Evaluation of Golden-Angle-Sampled Dynamic Contrast-Enhanced MRI Reconstruction Using Objective Image Quality Measures: A Simulated Phantom Study

Nathan Murtha[1], Allister Mason[2], Chris Bowen[3,4], Sharon Clarke[3,4], James Rioux[3,4], and Steven Beyea[3,4]

[1]Department of Physics, Carleton University, Ottawa, ON, Canada; [2]Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, Canada; [3]Department of Diagnostic Radiology, Dalhousie University, Halifax, NS, Canada; and [4]BIOmedical Translational Imaging Centre (BIOTIC), Halifax, NS, Canada

Corresponding Author:
Steven Beyea, PhD
BIOmedical Translational Imaging Centre (BIOTIC), 3900-1796 Summer St, Halifax, NS, B3H 3A7, Canada;
E-mail: steven.beyea@dal.ca

We aim to extend the use of image quality metrics (IQMs) from static magnetic resonance imaging (MRI) applications to dynamic MRI studies. We assessed the use of 2 IQMs, the root mean square error and structural similarity index, in evaluating the reconstruction of quantitative dynamic contrast-enhanced (DCE) MRI data acquired using golden-angle sampling and compressed sensing (CS). To address the difficulty of obtaining ground-truth knowledge of parameters describing dynamics in real patient data, we developed a Matlab simulation framework to assess quantitative CS-DCE-MRI. We began by validating the response of each IQM to the CS-MRI reconstruction process using static data and the performance of our simulation framework with simple dynamic data. We then extended the simulations to the more realistic extended Tofts model. When assessing the Tofts model, we tested 4 different methods of selecting a reference image for the IQMs. Results from the retrospective static CS-MRI reconstructions showed that each IQM is responsive to the CS-MRI reconstruction process. Simulations of a simple contrast evolution model validated the performance of our framework. Despite the complexity of the Tofts model, both IQM scores correlated well with the recovery accuracy of a central model parameter for all reference cases studied. This finding may form the basis of algorithms for automated selection of image reconstruction aspects, such as temporal resolution, in golden-angle-sampled CS-DCE-MRI. These further suggest that objective measures of image quality may find use in general dynamic MRI applications.

INTRODUCTION

Motivation for This Work

Magnetic resonance imaging (MRI) can not only provide static anatomical information but also capture dynamic processes such as the uptake of an intravenously injected contrast agent with dynamic contrast-enhanced MRI (DCE-MRI). Rather than using a fixed temporal resolution for such dynamic scans, emerging methods based on golden-angle sampling such as XD-GRASP (1) or CIRCUS (2) allow retrospective temporal resampling at the time of image reconstruction. This framework allows retrospective control of the temporal "footprint" by varying the number of k-space rays or spokes per undersampled image, without necessarily requiring the use of view sharing, which can introduce temporal blurring of fast dynamics (2,

3). Doing so allows one to flexibly optimize the trade-off between better spatial quality (ie, low undersampling factor) and better temporal quality (ie, high undersampling factor).

With additional reconstruction flexibility, however, comes the need to optimize image reconstruction of that specific data set and its associated spatiotemporal features. This is particularly true when one considers the large and intertwined joint optimization space associated with these reconstructions, in which the undersampling factor (ie, temporal footprint), the form of regularization enforced, and the degree of regularization must all be simultaneously examined.

Recently, it has become common to use objective image quality metrics (IQMs) (4, 5) as an outcome measure or within

the cost function of compressed sensing (CS) or with deep learning–based image reconstruction (6–11). A recent study by our group examined a broad range of IQMs to determine the varying ability of these metrics to correlate with expert radiologists' rating of diagnostic image quality (12). Although such an approach is sufficient in the reconstruction of a single image data set with fixed undersampling factor, in a data set with golden-angle sampling that allows multiple potential reconstructions, most IQMs will be maximal for the lowest undersampling factor possible, that is, with no ability to resolve temporal dynamics. Therefore, additional research is required to make use of IQMs in such a scenario, and the approach using expert radiologist opinion may be impractical.

Furthermore, the k-space sampling and image reconstruction factors that maximize the quality of diagnostic images from a dynamic data set may not be the same ones that minimize the parametric error when a model is subsequently fit to the data set in a quantitative dynamic application, such as the Tofts pharmacokinetic model in the case of quantitative DCE-MRI (13, 14). For example, while a radiologist may visually prefer a diagnostic image with less noise, a reconstruction with increased image noise but improved temporal frame rate may potentially result in more accurate parameter fits. It is also not feasible to present radiologists with thousands of images resulting from the numerous permutations of k-space sampling and reconstruction choices. The ability to automate the joint tuning of golden-angle-sampled dynamic data reconstruction, for example, through the use of an IQM-based approach, would, therefore, potentially improve the accuracy of recovered parameters in quantitative dynamic applications.

Goals of This Study

We determined the relationship between 2 well-established IQMs—root mean square error (RMSE) and structural similarity index (SSIM) (5)—and the subsequent fidelity of recovered dynamic model parameters when applied to DCE-MRI with CS reconstruction of golden-angle-sampled data.

We began by assessing the properties of RMSE and SSIM in static CS-MRI reconstructions using retrospectively undersampled data from pelvic scans as well as a numerical phantom. This verified that each IQM was sensitive to the CS reconstruction parameters on a patient-by-patient basis.

With verification of IQM response in basic CS-MRI application, we then evaluated their ability to predict the quality of golden-angle-sampled dynamic MRI data reconstructed with CS. To this end, we developed a simplified simulation framework for a dynamically evolving numerical phantom as a surrogate for the assessment of CS-MRI in dynamic MRI procedures. The use of a numerical phantom allowed for ground-truth knowledge of the instantaneous temporal evolution. To validate our simulation framework, we first analyzed cylindrical features that evolved according to a monoexponential decay within a digital phantom. Once validated on a simple closed-form signal evolution model, we analyzed the more physiologically relevant extended Tofts model of pharmacokinetics, which models the uptake of a contrast media in tissue.

METHODOLOGY

Modified CIRCUS k-Space Acquisition

Golden-angle sampling of k-space was simulated using a modified CIRCUS acquisition (15), with a repetition time of 3.7 milliseconds. Although multiple golden-angle sampling techniques are possible, we chose CIRCUS because it is a Cartesian trajectory, making for easy implementation and analysis. CIRCUS subdivides k-space into nested squares on a Cartesian grid, with rudimentary sampling patterns formed by selecting one point from each square according to a golden-angle rotation (2). We refer to these rudimentary sampling patterns as "quanta." The 2D CIRCUS patterns are extended for use in 3D imaging by extending the sampling mask along the third dimension (ie, the frequency encoded dimension). The shapes of each quanta are controlled by parameters b and c, which apply linear and nonlinear shifts to the position of successive samples and may loosely be thought of as contributing to the "shear" and "twist" of the quanta. The value of b must be positive, and c should be between 1 and 2; using $b = 0$ and $c = 1$ yields quanta that are straight lines, whereas $b \gg 1$ and $c = 2$ maximizes randomness. For this work, we set $b = 40$ and $c = 1.5$, which are consistent with the values recommended by the authors of CIRCUS.

