Abstract

Purpose: To assess the effect of a low-rank denoising algorithm on quantitative MRI-based measures of liver fat and iron.

Materials and Methods: This was an IRB-approved, HIPAA compliant, retrospective analysis of 42 consecutive subjects who were imaged at 3T using a multi-echo gradient echo sequence that was reconstructed using a multi-step adaptive fitting algorithm to obtain quantitative proton density fat fraction (PDFF) and R2* maps (original maps). A patch-wise low-rank denoising algorithm was then applied and PDFF and R2* maps were created (denoised maps). Three readers independently rated the PDFF maps in terms of vessel and liver edge sharpness and image noise using a 5-point scale. Two other readers independently measured mean and standard deviation of PDFF and R2* values for the original and denoised maps; values were compared using intraclass correlation coefficients (ICCs) and mean difference analyses.

Results: Qualitatively, the denoised maps were preferred by all three readers based on image noise ($p<0.001$) and by two of three readers based on vessel edge sharpness ($p<0.001$-0.99). No reader had a significant preference regarding liver edge sharpness ($p=0.16$-0.48). Quantitatively, agreement was near perfect between the original and denoised maps for PDFF (ICC=0.995) and R2* (ICC=0.995) values. Mean quantitative values obtained from the original and denoised maps were similar for PDFF (7.6+/-7.2% vs. 7.7+/-7.7%, $p=0.15$) and for R2* (53.8+/-4s^{-1} vs. 53.1+/-4.1s^{-1}, $p=0.11$).

Conclusion: Applying the low-rank denoising algorithm to liver fat and iron quantification reduces image noise in PDFF and R2* maps without adversely affecting mean quantitative values or subjective image quality.

Key words: Denoising; filtering; fat fraction; iron quantification; magnetic resonance imaging

Introduction

MRI-based measures of liver fat and iron are becoming more widely available and clinically important for the assessment of fatty liver disease and iron deposition (1-5). While liver biopsy is the current standard for the assessment of steatosis, it is an invasive procedure and is not without risk (6). MR spectroscopy (MRS) is a robust method to accurately and non-invasively quantify hepatic fat fraction with high sensitivity, but suffers from a limited sampling volume (single-voxel acquisition) and low spatial resolution, which can be problematic in the setting of heterogeneous hepatic fat or iron deposition (1,7). Chemical-shift-based MRI quantitative PDFF and R2* relaxometry have been shown to correlate well with MRS and biopsy, but also have higher spatial resolution than MRS which provides more anatomic information (8-11).

Performing region of interest (ROI)-based measurements on chemical-shift-based MRI PDFF and R2* maps can be challenging due to artifacts such as image noise, and noise can bias the PDFF and R2* measures themselves (12-14). Additionally, while imaging can be performed at low spatial resolution to reduce noise, resolution must be sufficient to allow visualization of vessels and liver lesions, so that these can be avoided during the measurement process. Low-rank denoising is a method which has been described for reducing noise in image reconstructions based on multiple data sets which share similarities, and it has been applied to MRI-based PDFF measures obtained from six-echo gradient echo sequences (15). However, when applying smoothing/denoising algorithms in general, there is a potential for edge blurring, which can make avoiding confounding structure such as blood vessels more difficult.

The purpose of this study is to assess the effect of a low-rank denoising algorithm on quantitative MRI-based measures of liver fat and iron, specifically in terms of changes in PDFF values, R2* values, and subjective image quality metrics.

Materials and Methods

This was an institutional review board-approved, health insurance portability and accountability act compliant retrospective analysis of prospectively acquired data. Written informed consent was obtained from all subjects. Siemens Healthcare provided support for the original subject recruitment and data acquisition. The authors who are not employees of Siemens Healthcare had control of all data and information that might have presented a conflict of interest for the duration of the study.

