R2* estimation performance in iron-overloaded livers: fit first or average first?

Debra E. Horng[1,2], Diego Hernando[1], and Scott B. Reeder[1,2]

[1]*Radiology, University of Wisconsin-Madison, Madison, WI, United States,* [2]*Medical Physics, University of Wisconsin-Madison, Madison, WI, United States*

Target audience: Researchers interested in measuring R2*(=1/T2*) in the liver.

Purpose: R2* (=1/T2*) relaxometry techniques based on multi-echo spoiled gradient echo (SGRE) acquisitions have the potential to provide accurate assessment of iron concentration for detection and quantitative staging of liver iron overload. R2*-based techniques are rapid and R2* has a linear relationship with iron concentration[1]. Estimation techniques in the presence of high R2* values can be generally classified into two main categories: "fit-first" (FF) techniques[2], where a parametric R2* map is generated at each voxel, followed by region-of-interest (ROI) measurement of R2*, and "average-first" (AF) techniques[3], where the signal for each echo time is averaged over an ROI, followed by fitting the averaged signal to obtain an R2* estimate. Therefore, ***the purpose of this work*** was to compare the performance of FF and AF techniques by comparing the bias and standard deviation of their R2* estimates over a wide range of R2* values, in simulations and in phantoms.

Methods: **Simulations**: Multi-echo signals were simulated with 16 echo times, TE1=0.92ms, ΔTE=1.4ms, 8 coils and a coil-combined SNR of 15 (at TE=0), over the range of R2* values 50-2000s^{-1} and assuming no B_0 inhomogeneity. An ROI of 100 pixels was created, and noise performance was simulated by 100 repetitions for SNR=15.

Phantom: A phantom was built with 9 glass vials containing 2% agar and aqueous dilutions of ferumoxytol (AMAG Pharmaceuticals, Waltham, MA, USA): 22, 44, 88, 132, 176, 221, 441, 662, and 882 mg/mL. These concentrations were chosen to mimic R2* values found in normal to severely iron-overloaded livers[4] (min R2*=49s^{-1}, max expected R2*~2000 s^{-1}).

Acquisition: The phantom was scanned 3 times on a 1.5T system (Signa HDxt, GE Healthcare, Waukesha, WI, USA), using an 8-channel head coil, a multi-echo 3D SGRE sequence, and no parallel imaging acceleration. Echo times were the same as in simulation. Other scan parameters included: FOV=34mm, slice thickness=2.1mm, 28 slices, matrix=128×128, TR=22.4ms, and flip angle=1°. Flip angle was chosen empirically to obtain SNR~15. A high SNR scan was also acquired for reference (4mm slices, flip angle=8°).

Processing: The scan data were reconstructed at a resolution of 1.3mm×1.3mm, using two methods used for coil combination: root sum-of-squares (RSOS), resulting in magnitude images, and matched filter (MF)[5], resulting in complex images. For RSOS images, two different models were used: the first and second moment models[6] of non-central chi distributed noise. Since the first moment model has been shown to have equivalent performance as the second moment model over, we only considered the second moment model in the range 50<R2*<1000s^{-1} [6]. Taylor series and asymptotic approximations to the hypergeometric function[7] were utilized for faster first moment computation over 1000<R2*<2000s^{-1}. The MF complex images were fitted as either magnitude (discarding the phase) or complex data, and fitted using either FF or AF, for a total of 4 MF models. In summary, we had a total of 6 models: *Model 1:* RSOS AF magnitude, 1^{st} moment correction (for 1000<R2*<2000s^{-1}), *Model 2:* RSOS AF magnitude, 2^{nd} moment correction, *Model 3:* MF AF complex, *Model 4:* MF FF complex, *Model 5:* MF AF magnitude, and *Model 6:* MF FF magnitude.

Performance comparison: The mean and standard deviation of R2* estimates for each of the 6 models were calculated across the 100 simulation trials and across the 3 phantom scans. The reference R2* for the phantom vials was calculated based the relaxivity of ferumoxytol, estimated from the high SNR scan.

