Quantitative Magnetic Resonance Imaging Phantoms: A Review and the Need for a System Phantom

Kathryn E. Keenan,[1]* Maureen Ainslie,[2] Alex J. Barker [ID],[3] Michael A. Boss [ID],[1]
Kim M. Cecil,[4] Cecil Charles,[2] Thomas L. Chenevert,[5] Larry Clarke,[6]
Jeffrey L. Evelhoch,[7] Paul Finn,[8] Daniel Gembris,[9] Jeffrey L. Gunter,[10] Derek L.G. Hill,[11]
Clifford R. Jack Jr.,[10] Edward F. Jackson,[12] Guoying Liu,[6] Stephen E. Russek,[1]
Samir D. Sharma,[12] Michael Steckner,[13] Karl F. Stupic,[1] Joshua D. Trzasko,[10]
Chun Yuan,[14] and Jie Zheng[15]

The MRI community is using quantitative mapping techniques to complement qualitative imaging. For quantitative imaging to reach its full potential, it is necessary to analyze measurements across systems and longitudinally. Clinical use of quantitative imaging can be facilitated through adoption and use of a standard system phantom, a calibration/standard reference object, to assess the performance of an MRI machine. The International Society of Magnetic Resonance in Medicine AdHoc Committee on Standards for Quantitative Magnetic Resonance was established in February 2007 to facilitate the expansion of MRI as a mainstream modality for multi-institutional measurements, including, among other things, multicenter trials. The goal of the Standards for Quantitative Magnetic Resonance committee was to provide a framework to ensure that quantitative measures derived from MR data are comparable over time, between subjects, between sites, and between vendors. This paper, written by members of the Standards for Quantitative Magnetic Resonance committee, reviews standardization attempts and then details the need, requirements, and implementation plan for a standard system phantom for quantitative MRI. In addition, application-specific phantoms and implementation of quantitative MRI are reviewed. **Magn Reson Med 79:48–61, 2018. © 2017 International Society for Magnetic Resonance in Medicine.**

Key words: phantom; quantitative; system consistency; quality assurance

[1]National Institute of Standards and Technology Boulder, Colorado, USA.

[2]Duke University, Durham, North Carolina, USA.

[3]Northwestern University, Evanston, Illinois, USA.

[4]Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA.

[5]University of Michigan, Ann Arbor, Michigan, USA.

[6]National Institutes of Health, Bethesda, Maryland, USA.

[7]Merck Research Laboratories, West Point, Pennsylvania, USA.

[8]University of California, Los Angeles, California, USA.

[9]University of Cooperative Education Dresden, Dresden, Germany.

[10]Mayo Clinic, Rochester, Minnesota, USA.

[11]King's College, London, United Kingdom.

[12]University of Wisconsin, Madison, Wisconsin, USA.

[13]Toshiba Medical Research Institute, Mayfield Village, Ohio, USA.

[14]University of Washington, Seattle, Washington, USA.

[15]Washington University in St. Louis, St. Louis, Missouri, USA.

*Correspondence to: Kathryn E. Keenan, Ph.D., National Institute of Standards and Technology, 325 Broadway, MC 818.03, Boulder, CO 80305, USA. E-mail: kathryn.keenan@nist.gov

INTRODUCTION: NEED FOR QUANTITATIVE PHANTOMS

Over the past two decades, interest in the use of MR biological markers (or "biomarkers") to provide information critical to the development of novel therapeutic agents and improved clinical diagnostics has grown. Biomarkers (1–3) are objectively measured parameters that indicate the biological state, biological/pathobiological processes, or pharmacologic responses to treatment. Examples of MR biomarkers include tumor volume (4–6), brain volume (7–10), functional network connectivity (11–13), isotropic (14,15) or anisotropic (16,17) water diffusion constants (18), local metabolite concentrations (10,15,19,20), blood flow fields (21–23), fat fraction (24–27), lung function (28,29), temperature (30–32), and tissue elasticity (33,34).