Subjects

Forty-two consecutive subjects undergoing clinical liver MRI (18 males, 24 females; mean age 52.8 years [range 20-80 years]) were prospectively enrolled. Mean subject weight was 85.3 kg (range, 49.9-163.3 kg), and mean body mass index was 29.3 kg/m^2 (range, 17.8-46.2 kg/m^2). Indications for imaging included oncologic workup (n=25), chronic liver disease (n=14), and pain/biliary obstruction (n=3).

MRI

Imaging was performed on one of two 3T MRI systems (Magnetom Skyra, Siemens Healthcare, Erlangen, Germany) with the subjects supine, using an 18-channel phased-array body coil centered over the upper abdomen combined with the table-mounted spine matrix. A whole-liver volume acquisition was performed using a prototypical six-echo 3D spoiled-gradient-echo acquisition. Two-dimensional parallel acceleration was used to allow whole liver coverage within a single breath-hold (controlled aliasing in parallel imaging results in higher acceleration) (16).

Pulse sequence parameters included: TR=8.9 ms; TE_1=1.23 ms, with six echoes collected at ΔTE 1.23 ms; flip angle=4°; receiver bandwidth=1085 Hz/pixel; field of view=42 x 32.8 x 24 cm^3; acquisition matrix=256 x 160 x 50 interpolated to an image matrix of 256 x 200 with 60 slices; spatial resolution=1.6 x 2.1 x 4 mm^3; parallel imaging factor=2 x 2; and acquisition time=21 seconds.

Image Reconstruction

The standard six-echo gradient echo sequence was reconstructed using a conventional image reconstruction, followed by the multi-step adaptive fitting algorithm to obtain PDFF and R2* maps (assuming the same R2* for fat and water); termed "original" maps (10,17). In order to obtain denoised maps, a patch-wise low-rank denoising algorithm including automated noise adjustment was applied retrospectively in the scanner environment to the reconstructed six-echo image series prior to applying the fitting routine (15). Briefly, this algorithm makes use of the local spectral redundancy of the multi-echo data by patch-wise promoting low-rank representations of the data. To this end, a risk-minimizing optimization yields an optimal denoising for every patch by taking both the local data as well as the global image noise estimation into account. This step was followed by the same reconstruction using the multi-step adaptive fitting algorithm to obtain PDFF and R2* maps; termed "denoised" maps.

For both the original and denoised data sets, goodness-of-fit (GOF) maps were reconstructed. In order to measure GOF, the PDFF and R2* maps were used to generate modeled 6-echo magnitude image sets, and GOF was defined on a per-voxel basis as the sum-of-squared residuals (modeled data – measured data) divided by the sum of squared values of the measured data (18).

Qualitative Analysis

The PDFF and R2* maps were de-identified and randomized pairwise, such that the original and denoised maps were presented to reviewers as a pair, but in random order. Three abdominal radiologists with 13, 5 and 4 years post-fellowship experience with liver MRI, respectively, independently reviewed the original and denoised maps side-by-side, blinded to which image sets had been denoised. Readers recorded their preference between images in terms of vessel edge sharpness, liver edge sharpness, and image noise. After recording these preferences for all map pairs, results were de-randomized to yield a scale of: 1 – strongly prefer denoised; 2 – somewhat prefer denoised; 3 – no preference; 4 – somewhat prefer original; 5 – strongly prefer original.

Quantitative Analysis

Using the original TE=6.15 ms images, two separate readers independently placed four regions of interest (ROIs) in the right hepatic lobe (anterior, posterior, medial, and lateral locations), careful to avoid hepatic vasculature and artifacts. These ROIs were automatically copied onto the PDFF, R2* maps, and GOF maps for both the original and denoised reconstructions. The mean and standard deviations for each ROI on the PDFF and R2* maps were recorded, as were the mean values of each ROI for the GOF maps.

Statistical Analysis

All statistical analyses were performed using SPSS software (version 20.0, IBM, Chicago, IL) as well as R (version 3.0.3, The R Foundation for Statistical Computing, Vienna, Austria).

