma is low (ie, when SNR is high). However, the signal intensity–based method (“ROI sigma”) is less robust to any inhomogeneity in the ROI, providing evidence for the use of the fitting-based method in this work. Note that the fitted sigma estimates slightly underestimate sigma due to a degree of overfitting, but that the use of the correction factor allowed for very accurate sigma estimation on unseen simulated data.

The impact of inaccuracies in the estimated sigma is shown in Supporting Information Figure S5. Even with relatively large inaccuracies (greater than those expected from the previous simulations), the performance

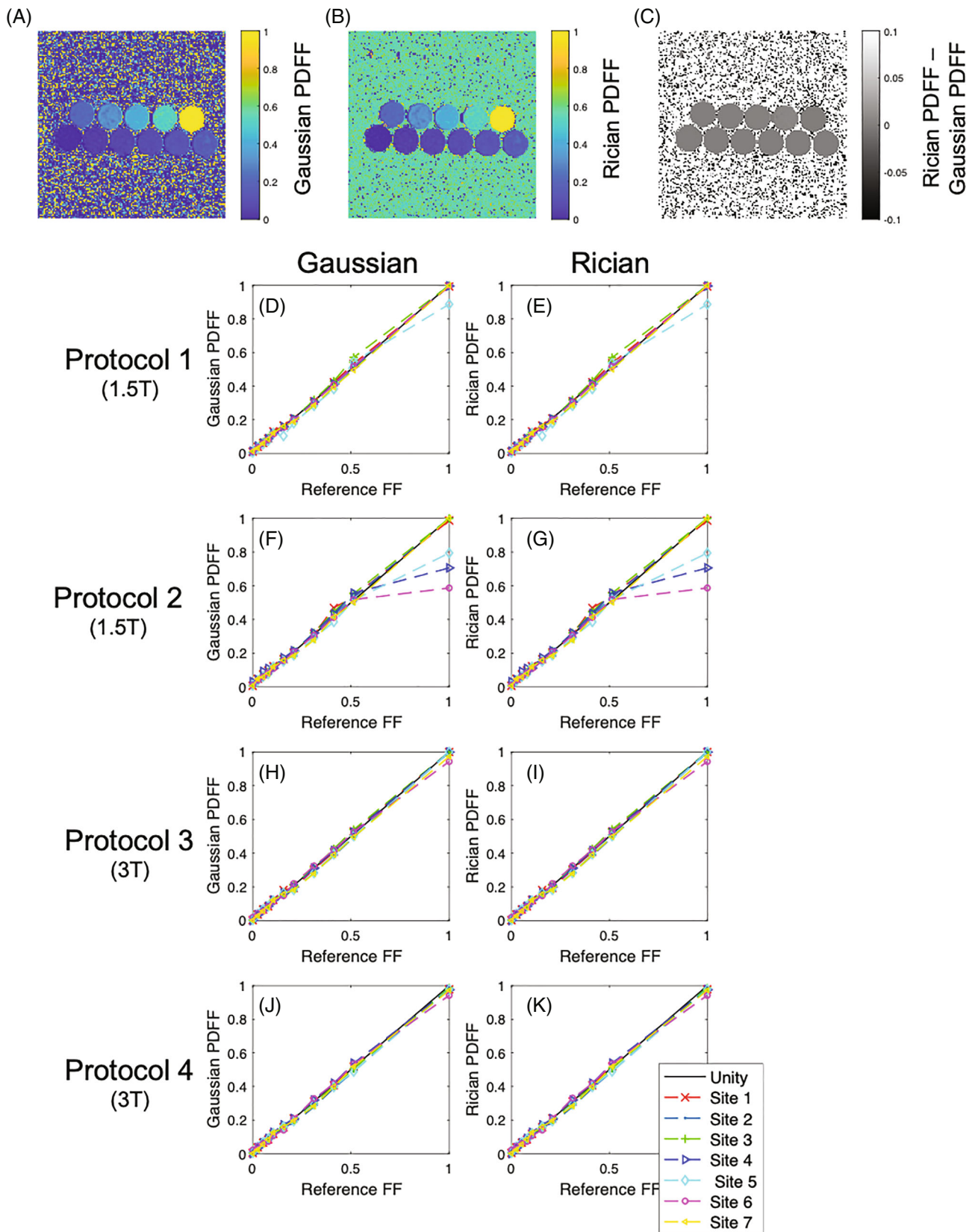


FIGURE 8 Agreement of Rician (MAGORINO) and Gaussian (MAGO) PDFF with reference values on a multisite, multivendor, multi-field-strength phantom data set. Agreement plots are shown for each of the four protocols, with individual points for each of the six sites. The black line indicates perfect agreement with reference PDFF values. "Site 7" refers to the repeat scans at site 1. The example images (A-C) are from site 1, 3T protocol 2, chosen to enable a direct visual comparison with Triay Bagur et al.²⁸