Individual quanta are too highly undersampled to yield useful images, so they were combined into groups of 10 to form "packets." Quantities of packets ranging from 3 to 21 in steps of 2 were then interleaved in various numbers when forming image reconstructions. The number of packets interleaved then dictated the temporal resolution of the resulting reconstructions, where more packets meant a larger temporal footprint.

The k-space data were sampled by taking the Fourier transform of the image and selecting the coefficients indicated by the CIRCUS packets. In addition, we opted to perform the studies presented herein without the addition of simulated noise to the signal. We acknowledge this to be a simplification of the underlying signal acquisition physics; however, it allows a preliminary investigation of the application of IQMs to the analysis of dynamically evolving data and to determine their fundamental response to such data in the absence of external factors, which is the goal of the present work.

Retrospective Static Pelvic CS-MRI

To show the performance of IQMs in static MRI data as a function of reconstruction parameters, 15 static pelvic images were retrospectively undersampled. The acquisition of patient data for this research was approved by the Nova Scotia Health Authority Research Ethics Board and complied with all ethical obligations. Data were acquired on a GE MR750 3 T system using a 32-channel abdominal radiofrequency coil array and a 3D spoiled gradient echo sequence (DISCO) (16) with fat separation. Data were acquired with auto-calibrated parallel imaging (acceleration factor 2) in the phase-encoding direction, with all omitted lines reconstructed before retrospective undersampling.

Retrospective undersampling of the static pelvic images was performed using the modified CIRCUS acquisition described previously. Combinations of packets were chosen resulting in undersampling factors ranging from 1 to 12 in steps of 0.5. CS

reconstructions used the Berkeley Advanced Reconstruction Toolbox (17) to solve the following system:

$$\arg \min_{\rho} \left(\|\mathcal{F}[\rho] - \boldsymbol{k}\|_2^2 + \lambda \|\Psi\rho\|_1 \right) \qquad (1)$$

where $\mathcal{F}$ is the Fourier transform operator, ρ is the estimated image reconstruction, $\boldsymbol{k}$ are the measured k-space data, and λ is the regularization weight given to enforcing sparsity in the transform domain Ψ at the expense of data consistency. The transform domain Ψ should be a basis within which the image can be represented by relatively few elements; said another way, the image must be compressible within Ψ (18). A limiting factor in the choice of Ψ is the ability of the basis to represent the image in this way. A regularization weight λ of 0 represents no enforcement of sparsity in the domain Ψ, and increasing values of λ represent a compromise wherein data consistency requirements are relaxed to allow for a solution to be found that is sparse in Ψ. Choosing $\lambda = 0$, therefore, results in no application of CS, and very large values of λ will result in reconstructions that are of poor quality owing to a lack of consistency with the acquired data. An appropriate choice of λ is typically relatively small, and the final value of λ to use in a CS reconstruction has frequently been an empirical choice.

For this reconstruction, we chose a wavelet basis for Ψ and used values of λ ranging from 0 to 0.08 in steps of 0.002. The quality of the resulting image reconstructions was assessed by RMSE and SSIM.

Dynamic CS-MRI Simulations

Simulations were written in Matlab 2016b (The MathWorks, Natick, MA). A digital phantom of size $200 \times 200 \times 32$ voxels was designed to contain features that evolve over time. The embedded features of the phantom are shown in Figure 1. The plane of the embedded cylindrical features at the very center of the phantom was chosen for assessment in this work, with the remaining features in the phantom intended to provide flexibility for future work. Within the plane of the embedded cylindrical features used, we placed regions of interest (ROIs) inside the inner cylindrical features of a row of cylinders, shown in yellow in Figure 1, for quantitative parameter reconstructions. The ROIs for each embedded cylindrical feature were combined and averaged for analysis. In total, 596 voxels were used in the combined ROIs.

For this work, we simulated 2 scenarios. To begin, a simple monoexponential decay model was simulated. This model consisted of 2 free parameters—the initial amplitude and the decay rate of the signal amplitude. In total, 200 packets were acquired during the simulated evolution of the phantom, with individual k-space samples acquired at a simulated interval of 3.7 milliseconds. This result corresponded to 2 minutes and 32 seconds of simulated data acquisition. From these simulations, we fit the model to the reconstructed time-courses extracted from each voxel to recover both the initial magnitude and decay rate parameters. The use of a simple model such as this one validated that the simulations were providing rational results before extending it to a more realistic and complex one such as the extended Tofts model. For validation, we expected the following:
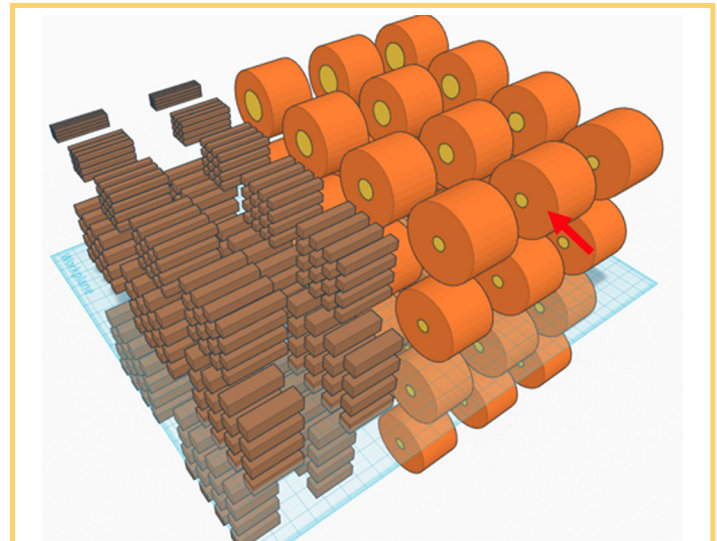

Figure 1. Internal features of the digital phantom embedded in a homogeneous static background. Within each plane of pin-grids (shown in brown), all grids can be made to evolve together according to a user-designed function. Within each plane of embedded cylinders, the outer cylinders (shown in orange) can be made to evolve together according to a user-designed function, and the inner cylinders (shown in yellow) can be made to evolve independently along each row. The interior cylinders in one plane of embedded cylindrical features, indicated by a red arrow, were used for this work; a 2D view of this slice is given in Figure 2F.

1. A trade-off between temporal resolution and image quality would be evident.
2. The temporal footprint (ie, the undersampling factor) used would have a larger effect than the degree of regularization applied.

Following the simple simulations described previously, features in an extended Tofts model were then simulated to evolve according to equation [2]:

$$C_t(t) = v_p C_p(t) + K^{trans} \int_0^t C_p(\tau) e^{-\frac{K^{trans}}{v_e}(t-\tau)} d\tau. \qquad (2)$$

The concentrations of contrast agent in the tissue and in the blood plasma are represented in equation [2] by $C_t(t)$ and $C_p(t)$, respectively. The terms v_e and v_p represent the volume fraction in tissue of the extravascular extracellular space and blood plasma, respectively. The term K^{trans} in equation [2] is a transfer constant, which is indicative of the permeability of the vasculature. From these simulations, we fit equation [2] to the reconstructed time-courses extracted from each voxel to recover K^{trans}. To implement equation [2], we used a population-averaged arterial input

function (AIF) to define $C_p(t)$ at each point in time (19). The Tofts model–defined time-courses evolved for 2 minutes and 30 seconds, with the contrast agent introduced after 30 seconds. This allowed for the collection of 197 packets.