For the qualitative analysis, inter-rater reliability was assessed using Fleiss' Kappa for each evaluation criterion. Then, a mixed-effects linear model was used to assess reader scores for differences from an overall value of 3 (no preference) for each visual assessment, with the reader treated as a fixed effect and the subject treated as a random effect. Post-hoc p-values were adjusted for multiple comparisons using the single-step method (19).

For the quantitative analysis, mean and standard deviation of PDFF and R2* values were calculated. Agreement between the mean values of ROIs measured by each reader was assessed using intraclass correlation coefficients (ICCs). Differences between PDFF, R2* and GOF values were compared between the original and denoised maps using the Mann-Whitney-U test as well as a linear model on the log-transformed data. The anatomical location and the use of denoising were modeled as fixed effects, and post-hoc comparisons were performed using Tukey's honest significant difference (HSD) test to correct for multiple comparisons.

Results

Representative PDFF and R2* maps demonstrate visual reduction in noise by the denoising algorithm in subjects with high (Fig. 1) and normal (Fig. 2) PDFF values.

Qualitative Analysis

Subjective reader preferences are summarized in Figure 3 and Table 1. None of the readers preferred the original PDFF maps over the denoised maps for any of the evaluated criteria. For vessel edge sharpness, readers 1 and 2 significantly preferred the denoised maps ($p<0.01$), while reader 3 had no significant preference ($p=0.99$). None of the readers had a significant preference with respect to liver edge sharpness ($p=0.86$-1.00). With regard to image noise, all three readers significantly preferred the denoised maps ($p<0.001$).

Quantitative Analysis

Agreement between the two readers for ROI-based quantitative PDFF and R2* measures was near perfect (ICC=0.96 for PDFF and 0.98 for R2*), and the readers were treated as equivalent for the remainder of the analysis.

For liver PDFF measurements, the mean value across all subjects was 7.7±7.7% (7.7±7.7% for the original maps and 7.7±7.8% for the denoised maps, $p=0.12$). In the linear model comparison of the ROI measurements on the original vs. denoised images, the mean PDFF was not significantly different ($p=0.12$) between original and denoised images. However, the standard deviation of values within ROIs was 41% smaller for the denoised images (from 2.6±1.9% for the original maps to

1.6±1.6% for the denoised maps, $p<0.001$), indicating that the heterogeneity of the liver within the ROIs was reduced by the denoising algorithm.

For the R2* measurements, the overall mean value was 49.0±44.3 s^{-1} (52.9±40.3 s^{-1} for the original maps and 52.8±41.1 s^{-1} for the denoised maps, $p=0.20$). In the linear model comparison of the ROI measurements on the original vs. the standard images, the mean R2* was not significantly different ($p=0.89$) between the original and denoised maps. However, the standard deviations were 29% smaller (from 12.1±10.5 s^{-1} for the original maps to 8.6±8.4 s^{-1} for the denoised maps, $p<0.001$), indicating that the heterogeneity of the liver within the ROIs was reduced by the denoising algorithm.

Mean GOF was improved by 46% using the denoising algorithm (from 3.08±1.51% for the original maps to 1.67±0.94% for the denoised maps), which was statistically significant ($p<0.0001$).

Discussion

In this study, a low-rank denoising algorithm was applied to gradient echo magnitude images prior to the multi-step adaptive fitting algorithm for calculation of quantitative proton density fat fraction and R2* maps. The quantitative PDFF and R2* measurements were not significantly affected, but liver heterogeneity was improved visually and quantitatively, while preserving or improving edge features such as vessel sharpness. These findings suggest that the described low-rank denoising algorithm improves image heterogeneity without causing edge blurring.

