Received: 16 March 2020 | Revised: 3 September 2020 | Accepted: 14 September 2020

DOI: 10.1002/mrm.28545

FULL PAPER

Magnetic Resonance in Medicine

Confounding factors in breast magnetic resonance fingerprinting: B_1^+, slice profile, and diffusion effects

Teresa Nolte[1] | **Hannah Scholten**[2] | **Nicolas Gross-Weege**[1] | **Thomas Amthor**[3] | **Peter Koken**[3] | **Mariya Doneva**[3] | **Volkmar Schulz**[1,4,5,6]

[1]Physics of Molecular Imaging Systems, Experimental Molecular Imaging, RWTH Aachen University, Aachen, Germany

[2]Department of Diagnostic and Interventional Radiology, University of Würzburg, Würzburg, Germany

[3]Tomographic Imaging Systems, Philips Research Europe, Hamburg, Germany

[4]Hyperion Hybrid Imaging Systems GmbH, Aachen, Germany

[5]Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany

[6]Physics Institute III B, RWTH Aachen University, Aachen, Germany

Correspondence
Volkmar Schulz, Physics of Molecular Imaging Systems, Experimental Molecular Imaging, Center For Biohybrid Medical Systems (CBMS), RWTH Aachen University, Forckenbeckstr. 55, 52074 Aachen, Germany.
Email: volkmar.schulz@pmi.rwth-aachen.de

Funding information
Horizon 2020 Framework Programme, Grant/Award Number: 667211

Purpose: Magnetic resonance fingerprinting (MRF) offers rapid quantitative imaging but may be subject to confounding effects (CE) if these are not included in the model-based reconstruction. This study characterizes the influence of in-plane B_1^+, slice profile and diffusion effects on T_1 and T_2 estimation in the female breast at 1.5T.

Methods: Simulations were used to predict the influence of each CE on the accuracy of MRF and to investigate the influence of electronic noise and spiral aliasing artefacts. The experimentally observed bias in regions of fibroglandular tissue (FGT) and fatty tissue (FT) was analyzed for undersampled spiral breast MRF data of 6 healthy volunteers by performing MRF reconstruction with and without a CE.

Results: Theoretic analysis predicts T_1 under-/T_2 overestimation if the nominal flip angles are underestimated and inversely, T_1 under-/T_2 overestimation if omitting slice profile correction, and T_1 under-/T_2 underestimation if omitting diffusion in the signal model. Averaged over repeated signal simulations, including spiral aliasing artefacts affected precision more than accuracy. Strong in-plane B_1^+ effects occurred in vivo, causing T_2 left–right inhomogeneity between both breasts. Their correction decreased the T_2 difference from 29 to 5 ms in FGT and from 29 to 9 ms in FT. Slice profile correction affected FGT T_2 most strongly, resulting in −22% smaller values. For the employed spoiler gradient strengths, diffusion did not affect the parameter maps, corresponding well with theoretic predictions.

Conclusion: Understanding CEs and their relative significance for an MRF sequence is important when defining an MRF signal model for accurate parameter mapping.

Teresa Nolte is added as co-corresponding author.

[Correction added January 7, 2021, after first online publication: Teresa Nolte was designated as corresponding author.]

KEYWORDS
aliasing, bias, breast, confounding effects, magnetic resonance fingerprinting, quantitative imaging

1 | INTRODUCTION

Quantitative magnetic resonance imaging (qMRI) aims for the quantitative evaluation of the tissue relaxation times T_1 and T_2,[1] and its clinical value as has been demonstrated, for example, in cardiac[2] and muscoskeletal[3] MRI. Moreover, qMRI promises vendor-independent data, which is of advantage during multi-center clinical trials and for deep learning-based analysis methods. In the female breast, qMRI has previously been explored for response assessment of neoadjuvant chemotherapy[4-6] and for distinction between invasive ductal carcinoma from healthy tissue[7] or between different types of lesions.[8,9] Adversely, gold standard qMRI sequences measuring either T_1 or T_2 are of slow acquisition speed, which complicates their routine clinical use and therefore limits qMRI investigations of breast malignancies. Recently, magnetic resonance fingerprinting (MRF) has been presented as a rapid quantitative sequence,[10,11] and qMRI in the breast within a clinical feasible scan time using MRF has been demonstrated.[7]

MRF extracts multiple quantitative parameters from a time series of undersampled images with varying acquisition parameters. By varying the flip angle and/or the sequence timings from image to image, MRF encodes a tissue's spin history over the image series without leaving the transient state, thereby yielding distinct signal evolutions for different tissues. Typically, the quantitative parameters of interest measured by MRF include the tissue relaxation times T_1 and T_2, although MRF, by design, is not limited to the relaxation times only. Parameter maps are reconstructed by matching the measured temporal signal evolution in every voxel to a pre-calculated dictionary of possible signal evolutions, from which the best-matching entry yields the quantitative parameters for that voxel. Hence, the accuracy of the parameter maps crucially depends on the extent to which the signal model replicates the physical reality behind the MRF measurement. In other words, MRF can only be accurate if all influences on the magnetization dynamics are taken into account. For unbalanced gradient-echo MRF sequences, extended phase graph (EPG) simulations are used for dictionary calculation.[12,13] The signal model commonly includes instantaneous RF excitations defined by the nominal MRF flip angle train, T_1 and T_2 relaxation between excitations, as well as re- and dephasing of magnetization induced by the unbalanced spoiling gradient moments.

However, other physical effects may influence the signal evolution and therefore act as confounding effects (CE). On the one hand, deviations from the nominal sequence parameters may lead to bias on the MRF results. B_1^+ effects (ie, both in-plane B_1^+ inhomogeneity and a non-rectangular slice excitation profile [SP]), were observed to affect the outcome of MRF measurements[14-17] and are often included into the matching reconstruction, either as prior information[15,16] or as additional free dictionary parameters.[14,18,19] However, different MRF sequences and anatomies may be more or less susceptible to a CE. A dedicated analysis of CEs on the accuracy of MRF has been conducted for cardiac MRF sequences, addressing bias because of a non-rectangular SP, in-plane B_1^+ inhomogeneity as well as preparation pulse efficiency.[20] On the other hand, tissue properties not included in the EPG signal model may lead to deviations of measured from modeled signal evolutions. Examples are diffusion or magnetization transfer (MT). Depending on the spoiling gradient strength, diffusion was found to explain bias on MRF T_2 estimates, most pronounced when spoiling gradient moment, T_2 value and ADC were high.[21] In the presence of a semi-solid pool, MT introduced bias on the T_1 and T_2 relaxation times that increased when additional off-resonant pulses were included in the MRF sequence design.[22] Along with physical effects, the presence of noise may have an impact on both accuracy and precision of MRF results. Whereas electronic noise of constant amplitude is present in all MRI data, aliasing artefacts additionally overlay undersampled spiral MRF data. Up to a certain amplitude, the inner product matching used during reconstruction can filter out noise and even the aliasing artefacts; however, they may also act as an additional confounder.[23,24]

Because strong dielectric effects are present in the female breast,[25] especially in-plane B_1^+ inhomogeneity is expected to induce bias on the MRF results. In fibroglandular tissue (FGT), moreover, diffusion may influence the measured MRF signal evolutions depending on the employed spoiling gradient moment, because an ADC of $(1.95 \pm 0.24) \times 10^{-3}$ mm^2/s is reported for healthy FGT,[26] being rather high in comparison to other tissues.[27,28] Concerning MT in the breast, the (qualitative) MT ratio has been assessed in different breast tissues including tumors.[29,30] There is also a first report on (quantitative) pool size ratios in the healthy breast.[31]

We have described breast MRF measurements on 6 healthy volunteers before,[32] in which MRF data were acquired at 1.5T using a 2D MRF-Dixon sequence. Adding to these results, the current study aims to theoretically investigate the potential bias because of in-plane B_1^+ inhomogeneity, SP, and diffusion effects by means of simulation and to evaluate their

importance with respect to the experimental MRF estimates in the breast. Additionally, the impact of both electronic noise and aliasing artefacts will be included in the analysis.

2 | METHODS

2.1 | MRF-Dixon measurements

MRF measurements in the female breast were performed on a 1.5T Achieva System (Philips, Best, The Netherlands) with a 4-channel breast coil (InVivo, Gainesville, FL), using a 2D single-slice MRF-Dixon sequence with spiral readout.[32] Data were acquired in 6 healthy female volunteers, age range 24-31 years, in accordance with the institutional ethics board. Spiral imaging leads to off-resonance blurring, being more severe for fatty tissue (FT) than for FGT because of the fat chemical shift. Therefore, the MRF-Dixon sequence acquires 3 MRF trains separated by a delay time (t_d) and differing solely in their echo time (TE). Here, the 1st/2nd/3rd flip angle (FA) train exhibited a constant TE = (4.6/6.9/9.2) ms and a TR of 20 ms. Likewise, water–fat separation and subsequent deblurring of spiral off-resonance artefacts were conducted before the MRF matching reconstruction. As depicted in Figure 1A, the MRF sequence used a FA train of 500 TR intervals preceded by a 180° inversion pulse that has shown good encoding capability.[33] The present study re-evaluates the 6 undersampled MRF measurements (ie, 1 data set per volunteer), which were realized with one rotating spiral interleave of 7 ms acquisition time and 20-fold undersampling. $N_{int} = 20$ spiral interleaves would be needed to fully sample *k*-space. Table 1 summarizes the sequence parameters used.