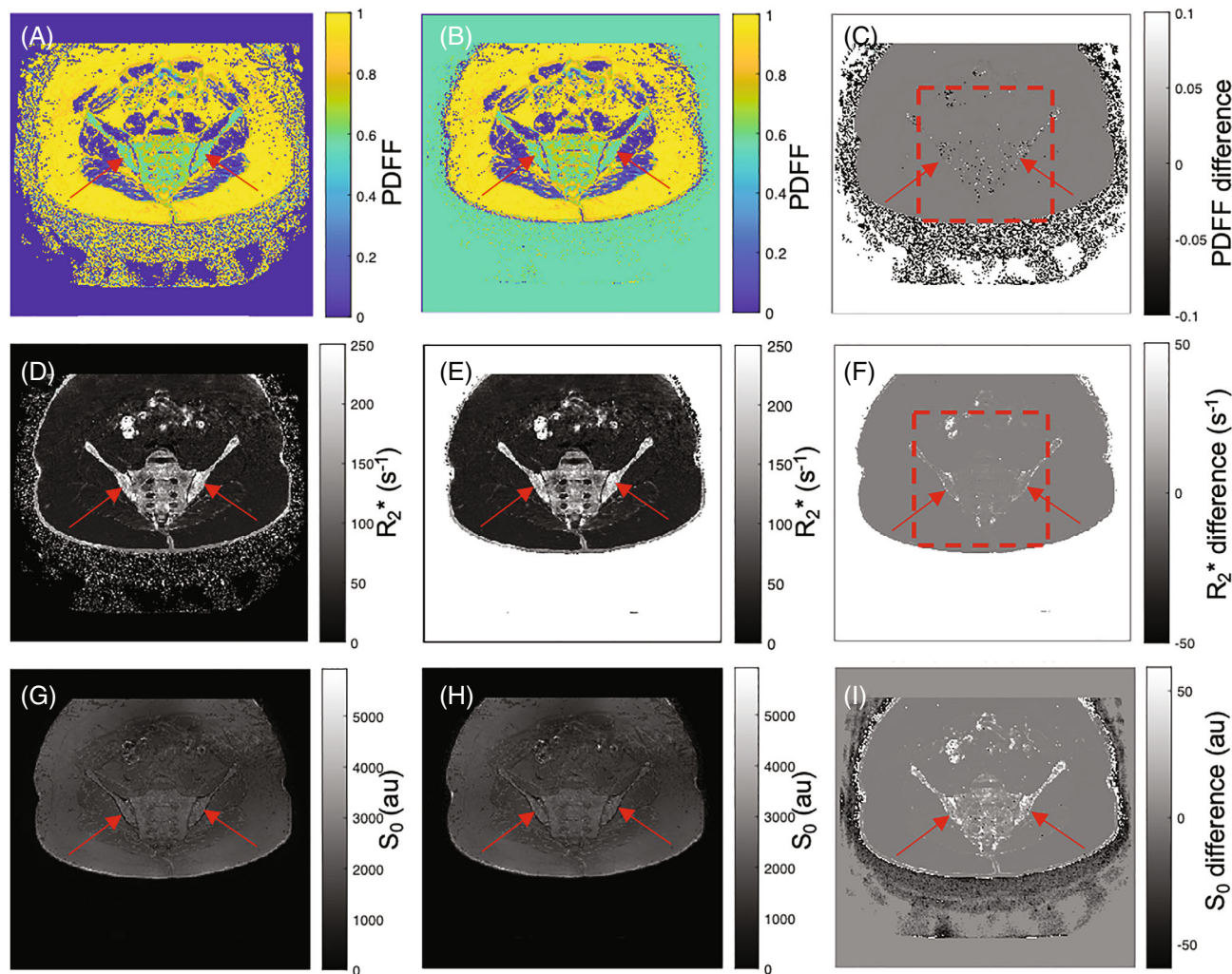


FIGURE 9 Example images of the pelvis (subject 1) for Gaussian fitting (left column) and Rician fitting (middle column). Both methods produce satisfactory fat–water separation across the image and good-quality R_2^* maps. However, as predicted by the simulations, the methods diverge in regions of high R_2^* /low SNR, which is particularly pronounced in the bone marrow and cortex (red arrows) as well as the skin. The difference maps (right column) show systematically higher R_2^* values in these regions, whereas differences in PDFF can be both positive or negative. In accordance with predictions from theory and simulation, Rician-estimated S_0 values are higher in regions of high R_2^* but lower in the air surrounding the patient than for Gaussian fitting. Figure 10 provides a more detailed graphical illustration of the differences in these Rician-regime regions within the bone marrow (the region analyzed is indicated by the red “dashed box” in [C] and [F])

of MAGORINO remains superior to MAGO for almost all plausible PDFF/ R_2^* combinations (the only exception being high PDFF and high R_2^* voxels, which is a very rare combination in vivo).