In order to obtain PDFF and R2* measurements that are independent of T1 bias, low flip angles and long relaxation times must be used, which can result in acquisitions with low signal-to-noise ration (SNR) (12,15,20-23). The resulting image noise can hinder ROI placement when measuring PDFF and R2*, and noise has been shown to bias quantitative measurements of PDFF and R2* (12,13). To improve image quality in low-SNR images in general, a variety of image post-processing techniques have been studied. While post-processing algorithms do not increase image acquisition time, images are subject to blurring and other artifacts introduced by the post-processing algorithms themselves. Simple spatial averaging, in which a pixel is given the value of the average of the adjacent pixels, will reduce noise, but also degrades sharp details (lines and edges), as these structures are inadvertently averaged with the adjacent pixels of interest (24). Edge-preserving filters, such as the anisotropic diffusion filter, can distort small features and make images look unnatural (24,25). Wavelet-based filters are subject to the introduction of artifacts that can affect subjective image quality (25,26). In our study, the described low-rank denoising algorithm appears to have no adverse effect on subjective liver and vessel edge definition, and all three readers preferred the denoised datasets with respect to image

noise. These data suggest that the low-rank denoising algorithm does not adversely affect subjective edge sharpness, unlike many other smoothing algorithms. With decreased subjective image noise and similar edge sharpness, ROI placement on PDFF and R2* maps could theoretically be facilitated.

In addition to hindering ROI placement when measuring PDFF and R2*, poor SNR has been shown to bias quantitative measurements of PDFF and R2* (12,13). In our quantitative analysis, the measured PDFF and R2* values were not significantly changed by application of the low-rank denoising algorithm. The accuracy of PDFF values using the acquisition and image reconstruction chain utilized in this study has been previously validated, thus a lack of change due to application of the algorithm is desirable (17). It is unclear what effect the low-rank denoising algorithm would have on MRI data affected by bias due to lower SNR; we were unable to explore this scenario using the available data.

There are limitations to this study. It was a retrospective analysis of prospectively acquired data, and a relatively small number of patients were included. In addition, all imaging was performed on one of two identical MRI systems from a single vendor, and may not be applicable to other vendors or scanner platforms. As described, our data had adequate SNR for accurate PDFF estimation, and we were unable to investigate the potential effect of the low-rank denoising algorithm on data with lower SNR.

In conclusion, applying the described low-rank denoising algorithm to a liver fat/iron quantification technique reduces image noise in PDFF and R2* maps without adversely affecting mean values of the quantitative measures or subjective edge sharpness.

References

1. 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(5):1003-1021.
2. Reeder SB, Cruite I, Hamilton G, Sirlin CB. Quantitative Assessment of Liver Fat with Magnetic Resonance Imaging and Spectroscopy. J Magn Reson Imaging 2011;34(4):spcone.
3. Noureddin M, Lam J, Peterson MR, et al. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. Hepatology 2013;58(6):1930-1940.
4. Arulanandan A, Ang B, Bettencourt R, et al. Association Between Quantity of Liver Fat and Cardiovascular Risk in Patients With Nonalcoholic Fatty Liver Disease Independent of Nonalcoholic Steatohepatitis. Clin Gastroenterol H 2015;13(8):1513-+.
5. Gandon Y, Olivie D, Guyader D, et al. Non-invasive assessment of hepatic iron stores by MRI. Lancet 2004;363(9406):357-362.
6. Boyum JH, Atwell TD, Schmit GD, et al. Incidence and Risk Factors for Adverse Events Related to Image-Guided Liver Biopsy. Mayo Clin Proc 2016.
7. Longo R, Pollesello P, Ricci C, et al. Proton MR spectroscopy in quantitative in vivo determination of fat content in human liver steatosis. J Magn Reson Imaging 1995;5(3):281-285.
8. Idilman IS, Aniktar H, Idilman R, et al. Hepatic steatosis: quantification by proton density fat fraction with MR imaging versus liver biopsy. Radiology 2013;267(3):767-775.