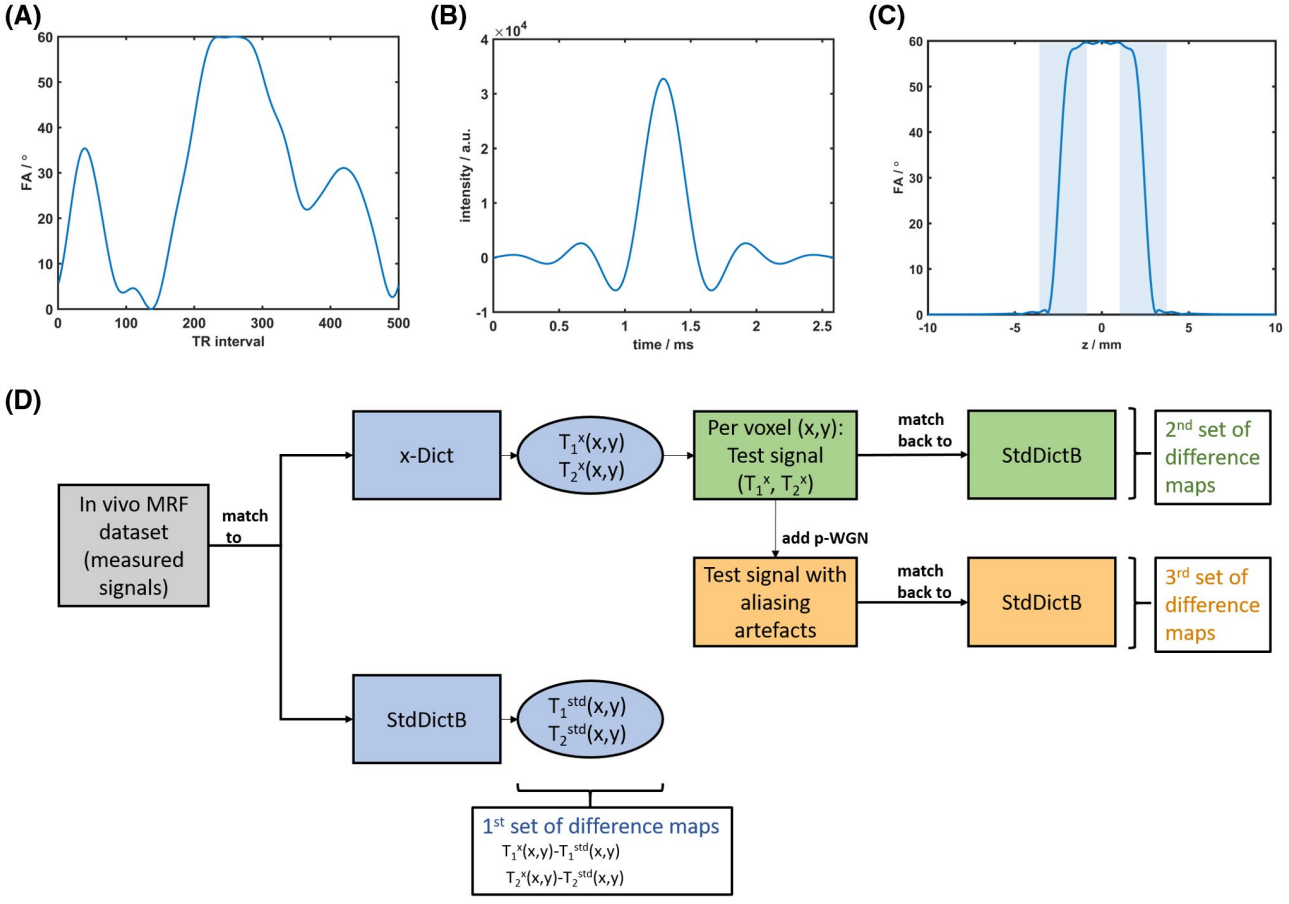

FIGURE 1 (A) Flip angle train of the employed MRF sequence, which was preceded by a 180° inversion prepulse (not depicted).[34] (B) Employed RF pulse shape and (C) resulting slice profile, shown for a flip angle of 60°. (D) Schematics explaining the in vivo bias analysis of the breast MRF data sets. Here, "x" stands for either one of the analyzed CEs (ie, in-plane B_1^+ inhomogeneity, the non-rectangular slice profile, or diffusion). "x-Dict," the dictionary in which CE x was included; "StdDict," the standard dictionary. Per C, 3 sets of difference maps were computed: (1) the experimentally observed T_1 and T_2 difference map between matching to the x-Dict and the StdDict, (2) the difference maps predicted by simulation (without including any noise or artefacts), and (3) the difference maps predicted by simulation including spiral aliasing artefacts. Spiral aliasing artefacts were simulated based on periodic white Gaussian noise (p-WGN) scaled to the signal strength

TABLE 1 Sequence parameters for the in vivo breast MRF measurements

Sequence parameter	Value
Number of *TR* intervals	500
$TE_1/TE_2/TE_3$	4.6/6.9/9.2 ms
TR	20 ms
Spiral T_{acq}	7 ms
Undersampling	20-fold
FOV	430 cm
Matrix	224×224
Voxel size	$1.9 \times 1.9 \times 5$ cm^3
Delay time t_d	7.5 seconds
Spoiling moment	$4 \times 2\pi$ along slice direction
Scan time	53 seconds

2.2 | Theoretic bias analysis including noise and aliasing artefacts

To analyze the theoretic amount of bias that a CE adds to the MRF dictionary matching reconstruction, a standard dictionary (StdDict) was calculated using EPG simulations. Signal evolutions within the StdDict take into account the tissue relaxation times T_1 and T_2, dephasing because of spoiler gradients (G), the flip angle series (FA(t)) and the timings TR and TE of the MRF sequence; however, all of the investigated CEs are excluded. Signals in the StdDict covered a T_1 range of [5, 2000] ms in steps of 5 ms and a T_2 range of [2, 200] in steps of 2 ms. Unphysical signals with $T_1 < T_2$ were not included. Subsequently, test signals were calculated over a grid of values T_1^{test} and T_2^{test} with spacing $\Delta T_1^{test} = 40$ ms and $\Delta T_2^{test} = 20$ ms, respectively, including either B_1^+, SP, or diffusion effects, explained below. Each test signal was matched to the StdDict based on maximum inner product matching.[11] By plotting the deviations of the test signal's nominal T_1 (T_2) value from the matched T_1 (T_2) value, ie, ($T_1^{matched} - T_1^{test}$) and ($T_2^{matched} - T_2^{test}$), respectively, we assessed the expected theoretic matching error that results from omitting a CE. To test different noise conditions, the above analysis was conducted three times. First, the test signals were computed without adding noise. Electronic noise or spiral aliasing artefacts were then added to the test signals before matching. A noisy signal $\boldsymbol{m}$ was simulated by adding a noise vector $\boldsymbol{n}$ to a simulated "true" signal $\boldsymbol{s}$where $(\mathbf{m}, \mathbf{n}, \mathbf{s}) \in \mathbb{C}^{1\times N}$ are defined in image space, with N the number of TR intervals of the MRF sequence:

$$\mathbf{m} = \mathbf{s} + \mathbf{n}. \tag{1}$$

In turn, $\boldsymbol{n}$ may be extracted from a measured signal $\boldsymbol{m}$ by subtracting the best-matching signal $\boldsymbol{s}$. Following,[34] we define SNR for a noisy MRF signal as

$$r = \frac{|\mathbf{s}|_2}{|\mathbf{n}|_2} = \frac{|\mathbf{s}|_2}{|\mathbf{m} - \mathbf{s}|_2}, \tag{2}$$

with the complex norm $|\ |_2$. Before calculating r, $\boldsymbol{m}$ and $\boldsymbol{s}$ are normalized to 1. Given a noise vector $\mathbf{n}' \in \mathbb{C}^{1\times N}$ of arbitrary magnitude, we obtain a noisy test signal $\boldsymbol{m}$ with adjustable r as[34]

$$\mathbf{m} = \mathbf{s} + \frac{1}{r}\frac{\mathbf{n}'}{|\mathbf{n}'|_2}. \tag{3}$$

To investigate the effects of electronic noise, we set $\mathbf{n}'$ to be complex white Gaussian noise (WGN). Test signals with electronic noise levels between $r = 10$ and $r = 50$ were simulated. For each test signal, adding noise and matching was repeated 200 times, so that a mean matched T_1 (T_2) value and SD were obtained to predict accuracy and precision.

To investigate the effects of spiral aliasing artefacts that resemble noise-like fluctuations around the true signal evolutions,[11] a different "noise" model was chosen for $\boldsymbol{n}'$. In the literature, such aliasing artefacts have been mimicked by adding complex WGN scaling with the signal strength to simulated MRF signals.[23,35] We further adapted this artefact model to our undersampled MRF data: because the artefact patterns exhibit a periodicity of $N_{int} = 20$, caused by the rotation of the spiral interleaf with $(360/N_{int})°$ between successive TR intervals, we defined periodic WGN vectors before scaling them with the signal strength. More specifically, we composed a periodic noise vector $\mathbf{n^p} \in \mathbb{C}^{1\times N}$ in image space of 25 replicates of a shorter complex WGN pattern of length 20. Then, we set

$$\mathbf{n}' = \mathbf{s} \odot \mathbf{n^p}, \tag{4}$$

where $\odot$ denotes the Hadamard product. In the context of spiral aliasing artefacts, we will refer to r as the "signal-to-artefact ratio." Test signals with signal-to-artefact ratios between $r = 1$ and $r = 5$ were simulated. Again, adding aliasing artefacts and matching was repeated 200 times for each test signal to predict accuracy and precision.

2.3 | Experimental bias analysis

The experimental bias analysis on in vivo breast MRF data was conducted by means of difference maps as summarized in Figure 1D. In addition to the StdDict described above, we calculated one dictionary per CE, ie, an in-plane B_1^+ dictionary (B_1^+Dict), SP dictionary (SPDict), and diffusion dictionary (DiffDict), using the same range and resolution as StdDict, but including the respective CE into the signal model as described below. In addition, a dictionary combining SP and in-plane B_1^+ effects (SP&B_1^+Dict) was calculated. To evaluate the experimentally observed bias, the breast MRF data sets were matched to the StdDict as well as to the B_1^+

Dict, SPDict, SP&B_1^+Dict, and DiffDict, respectively. T_1 and T_2 difference maps comparing the 2 MRF reconstructions were computed, based on which mean difference and SD in regions of interest (ROIs) of FGT and FT were evaluated for all 6 volunteers. To define the ROIs, maps of the mean water ($\overline{W}$) and fat signal ($\overline{F}$) over the MRF train were retrieved and the mean fat signal fraction $\overline{FSF} = \overline{F} / \left(\overline{W} + \overline{F}\right)$ was computed. Voxels with a $\overline{FSF}$ of <15% were counted as FGT and voxels with a $\overline{FSF}$ > 85% were counted as FT. For the in-plane B_1^+ analysis, mean T_1 and T_2 values and SDs were separately calculated for the left and right breast. Left–right differences were compared between StdDict and B_1^+Dict reconstructions.

To compare the experimentally observed bias with the predictions from simulation, we calculated a second set of difference maps. Potential differences between the experimentally observed bias and the predicted bias may point out an influence of noise or the simultaneous presence of another CE. To predict the in-plane B_1^+ difference map, for example, the respective (T_1, T_2, B_1^+) combination resulting from the reconstruction with the B_1^+Dict was used to compute a noise-free test signal for each voxel, which was subsequently matched back to the StdDict to obtain the predicted bias for that voxel (and analogously for the other CEs). To further investigate whether any observed discrepancies are attributable to spiral aliasing artefacts, a third set of difference maps was calculated by adding, once for each voxel, simulated aliasing artefacts to the test signal before matching it back to the StdDict. Test signals including aliasing artefacts were simulated using periodic complex scaled WGN as described by Equations (3) and (4). The signal-to-artefact ratio in every voxel was estimated beforehand from the data using Equation (2).