Medical imaging modalities are now expanding to include quantitative mapping of biomarkers in addition to qualitative imaging. Although quantitative mapping of biomarkers can greatly increase the amount, reliability, and comparability of the data obtained from medical imaging, it requires careful standardization of protocols and the development of phantoms (standard reference objects or calibration structures) to validate the accuracy of these in vivo measurements, and to assess the repeatability and reproducibility of the measurements across imaging platforms and time.

Despite substantial recent advances in biomedical science, the process of developing more effective and safer therapeutics for patients has become increasingly challenging and costly (35). Magnetic resonance biomarkers are one potential way to address these problems, such as

in clinical trials that evaluate novel therapeutic agents or establish efficacy and/or safety for regulatory approval as a substitute for a clinical characteristic reflecting patient condition, function, or survival (i.e., surrogate endpoint). The expectation is that information provided by biomarkers will improve predictability and efficiency along the path from laboratory concept to commercial product (36). Another motivation for the implementation of image-based biomarkers is their use in safe, noninvasive diagnostics replacing biopsy-based diagnostics. Examples include the use of MR elastography to diagnose and stage liver cirrhosis and fibrosis (34,37) and dynamic contrast MRI to measure key tissue parameters of tumors (38) to assess the effectiveness of cancer treatment. Finally, the development of accurate and sensitive MR-based biomarkers may lead to physical diagnostics of conditions such as mild traumatic brain injury and many types of neural diseases for which there are no adequate physical diagnostics and which rely on neuropsychological assessment (17,39).

An important part of precision measurement of primary MR parameters is to develop rigorous definitions of the measurands. For example, the static magnetic field, B_0, contains contributions from the scanner magnet and from diamagnetic and/or paramagnetic components arising from the radiofrequency (RF) coil assembly and sample under study. Although these effects are relatively small, the impact of their presence can be significant on image quality and quantitative parameter accuracy and precision. For high-performance quantitative imaging, careful definitions and recommended procedures for measuring such effects are required. Perhaps more importantly, careful definitions of the proton spin relaxation times are required. Although exponential relaxation of the proton magnetization is often observed, multi-exponential or nonexponential relaxation can be present in many materials, including biological tissue. Different apparent relaxation times for complex materials are measured by different pulse sequences on different platforms (e.g., NMR spectrometer systems versus MRI scanners). Pragmatic definitions of apparent T_1 and T_2 relaxation times are required if one desires to use relaxation times as biomarkers of tissue type and disease processes.

The Radiological Society of North America (RSNA) established the Quantitative Imaging Biomarkers Alliance (QIBA) (40) to address this issue by developing quantitative imaging protocols, phantoms, and technical standards documents, referred to as profiles. In addition, the National Cancer Institute established the Quantitative Imaging Network (41) to help validate quantitative imaging through the use of standard protocols and phantoms. The December 2016 issue of *Tomography—A Journal for Imaging Research* was devoted to the work of the Quantitative Imaging Network, which is bringing quantitative imaging methods into clinical utility, measuring response to therapy, and supporting clinical decision making during clinical trials (42). The development of a standard system phantom by the Standards for Quantitative Magnetic Resonance committee is meant to support these efforts by establishing a procedure to develop MRI phantoms with traceable, validated, and monitored components.

Biomarkers must provide quantitative measures of anatomical, physiological, and/or biochemical characteristics that are comparable over time, between subjects, between scanner locations, between manufacturers, across protocols, and across field strengths. Such comparisons can be difficult because of a variety of purely technical factors ranging from subtle variations in hardware performance influencing the MR signal to differences in hardware and software between and within manufacturers, differences in image acquisition and reconstruction protocols, and differences in data processing and analysis. As with any analytical instrument, regular quality assurance of the MRI scanner allows many of these factors to be characterized and, when feasible, considered when extracting quantitative measurements from the MR data. Existing phantoms are designed for accreditation (see Supporting Information) or for measurement of a single specific scanner property or for a specific application (see "Application of Specific Phantoms" section). Few of the existing phantoms contain SI-traceable components, few are monitored for long-term stability, and few have been validated by a national metrology institute (43,44). In this paper, we propose that a standard system phantom, with stable and traceable properties enabling evaluation of as many critical aspects of the MRI system as possible, would facilitate the use of quantitative MRI measurements as a biomarker.