We interleaved differing numbers of packets before CS-MRI image reconstruction when making reconstructions of both the monoexponential decay and the Tofts model scenarios. Combinations of 3–21 packets, in steps of 2 packets, were supplied to Berkeley Advanced Reconstruction Toolbox's *pics* command, which performs an eSPIRiT-based parallel imaging and CS image reconstruction (20), enforcing a wavelet sparsity constraint with regularization weights of 0–0.05 in steps of 0.005. This corresponded to undersampling factors ranging from 12.5 to 2, where the undersampling factor is defined as the reciprocal of the fraction of the k-space matrix that is sampled, for example, sampling 50% of the k-space coefficients results in an undersampling factor of 2. We chose to present the results of this portion of the study as a function of the number of packets, rather than of the undersampling factor, because it is the number of packets that the user has control over. These packet combinations corresponded to simulated temporal windows of ~2–13 seconds.

For the simple dynamic simulations consisting of a monoexponential decay, the reference image was taken to be the instantaneous value of the numerical phantom at the average time of all constituent packets; this is referred to as the instantaneous reference image. For the dynamic simulations using the Tofts model of signal evolution in equation [2], 3 additional reference images were generated from the simulated data. In practice, an instantaneous reference image would not be available, so this step allowed us to test the feasibility of using a data-generated reference image. The following were the 3 data-generated references:

1. A precontrast reference image made by acquiring an extra 30 seconds of data before contrast enhancement and combining all packets up until 5 seconds before contrast enhancement.
2. A peak-AIF reference image, made by combining all packets in a 20-second window starting at the peak of the AIF used to get $C_p(t)$ in equation [2].
3. A late-contrast reference image, made by combining all packets in the last 20 seconds of data acquisition where the contrast dynamics are relatively slow.

In testing these 3 additional data-generated references, reconstructions were made using 3–21 packets as described previously, but only a single regularization weight of 0.05 was used.

Image Quality Assessment
RMSE and SSIM were implemented as 3D IQMs. RMSE was implemented by calculating the voxelwise errors between the reconstructed image I_D and the reference image I_R, which is given by equation [3]:

$$RMSE = \sqrt{\frac{1}{N_x N_y N_z} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{k=1}^{N_z} [I_R(i,j,k) - I_D(i,j,k)]^2}.$$

(3)

The terms N_x, N_y, and N_z in equation [3] represent the number of voxels in the images along each dimension. We then normalized the value given by equation [3] by dividing by the average intensity of the reference image. A perfect RMSE score is 0, indicating no difference at all between the voxel contents in I_D and I_R.

The image processing toolbox in Matlab includes a 3D implementation of SSIM (21), which we used for this work. SSIM was calculated by equation [4]:

$$SSIM = \frac{(2\mu_R\mu_D + C_1)(2\sigma_{RD} + C_2)}{(\mu_R^2 + \mu_D^2 + C_1)(\sigma_R^2 + \sigma_D^2 + C_2)}$$

(4)

where μ_R and μ_D represent the average of the reference image and the reconstructed image in an $11 \times 11 \times 11$ voxel window, , respectively; σ_R and σ_D represent the standard deviations of voxel values within the window in each image; and σ_{RD} represents the cross-covariance for the images in the window. The constants were chosen as $C_1 = 0.01L$ and $C_2 = 0.03L$, where L is the maximum value of the image. The window size and the parameters specified were chosen for calculating SSIM because they represented the default values suggested by Wang (5) and implemented within Matlab (21). SSIM ranges between -1 and 1, with the score of 1 indicating optimal structural similarity between the I_D and I_R.

For IQM calculations during the retrospective static pelvic MRI investigation, IQM scores were calculated by comparing the CS-MRI reconstruction with the original reference image.

For IQM calculations during the dynamic simulations, RMSE and SSIM were calculated for each reconstructed image by comparing the reconstruction with the references outlined previously. This resulted in many IQM scores over time for a single simulation, which were then averaged to output a single IQM score that was representative of the entire simulation. Time-courses were extracted voxel by voxel from the expected feature locations in the reconstructed images. The exponential model parameters were recovered using the *lsqcurvefit* function in the Matlab optimization toolbox (22). The Tofts model parameters were recovered using a modified version of the ROCKETSHIP toolbox (23). The accuracy of recovered model parameters was compared with the known input parameters to assess the quantitative accuracy of the reconstructed temporal data.

RESULTS
Retrospective Static Pelvic CS-MRI
Figure 2 shows the effect of varying the regularization weight for a wavelet sparsity constraint and the undersampling factor for static CS-MRI reconstructions of retrospectively undersampled data, from a representative patient and from static numerical phantom data. In general, higher SSIM score corresponded to lower RMSE score. The regularization weight resulting in preferential IQM score (defined as the maximal SSIM or the minimal RMSE) tended to differ between the IQMs at each undersampling factor. This is shown in Figure 3, which displays the regularization weight resulting in preferential IQM score for static CS-MRI reconstructions at each undersampling factor for all 15 patients in the study and our numerical phantom. It was additionally observed that the preferential regularization weight varied among patients in the study. The preferential regularization
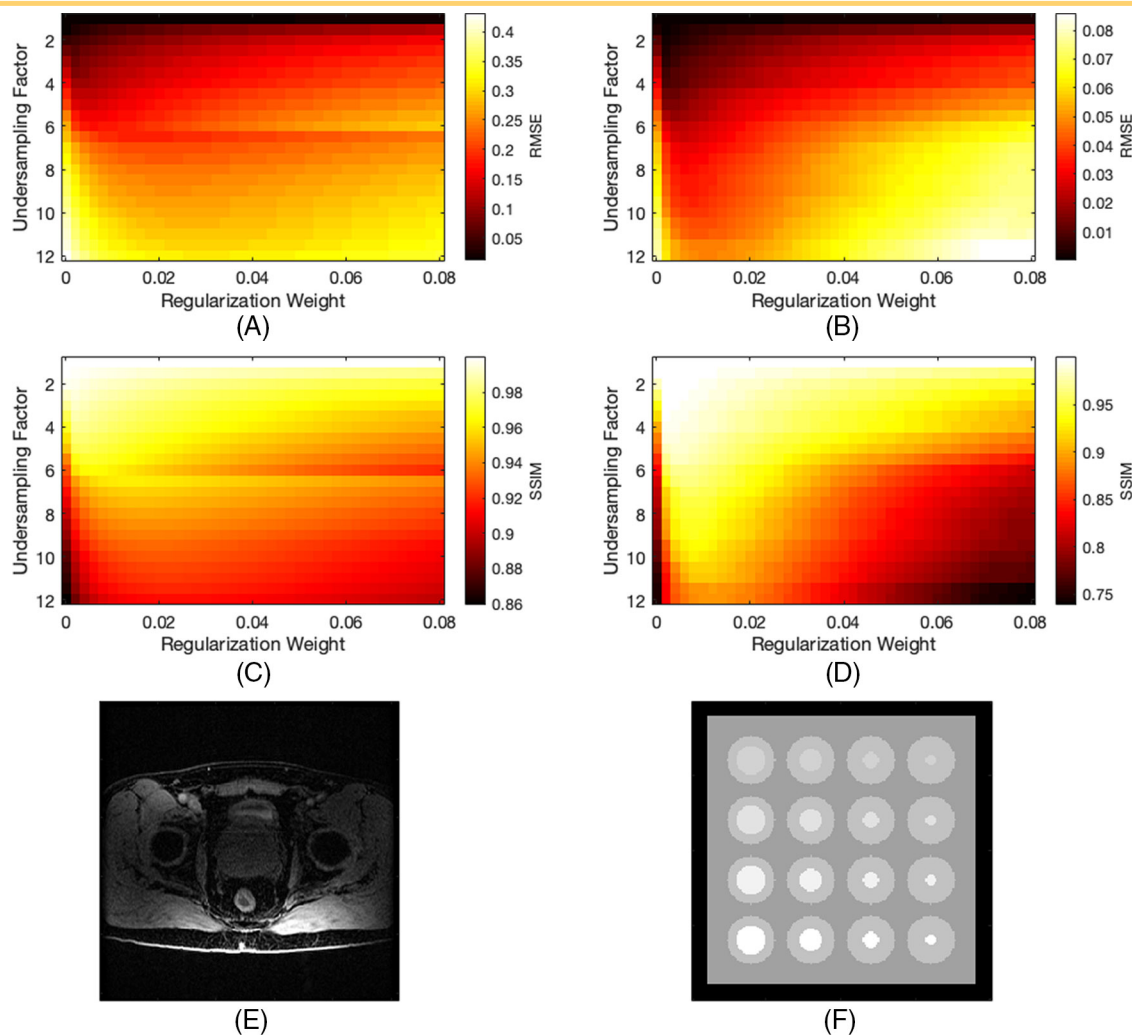
Figure 2. Examples of the effect of under-sampling factor and regularization weight on the root mean square error (RMSE) and structural similarity index (SSIM) scores for static compressed sensing–magnetic resonance imaging reconstructions. RMSE and SSIM scores for representative retrospectively undersampled patient data are shown in (A) and (C), respectively, whereas RMSE and SSIM scores for numerical phantom data are shown in (B) and (D), respectively. Unprocessed images of the patient and of the numerical phantom are shown in (E) and (F), respectively.