2.4 | EPG-simulations including confounding effects

In the following, we outline briefly how each of the CEs was included into the EPG signal calculations.

2.4.1 | In-plane B_1^+ inhomogeneity

In-plane B_1^+ inhomogeneity causes deviations from the nominal FA train

$$FA\,(x, y, t) = f_{B1+}\,(x, y) \cdot FA_{\text{nom}}(t), \tag{5}$$

ie, it acts as a multiplicative factor $f_{B_1^+}$ on the nominal flip angle train. The correction factor in a voxel (x, y) may be measured using a B_1^+ mapping technique. Along with the MRF scans, we acquired B_1^+ maps of the MRF slice using 3D actual flip angle imaging (AFI) with FOV, slice thickness, and resolution matching the MRF slice (scan duration = 3:40 minutes).[36] The B_1^+Dict included correction factors $f_{B_1^+}$ between 0.7 and 1.3 with a step size of 0.025. In comparison to the StdDict, its size and computation time therefore increased by a factor of 25. Matching the measured MRF signal of a voxel to the B_1^+Dict was restricted to the subset of simulated signal evolutions with $f_{B_1^+}$ closest to the measured correction factor in that voxel.

2.4.2 | Slice profile effects

A non-rectangular SP results in a flip angle distribution along the slice direction z for each RF excitation. To minimize this effect during the in vivo breast MRF measurements, an RF pulse shape of high time-bandwidth product (TBW = 10.2) was used (Figure 1B). For each FA, the SP was calculated by Shinnar–Le Roux transformation[37] of the RF pulse shape. Figure 1C shows the resulting excitation profile for a FA of 60°, which is indeed close to, but not perfectly block-shaped. To account for the SP in the MRF dictionary, the z-axis was subdivided into $n_b = 77$ discrete bins. The bin size was chosen larger in regions of constant or slowly varying SP amplitude to reduce computation time. One MRF signal was first calculated for each bin. Subsequently, a weighted sum of signals over the slice was performed, with weighting factors taking into account the individual bin sizes. The SP&B_1^+Dict additionally included the correction factor $f_{B_1^+}$ on the FA train in each bin. Although the SPDict (SP&B_1^+Dict) is of equal size than the StdDict (B_1^+), the respective computation time increased by n_b, (or by $(n_b - 1)/2 + 1$ for a symmetric SP), which required parallel computing.

2.4.3 | Diffusion effects

The diffusion sensitivity of any MRI sequence increases when adding de- and rephasing gradients, in between which microscopic motion of nuclear spins leads to an additional signal decay.[38] For isotropic diffusion in 1D, the associated signal attenuation S/S_0 is commonly expressed as[38,39]

$$\frac{S}{S_0} = \exp\left(-b \cdot D\right), \tag{6}$$

with the isotropic diffusion constant D b takes the effect of all diffusion gradients over 1 TR into account

$$b = \int_0^{TR} k(t)^2 dt. \tag{7}$$

Here, $k(t) = \gamma \int_0^t G(\tau) d\tau$ is the gradient moment and γ is the gyromagnetic ratio in rad/(s × T). Gradient-spoiled

MRF sequences commonly exhibit 1 unbalanced spoiler gradient per TR interval. Much like the diffusion-sensitizing gradients in conventional diffusion-weighted sequences, the consecutive spoiler gradients within the MRF sequence de- and rephase the spin system and thereby add diffusion weighting to MRF signals. Numeric equivalents of Equations (6) and (7) can be included in the EPG framework as a diffusion operator

$$\mathbf{D}=\begin{pmatrix} \exp\left(-b^T\cdot D\right) & 0 & 0 \\ 0 & \exp\left(-b^T\cdot D\right) & 0 \\ 0 & 0 & \exp\left(-b^L\cdot D\right) \end{pmatrix}, \quad (8)$$

with b-values $b_i^L=(i\cdot\Delta k)^2\cdot T$ and $b_i^T=\left(\left(i+0.5\cdot gr_{\text{on}}\right)^2+\frac{gr_{\text{on}}}{12}\right)\cdot\Delta k^2\cdot T$. Here, i: order of de-/rephasing state (with $i>0$ for dephasing and $i<0$ for rephasing states), $\Delta k=\gamma GT$: amount of de- and rephrasing introduced by the application of one spoiler gradient, D: gradient strength, T: gradient duration, gr_{on}: 1 if gradient is being played out, 0 otherwise.[13,40] The spoiler gradients of our breast MRF sequence exhibited a gradient moment of $\Delta k = 4\times 2\pi$ over the 5 mm slice per TR. We assumed $D = 1.95\times 10^{-3}$ mm^2/s for FGT.[26] We neglected diffusion in FT, because diffusion constants are reported to be ~2 orders of magnitude smaller than in aqueous tissues.[41] The DiffDict is of equal size and similar computation time as the StdDict.

3 | RESULTS

3.1 | Theoretic bias analysis

Figures 2 and 3 show the theoretic bias analysis for T_1 and T_2, respectively. The difference plots show, for the 3 investigated CEs and our MRF sequence, ($T_1^{\text{matched}}-T_1^{\text{test}}$) and ($T_2^{\text{matched}}-T_2^{\text{test}}$) over the grid of investigated test signals, with T_1^{matched} and T_2^{matched} resulting from matching the test signals to the StdDict. Hence, these plots predict the impact of a CE on MRF accuracy. The analysis was conducted (1) without noise, (2) with electronic noise of SNR $r = 50$ that we selected as a representative value for a fully sampled MRF scan, and (3) with aliasing artefacts of $r = 1.5$, resembling the experimentally observed signal-to-artefact ratios. The control plots comprise test signals without any CE, but with added noise or aliasing artefacts, and therefore show bias introduced by noise/aliasing artefacts alone. Supporting Information Figure S1 presents the SD of T_1 and T_2 over the 200 matching steps per test signal as a measure of the predicted precision when including electronic noise or spiral aliasing artefacts along with a CE. Supporting Information Figure S2 presents the predicted accuracy and precision for more signal-to-noise/signal-to-artefact ratios.

For 2 example correction factors of $f_{B_1^+}=0.9$ and $f_{B_1^+}=1.1$, corresponding to actual FAs smaller and larger than the nominal ones, it is visible that omitting in-plane B_1^+ effects in the MRF dictionary leads to T_1 under-/T_2 overestimation if $f_{B_1^+}<1$ and to T_1 over-/T_2 underestimation if $f_{B_1^+}>1$. Omitting the SP during dictionary calculation results in T_1 under-/T_2 overestimation. For the employed spoiling gradient moment of $4\times 2\pi$ over the slice, omitting diffusion in the MRF dictionary causes T_1 underestimation toward longer T_1 and T_2, as well as T_2 underestimation toward longer T_2. Yet, for the here-investigated spoiling moment, diffusion constant D and dictionary resolution, an effect of diffusion becomes visible from $T_2 > 100$ ms onward only. Simulation results for more values of $f_{B_1^+}$ and D are presented in Supporting Information Figure S3.

Adding WGN with an SNR of $r = 50$ to the test signals has a smoothing effect on the T_1 and T_2 bias prediction, but overall leaves the mean predicted T_1 and T_2 values unchanged. As visible from the control plots, matching accuracy is not affected. Adding aliasing artefacts with $r = 1.5$ to the test signals has a slightly stronger smoothing effect on both T_1 and T_2 and, moreover, adds some granularity to the T_1 bias predictions. However, the corresponding control plots in Figure 2 reveal no systematic effect, but rather a fluctuation ~0. It should be noted that the already small bias caused by diffusion is hardly noticeable anymore after aliasing artefacts were added. As shown in Supporting Information Figures S1 and S2, precision is more strongly affected for T_1 than for T_2. For spiral aliasing artefacts with $r = 1.5$, predicted T_1 SDs are larger than for electronic noise with $r = 50$.

Supporting Information Figure S4 presents simulated MRF signal evolutions of an example test signal for the different CEs. In line with the previous observations, the signal evolutions differ hardly when including diffusion, but notably when including in-plane B_1^+ inhomogeneity or the SP.

3.2 | Experimental bias analysis

Next, we analyzed the experimental bias on the in vivo MRF results. Figure 4A,B show, for volunteer 5, the MRF T_1 and T_2 map after matching to the StdDict (ie, without accounting for any of the CEs during reconstruction). Figure 4C shows the measured B_1^+ map. A strong left-right inhomogeneity is visible between both breasts, which visibly transmits into the T_2 map in Figure 4B, as marked by arrows indicating regions of $f_{B_1^+}$ smaller/larger than 1 and the corresponding areas of higher/lower T_2 values. Figure 4D,E summarize the effect of a combined correction for in-plane B_1^+ inhomogeneity and the