SYSTEM PHANTOM DESIGN CONSIDERATIONS

Magnetic Resonance Quantities to Be Evaluated by the System Phantom

The system phantom should be designed to assess basic system parameters such as signal-to-noise ratio (SNR), resolution, relaxation times, proton density, and geometric distortion, and to compare results across manufacturers, hardware and software versions, time, and physiologic ranges at 1.5 T and 3 T. A standard system phantom package requires, in addition to the physical phantom, standard imaging protocols, standard image analysis procedures, a full description of the phantom including field and temperature-dependent material properties, numerical description of the phantom to allow simulations, setup and imaging instructions, and a data archive to allow comparison of data and scanner assessment.

The American College of Radiology (ACR) phantom, described in the Supporting Information, was designed to partially meet the goals of a system phantom, but it has certain limitations. For example, certain measurements (e.g., section thickness and high contrast resolution) can only be obtained in one orientation. In addition, there is no organized long-term monitoring of the phantom components for stability. Over time, it is known that the acrylic will warp, which can render geometric components unusable (45).

Here we describe an MRI system phantom to meet the aforementioned goals, which will contain SI-traceable components, be monitored for long-term stability, and be validated by a national metrology institute. The system phantom can address several concerns with implementation of quantitative MRI, including system constancy and assessment of the data acquisition and analysis pipeline.

The following quantities can affect the accuracy or precision of measurement in MRI studies, and are candidate quantities to be evaluated by a system phantom. Some

quantities such as B_0, B_1, and gradient nonuniformity are primary factors that affect other quantities, such as SNR, and slice profile.

B_1 (Transmit) Nonuniformity

B_1 transmit (B_1^+) nonuniformity is a major confounding factor, especially with the use of high magnetic fields and surface coils for transmission. The accuracy of the flip angle achieved at any position depends on the B_1^+ inhomogeneity and can be determined by B_1^+ mapping (46). For example, the B_1^+ map may be used during T_1 mapping model fits to correct the desired versus achieved flip angle (47).

B_1 (Receive) Nonuniformity

With the increasing use of high channel count phased array and anatomy-specific surface coils for MR signal reception, B_1 receive (B_1^-) inhomogeneity must be assessed and addressed. Characterized by spatial variations in image intensity and SNR, it confounds both the accuracy and precision of many quantitative MRI applications if not prospectively accounted for during parameter estimation. The validity of the reciprocity principle that allows B_1 receive nonuniformity to be measured from transmit nonuniformity (B_1 map) has been challenged at very high field strengths.

B_0 Nonuniformity

Assuming good shimming, the main magnetic field, B_0, can generally be considered uniform over standard clinical imaging fields of view (FOVs). However, for extended FOVs (e.g., breast imaging), nontrivial deviations from uniformity can occur in the periphery. Susceptibility effects introduced by air–tissue interfaces or other non-tissue materials (e.g., gadolinium contrast media) will also induce local changes in the effective B_0 field. For certain applications, such as proton-density fat fraction estimation, B_0 inhomogeneity causes spatially dependent phasing of chemical species signals, whereas for other applications, such as echo-planar imaging–based diffusion imaging, it can lead to geometric image distortion. The effective B_0 field can be mapped using multi-echo procedures, or indirectly characterized using proxy measures like SNR and image uniformity.

Signal-to-Noise Ratio

The SNR is known to be influenced by several system factors, such as resonance frequency, flip angle accuracy, transmitter gain, coil loading (fill factor) and tuning, scan parameters, slice profile and shape, scan acceleration (e.g., use of parallel imaging), image reconstruction method, postprocessing, and parameter fitting strategy. Quantitative measurement of SNR can give a general indication of the state of the imaging pipeline.