weight increased monotonically with an increase in the under-sampling of k-space for most patients, but not for all.

Simulations of Dynamic Digital Phantom With Monoexponential Signal Evolution

Figure 4 shows the effect of changing the regularization weight and the number of packets used in a simple CS-DCE-MRI scenario, where features evolve in intensity according to a monoexponential decay. While choosing a regularization weight of 0 consistently resulted in poorer performance, nonzero regularization weights were observed to improve the time-course fidelity. Using too few packets resulted in poor time-course fidelity owing to image degradation, and it was generally observed that proper choice of the number of packets included in a reconstruction had a larger effect than tuning of nonzero regularization weights.

Figure 5 shows the relationship between the recovery accuracy of quantitative parameters describing a monoexponentially

decaying feature and the image quality as assessed via RMSE and SSIM. Within Figure 5, reconstructions made using different numbers of packets have been color coded, and each color-coded group contains reconstructions made with regularization weights ranging from 0 to 0.05. Within each color-coded group, an outlier data point may be observed at a worse IQM score. These outliers corresponded to reconstructions made with a regularization weight of 0. Increasing the number of packets used in image reconstruction was only beneficial to a point, beyond which the error in each parameter again grew in magnitude. This outcome shows the expected trade-off between temporal resolution and image quality.

For both the amplitude and the decay parameter, a correlation between the error of the parameter recovery and each IQM score was observed. These correlations were quantified using the Pearson linear correlation coefficient (PLCC), which measures linear correlations, and the Spearman rank correlation coefficient
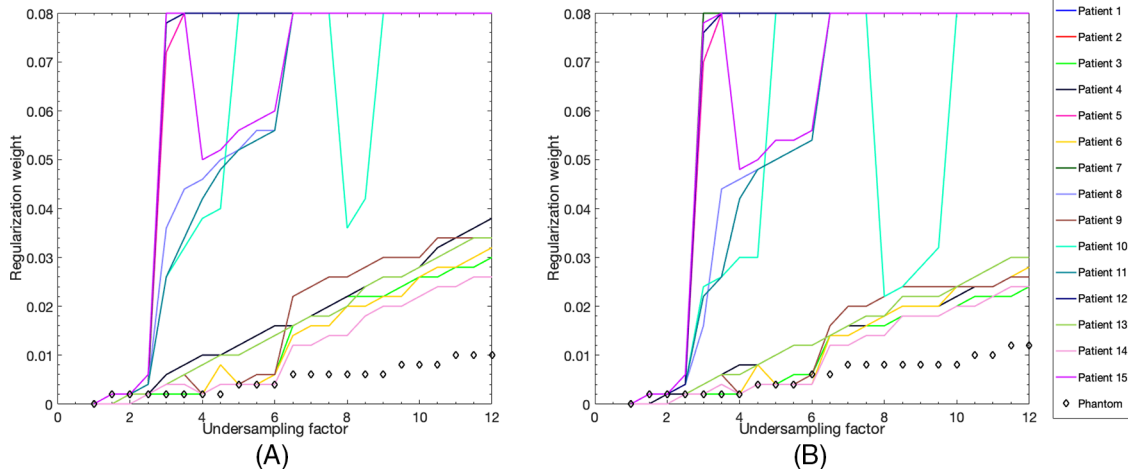
Figure 3. Preferential regularization weights with increasing undersampling factor are shown for the RMSE and the SSIM in (A) and (B), respectively. The preferential score was defined as the maximal SSIM or the minimal RMSE at each undersampling factor.

(SRCC), which measures the extent of a more general monotonic relationship. The values of the PLCC and SRCC for the data in Figure 5 are shown in Table 1.

Simulations of Dynamic Digital Phantom With Tofts Model Signal Evolution

Figure 6 shows the effect of changing the regularization weight and number of packets in a CS-MRI reconstruction on the recovered time-course fidelity of features whose intensity evolves according to equation [2]. The same observations seen in the monoexponentially evolving phantom simulation were observed

in this case. In addition, the choice of temporal resolution, made through the choice of the number of packets to include in an image reconstruction, most affected the ability to resolve the period of rapid contrast enhancement.