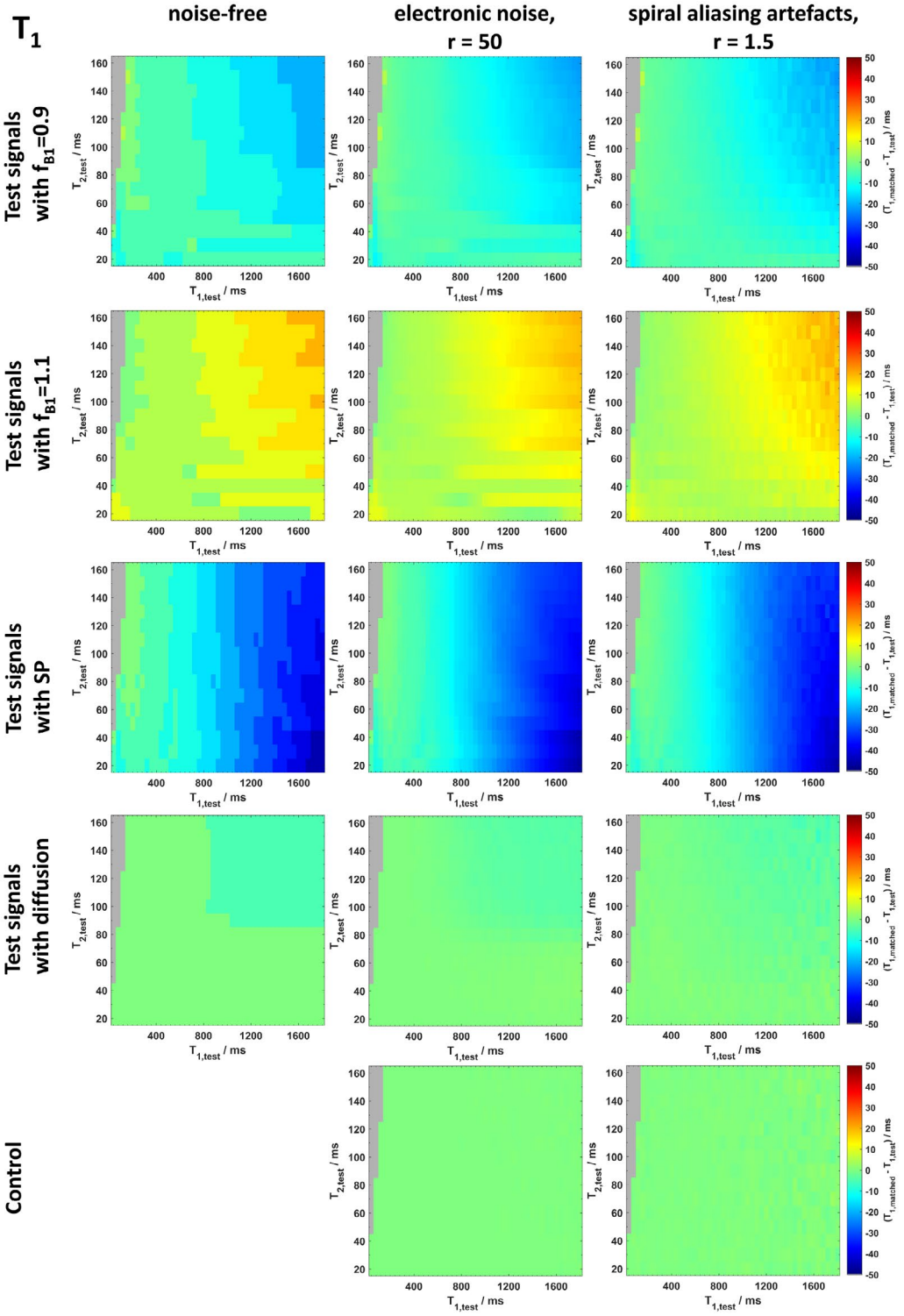

FIGURE 2 Simulated bias on T_1 for different CEs. Test signals including one of the CEs were matched to the StdDict. The plots visualize $T_1^{matched} - T_1^{test}$, ie, the bias that results from omitting a CE from the MRF dictionary, over a grid of test signals with different relaxation times T_1^{test} and T_2^{test}. The first and second row present the expected bias on T_1 because of in-plane B_1^+ effects, for 2 example correction factors ($f_{B_1^+} = 0.9$ and $f_{B_1^+} = 1.1$), corresponding to under- or overestimation of the nominal flip angle train, respectively. The third row shows the expected bias on T_1 because of slice profile effects. In the fourth row, the expected bias on T_1 because of diffusion effects is shown for a diffusion coefficient of $D = 1.95 \times 10^{-3}$ mm^2/s. The 3 columns correspond to different noise conditions. While the test signals are left noise-free in the first column, white Gaussian noise with an SNR of $r = 50$ was superimposed for the second column, and spiral aliasing artefacts with a signal-to-artefact ratio of $r = 1.5$ were superimposed for the third column before matching the test signals to the StdDict. Adding electronic noise/spiral aliasing artefacts was repeated 200 times. The plots therefore show the mean bias. For the control plots in the fifth row, no confounding effects have been included to test the effect of electronic noise or spiral aliasing artefacts alone on the accuracy of the matching procedure

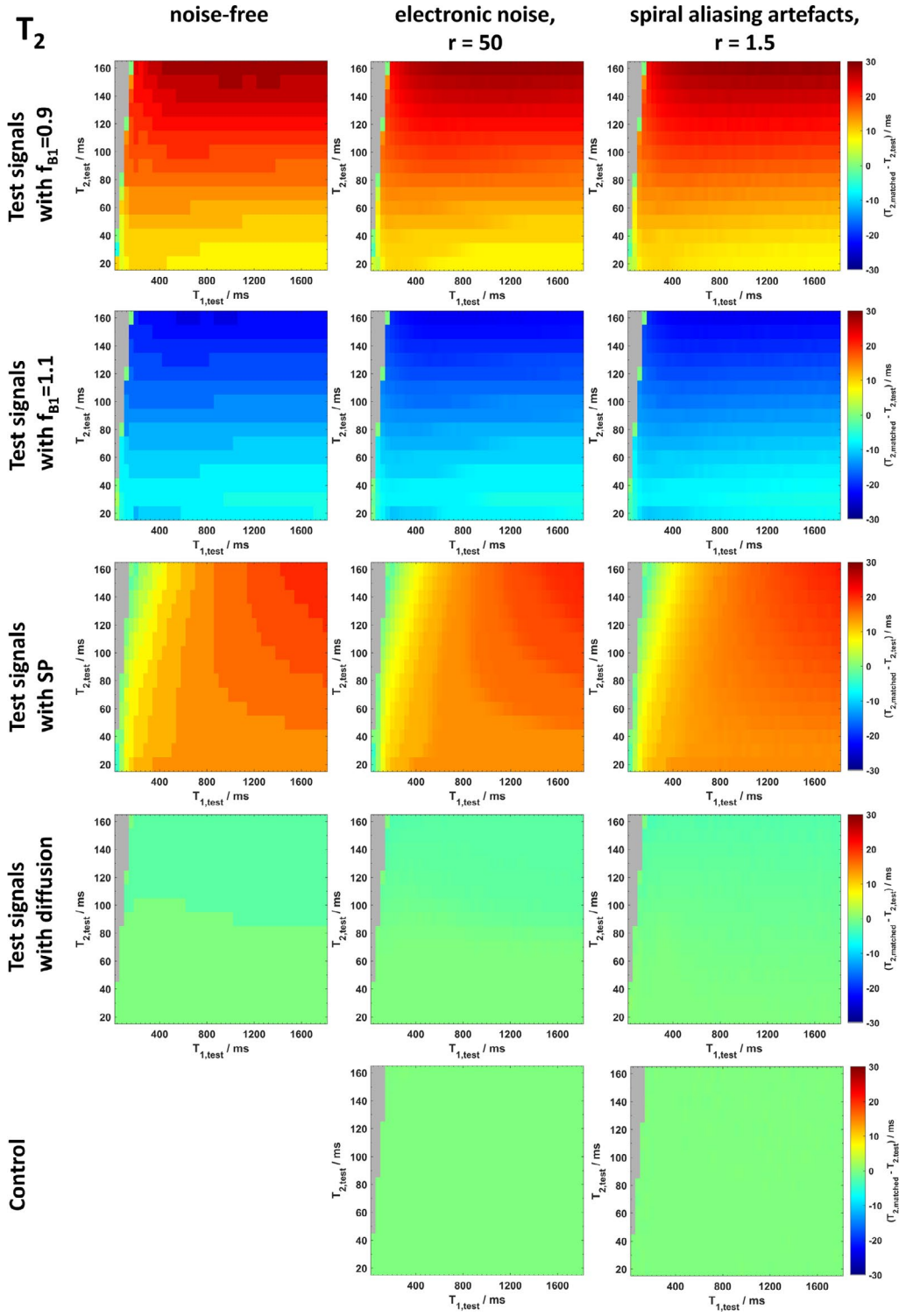

FIGURE 3 Simulated bias on T_2 for different CEs. Test signals including one of the CEs were matched to the StdDict. The plots visualize $T_2^{matched}-T_2^{test}$ (ie, the bias that results from omitting a CE from the MRF dictionary) over a grid of test signals with different relaxation times T_1^{test} and T_2^{test}. Similarly to Figure 2, the first and second row present the expected bias on T_2 because of in-plane B_1^+ effects, for 2 example correction factors ($f_{B_1^+} = 0.9$ and $f_{B_1^+} = 1.1$), corresponding to under- or overestimation of the nominal flip angle train, respectively. The third row shows the expected bias on T_2 because of slice profile effects. In the fourth row, the expected bias on T_2 because of diffusion effects is shown for a diffusion coefficient of $D = 1.95 \times 10^{-3}$ mm^2/s. Again, the 3 columns correspond to different noise conditions. Although the test signals are left noise-free in the first column, white Gaussian noise with an SNR of $r = 50$ was superimposed for the second column, and spiral aliasing artefacts with a signal-to-artefact ratio of $r = 1.5$ were superimposed for the third column before matching the test signals to the StdDict. Adding electronic noise/spiral aliasing artefacts was repeated 200 times. The plots therefore show the mean bias. In analogy to Figure 2, no confounding effects have been included for the control plots in the fifth row, to test the effect of electronic noise or spiral aliasing artefacts alone on the accuracy of the matching procedure