Results: The mean R2* in each trial (of 100 repetitions) is plotted against the true R2* and shown in Fig. 1. The standard deviation of each trial is plotted against the true R2* and shown in Fig. 2. Simulations indicate that the best performing methods are Models 3 and 5, both MF average-first models: they have the lowest error in the mean and low standard deviation up to the highest true R2*. The mean R2* estimated in the phantom vials for each model is plotted against the true R2* (calculated from ferumoxytol relaxivity of 2.22 s^{-1}·mL/mg) and shown in Fig. 3. In the phantom scans, most methods underestimate the R2* above 500s^{-1}.

Discussion and Conclusion: Although simulations indicate matched-filter average-first models have the best performance, in practice the field map is inhomogeneous and averaging across a voxel leads to dephasing error. The combination of high R2* and low SNR leads to poor performance for all methods in the phantom, likely requiring higher SNR/shorter TEs for accurate R2* relaxometry.

References: **1.** Wood JC et al. Blood. 2005;106:1460-1465. **2.** Hankins JS et al. Blood. 2009:113;4853-4855. **3.** Ferguson MR et al. JMRI. 2013;38:987-990. **4.** Schenck JF. Med Phys. 1996;23(6):815-50. **5.** Walsh DO et al. MRM. 2000;43(5):682-90. **6.** Feng Y et al. MRM. 2013;70(6):1765-74. **7.** Pearson J. 2009. MA thesis, University of Oxford.

Acknowledgements: We acknowledge the support of the NIH (R01DK083380, R01DK088925, R01DK100651, UL1TR00427). We also wish to thank GE Healthcare for departmental research support.

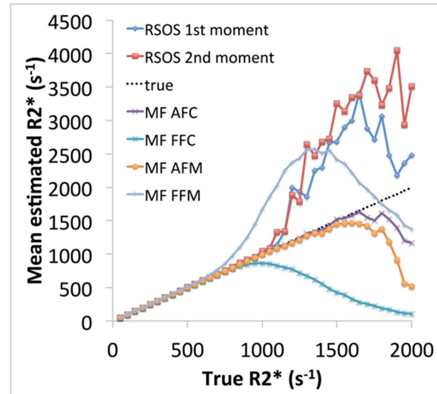

Figure 1. Simulations show that all models perform well up to ~1000s^{-1}, but most have severe bias from 1000s^{-1} to 2000s^{-1}. The simulations were done assuming a homogeneous magnet, i.e. the best possible case for complex averaging.

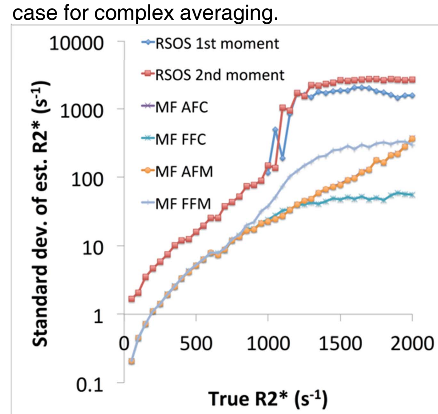

Figure 2. Simulated standard deviation increases logarithmically with increasing R2*. RSOS methods have higher standard deviations over the range than MF methods. MF AFM and MF AFC methods have the same standard deviation curve in this graph.

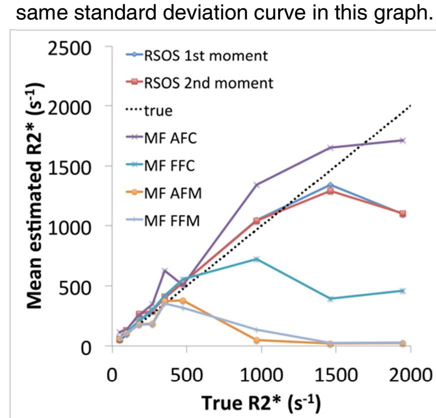

Figure 3. Phantom scans at SNR=15 have good performance (low bias) up to 500s^{-1}, but quickly deteriorate above 500s^{-1}. MF magnitude methods severely underestimate the R2* at higher R2*.