Figure 7 shows the relationship between the recovery accuracy of K^{trans} and the image quality as assessed via RMSE and SSIM. Color coding follows the same convention as described for Figure 5. Both RMSE and SSIM show the ability to define a range of IQM scores within which K^{trans} error is minimized, similar to the results for the simple simulation with a monoexponentially evolving phantom. The correlations
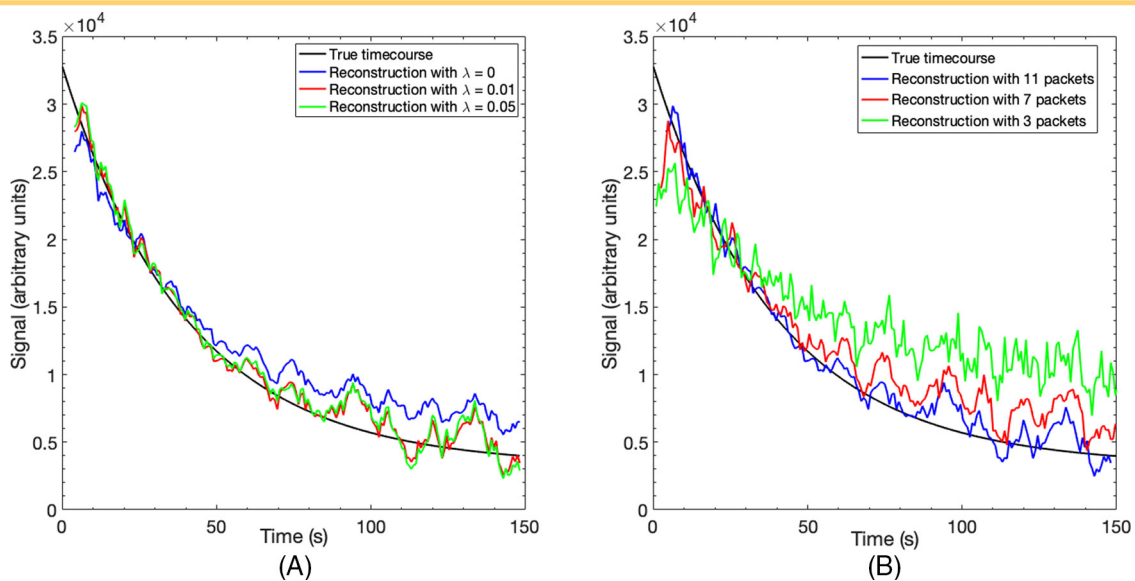

Figure 4. Examples of the fidelity of the average reconstructed time course for an exponentially decaying feature. There were 11 packets used in (A) as the regularization weight was varied, whereas (B) used a regularization weight of 0.01 and varied the number of packets included in image reconstruction.
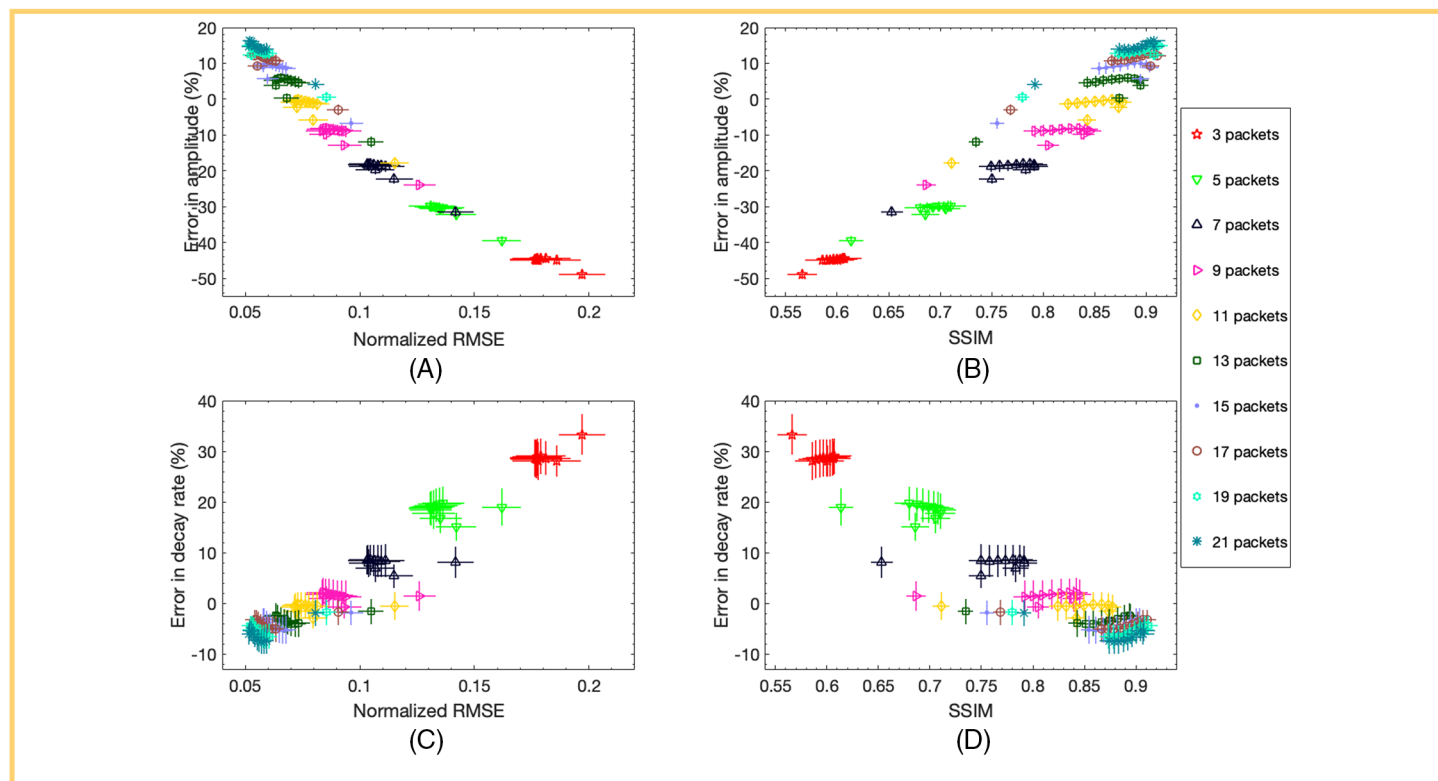
Figure 5. Relationship between RMSE scores with the recovery accuracy of quantitative parameters in simulated data of a monoexponentially decaying feature are shown in (A) and (C), whereas the relationship with the SSIM are shown in (B) and (D).

between the recovery accuracy of K^{trans} and RMSE and SSIM scores are shown in Table 2.

Figure 8 shows the relationship between IQM scores and the recovery accuracy of K^{trans} for the 4 different reference image scenarios described earlier, reconstructed with a regularization weight of 0.05. The choice of reference image was found to change the absolute values of the IQM score corresponding to a given parameter recovery accuracy and, in particular, the score corresponding to minimum parameter error. However, a clear relationship between IQM score and parameter accuracy was observed for all choices, and the observed correlations are shown in Table 3.

Table 1. Correlations between IQM Scores and Error in the Recovered Monoexponential Decay Parameters for the Data in Figure 5

		RMSE	SSIM
Initial Amplitude	PLCC	−0.973	0.982
	SRCC	−0.986	0.949
Decay Rate	PLCC	0.959	−0.963
	SRCC	0.937	−0.882

Abbreviations: PLCC, Pearson linear correlation coefficient; SRCC, Spearman rank correlation coefficient; RMSE, root mean squared error; SSIM, structural similarity index

DISCUSSION

Retrospective Static Pelvic CS-MRI

We assessed the response of RMSE and SSIM to wavelet regularization weight and k-space undersampling factor. Both IQMs show that there is a general decline in image quality as the undersampling factor is increased. Furthermore, at each undersampling factor, the effects of under or overregularizing the image are captured by the IQMs as a variation in their reported scores. There was also a tendency for the preferential regularization weight at each undersampling factor to differ slightly between patients and to not always change monotonically. This was true for both RMSE and SSIM, showing their sensitivity to CS-MRI reconstruction process for individual patients. The patient results potentially suggest the presence of 2 subpopulations of preferred CS regularization weights, evident in Figure 3; future research can be conducted to assess which image properties caused differences between the responses of the IQMs to each patient on an individual basis that may lead to such subpopulations in CS-MRI reconstruction parameters. The present results serve to confirm that each IQM is sensitive to the CS-MRI reconstruction process on a patient-by-patient basis when reconstructing golden-angle-sampled data.