9. Meisamy S, Hines CD, Hamilton G, et al. Quantification of hepatic steatosis with T1-independent, T2-corrected MR imaging with spectral modeling of fat: blinded comparison with MR spectroscopy. Radiology 2011;258(3):767-775.
10. Bashir MR, Zhong X, Nickel MD, et al. Quantification of hepatic steatosis with a multistep adaptive fitting MRI approach: prospective validation against MR spectroscopy. AJR Am J Roentgenol 2015;204(2):297-306.
11. Kuhn JP, Hernando D, Mensel B, et al. Quantitative chemical shift-encoded MRI is an accurate method to quantify hepatic steatosis. J Magn Reson Imaging 2014;39(6):1494-1501.
12. 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(2):354-364.
13. Sharma P, Altbach M, Galons JP, Kalb B, Martin DR. Measurement of liver fat fraction and iron with MRI and MR spectroscopy techniques. Diagn Interv Radiol 2014;20(1):17-26.
14. Hansen KH, Schroeder ME, Hamilton G, Sirlin CB, Bydder M. Robustness of fat quantification using chemical shift imaging. Magn Reson Imaging 2012;30(2):151-157.
15. Lugauer F, Nickel D, Wetzl J, Kannengiesser SAR, Maier A, Hornegger J. Robust Spectral Denoising for Water-Fat Separation in Magnetic Resonance Imaging. Lect Notes Comput Sc 2015;9350:667-674.
16. Breuer FA, Blaimer M, Mueller MF, et al. Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA). Magn Reson Med 2006;55(3):549-556.

17. Zhong X, Nickel MD, Kannengiesser SA, Dale BM, Kiefer B, Bashir MR. Liver fat quantification using a multi-step adaptive fitting approach with multi-echo GRE imaging. Magn Reson Med 2014;72(5):1353-1365.

18. Bacher MA, Zhong X, Dale BM, et al. Signal model consistency analysis of different protocols and spectral models in multi gradient echo liver PDFF and R2* quantification. International Society of Magnetic Resonance in Medicine 22nd Annual Meeting and Exhibition. Milan, Italy; Preceedings of the ISMRM 2014.

19. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J 2008;50(3):346-363.

20. Yokoo T, Bydder M, Hamilton G, et al. Nonalcoholic fatty liver disease: diagnostic and fat-grading accuracy of low-flip-angle multiecho gradient-recalled-echo MR imaging at 1.5 T. Radiology 2009;251(1):67-76.

21. Hines CD, Frydrychowicz A, Hamilton G, et al. T(1) independent, T(2) (*) corrected chemical shift based fat-water separation with multi-peak fat spectral modeling is an accurate and precise measure of hepatic steatosis. J Magn Reson Imaging 2011;33(4):873-881.

22. Bydder M, Yokoo T, Hamilton G, et al. Relaxation effects in the quantification of fat using gradient echo imaging. Magn Reson Imaging 2008;26(3):347-359.

23. Kuhn JP, Jahn C, Hernando D, et al. T1 bias in chemical shift-encoded liver fat-fraction: role of the flip angle. J Magn Reson Imaging 2014;40(4):875-883.

24. Gerig G, Kubler O, Kikinis R, Jolesz FA. Nonlinear anisotropic filtering of MRI data. IEEE Trans Med Imaging 1992;11(2):221-232.

25. Manjon JV, Carbonell-Caballero J, Lull JJ, Garcia-Marti G, Marti-Bonmati L, Robles M. MRI denoising using non-local means. Med Image Anal 2008;12(4):514-523.

26. Wood JC, Johnson KM. Wavelet packet denoising of magnetic resonance images: importance of Rician noise at low SNR. Magn Reson Med 1999;41(3):631-635.

Tables

Table 1. Reader preference for each qualitative image feature on the PDFF maps based on the following rating scale: 1=strongly prefer denoised; 2=somewhat prefer denoised; 3=no preference; 4=somewhat prefer original; 5=strongly prefer original. Ratings are summarized as mean +/- standard deviation.