4.2 | Phantom experiments

Results of the analysis of the Hernando multisite phantom data set (where the R_2^* values are close to 0) are shown in Figure 8. The difference image shows excellent voxel-wise agreement among the methods. Median PDFF values for all 11 phantom vials are plotted against reference fat-fraction values for all sites, acquisition protocols, and field strengths. Linear regression results are

found in Supporting Information Table S2 (slope, intercept, and R-squared) for Hernando PDFF, MAGO PDFF, and MAGORINO PDFF. The results show excellent agreement among the methods, with high accuracy, high linearity, and small bias (reflected in R-squared coefficients close to 1, slope close to 1, and intercept close to 0). As expected, the performance of MAGO and MAGORINO is very similar in this data set, where R_2^* is close to 0.

4.3 | In vivo imaging

Images from the 2 subjects are shown in Figure 9 and Supporting Information Figure S6, and a more detailed interrogation of the differences in parameter values between MAGO and MAGORINO is found in Figure 10.

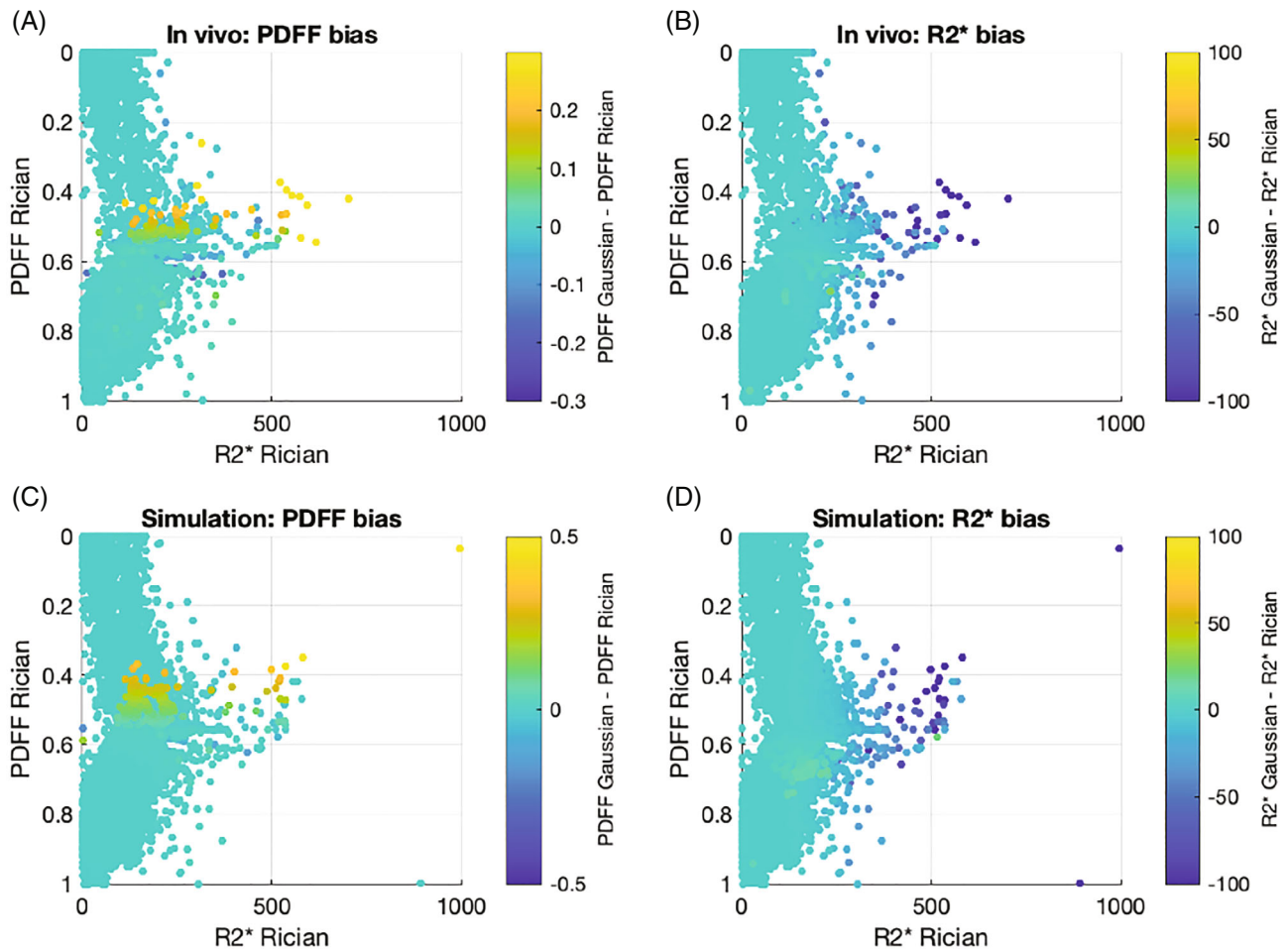


FIGURE 10 The pattern of bias for Gaussian versus Rician fitting in vivo (A, B) is similar to that shown by simulation (C, D) in regions where Rician noise dominates, suggesting that the simulation experiments can accurately predict performance in vivo. Parameter estimates obtained for the central region of the image (indicated in Figure 9 by the dotted red line) are shown for subject 1 as 3D scatterplots. The position of each point is dictated by the PDFF and R_2^* estimates from Rician fitting, and the color of each point represents the bias (Gaussian estimate–Rician estimate). To facilitate direct comparison with simulation, the in vivo values shown in (A) and (B) were used to simulate and fit noisy signals, thus generating the plots in (C) and (D). The pattern of bias is very similar for in vivo results (top row) and simulation results (bottom row), with positive bias in PDFF values < 0.5 and R_2^* above 0.2 for Gaussian fitting and substantial negative bias for Gaussian fitting at higher R_2^* . This suggests that the simulation experiments can accurately predict performance in vivo

Both methods produced satisfactory fat–water separation across the image and good-quality R_2^* maps. In both cases, there is some “speckle” in the images but no structured fat–water swaps. However, as predicted by the simulations, the methods diverged in regions of high R_2^* /low SNR such as bone marrow and cortex (where proton density is lower and the presence of calcium hydroxyapatite causes rapid dephasing). The difference maps in Figure 9 and Supporting Information Figure S6 and the plots in Figure 10 show that Gaussian R_2^* measurements are negatively biased with respect to Rician R_2^* measurements, and Figure 10 shows that the pattern of PDFF bias observed in those high R_2^* tissues is extremely similar to the bias predicted by simulation experiments. This suggests that the