These IQMs could thus be used on a patient-by-patient basis to guide the selection of image reconstruction parameters. These results motivate the assessment of IQMs in selecting image acquisition parameters for CS-DCE-MRI, where the appropriate choice of temporal resolution (ie, undersampling factor) and CS regularization weight is essential.
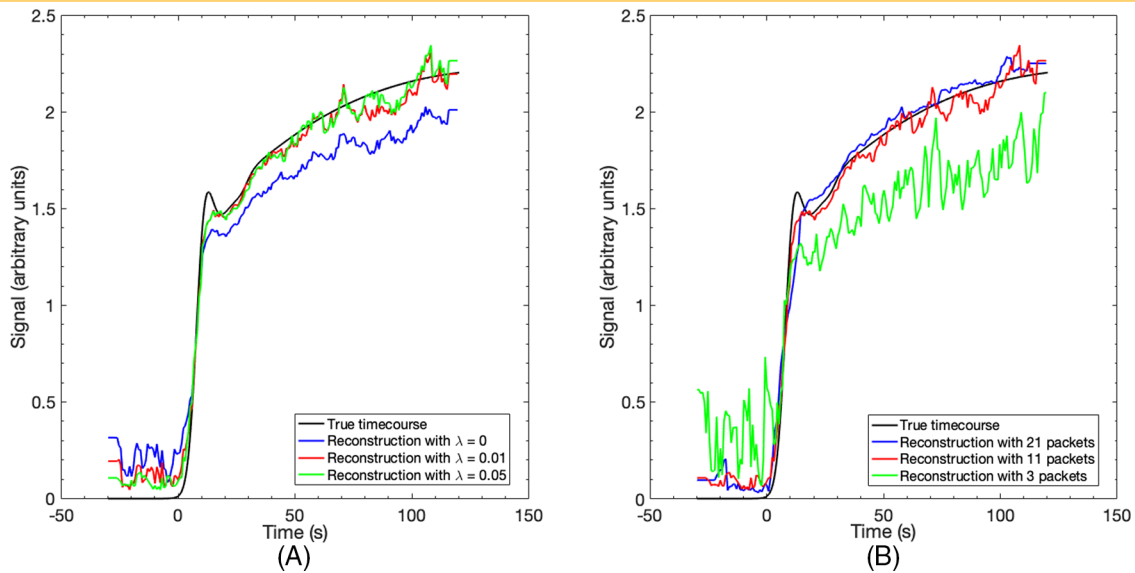
Figure 6. Examples of the fidelity of the average reconstructed time course for a voxel evolving according to the Tofts model described by equation [2], where the number of packets was held constant in (A) and the regularization weight was held constant in (B).

We note that the magnitude of each IQM, as well as the preferential regularization weight at each undersampling factor, tended to differ in magnitude between the synthetic phantom and the patient data. This is in part owing to the highly nonanthropomorphic nature of our simple synthetic phantom. Although our present work has established a basic understanding of the expected trends in RMSE and SSIM scores as a function of CS regularization weight and undersampling factor, we stress that our current work does not offer a quantitative rule to select these parameters. Future work may use a more anthropomorphically correct phantom that is expected to more closely mirror the reconstruction parameters desirable for real patient data.

We acknowledge that in practice, RMSE and SSIM may not be suitable for the assessment of prospectively accelerated static images. This is owing to the lack of a suitable reference image and the redundancy in generating one. However, the response of RMSE and SSIM IQMs in our simulated study indicates that alternative reference-free IQMs might be a worthwhile area of investigation for this purpose.
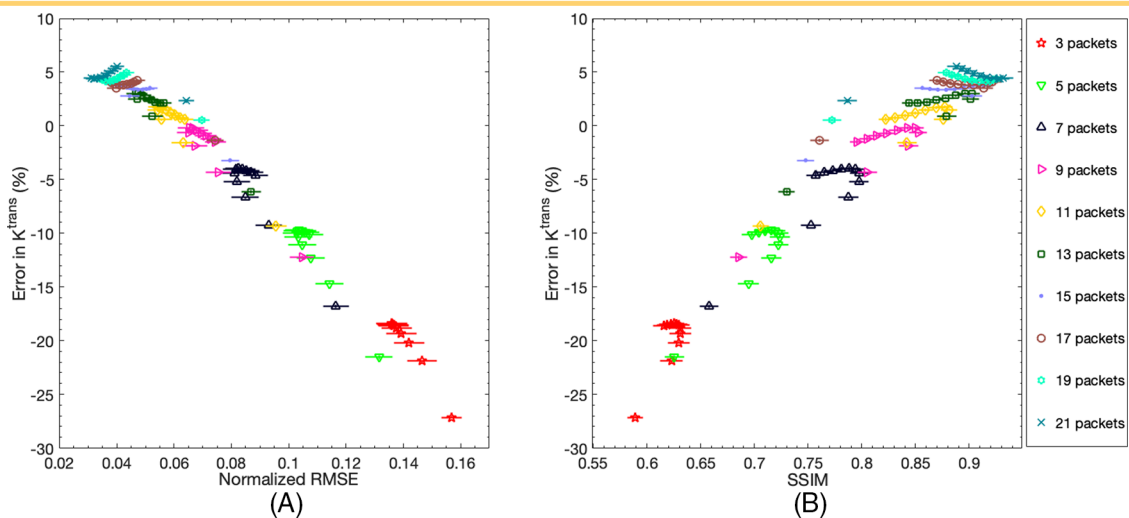

Figure 7. Relationship between RMSE scores with the recovery error in quantitative parameters in simulated data of a feature evolving according to the Tofts model described by equation [2] are shown in (A), whereas the relationship with the SSIM is shown in (B).

Table 2. Correlations between IQM Scores and the Error in the Recovered Tofts Model Parameter for the Data in Figure 7

		RMSE	SSIM
K^{trans}	PLCC	−0.986	0.968
	SRCC	−0.975	0.929

Abbreviations: PLCC, Pearson linear correlation coefficient; SRCC, Spearman rank correlation coefficient; RMSE, root mean squared error; SSIM, structural similarity index

Simulations of Dynamic Digital Phantom With Monoexponential Signal Evolution

To validate our simulation framework in the context of CS-DCE-MRI, we first simulated a phantom containing features that evolved according to a simple monoexponential function. This simple model allowed us to determine the effects of CS-DCE-MRI reconstructions on the accuracy of recovery of parameters describing contrast enhancement.

Image reconstruction parameters affected the recovered time-courses as we expected they should have. The reconstructed time-courses were most affected by the number of packets included in the reconstruction. While the choice of regularization weight was also found to influence the accuracy of the reconstructed time-course, in practice, any reasonable nonzero regularization weight was found to be beneficial.

Choosing too high of a temporal resolution (ie, a lower number of packets per reconstruction) degraded the image quality owing to the increased presence of incoherent aliasing artifacts, whereas choosing too low of a temporal resolution (ie, a larger number of packets per reconstruction) reduced the ability to capture the dynamic signal evolution. This outcome shows the expected compromise between temporal resolution and image quality.