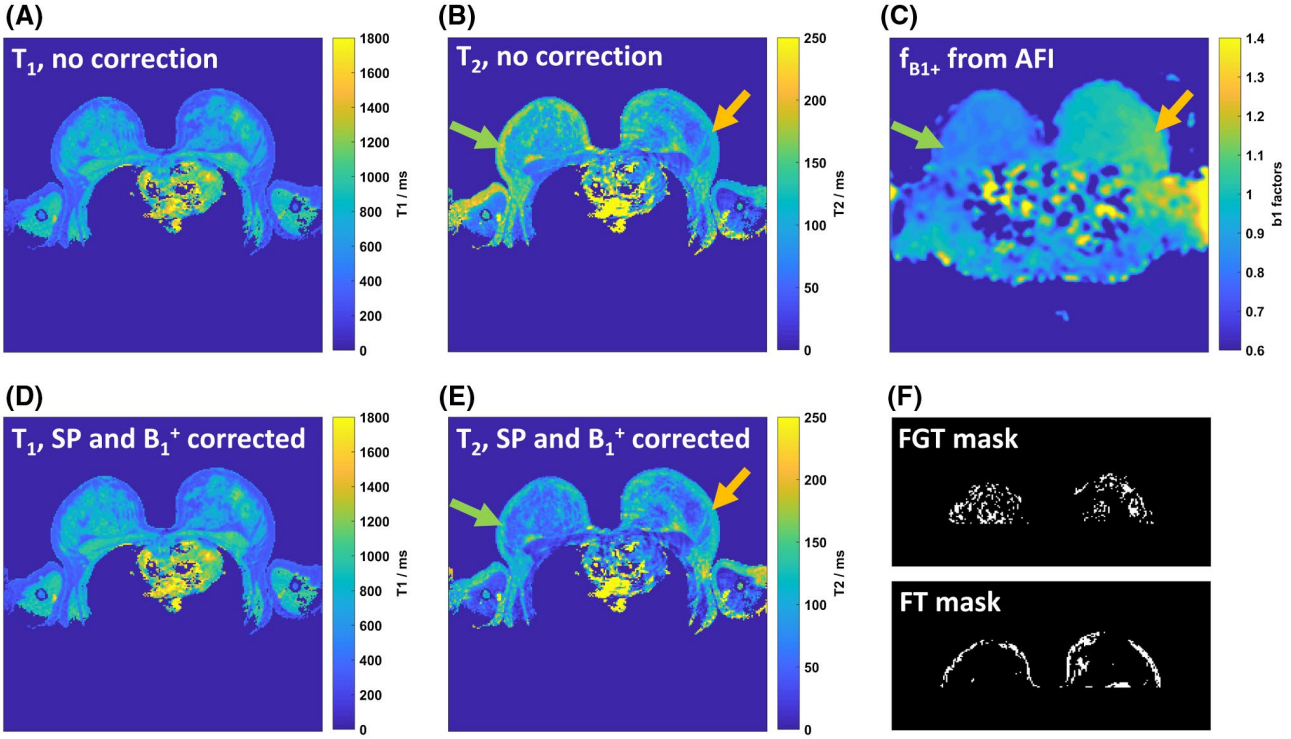

FIGURE 4 Measured quantitative (A) T_1 and (B) T_2 map, obtained by matching the measured MRF data to the StdDict. The StdDict excludes any of the 3 investigated CEs (ie, they are not corrected for during the MRF dictionary matching). (C) In-plane B_1^+ correction factors as obtained from the AFI B_1^+ measurement. A pronounced left–right inhomogeneity is visible, which causes bias in the uncorrected T_2 map (B). The orange and green arrow visualize this: the former one points at a region in the left breast where $f_{B_1^+} > 1$, while the latter one points at a region in the right breast where $f_{B_1^+} > 1$. In result, lower T_2 values are found in the left breast than in the right one before in-plane B_1^+ correction is applied. (D) and (E) show the T_1 and T_2 maps after correction for SP and in-plane B_1^+ effects. Including in-plane B_1^+ into the MRF dictionary results in a T_2 map of improved left-right homogeneity. (F) Segmentation of FGT and FT regions, which have been used throughout the in vivo bias analysis

non-rectangular SP, which permits to compare the corrected T_1 and T_2 map with their uncorrected counterparts in Figure 4A,B. Next to the improved in-plane B_1^+ homogeneity, overall lower T_2 values are visible in the corrected T_2 map (Figure 4D). T_1 maps before and after correction look similar. Figure 4F shows the ROI masks for segmentation of FGT and FT. Table 2 presents mean inner product (IP) values and SDs in FGT and FT regions that result from matching the data set to the different dictionaries. IPs closer to 1 indicate a better matching quality. Although changes in mean IP values between different dictionaries are subtle, they increase when including in-plane B_1^+ or the SP into reconstruction. The SP&B_1^+Dict fits the measured signal evolutions best.

Difference maps between MRF results with and without including a CE "x" in the MRF dictionary, ie, ($T_1^x - T_1^{std}$) and (T_2^x-T_2^{std}), are shown in the first column of Figures 5 and 6 for T_1 and T_2, respectively. The second column predicts, based on the observed values T_1^x and T_2^x, the difference map expected from the theoretic simulations. For the third column, the prediction additionally includes spiral aliasing artefacts.

3.2.1 | In-plane B_1^+ inhomogeneity

As predicted by simulations, factors $f_{B_1^+} < 1$ result in an overestimation of T_2 and vice versa when reconstructing with the StdDict instead of the B_1^+Dict, whereas the inverse is seen to a less strong extent for T_1. When comparing the measured T_1 difference map to the 2 predicted ones, the predicted bias without aliasing artefacts is similar, but slightly less important than the measured bias. Adding aliasing artefacts results in a more noisy appearance. For T_2, including aliasing artefacts has no visually obvious effect. Local minima in the right breast of the experimental difference map (Figure 6 marked by white arrows) are not reproduced by any of the predicted difference maps and are therefore not attributable to the presence of spiral aliasing artefacts. Figure 7 compares the mean $\overline{T_1}$ and $\overline{T_2}$ values within FGT and FT resulting from reconstruction with either the StdDict or the B_1^+Dict. Mean values were separately calculated for the left (L) and right (R) breast. For all volunteers, differences between T_1^L and T_1^R are barely altered when considering in-plane B_1^+ inhomogeneity,

TABLE 2 Mean inner product (IP) values and standard deviations in masked areas of FGT and FT, obtained for the presented data set of volunteer 5 after matching to the different dictionaries

	IP in FGT	IP in FT
StdDict	0.5325 ± 0.1546	0.7856 ± 0.1253
DiffDict	0.5325 ± 0.1546	0.7856 ± 0.1254
B_1^+Dict	0.5327 ± 0.1550	0.7857 ± 0.1252
SPDict	0.5335 ± 0.1549	0.7863 ± 0.1256
SP&B_1^+Dict	0.5337 ± 0.1554	0.7864 ± 0.1256

which is in line with the overall small T_1 differences observed in the top left map in Figure 5. Averaged over all 6 volunteers, the absolute difference $|\overline{T_1^L} - \overline{T_1^R}|$ changes from 42 to 40 ms in FGT (mean relative difference stays at 4%) and from 13 to 18 ms in FT (mean relative difference changes from 4% to 6%), if reconstructing with the B_1^+Dict instead of the StdDict. Here, we calculated the mean relative difference for each volunteer as $\left|\overline{T_1^L} - \overline{T_1^R}\right| / \overline{T_1^{\text{both}}}$, (ie, the absolute left–right difference was normalized to the mean value calculated over both breasts). Averaged over all 6 volunteers, the absolute difference $\left|\overline{T_2^L} - \overline{T_2^R}\right|$ is reduced from 29 to 5 ms in FGT (mean relative difference changes from 38% to 8%) and from 29 to 9 ms in FT (mean relative difference changes from 21% to 7%). A separately conducted phantom study, presented in Supporting Information Figure S5, showed that reconstructing with the B_1^+Dict improved both the accuracy of the MRF T_2 estimates with respect to reference values and the repeatability of the MRF experiment on repositioning. MRF T_1 estimates were barely affected by the in-plane B_1^+ correction. These findings support the in vivo analysis presented here.

3.2.2 | Slice profile effects

Matching to the SPDict resulted in increased T_1 values and smaller T_2 values than matching to the StdDict. The predicted difference maps without aliasing artefacts look more homogeneously structured than the measured ones. Adding aliasing artefacts results in more noisy difference maps of improved resemblance with the experimental ones. Table 3 lists the mean deviations in T_1 and T_2 in areas of FGT and of FT and their SD for all volunteers. With average relative deviations of (2, −22, 2, −6)% for ($T_{1,FGT}$, $T_{2,\,FGT}$, $T_{1,FT}$, $T_{2,FGT}$), SP corrections affected T_2 more strongly than T_1 and FGT more strongly than FT. This is in line with the theoretic analysis, which predicted a stronger SP effect on T_2 for tissues with higher T_1 (Figure 3). For comparison, difference maps visualizing the combined effect of in-plane B_1^+ inhomogeneity and the SP are shown in the third row. Qualitatively, the two CEs seem to add up.

3.2.3 | Diffusion effects

Apart from some noisy pixels that are visible in the experimentally observed T_1 difference map and, more strongly, also in its predicted counterpart including aliasing artefacts, no difference between matching to the DiffDict and StdDict is visible. From Table 3, matching including diffusion resulted in slightly higher T_2 in areas of FGT than matching to the StdDict, corresponding to average relative deviations of 0.1% for $T_{1,FGT}$ and 1% for $T_{2,FGT}$. Yet, the observed mean differences are within the SDs, which is in line with the simulations.

4 | DISCUSSION

This study explored the extent to which bias from different CEs affected the outcome of MRF measurements in the female breast. In vivo, including in-plane B_1^+ inhomogeneity into the MRF dictionary reduced the left–right asymmetry in T_2 maps of the breast, whereas correcting for slice profile effects yielded decreased T_2 values especially in FGT. For the spoiling moment of $4 \times 2\pi$ over the slice used here, diffusion effects in FGT proved not to be significant, corresponding well to the theoretic predictions.

Our analysis underlines the need to consider CEs during MRF reconstruction. Some effects, such as the shape of the SP, are independent of the scanned subject and can, therefore, readily be included into the dictionary, albeit at an increase in computation time. Other effects like in-plane B_1^+ inhomogeneity commonly need a calibration measurement (ie, the B_1^+ map). The same would have held true for the ADC, if our MRF sequence were sensitive to this effect. Such prescans add to the scan time of MRF, and compromise its main promise: to offer rapid quantitative imaging. Ideally, MRF sequences would allow for inclusion of a CE as a free parameter in the dictionary and result in unambiguous matching. For example, adding rapid variations in the end of the flip angle series reduced matching errors because of in-plane B_1^+ inhomogeneity,[18] and including different coil modes into the MRF framework permitted to estimate in-plane B_1^+ during matching reconstruction at high field strength.[19] However, undersampling may pose a challenge to robust parameter estimation when the number of free parameters augments.