simulation experiments provide an accurate prediction of performance in vivo.

5 | DISCUSSION

The choice of fitting algorithm for PDFF and R_2^* estimation represents a tradeoff. Complex-based fitting enables resolution of fat–water ambiguity based on phase data, has a greater number of datapoints, and avoids Rician noise–related parameter bias, but dictates that phase information must be accessible and reliable (an important limitation for standard care and clinical trials, where a range of scanners may be used) and can fail in areas of

large B_0 inhomogeneity. Conversely, magnitude-based fitting can be performed without reliable phase data but suffers from bias due to the Rician nature of the noise distribution in the magnitude signal and requires fat–water ambiguity to be solved by another method. The strengths of the magnitude-based approach have led to its use as a “final step” in processing of data from multisite studies³⁷ and recently motivated the development of a pure magnitude-only algorithm known as MAGO.²⁸ Despite producing good agreement values with complex-based fitting, this method still suffers from noise-related bias, which is an important limitation when imaging at low SNR and/or high R_2^* . Here, we describe a new fitting algorithm known as MAGORINO that addresses a key limitation of MAGO, namely, its vulnerability to PDFF and R_2^* bias arising from the Rician distribution of noise.

The key results of our study are as follows. First, Rician noise modeling increases the chance that the global optimum corresponds to ground truth, thus increasing the chance of finding the “true” solution and producing more accurate PDFF and R_2^* estimates than with the erroneous Gaussian noise assumption. The advantage of MAGORINO over MAGO becomes apparent with increasing R_2^* and/or low SNR, where the performance of MAGO begins to deteriorate but MAGORINO retains its performance. We show that this behavior arises because the difference in likelihood between the true (nonswapped) and swapped optima is, on average, greater for MAGORINO than for MAGO because of the use of Rician noise modeling. The use of the Rician noise model dictates that the true (nonswapped) optimum is selected by the algorithm in a greater proportion of cases, resulting in a reduction in PDFF bias and variance. Furthermore, our results (Figure 7) demonstrate that the true and swapped maxima typically occur at different R_2^* values: Effectively, the R_2^* measurement can also be “swapped,” further exacerbating bias. This problem is also mitigated by the MAGORINO algorithm.

Second, even when the global optimum corresponds to the ground truth for both noise models, the optimum from Rician noise modeling is closer to the ground truth. This reflects the ability of the Rician noise–based fitting to correctly attribute the nonzero signal intensities at longer TEs to noise rather than a reduction decay rate.