We observed that the scale-related (ie, amplitude) and rate-related (ie, decay constant) parameters in our monoexponential model responded differently to the adjustment of CS regularization weight. In general, the use of a nonzero regularization weight was found to improve the IQM scores without affecting the recovery accuracy of the rate parameter. When recovering the scale parameter with a nonzero regularization weight, the recovery accuracy was found to have a positive offset compared with recovery with a regularization weight of 0. This shows the sensitivity of various model parameters to aspects of the image reconstruction process. For example, the scale parameter should be sensitive to losses in overall image intensity, whereas the rate parameter should be relatively robust to such an effect.

We also note that in the limit of a perfect IQM score (SSIM = 1 or RMSE = 0), the results would not yield perfect recovery accuracy of quantitative parameters. To achieve a perfect IQM score, assuming a noiseless simulated acquisition, a complete sampling of k-space would be required for a given image reconstruction. The temporal span of all the interleaved data for each image reconstruction would then span a period long enough to introduce blurring of the signal evolution. This trade-off in image reconstruction quality for the appropriate temporal resolution is captured by each IQM in our results, where continuously increasing the number of packets included in each image reconstruction was found to improve the IQM scores while degrading parameter recovery accuracy past a certain point.

The results from the monoexponentially evolving features thus agree with predictable outcomes, verifying the utility of our simulation framework for CS-DCE-MRI investigations. In particular, these results allow us to confidently determine a more complex contrast-enhancement model.
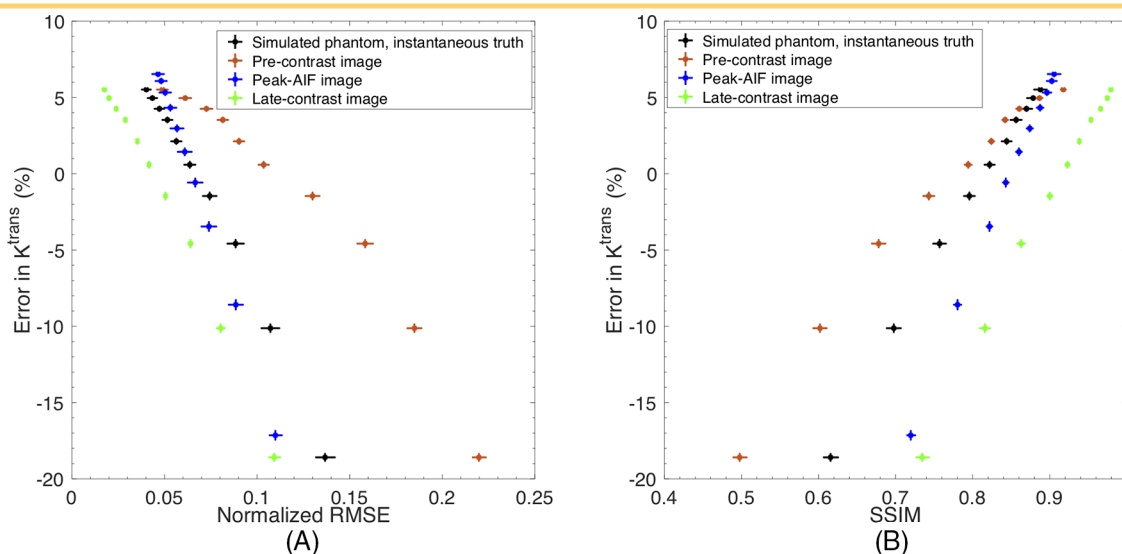

Figure 8. Effect of choosing each of the different reference images, and the relationship that results between each image quality metric and the recovery accuracy of K^{trans}. Results for the RMSE are shown in (A), whereas results for the SSIM are shown in (B). Images were reconstructed with a regularization weight of 0.05.

Table 3. Correlations between IQM Scores and Error in the Recovered K^{trans} for the Data Shown in Figure 8

		RMSE	SSIM
Instantaneous Truth	PLCC	−0.997	0.998
	SRCC	−1.000	1.000
Precontrast	PLCC	−0.976	0.988
	SRCC	−1.000	1.000
Peak AIF	PLCC	−1.000	0.999
	SRCC	−1.000	1.000
Late-contrast	PLCC	−0.995	0.998
	SRCC	−1.000	1.000

Abbreviations: PLCC, Pearson linear correlation coefficient; SRCC, Spearman rank correlation coefficient; RMSE, root mean squared error; SSIM, structural similarity index

Simulations of Dynamic Digital Phantom with Tofts Model Signal Evolution

The Tofts model exhibits an underlying signal evolution with more complex and physiologically relevant features. This increased temporal signal complexity further motivates the requirement to properly select the image reconstruction parameters in a physiologically realistic CS-DCE-MRI application.

Similar to the monoexponentially evolving signal model, the reconstructed time-course for the Tofts model was most sensitive to the number of packets included in the reconstruction. The effect of choosing too low or too high of a temporal resolution was especially pronounced on the recovered time-courses owing to their more complex structure; the period of rapid signal enhancement between 0 and 10 seconds was especially prone to losing its structure.

If RMSE or SSIM were to be implemented in a clinical setting, there would be no "instantaneous reference image," unlike in our simulated data, and some appropriate choice of a reference image would need to be made. In addition to IQM scores calculated by using the instantaneous reference image, we determined the possible use of 3 realistic methods of obtaining a data-generated reference image. Despite the more complex model for signal evolution under the Tofts model, the recovery accuracy of K^{trans} was found to correlate very well with both IQMs for all 4 reference images tested. This suggests that RMSE and SSIM may be used to select the appropriate reconstruction parameters in CS-DCE-MRI to reach a minimum tolerance of parameter recovery accuracy. These results motivate further assessment of the use of these IQMs to guide CS-DCE-MRI reconstructions for patients in clinical settings. For example, algorithms incorporating RMSE or SSIM could be developed for the automated selection of image reconstruction parameters in golden-angle-sampled CS-DCE-MRI.

We note that we have made direct use of a population-averaged AIF in our present study. Patient-specific measurement of pharmacokinetic parameters should ideally make use of a patient-specific AIF, which affects the $C_p(t)$ term in equation [2] through measurement uncertainties introduced by inadequate temporal resolution. Our direct use of a population-averaged AIF ignores the extra error that would be introduced into parameter reconstructions from a patient-specific AIF measurement; however, the goal of our present study was to establish a baseline from which such factors could be studied. Future work may begin to determine the effect of simulated "in vivo" measurements of AIF on the accuracy of the reconstructed parameters and the resulting correlations with IQMs such as RMSE or SSIM.

Although the extended Tofts model (equation [2]) is parametrized by 3 parameters, namely, v_e, v_p, and K^{trans}, we chose to determine the relationship between RMSE and SSIM scores and the accuracy of recovery of K^{trans} only. We chose to focus exclusively on K^{trans} because previous work has found it to be of primary clinical performance. In the diagnosis of prostate cancer, for example, it was found that the inclusion of pharmacokinetic parameters aside from K^{trans} does not improve performance of data-driven cancer detection using DCE-MRI techniques (24), and several successful computer-aided diagnosis tools exist that make use of only K^{trans} (25, 26). A subject for future assessment would be the characterization of the relationship of different pharmacokinetic parameters to objective measures of image quality, such as RMSE or SSIM.