To measure SNR, a relatively large, uniform compartment of a phantom filled with a solution that has well-characterized, stable proton density and T_1 and T_2 relaxation times is recommended to properly follow the National Electrical Manufacturers Association and International Electrotechnical Commission (IEC) methods

(48,49). This allows for the use of one of the four protocols described in the National Electrical Manufacturers Association standard MS1-2008 (R2014) to determine SNR (48). The first two protocols use image subtraction from a pair of nominally identical images to determine image noise; these methods are sensitive to system drift artifacts and suggest that the images be obtained within a minimum of elapsed time. The third protocol uses a single k-space scan to produce two images, which are subtracted; this method reduces the sensitivity to system drift. The fourth protocol measures the noise on a zero-signal region outside the phantom. It is imperative to reference the SNR measurement method used to assess the data to allow comparisons. We do recognize that given the multiple goals of the proposed system phantom, it is space-limited, and therefore may not be able to include large volumes recommended for SNR measurements.

Image Uniformity

Ideal homogenous signal across the FOV can be affected by many factors, including, but not limited to, B_0, B_1 nonuniformities (transmit and receive), gradient linearity, eddy currents, and postprocessing. Image uniformity is also a general indicator of the performance of the imaging pipeline.

For measuring image uniformity, National Electrical Manufacturers Association methods in standard MS3-2008 (R2014) are recommended (50). The method recommends a phantom that covers at least 85% of the specification area. The proposed system phantom should cover at least 85% of most head coils. The method also recommends that the fill solution have physiologic T_1, T_2, and spin density values. Multiple methods are outlined to calculate the image uniformity and assess the image contrast.

Gradient Amplitude

Inaccuracies in gradient amplitude can affect measurements of object size, which may be critical for studies requiring accurate registration. This parameter can be measured by comparing image measurements of phantoms with known properties.

Geometric Linearity

Geometric linearity can be affected either by B_0 nonuniformity and/or gradient nonlinearity. Gradient nonlinearity typically leads to geometric distortion and needs to be measured in all three axes. These nonlinearities are more pronounced at the edges of the FOV. Manufacturers apply at least a 2-dimensional gradient nonlinearity correction before image display, but this correction may not be sufficient for some applications (51) and can degrade image resolution (52).

To assess the geometric accuracy for head size volumes, the Alzheimer's Disease Neuroimaging Initiative (ADNI) MagPhan phantom (The Phantom Laboratory, Salem, New York, USA) can be used (53,54). Any phantom design for such application must make it possible to identify the phantom orientation from the image itself. This requires that there is sufficient asymmetry in the phantom or built-in fiduciary features to enable the

orientation to be automatically determined. This can be in conflict with the design of a geometric distortion phantom, which requires a regular set of points. When using the ADNI phantom, the ADNI software should be used to assess the geometric distortions, to minimize any measurement variance caused by the use of varying analysis software packages (54).

Slice Position and Profile

B_0 variations, RF amplifier nonlinearities, and gradient nonlinearity problems can affect slice position and profile. Hence, accurate determination of the slice profile is required in MRI. The B_1 variation across the slice warrants accurate measurement of slice profile and slice crosstalk effects. Accuracy of slice separation is a related factor.

Contrast Compartments

Contrast response, such as T_1, T_2 and proton density measurements, can be measured through any number of experiments and signal models. Contrast compartments allow testing of the entire measurement protocol, including data acquisition and analysis. These components can reveal issues from the scanner or acquisition and from the processing pipeline, such as an incomplete signal model.

The recommended system phantom should have three groups of at least 10 compartments. Within each group, a single parameter (proton density, T_1, or T_2) changes in a manner that is as independent as possible with respect to the other two parameters. Care must be taken to ensure that all spheres required for geometric accuracy assessment using automated analysis procedures have appropriate signal characteristics on at least one specific pulse sequence and set of acquisition parameters. If the contrast compartments are temperature dependent, the temperature dependence should be measured, and temperature correction coefficients should be reported.

In the proton density group, at least 10 compartments in which the proton density varies linearly over the range of 50 to 100% relative to pure water is desired.

In the T_1 relaxation time group, at least 10 compartments in which the T_1 relaxation time varies linearly over the range of 100 to 2000 ms at 1.5 T is desired. The T_1 relaxation time values included should cover physiologic ranges at both 1.5 T and 3 T, including short relaxation times typically encountered intravascularly following the administration of gadolinium contrast media.