In breast MRI, the potential of the ADC as a native biomarker for the distinction of malignant from benign breast lesions and therapy response[42-46] is increasingly investigated. This encourages the incorporation of the ADC as a free parameter into MRF reconstruction; yet, few MRF studies so far showed successful matching of T_1, T_2, and ADC.[47,48] This challenge is probably present because diffusion acts very similarly to T_2 because of its signal-attenuating nature.[21] Hence, as long as an MRF sequence is not able to distinguish

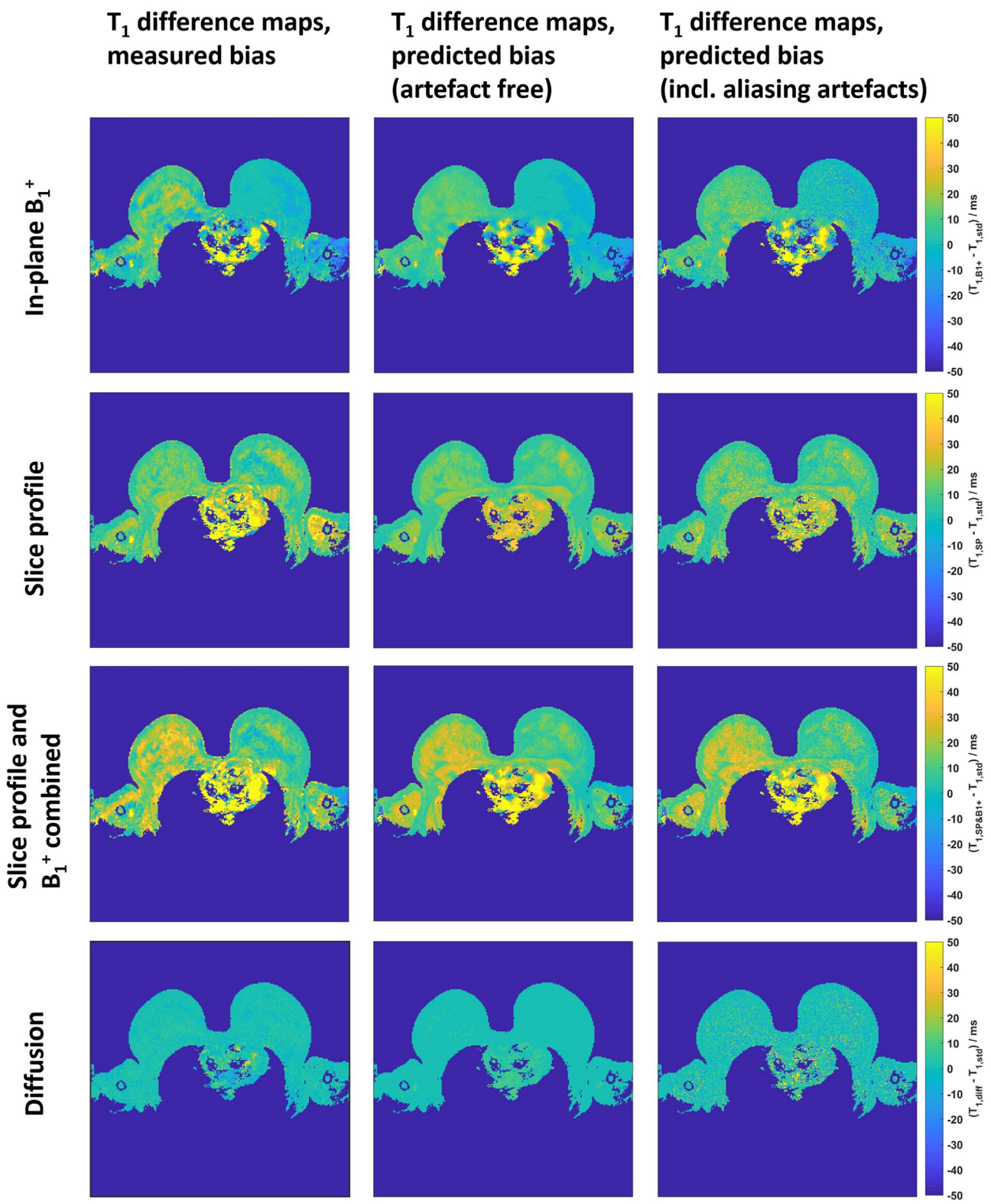

FIGURE 5 Experimental bias analysis for T_1 by means of T_1 difference maps. Presented from top to bottom, the CEs under investigation are in-plane B_1^+ inhomogeneity, the non-rectangular slice profile, the combination thereof, as well as diffusion. The difference maps presented in the first column show the experimentally observed bias. For their calculation, the T_1 map obtained by matching the experimental data to the StdDict was subtracted from the T_1 map obtained by matching the same data set to the dictionary including a CE. The second column shows the bias predicted by simulations. To calculate this second set of difference maps, one test signal per voxel including a CE was generated, with T_1 and T_2 as obtained by matching to the corresponding dictionary, and matching back to the StdDict. The resulting matched T_1 value was subtracted from the test signal's T_1 value to calculate the predicted difference map. For the third column, aliasing artefacts with a signal-to-artefact ratio of $r = 1.5$ were added to the test signals before matching them to the StdDict

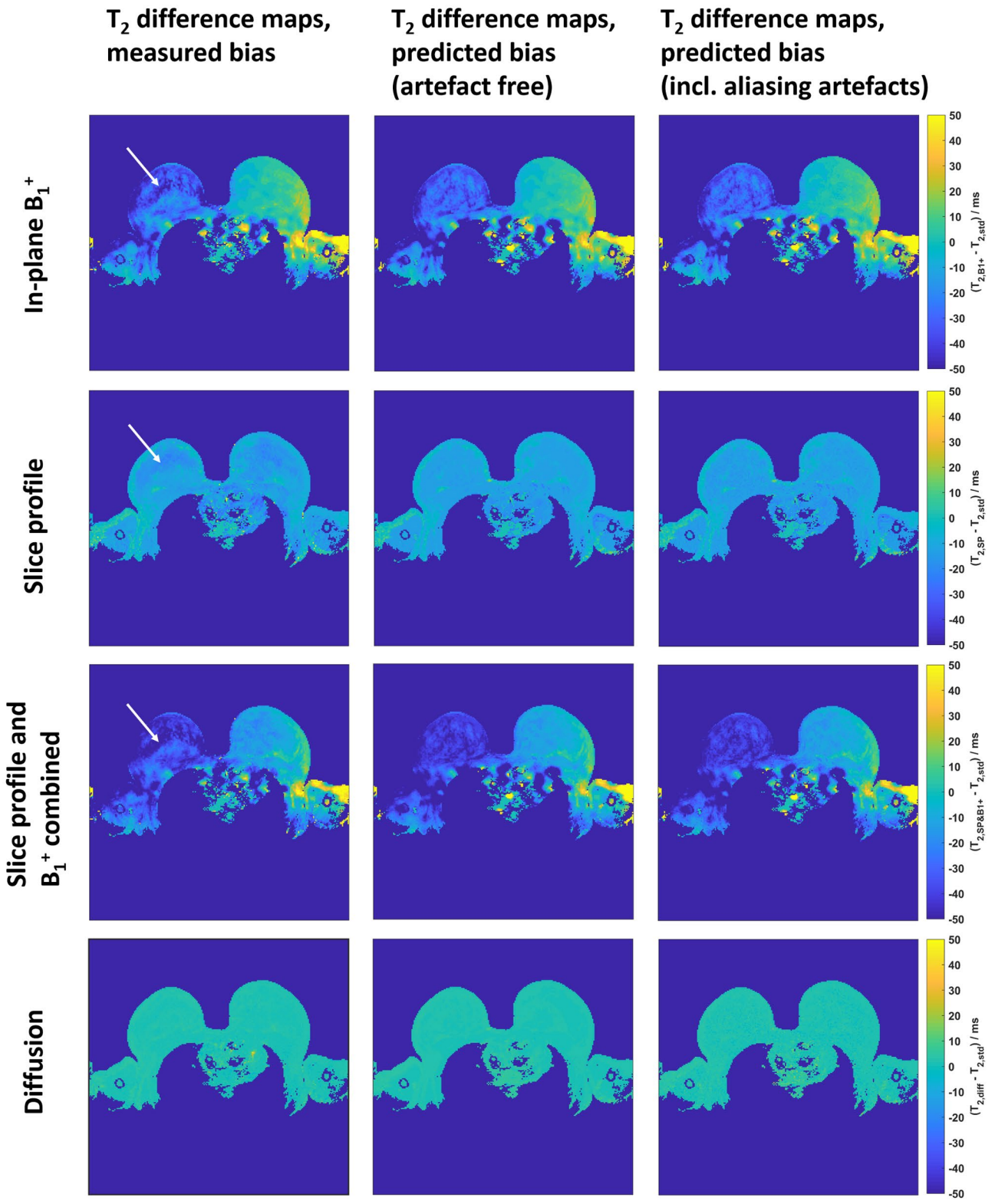

FIGURE 6 Experimental bias analysis for T_2 by means of T_2 difference maps. Presented from top to bottom, the CEs under investigation are in-plane B_1^+ inhomogeneity, SP effects, the combination thereof as well as diffusion effects. The difference maps presented in the first column show the experimentally observed bias. For their calculation, the T_2 map obtained by matching the experimental data to the StdDict was subtracted from the T_2 map obtained by matching the same data set to the dictionary including a CE. The second column shows the bias prediction by simulations. To calculate this second set of difference maps, 1 test signal per voxel including a CE was generated, with T_1 and T_2, as obtained by matching to the corresponding dictionary, and matched back to the StdDict. The resulting matched T_2 value was subtracted from the test signal's T_2 value to calculate the predicted difference map. For the third column, aliasing artefacts with a signal-to-artefact ratio of $r = 1.5$ were added to the test signals before matching them to the StdDict. White arrows indicate minima in the experimental difference maps that are not reproduced by the predicted difference maps