	Vessel Edge		Liver Edge		Image Noise	
	Rating	p-value*	Rating	p-value*	Mean	p-value*
Reader 1	2.12 +/- 0.67	<0.001	2.98 +/- 0.35	0.86	1.45 +/- 0.80	<0.001
Reader 2	2.76 +/- 0.48	<0.01	3.0 +/- 0.0	1.000	2.02 +/- 0.81	<0.001
Reader 3	2.98 +/- 0.27	0.99	2.98 +/- 0.15	0.86	2.45 +/- 0.71	<0.001

* p-values are for an adjusted linear effects model with post-hoc test comparing the reader ratings against a rating of 3 (no preference).

Figure Legends

Figure 1. Representative images from a patient with a relatively high PDFF value of 23.7% and an R2* value of 47.4 s^{-1} on the original image reconstructions. (a) original PDFF map; (b) denoised PDFF map; (c) original R2* map; (d) denoised R2* map. Readers strongly preferred the denoised PDFF map based on image noise (average rating among 3 readers = 2.0).

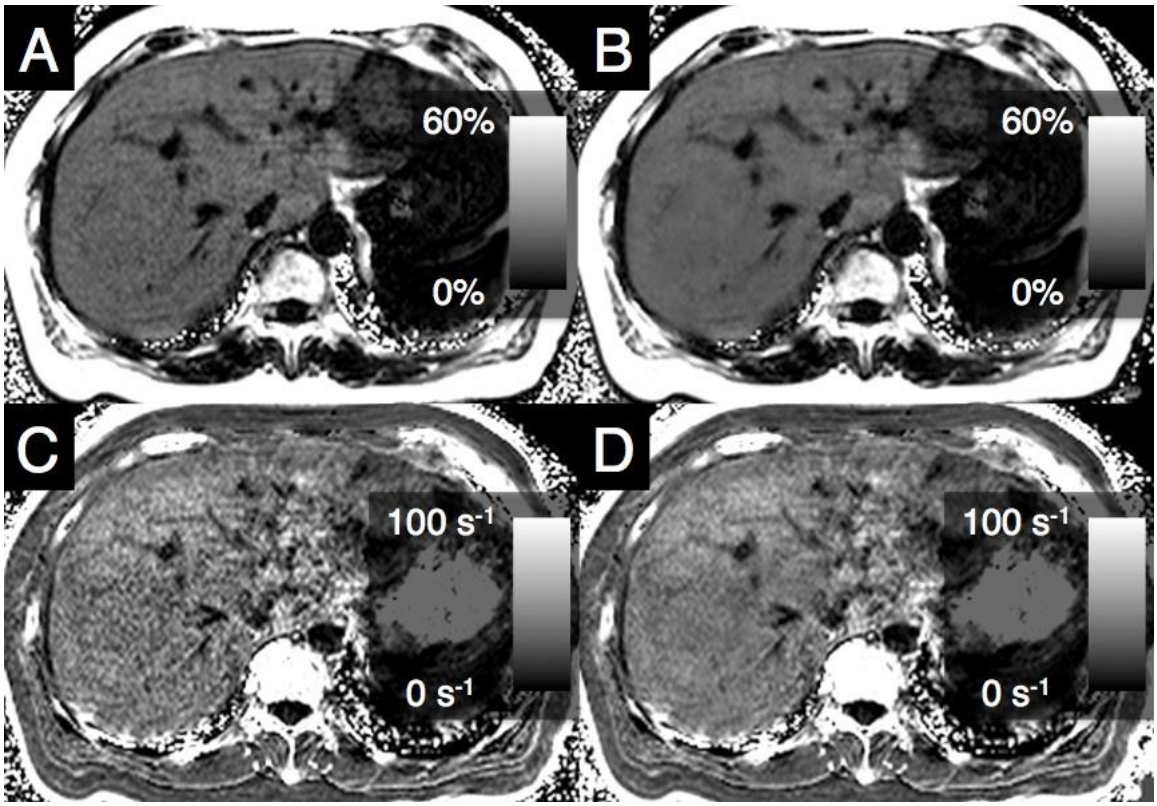

Figure 2. Representative images from a patient with a normal PDFF value of 4.6% and an R2* value of 42.4 s^{-1} on the original image reconstructions. (a) original PDFF map; (b) denoised PDFF map; (c) original R2* map; (d) denoised R2* map. Readers strongly preferred the denoised PDFF map based on image noise and slightly preferred the denoised images based on vessel edge sharpness (average rating among 3 readers = 2.62).

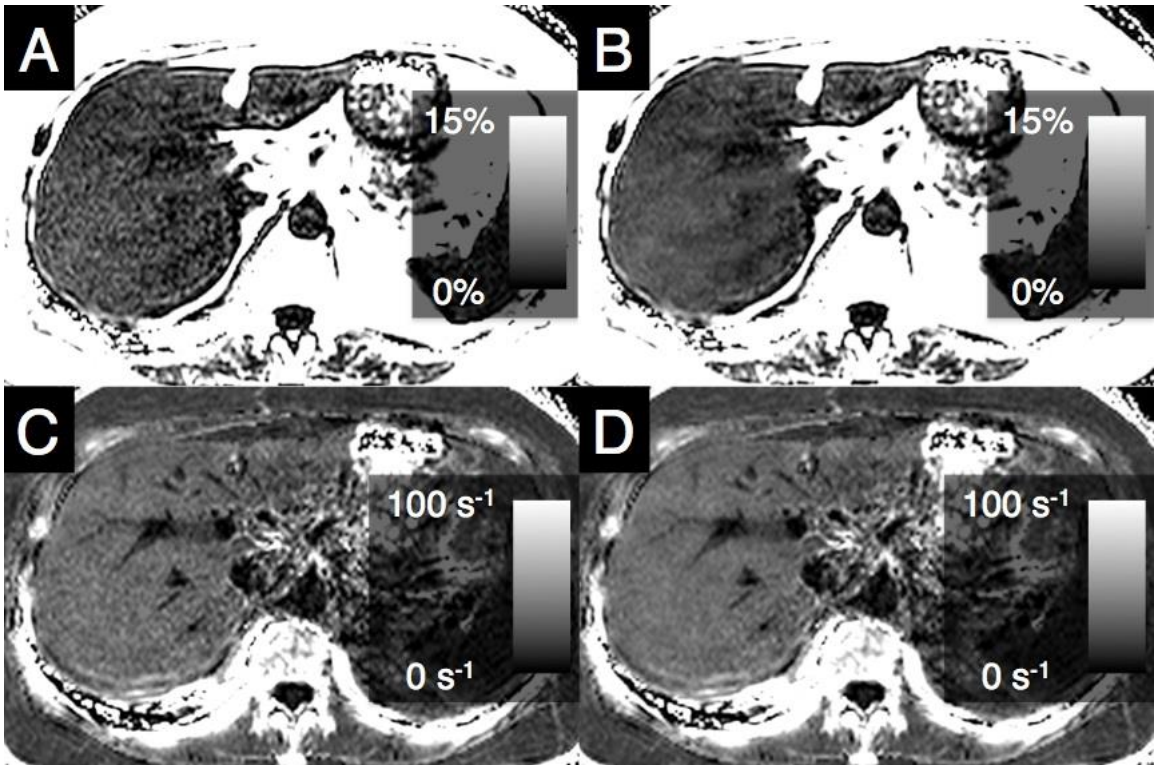

Figure 3. Summary of reader preferences for each of the three visual assessments performed in this study. * indicates statistically significant preferences for the denoised images for 2/3 readers based on vessel edge sharpness and all three readers based on image noise.