Third, we demonstrate the feasibility of using MAGORINO in phantom data and in vivo. In the phantom data, MAGORINO showed excellent agreement with reference values and, as predicted from simulations (because R_2^* is low in the phantoms), was equivalent to MAGO. However, in vivo, at higher R_2^* (demonstrated in regions of bone marrow and cortex), we observed bias in both R_2^* and PDFF measurements for MAGO relative to MAGORINO, with a pattern entirely consistent with that predicted by

simulation. This suggests that the simulation experiments provide an accurate indication of performance in vivo.

The improvement in PDFF and R_2^* quantification provided by MAGORINO at high R_2^* may be particularly important when imaging tissues with “native” high R_2^* , such as bone marrow and cortex (particularly at low field strength/SNR), in pathological states such as iron overload, or when using iron-based contrast agents.^{12,14–16} For example, in severe iron overload, R_2^* values are typically greater than $.580\text{ s}^{-1}$ and can be as high as 2000 s^{-1} ,³⁸ beyond the upper end of the range of values simulated in this study, meaning that the biases observed here are biologically and clinically relevant.

To our knowledge, this is the first study combining explicit modeling of Rician noise with magnitude-based fitting for CSE-MRI. Previous studies have investigated the use of Rician noise and noncentral chi modeling in R_2^* estimation in the liver^{38,39} but did not consider the effect of fat, whereas Hernando et al used Rician noise modeling to suppress chemical shift artifact due to olefinic fat suppression, but did not apply this to multi-echo gradient-echo imaging.⁴⁰ Triay Bagur et al’s MAGO algorithm described a two-point search approach to magnitude-based fitting but did not include Rician noise modeling.²⁸ An important contribution of our study is that combining two-point search with Rician noise modeling has a synergistic effect, particularly with regard to resolution of fat–water ambiguity.

This work has several important limitations in terms of the method itself and the validation studies. First, the current implementation of MAGORINO assumes a single R_2^* term for water and fat, whereas in some tissues the true behavior may be more complex. However, the mono-exponential model is broadly accepted to be a good approximation in various tissues including liver and bone marrow. Second, we did not explore the effect of variations in imaging parameters such as the choice of field strength, number of TEs, acquisition geometry, or volumetric imaging. As these parameters affect SNR, they are likely to impact on the success with which fat and water can be resolved. It would be desirable to evaluate the robustness of MAGORINO compared with MAGO using acquisitions with different SNR (such as by modulating slice thickness or number of averages), but this is beyond the scope of the present study. Third, even with the Rician noise model, MAGORINO cannot correctly resolve fat–water ambiguity in all voxels; this generally results in a nonstructured “speckle” across the image rather than fat–water swaps in specific image locations. Incorporation of spatial regularization may therefore be of value for improving the homogeneity of parameter estimates and would be a simple addition to the method. Fourth, the current implementation of MAGORINO requires manual definition of a small

ROI for sigma estimation, which introduces some subjectivity. Automation of this step, such as using deep learning, could help to address this issue. Finally, the phantom experiments and in vivo validation performed in this work are preliminary. In particular, the current in vivo results were obtained from healthy subjects, meaning that “extreme” R_2^* values observed in states such as iron overload were not present. Although our results do show an advantage of MAGORINO in normal tissues with higher R_2^* (particularly bone marrow), disease states (e.g., iron overload) or imaging with iron-based contrast agents, where MAGORINO could show a more dramatic performance benefit, have not yet been studied.

6 | CONCLUSIONS

The MAGORINO algorithm reduces Rician noise-related bias in PDFF and R_2^* estimation, thus addressing a key limitation of traditional Gaussian noise-based magnitude-only fitting and removing a potential barrier to wider implementation of CSE-MRI in clinical care and trials.

ACKNOWLEDGMENT

Timothy J.P. Bray was supported by an NIHR Clinical Lectureship (CL-2019-18-001). MHC is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC). This work was undertaken at UCLH/UCL, which receives funding from the UK Department of Health’s NIHR BRC funding scheme. The views expressed in this publication are those of the authors and not necessarily those of the UK Department of Health.

DATA AVAILABILITY STATEMENT

The *MATLAB* code used to generate the results in this paper has been preserved in its original form and is available at doi: 10.5281/zenodo.6483430. The latest version of the software is available at <https://github.com/TJPBray/Rician-FW-MRI>.