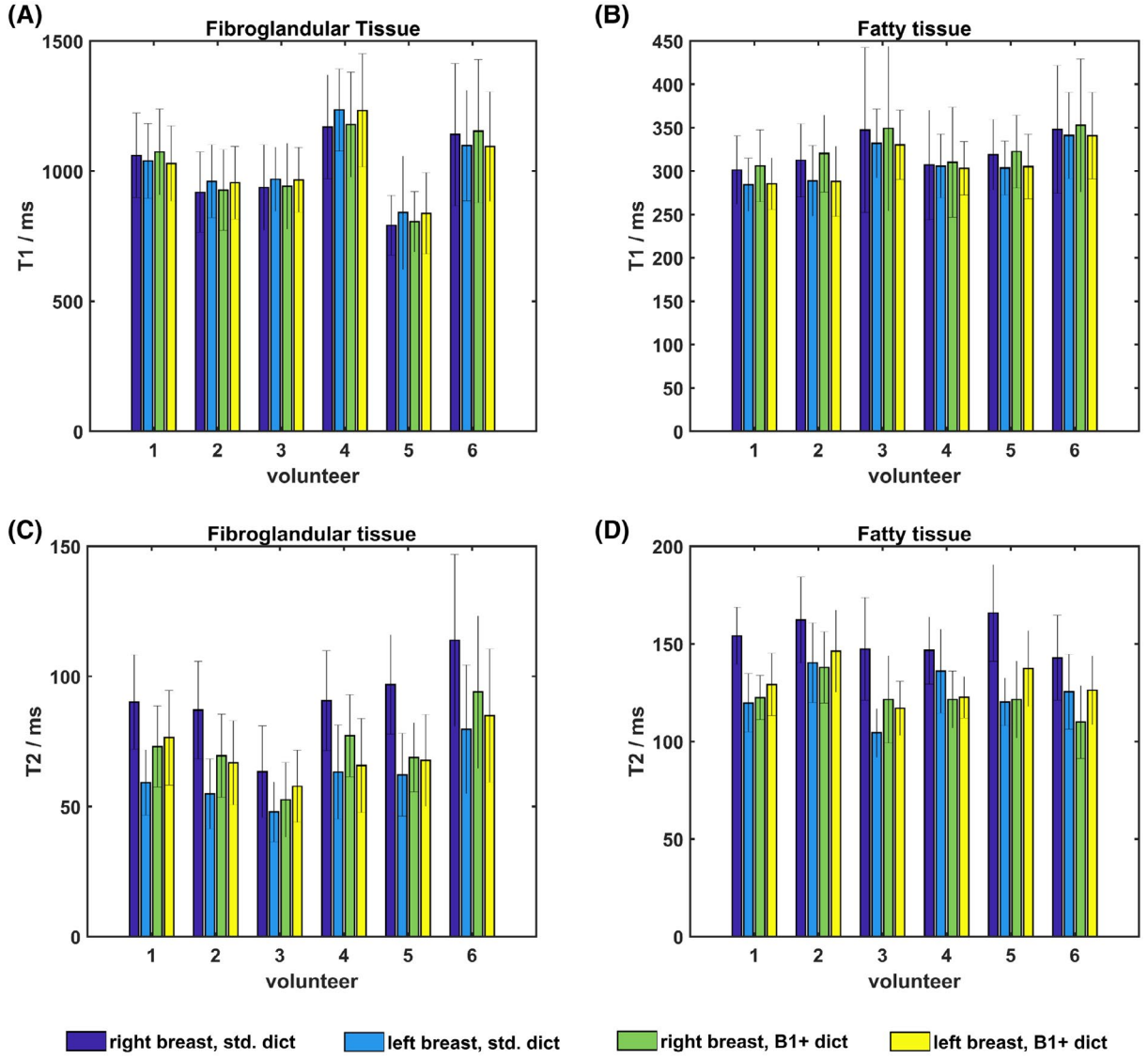

FIGURE 7 Left–right comparison of (A and B) mean T_1 and (C and D) mean T_2 values for FGT (A and C) and FT (B and D). The mean T_1 values and SDs in right and left breast obtained from matching to the StdDict are depicted in dark and light blue, whereas the corresponding values obtained from matching to the B_1^+Dict are depicted in light green and yellow. Correcting for in-plane B_1^+ predominantly affected the mean T_2 values and resulted in improved homogeneity in both fibroglandular and fatty tissue

TABLE 3 Experimental bias because of slice profile and diffusion effects

	Slice profile				Diffusion	
	$\Delta T_{1,FGT}$/ ms	$\Delta T_{2,FGT}$/ ms	$\Delta T_{1,FT}$/ ms	$\Delta T_{2,FT}$/ ms	$\Delta T_{1,FGT}$/ ms	$\Delta T_{2,FGT}$/ ms
Vol 1	23 ± 6	-16 ± 2	2 ± 5	-8 ± 3	2 ± 3	1 ± 1
Vol 2	20 ± 7	-15 ± 2	7 ± 4	-3 ± 4	1 ± 3	1 ± 1
Vol 3	18 ± 7	-15 ± 2	5 ± 3	-9 ± 4	1 ± 3	0.5 ± 0.9
Vol 4	26 ± 8	-15 ± 3	5 ± 3	-6 ± 3	2 ± 3	1 ± 1
Vol 5	16 ± 6	-15 ± 3	4 ± 3	-6 ± 3	1 ± 3	1 ± 1
Vol 6	24 ± 8	-17 ± 3	5 ± 4	-10 ± 3	2 ± 3	1 ± 1

Note: For the slice profile, mean deviations in T_1 and T_2 and their SD were computed in FGT and FT. For diffusion, mean deviations were computed in FGT only.

between diffusion and relaxation effects, little diffusion bias is, at first, a positive attribute for the MRF sequence.

Repeated signal simulations including spiral aliasing artefacts, implemented as periodic complex WGN patterns scaled to the signal strength, suggest that spiral aliasing artefacts have a larger impact on precision than on accuracy of the matched T_1 and T_2 values, while affecting T_1 precision more strongly than T_2. Likewise, the aliasing artefacts themselves may act as a confounder, at least if low signal-to-artefact ratios are present such as in highly undersampled MRF scans. It should be emphasized that our periodic WGN simulations do not equal a full simulation of the k-space encoding based on a given object shape, sampling trajectory and MRF sequence.[24] Using pWGN patterns to simulate spiral aliasing artefacts may, however, give an estimate on the possible strength of spiral undersampling artefacts for a given measurement without knowledge about the exact trajectory design or the object shape. Importantly, the predicted artefacts patterns in our difference maps are spatially uncorrelated, as the noise patterns are generated independently for each voxel. Therefore, approximating spiral aliasing artefacts by pWGN will fail to capture any global spatial patterns that may be occur at high undersampling factors, which is an important limitation. An additional simulation study, comparing the pWGN approximation of spiral aliasing artefacts to a full simulation of k-space encoding in the breast, is available to the interested reader in Supporting Information Figure S6.

The in vivo bias analysis of left–right inhomogeneity caused by in-plane B_1^+ effects was based on mean T_1 and T_2 values, calculated over ROIs containing mostly FGT or FT. Left and right breast were assigned separate ROIs; yet, they enclosed a range of different B_1^+ values and, in the case of FGT, also heterogeneous tissue. Therefore, the calculated mean values comprised voxels with different amounts of bias. This explains why the bar plots in Figure 7 show large SD. Tissue heterogeneity may further be the reason for any residual differences observed between left and right breast FGT. However, we would not have expected left–right discrepancies to remain at least for FT after including in-plane B_1^+ effects into reconstruction, as FT should be relatively homogeneous throughout the breasts. The residual discrepancies might indicate inaccuracies in the AFI B_1^+ mapping method.

We have not analyzed magnetization transfer (MT) effects as a fourth, potentially CE, although there is evidence for MT to be present in FGT.[29-31] In FT, no MT effects are expected.[49] The missing analysis of bias because of MT in this work is because of a lack of quantitative MT literature values for the breast, as next to the fractional pool size also T_1 and T_2 of the semisolid pool as well as the exchange rate between the 2 pools are required for EPG simulations including MT.[50] Importantly, a recent study on MRF and MT has shown that MT is expected to affect MRF signal evolutions.[22] Although our MRF sequence did not include additional off-resonant pulses, the RF excitation pulses are nevertheless of a finite bandwidth and power,. Therefore, we cannot exclude the occurrence of residual MT effects at present and recommend a future bias analysis of MT for MRF in the breast.

In conclusion, we have shown that it is important to correct MRF measurements in the breast for 2 B_1^+-related effects, ie, for in-plane inhomogeneity and the SP shape, while the employed MRF sequence was not sensitive to diffusion. The analysis underlines that MRF has the power to overcome system imperfections, because their effect on the signal evolution can be understood, modeled, and hence, incorporated into MRF reconstruction. Yet, at the same time, it points out a major challenge for MRF: if an MRF sequence is sensitive to a CE, efforts need to be taken to desensitize the sequence—or to optimize it in such a way that the effect is no longer a confounding one, but becomes separable from other dictionary parameters during reconstruction.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 667211. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

M.D., T.A., and P.K. are employees of Philips Research Europe.

ORCID

Teresa Nolte https://orcid.org/0000-0003-3167-7121
Hannah Scholten https://orcid.org/0000-0002-9049-7668
Nicolas Gross-Weege https://orcid.org/0000-0003-3507-6980
Thomas Amthor https://orcid.org/0000-0002-1456-6131
Volkmar Schulz https://orcid.org/0000-0003-1341-9356

REFERENCES

1. Tofts P. *Quantitative MRI of the Brain: Measuring Changes Caused by Disease*. Chichester, UK: John Wiley & Sons Ltd; 2005:650.
2. Seraphim A, Knott K, Augusto J, Bhuva A, Manisty C, Moon JC. Quantitative cardiac MRI. *J Magn Reson Imaging*. 2019;51:693-711.
3. de Mello R, Ma Y, Ji Y, Du J, Chang EY. Quantitative MRI musculoskeletal techniques: an update. *AJR Am J Roentgenol*. 2019;213:524-533.
4. Manton DJ, Chaturvedi A, Hubbard A, et al. Neoadjuvant chemotherapy in breast cancer: early response prediction with quantitative MR imaging and spectroscopy. *Br J Cancer*. 2006;94:427-435.
5. Tan PC, Pickles MD, Lowry M, Manton DJ, Turnbull LW. Lesion T2 relaxation times and volumes predict the response of malignant breast lesions to neoadjuvant chemotherapy. *Magn Reson Imaging*. 2008;26:26-34.
6. Liu L, Yin B, Deng GY, Lu YP, Peng WJ. Changes of T2 relaxation time from neoadjuvant chemotherapy in breast cancer lesions. *Iran J Radiol*. 2016;13:e24014.
7. Chen Y, Panda A, Pahwa S, et al. Three-dimensional MR fingerprinting for quantitative breast imaging. *Radiology*. 2019;290:33-40.