In general, our current work has shown preliminary results showing that objective measures of image quality can serve as predictors of dynamic MRI data quality, which opens new research avenues. Future work may determine the use of other IQMs in CS-DCE-MRI applications or in alternative dynamic MRI applications. This includes both IQMs requiring reference images and those that do not require a reference image, such as the perception-based image quality evaluator IQM (27) or the blind/referenceless image spatial quality evaluator IQM (28).

CONCLUSION

We have determined the relationship between 2 IQMs, RMSE and SSIM, and their correlation with the recovery accuracy of K^{trans} in CS-DCE-MRI. To accomplish this, we began with a study of the properties of each IQM in retrospectively undersampled static pelvic CS-MR images. Each IQM was found to be sensitive to changes in the CS-MRI reconstruction process on a patient-by-patient basis.

We then studied RMSE and SSIM in CS-DCE-MRI by designing a simulation framework in Matlab. To validate our simulation framework, we assessed a numerical phantom containing features that evolved according to a simple monoexponential decay. The simulation results agreed with expected outcomes, validating the use of our framework for more complex contrast evolution models such as the extended Tofts model.

Using our simulation framework, we determined the ability of RMSE and of SSIM to predict the recovery accuracy of K^{trans}. Both IQMs require the use of a defined reference image. For 4 different methods of selecting a reference image (one based on the instantaneous image of the simulated phantom, one based on precontrast imaging, one based on late-contrast imaging, and one based on the peak of the AIF used in the Tofts model), we found that both RMSE and SSIM consistently correlate well with the recovery accuracy of K^{trans}. This motivates further assessment into the use of IQMs to guide CS-DCE-MRI in clinical settings,

such as in the design of automated algorithms for tuning the reconstruction parameters of CS-DCE-MRI acquisitions. More generally, these results suggest that objective measures of image quality have utility in the guidance of dynamic MRI techniques.

REFERENCES

1. Feng L, Axel L, Chandarana H, Block KT, Sodickson DK, Otazo R. XD-GRASP: golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. Magn Reson Med. 2016;75:775–788.
2. Liu L, Saloner D. Accelerated MRI with CIRcular Cartesian Under-sampling (CIRCUS): a variable density Cartesian sampling strategy for compressed sensing and parallel imaging. Quant Imaging Med Surg. 2014;4:57–67.
3. Winkelmann S, Schaeffter T, Koehler T, Eggers H, Doessel O. An optimal radial profile order based on the golden ratio for time-resolved MRI. IEEE Trans Med Imaging. 2007;26:68–76.
4. Wang Z, Bovik A. Mean squared error: love it or leave it? IEEE Signal Process Mag. 2009;26:98–117.
5. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13:600–612.
6. Kojima S, Shinohara H, Hashimoto T, Suzuki S. Under-sampling patterns in k-space for compressed sensing MRI using two-dimensional Cartesian sampling. Radiol Phys Technol. 2018;11:303–319.
7. Zheng H, Zeng K, Guo D, Ying J, Yang Y, Peng X, Huang F, Chen Z, Qu X. Multi-contrast brain MRI image superresolution with gradient-guided edge enhancement. IEEE Access. 2018;6:57856–57867.
8. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, Knoll F. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med. 2018;79:3055–3071.
9. Gözcü B, Mahabadi RK, Yen-Huan L, Ilicak E, Çukur T, Scarlett J, Cevher V. Learning-based compressive MRI. IEEE Trans Med Imaging. 2018;37:1394–1406.
10. Duffy BA, Zhang W, Tang H, Zhao L, Law M, Toga Aw Kim H. Retrospective correction of motion artifact affected structural MRI images using deep learning of simulation motion. Conference on Medical Imaging with Deep Learning. 2018.
11. Jeelani H, Martin J, Vasquez F, Salerno M, Weller DS. Image quality affects deep learning reconstruction of MRI. IEEE 15th annual symposium on biomedical imaging. 2018:357–360.
12. Mason A, Rioux J, Clarke SE, Costa A, Schmidt M, Keough V, Huynh T, Beyea S. Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of MR images. IEEE Trans Med Imaging. 2020;39:1064–1072.
13. Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, Larsson HBW, Lee TY, Mayr NA, Parker GJM, Port RE, Taylor J, Weisskoff RM. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. Magn Reson Med. 1999;10:223–232.
14. Nguyen VL, Kooi ME, Backes WH, van Hoof RHM, Saris AECM, Wishaupt MCJ, Hellenthal F, van der Geest RJ, Kessels AGH, Schurink GWH, Leiner T. Suitability of pharmacokinetic models for dynamic contrast-enhanced MRI of abdominal aortic aneurysm vessel wall: a comparison. PLoS One. 2013;8:6–12.
15. Rioux JA, Murtha NJ, Mason A, Bowen CV, Clarke S, Beyea S. Proc Intl Soc Mag Reson Med Honolulu, HI, USA. 2017:1426.
16. Saranathan M, Rettmann DW, Hargreaves BH, Clarke SE, Vasanawala SS. Differential subsampling with cartesian ordering (DISCO): A high spatio-temporal resolution Dixon imaging sequence for multiphasic contrast enhanced abdominal imaging. J Magn Reson Imaging. 2012;35:1484–1492.
17. Uecker M, Ong F, Tamir JI, Bahri D, Virtue P, Cheng JY, Zhang T, Lustig M. Berkely Advanced Reconstruction Tool-Box (BART). Proc Intl Soc Mag Reson Med Toronto, ON, Canada. 2015:2486.
18. Lustig M, Donoho DL, Santos JM, Pauly JM. Compressed Sensing MRI. IEEE Signal Process Mag. 2008;25:72–82.
19. Parker GJM, Roberts C, Macdonald A, Buonaccorsi GA, Cheung S, Buckley DL, Jackson A, Watson Y, Davies K, Jayson GC. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. Magn Reson Med. 2006;56:993–1000.
20. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, Lustig M. ESPIRiT – An eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. Magn Reson Med. 2014;71:990–1001.
21. The Mathworks Inc. Structural Similarity Index. Image Processing Toolbox Reference. 2601–2602.
22. The Mathworks Inc. Least Squares. Optimization Toolbox User's Guide. 1161–1162.
23. Barnes SR, Ng TSC, Santa-Maria N, Montagne A, Zlokovic BV, Jacobs RE. ROCKETSHIP: a flexible and modular software tool for the planning, processing and analysis of dynamic MRI studies. BMC Med Imaging. 2015;15:19.
24. Haq NF, Kozlowski P, Jones EC, Chang SD, Goldenberg SL, Moradi M. A data-driven approach to prostate cancer detection from dynamic contrast enhanced MRI. Comput Med Imaging Graph. 2015;41:37–45.
25. Liu S, Zheng H, Feng Y, Li W. Prostate cancer diagnosis using deep learning with 3D multiparametric MRI. Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis, 1013428. 2017.
26. Wang X, Yang W, Weinreb J, Han J, Li Q, Kong X, Yan Y, Ke Z, Luo B, Liu T, Wang L. Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. Sci Rep. 2017;7:15415.
27. Venkatanath N, Praneeth D, Maruthi CB, Channappayya SS, Medasani SS. Blind image quality evaluation using perception based features. Twenty First National Conference on Communications, Mumbai. 2015;2015:1–6.
28. Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Trans Image Process. 2012;21:4695–4708.