ORCID

Timothy J. P. Bray  <https://orcid.org/0000-0001-8886-5356>

Emma Lim  <https://orcid.org/0000-0001-5577-9427>

Margaret A. Hall-Craggs  <https://orcid.org/0000-0001-8734-4065>

Hui Zhang  <https://orcid.org/0000-0002-5426-2140>

REFERENCES

- Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with IDEAL gradient echo imaging: correction of bias from T_1 and noise. *Magn Reson Med*. 2007;58:354-364.
- Reeder SB, Wen Z, Yu H, et al. Multicoil Dixon chemical species separation with an iterative least-squares estimation method. *Magn Reson Med*. 2004;51:35-45.
- Reeder SB, Pineda AR, Wen Z, et al. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): application with fast spin-echo imaging. *Magn Reson Med*. 2005;54:636-644.
- Yokoo T, Serai SD, Pirasteh A, et al; for the RSNA-QIBA PDFF Biomarker Committee. Linearity, bias, and precision of hepatic proton density fat fraction measurements by using MR imaging: a meta-analysis. *Radiology*. 2018;286:486-498.
- Middleton MS, van Natta ML, Heba ER, et al; for the NASH Clinical Research Network. Diagnostic accuracy of magnetic resonance imaging hepatic proton density fat fraction in pediatric nonalcoholic fatty liver disease. *Hepatology*. 2018;67:858-872.
- Middleton MS, Heba ER, Hooker CA, et al; for the NASH Clinical Research Network. Agreement between magnetic resonance imaging proton density fat fraction measurements and pathologist-assigned steatosis grades of liver biopsies from adults with nonalcoholic Steatohepatitis. *Gastroenterology*. 2017;153:753-761.
- Noureddin, M, Lam, J, Peterson, MR, Middleton, M, Hamilton G, Le T. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. *Hepatology*. 2013;58:1930-1940.
- Yoon JH, Lee JM, Lee KB, et al. Pancreatic steatosis and fibrosis: quantitative assessment with preoperative multiparametric MR imaging. *Radiology*. 2016;279:140-150.
- Kühn JP, Berthold F, Mayerle J, et al. Pancreatic steatosis demonstrated at MR imaging in the general population: clinical relevance. *Radiology*. 2015;276:129-136.
- Morrow JM, Sinclair CDJ, Fischmann A, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. *Lancet Neurol*. 2015;15:65-77.
- Janiczek RL, Gambarota G, Sinclair CDJ, et al. Simultaneous T2 and lipid quantitation using IDEAL-CPMG. *Magn Reson Med*. 2011;66:1293-1302.
- Bray TJP, Bainbridge A, Punwani S, Ioannou Y, Hall-Craggs MA. Simultaneous quantification of bone edema/adiposity and structure in inflamed bone using chemical shift-encoded MRI in Spondyloarthritis. *Magn Reson Med*. 2018;79:1031-1042.
- Latifoltojar A, Hall-Craggs M, Bainbridge A, et al. Whole-body MRI quantitative biomarkers are associated significantly with treatment response in patients with newly diagnosed symptomatic multiple myeloma following bortezomib induction. *Eur Radiol*. 2017;27:5325-5336.
- Hernando D, Kramer JH, Reeder SB. Multiplex fat-corrected complex R_2^* relaxometry: theory, optimization, and clinical validation. *Magn Reson Med*. 2013;70:1319-1331.
- Hernando D, Levin YS, Sirlin CB, Reeder SB. Quantification of liver iron with MRI: state of the art and remaining challenges. *J Magn Reson Imaging*. 2014;40:1003-1021.
- Wells SA, Schubert T, Motosugi U, et al. *Radiology*. 2019;294:108-116.
- Reeder SB, McKenzie CA, Pineda AR, et al. Water-fat separation with IDEAL gradient-echo imaging. *J Magn Reson Imaging*. 2007;25:644-652.
- Reeder SB, Robson PM, Yu H, et al. Quantification of hepatic steatosis with MRI: the effects of accurate fat spectral modeling. *J Magn Reson Imaging*. 2009;29:1332-1339.