8. Liu LI, Yin BO, Shek K, et al. Role of quantitative analysis of T2 relaxation time in differentiating benign from malignant breast lesions. *J Int Med Res*. 2018;46:1928-1935.
9. Seo M, Ryu JK, Jahng G-H, et al. Estimation of T2* relaxation time of breast cancer: correlation with clinical, imaging and pathological features. *Korean J Radiol*. 2017;18:238-248.
10. Ma D, Gulani V, Seiberlich N, et al. Magnetic resonance fingerprinting. *Nature*. 2013;495:187-193.
11. Jiang Y, Ma D, Seiberlich N, Gulani V, Griswold MA. MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout. *Magn Reson Med*. 2015;74:1621-1631.
12. Scheffler K. A pictorial description of steady-states in rapid magnetic resonance imaging. *Concepts Magn Reson*. 1999;11:291-304.
13. Weigel M. Extended phase graphs: dephasing, rf pulses, and echoes—pure and simple. *J Magn Reson Imaging*. 2015;41:266-295.
14. Buonincontri G, Schulte RF, Cosottini M, Tosetti M. Spiral MR fingerprinting at 7 T with simultaneous B1 estimation. *Magn Reson Imaging*. 2017;41:1-6.
15. Ma D, Coppo S, Chen Y, et al. Slice profile and B1 corrections in 2D magnetic resonance fingerprinting. *Magn Reson Med*. 2017;78:1781-1789.
16. Hong T, Han D, Kim M, Kim D. RF slice profile effects in magnetic resonance fingerprinting. *Magn Reson Imaging*. 2017;41:73-79.
17. Gao Y, Chen Y, Ma D, et al. Preclinical magnetic resonance fingerprinting (MRF) at 7 T: effective quantitative imaging for rodent disease models. *NMR Biomed*. 2015;28:384-394.
18. Buonincontri G, Sawiak SJ. MR fingerprinting with simultaneous B1 estimation. *Magn Reson Med*. 2016;76:1127-1135.
19. Cloos MA, Knoll F, Zhao T, et al. Multiparametric imaging with heterogeneous radiofrequency fields. *Nat Commun*. 2016;7:12445.
20. Hamilton JI, Jiang Y, Ma D, et al. Investigating and reducing the effects of confounding factors for robust T1 and T2 mapping with cardiac MR fingerprinting. *Magn Reson Imaging* 2018;53:40-51.
21. Kobayashi Y, Terada Y. Diffusion-weighting caused by spoiler gradients in the fast imaging with steady-state precession sequence may lead to inaccurate T2 measurements in MR fingerprinting. *Magn Reson Med Sci*. 2018;18:96-104. https://doi.org/10.2463/mrms.tn.2018-0027
22. Hilbert T, Xia D, Block KT, et al. Magnetization transfer in magnetic resonance fingerprinting. *Magn Reson Med*. 2020;84:128-141. https://doi.org/10.1002/mrm.28096
23. Kara D, Fan M, Hamilton J, Griswold M, Seiberlich N, Brown R. Parameter map error due to normal noise and aliasing artifacts in MR fingerprinting. *Magn Reson Med*. 2019;81:3108-3123.
24. Körzdörfer G, Pfeuffer J, Kluge T, et al. Effect of spiral undersampling patterns on FISP MRF parameter maps. *Magn Reson Imaging*. 2019;62:174-180. https://doi.org/10.1016/j.mri.2019.01.011
25. Winkler SA, Rutt BK. Practical methods for improving B_1^+ homogeneity in 3 tesla breast imaging. *J Magn Reson Imaging*. 2015;41:992-999.
26. Partridge SC, Murthy RS, Ziadloo A, et al. Diffusion tensor magnetic resonance imaging of the normal breast. *Magn Reson Imaging*. 2010;28:320-328.
27. Choi Y, Hwang E-J, Nam Y, et al. Analysis of apparent diffusion coefficients of the brain in healthy controls: a comparison study between single-shot echo-planar imaging and read-out-segmented echo-planar imaging. *Korean J Radiol*. 2019;20:1138-1145.
28. Gawande RS, Gonzalez G, Messing S, Khurana A, Daldrup-Link HE. Role of diffusion-weighted imaging in differentiating benign and malignant pediatric abdominal tumors. *Pediatr Radiol*. 2013;43:836-845.
29. Bonini RHM, Zeotti D, Saraiva LAL, et al. Magnetization transfer ratio as a predictor of malignancy in breast lesions: preliminary results. *Magn Reson Med*. 2008;59:1030-1034.
30. Virostko J, Sorace AG, Wu C, et al. Magnetization transfer MRI of breast cancer in the community setting: reproducibility and preliminary results in neoadjuvant therapy. *Thomography*. 2018;5:44-52.
31. Arlinghaus LR, Dortch RD, Whisenant JG, et al. Quantitative magnetization transfer imaging of the breast at 3.0 T: reproducibility in healthy volunteers. *Tomography*. 2016;2:260-266.
32. Nolte T, Gross-Weege N, Doneva M, et al. Spiral blurring correction with water–fat separation for magnetic resonance fingerprinting in the breast. *Magn Reson Med*. 2020;83:1192-1207.
33. Sommer K, Amthor T, Koken P, Meineke J, Doneva M. Determination of the optimum pattern length of MRF sequences. In: *Proceedings of the 25th Annual Meeting of ISMRM*, Honolulu, HI, 2017. (abstract 1491).
34. Virtue P, Tamir JI, Doneva M, Yu SX, Lustig M. Learning contrast synthesis from MR Fingerprinting. In: *Proceedings of the Joint Annual Meeting of ISMRM-ESMRMB*, Paris, France, 2018. (abstract 0676).
35. Sommer K, Amthor T, Doneva M, Koken P, Meineke J, Boernert P. Towards predicting the encoding capability of MR fingerprinting sequences. *Magn Reson Imaging*. 2017;41:7-14.
36. Yarnykh VL. Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magn Reson Med*. 2007;57:192-200.
37. Pauly J, Le Roux P, Nishimura D, Macovski A. Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm. *IEEE Trans Med Imaging*. 1991;10:53-65.
38. Guellmar D, Haueisen J, Reichenbach JR. Analysis of b-value calculations in diffusion weighted and diffusion tensor imaging. *Concept Magn Reson A*. 2005;25A:53-66.
39. Weigel M, Hennig J. Diffusion sensitivity of turbo spin echo sequences. *Magn Reson Med*. 2012;67:1528-1537.
40. Weigel M, Schwenk S, Kiselev VG, Scheffler K, Henning J. Extended phase graphs with anisotropic diffusion. *J Magn Reson*. 2010;205:276-285.
41. Lehnert A, Machann J, Helms G, Claussen CD, Schick F. Diffusion characteristics of large molecules assessed by proton MRS on a whole-body MR system. *Magn Reson Imaging*. 2004;22:39-46.
42. Partridge SC, Nissan N, Rahbar H, Kitsch AE, Sigmund EE. Diffusion-weighted breast MRI: clinical applications and emerging techniques. *J Magn Reson Imaging*. 2017;45:337-355.
43. Iima M, Honda M, Sigmund EE, Kishimoto AO, Kataoka M, Togashi K. Diffusion MRI of the breast: current status and future directions. *J Magn Reson Imaging*. 2020;52:70-90.
44. Sharma U, Sah RG, Agarwal K, et al. Potential of Diffusion-Weighted imaging in the characterization of Malignant, Benign, and healthy Breast Tissues and Molecular subtypes of Breast cancer. *Front Oncol*. 2016;6:126.
45. Kuroki-Suzuki S, Kuroki Y, Nasu K, Nawano S, Moriyama N, Okazaki M. Detecting breast cancer with non-contrast MR imaging: combining diffusion-weighted and STIR imaging. *Magn Reson Med Sci*. 2007;6:21-27.
46. Khalil R, Osman NM, Chalabi N, Ghany EA. Unenhanced breast MRI: could it replace dynamic breast MRI in detecting and characterizing breast lesions? *Egypt J Radiol Nucl Med*. 2020;51:10.

47. Jiang Y, Ma D, Wright K, Seiberlich N, Gulani V, Griswold MA. Simultaneous T1, T2, diffusion and proton density quantification with MR fingerprinting. In: *Proceedings of the Joint Annual Meeting of ISMRM-ESMRMB*, Milan, Italy, 2014. (abstract 28).
48. Jiang Y, Wright KL, Hamilton J, et al. Rapid and simultaneous T1, T2 and diffusion quantification using MR fingerprinting in the breast. In: *Proceedings of the Joint Annual Meeting of ISMRM-ESMRMB*, Paris, France, 2018. (abstract 2419).
49. De Boer RW. Magnetization transfer contrast. Part 2: Clinical applications. Philips Medical Systems. *MedicaMundi*. 1995;40:74-83.
50. Malik SJ, Teixeira RPAG, Hajnal JV. Extended phase graph formalism for systems with magnetization transfer and exchange. *Magn Reson Med*. 2018;80:767-779.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

FIGURE S1 Effect of electronic noise and spiral aliasing artefacts on the precision of (A) the T_1 and (B) the T_2 values. Test signals including one of the confounding effects and either electronic noise or spiral aliasing artefacts were matched to the StdDict. For each test signal, the matching process was repeated 200 times, superimposing newly generated noise or spiral aliasing artefact patterns in each matching step. Precision was then computed as the standard deviation over the matched T_1 and T_2 values. Electronic noise was simulated as complex white Gaussian noise with an SNR of $r = 50$, while spiral aliasing artefacts were simulated as periodic complex white Gaussian noise patterns with a signal-to-artefact ratio of $r = 1.5$. The precision plots correspond to the simulation results presented in Figures 2 and 3 of the manuscript, which visualize the expected bias on the matched T_1 and T_2 values. No precision plots are shown for the simulations with noise-free test signals, as the standard deviation is always zero in that case

FIGURE S2 Predicted accuracy and precision on T_1 and T_2 for test signals including (A) electronic noise at 3 different SNRs and for (B) spiral aliasing artefacts at 3 different signal-to-artefact ratios. Each row has an individual scale bar

FIGURE S3 Expected bias on T_1 and T_2 for different in-plane B_1^+ correction factors $f_{B_1^+}$ between 0.7 and 1.3 and diffusion constants D between 0.5×10^{-3} and 3.0×10^{-3} mm^2/s. Test signals including a confounding effect were matched to the StdDict. The plots then visualize $T_1^{\text{matched}} - T_1^{\text{test}}$ (or $T_2^{\text{matched}} - T_2^{\text{test}}$, respectively), ie, the bias that results from omitting the confounding effect from the MRF dictionary. (A) Visualizes the expected bias on T_1 caused by in-plane B_1^+ effects, (B) visualizes the expected bias in T_2 accuracy caused by in-plane B_1^+ effects, (C) visualizes the expected bias in T_1 accuracy caused by diffusion effects, and (D) visualizes the expected bias in T_2 accuracy caused by diffusion effects

FIGURE S4 Influence of different confounding effects on the MRF signal evolutions. A test signal with nominal relaxation times of $T_1 = 1200$ ms and $T_2 = 70$ ms was computed (A) for different correction factors $f_{B_1^+}$ between 0.7 and 1.3, (B) with and without including the imperfect slice profile, and (C) for different diffusion constants (values in mm^2/s)

FIGURE S5 Impact of in-plane B_1^+ inhomogeneity on accuracy and repeatability of MRF estimates investigated in a phantom experiment at 3T

FIGURE S6 Spiral (under-) sampling simulation in a 2-tissue breast phantom to evaluate whether or not the spiral aliasing artefacts can be approximated by periodic scaled white Gaussian (pWGN) noise patterns

How to cite this article: Nolte T, Scholten H, Gross-Weege N, et al. Confounding factors in breast magnetic resonance fingerprinting: B_1^+, slice profile, and diffusion effects. *Magn Reson Med*. 2021;85:1865–1880. https://doi.org/10.1002/mrm.28545