19. Yu H, McKenzie CA, Shimakawa A, et al. Multiecho reconstruction for simultaneous water-fat decomposition and T_2^* estimation. *J Magn Reson Imaging*. 2007;26:1153-1161.
20. Hernando D, Haldar JP, Sutton BP, Ma J, Kellman P, Liang ZP. Joint estimation of water/fat images and field inhomogeneity map. *Magn Reson Med*. 2008;59:571-580.
21. Hernando D, Kellman P, Haldar JP, Liang ZP. Robust water/fat separation in the presence of large field inhomogeneities using a graph cut algorithm. *Magn Reson Med*. 2010;63:79-90.
22. Ghugre NR, Enriquez CM, Coates TD, Nelson MD, Wood JC. Improved R_2^* measurement in myocardial iron overload. *J Magn Reson Imaging*. 2006;23:9-16.
23. Wood JC, Enriquez C, Ghugre N, et al. MRI R_2 and R_2^* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood*. 2005;106:1460-1465.
24. He T, Gatehouse PD, Smith GC, Mohiaddin RH, Pennell DJ, Firmin DN. Myocardial T_2^* measurements in iron-overloaded thalassemia: an *in vivo* study to investigate optimal methods of quantification. *Magn Reson Med*. 2008;60:1082-1089.
25. Yin X, Shah S, Katsaggelos AK, Larson AC. Improved R_2^* measurement accuracy with absolute SNR truncation and optimal coil combination. *NMR Biomed*. 2010;23:1127-1136.
26. Bydder M, Yokoo T, Hamilton G, et al. Relaxation effects in the quantification of fat using gradient echo imaging. *Magn Reson Imaging*. 2008;26:347-359.
27. Hernando D, Kühn JP, Mensel B, et al. R_2^* estimation using “in-phase” echoes in the presence of fat: the effects of complex spectrum of fat. *J Magn Reson Imaging*. 2013;37:717-726.
28. Triay Bagur A, Hutton C, Irving B, Gyngell ML, Robson MD, Brady M. Magnitude-intrinsic water-fat ambiguity can be resolved with multipeak fat modeling and a multipoint search method. *Magn Reson Med*. 2019;82:460-475.
29. Bonny JM, Zanca M, Boire JY, Veyre A. T_2 maximum likelihood estimation from multiple spin-Echo magnitude images. *Magn Reson Med*. 1996;36:287-293.
30. Sijbers J, den Dekker AJ, Scheunders P, van Dyck D. Maximum-likelihood estimation of rician distribution parameters. *IEEE Trans Med Imaging*. 1998;17:357-361.
31. Karlsten OT, Verhagen R, Bovée WMMJ. Parameter estimation from Rician-distributed data sets using a maximum likelihood estimator: application to T_1 and perfusion measurements. *Magn Reson Med*. 1999;41:614-623.
32. Sijbers J, den Dekker AJ. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magn Reson Med*. 2004;51:586-594.
33. Raya JG, Dietrich O, Horng A, Weber J, Reiser MF, Glaser C. T_2 measurement in articular cartilage: impact of the fitting method on accuracy and precision at low SNR. *Magn Reson Med*. 2010;63:181-193.
34. Veraart J, van Hecke W, Sijbers J. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. *Magn Reson Med*. 2011;66:678-686.
35. Bouhrara M, Reiter DA, Celik H, et al. Incorporation of Rician noise in the analysis of Biexponential transverse relaxation in cartilage using a multiple gradient echo sequence at 3 and 7 Tesla. *Magn Reson Med*. 2015;73:352-366.
36. Dietrich O, Raya JG, Reeder SB, Ingrisch M, Reiser MF, Schoenberg SO. Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. *Magn Reson Imaging*. 2008;26:754-762.
37. Hernando D, Sharma SD, Aliyari Ghasabeh M, et al. Multisite, multivendor validation of the accuracy and reproducibility of proton-density fat-fraction quantification at 1.5T and 3T using a fat-water phantom. *Magn Reson Med*. 2017;77:1516-1524.
38. Yokoo T, Yuan Q, Sénégas J, Wiethoff AJ, Pedrosa I. Quantitative R_2^* MRI of the liver with Rician noise models for evaluation of hepatic iron overload: simulation, phantom, and early clinical experience. *J Magn Reson Imaging*. 2015;42:1544-1559.
39. Feng Y, He T, Gatehouse PD, et al. Improved MRI R_2^* relaxometry of iron-loaded liver with noise correction. *Magn Reson Med*. 2013;70:1765-1774.
40. Hernando D, Karampinos DC, King KF, et al. Removal of olefinic fat chemical shift artifact in diffusion MRI. *Magn Reson Med*. 2011;65:692-701.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Figure S1. Accuracy of sigma-estimation methods over the range of plausible SNR values. A–C, Estimated sigma on the y-axis against the true sigma on the x-axis over a range of values corresponding to the expected SNR at typical clinical field strengths *in vivo* (the SNR ranged from 20 to 70). The plots were generated by simulating and fitting 100 pure-water, low R_2^* voxels with three different degrees of inhomogeneity in the signal intensities within the region of interest (ROI). In (A), note that the fitted sigma estimates (red line) consistently slightly underestimate sigma due to a degree of overfitting. Note also that the degree of overfitting is very consistent, even in inhomogeneous voxels. To correct for this overfitting, we used linear regression to calculate and correct the slope of the red line. Having derived this slope, we calculated a correction factor k , which could be applied to sigma estimates to produce the “corrected fitted sigma” estimates. To allow a fair estimate of the performance of this correction, the slope was derived on a different set of noise instantiations to those used for testing (ie, different “training” and “test” simulation data were used). The “corrected fitted sigma” method provides a very accurate estimation on the test data set.

In (B) and (C), the effect of inhomogeneity on sigma estimates is assessed. The inhomogeneity factor specifies the proportional difference between the largest and smallest S_0 (where $S_0 = \rho_W + \rho_F$) in the voxel, with all other values evenly spaced between the largest and smallest values. In a homogenous voxel (A), all methods produce accurate sigma estimates. As the voxel becomes more inhomogeneous (B, C), the fitting sigma estimates remain accurate but the ROI-based sigma estimated becomes increasingly biased. The most accurate method is the fitting-based

sigma estimate with correction, referred to in the legend as “corrected fitted sigma”

Figure S2. Parameter SD for SNR = 60. The plots show the color-coded SD in PDFF (a-d), R_2^* (E-H) and S_0 (I-L) estimates for each combination of proton density fat fraction (PDFF) and R_2^* values over all simulations, with SNR = 60. Note that parameter SD generally increased with increasing R_2^* because fat–water swaps become more frequent (as shown in the figure above). At low R_2^* , both Gaussian and Rician magnitude fitting (A, B) show lower PDFF SD than complex fitting (C); the figures below show that this is because complex fitting does not reach the true (non-swapped) likelihood maximum in every case, resulting in a small positive bias and increased parameter SD. Note that this behavior is eliminated by fixing f_B (right-hand column), although this step is likely to be unrealistic in practice

Figure S3. Fitting error for SNR = 60. The plots show the grayscale-coded sum of sum of squared errors (SSE; A–D), “true SSE” (ie, SSE calculated relative to the ground truth) (E–H), and estimated noise (SSE/simulated noise SSE) (I–L) for each combination of PDFF and R_2^* values over all simulations, with SNR = 60. For Gaussian fitting, the “true SSE” (E) increases substantially at higher R_2^* values, indicating overfitting to the noise. This problem is substantially reduced by Rician magnitude fitting and complex fitting. For complex fitting (third column), SSE and noise estimates are highest at low R_2^* values because the two-point initialization does not reach the true (non-swapped) likelihood maximum in every case. Note that this behavior is eliminated by fixing f_B (right-hand column), although this step is likely to be unrealistic in practice

Figure S4. Parameter error for SNR = 20. The plots show the color-coded error in PDFF (A–D), R_2^* (E–H), and S_0 (I–L) estimates for each combination of PDFF and R_2^* values over all simulations. Note that the benefit of Rician fitting is more pronounced than for SNR = 60

Figure S5. Magnitude-only PDFF and R_2^* estimation with Rician noise modeling (MAGORINO) is robust to inaccurate sigma estimation. The plots show parameter error maps for PDFF (top row) and for R_2^* (bottom row) for Gaussian fitting (left column) and for Rician fitting with different sigma accuracies; the second column shows Rician fitting with sigma underestimated; the third column shows Rician fitting with correct sigma; and the right-hand column shows Rician fitting with sigma overestimated. Observe that the effect of the Rician noise model becomes

more pronounced from left to right, as the estimated sigma increases (note that Gaussian fitting is equivalent to a sigma assumption of 0). If sigma is underestimated by 30% (second column), the performance becomes closer to Gaussian fitting (MAGO) (left column) than for the correct sigma estimate (third column). If sigma is overestimated by 30% (right-hand column), the differences between MAGO and MAGORINO are exaggerated. Importantly, there is very little deterioration in performance for either underestimated or overestimated sigma, even with the substantial sigma errors assumed here

Figure S6. Example images of the lower legs (subject 2) for Gaussian fitting (left column) and Rician fitting (middle column). Note that this is a challenging case in which complex fitting produced a complete fat–water swap in one leg. Both Gaussian (MAGO) and Rician (MAGORINO) fitting produce satisfactory fat–water separation across the image and good-quality R_2^* maps, albeit with some nonstructured swapping in the subcutaneous fat. As with subject 1, the methods diverge in regions of high R_2^* /low SNR, which is particularly pronounced in the bone marrow and cortex (red arrows) as well as the skin. The difference maps (right column) show systematically higher R_2^* values in these regions, while differences in PDFF can be both positive or negative

Table S1. Summary of models and fitted parameters. All models were initialized using both fat-dominant and water-dominant initializations, as specified in the second column from the right. The objective functions and estimation of f_B varied among methods (i)–(iii). The constant C effectively compensates for reduction in the signal magnitude due to R_2^* decay and chemical shift and avoids the need for empirical manual adjustment of initial values depending on scanner gain, as performed in Triay Bagur et al.²⁸ *For complex fitting, the Gaussian log likelihood is computed separately for real and imaginary channels before summation

Table S2. Summary of linear regression parameters for multisite phantom data set (agreement between measured PDFF values and reference values)

How to cite this article: Bray TJP, Bainbridge A, Lim E, Hall-Craggs MA, Zhang H. MAGORINO: Magnitude-only fat fraction and R_2^* estimation with Rician noise modeling. *Magn Reson Med*. 2023;89:1173–1192. doi: 10.1002/mrm.29493