In the T_2 relaxation time group, at least 10 compartments in which the T_2 relaxation time varies linearly over the range of 20 to 200 ms at 1.5 T is desired. The T_2 relaxation time values included should cover physiologic ranges at both 1.5 T and 3 T.

The software should be open source to allow users to include their quantitative analysis for T_1, T_2, and proton density measurements. Software that enables researchers to use their own model for analysis encourages comparison of the acquisition and processing procedures for relaxometry. The software should include models for the recommended imaging protocol and allow for new relaxometry experiments and models.

High-Contrast Resolution

Point spread function, line spread function, or modulation transfer function could provide quantitative measures of spatial resolution (49). These are in addition to the "number of objects resolved" metric used to assess the ACR phantom high-contrast resolution insert. (Current ACR guidelines state "one visually determines the number of individual small bright spots" (55,56).)

System Constancy

Scanner performance and stability should be tracked over time for a range of the parameters described in this section and others, such as transmitter and receiver gain, receiver bandwidth, image ghosting (49), and eddy currents (57). Deviations in the system constancy measurements detected using the system phantom can reveal equipment failure or an underlying issue before it is noticed in clinical imaging (45).

Specific Design Criteria

1. All components should be in the public domain, including the phantom design, solid models, and material properties.
2. The standard system phantom should allow characterization of bias and variance of most of the desired quantities listed in "Quantities to be Evaluated," with the caveat that no single phantom will be optimal for all of the items listed.
3. No specific recommendation is provided with respect to spherical versus cylindrical geometry. Given the large variety of MR coils, including multichannel head coils, breast coils, and knee coils, a system phantom design will never be suitable for imaging in all current configurations.
4. All filling materials should be well-characterized with respect to physical NMR properties and stability. No user-fillable compartments should be included to maintain consistent and traceable phantom characteristics.
5. The phantom should be handled easily and positioned by MRI technologists. A positioning device should be designed to allow the system phantom to be accurately and precisely positioned at locations off-isocenter, including the volume of a typical thoracic or abdominal cavity (z-direction) and coverage of the shoulders and/or hips (x-direction). It is recognized that this will require a manufacturer-specific design component or will require that the site can independently and reproducibly provide an appropriate surface upon which the proposed positioning device can be placed.
6. The basic imaging protocol duration should be less than 1 h. Although more in-depth imaging protocols may be included, the general use of the phantom will require fitting the image acquisitions into tight schedules. There should be an automated analysis of the measured data to encourage regular use of the phantom for quality control purposes. The algorithms must perform well over prescribed ranges of SNR and artifact levels (e.g., geometric distortion, B_1

nonuniformity). The software should have well-defined regions of interest to enable automated selection of signals of interest.

7. The system phantom should have cost commensurate with existing phantoms. A complex, expensive system phantom may have diminishing value.

8. The phantom should be easy and safe to ship (such as if the phantom is dropped), and any hazardous materials should be contained so that they would not leak and require hazardous clean-up.

9. The phantom should be robust, with at least a 5-year stability, ideally close to 10 years.

10. The phantom should have well-defined accuracy and SI-traceability of important properties, such as dimensional parameters and composition of contrast compartments, such as using inductively coupled plasma mass spectroscopy. A metrology institution should be enlisted to verify accuracy and monitor stability of the system phantom.

11. The design should allow for the development and implementation of automated evaluation software tools. For example, the orientation of the phantom should be uniquely determined from the images; it does not need to be known beforehand. Ideally, it should be possible to analyze the images even if the scans are partially truncated or incorrectly oriented in the FOV, as human error will lead to such issues, particularly in large multicenter trials.

12. The phantom should come with an open-source analysis package to allow consistent analysis of imaging data using common DICOM format. Alternative or more extensive analysis should be encouraged by including all information required for analysis, such as region of interest positions. The image analysis software must be able to read images in the DICOM file format and interpret information specific to each manufacturer, even from private tags that are sometimes required for analysis. The software should be open source to allow users to test new algorithms or fitting methods.

13. The phantom should come with required environmental monitoring, such as a thermometer, to adequately assess potential non-system/environment-dependent effects (e.g., temperature).

14. The image analysis software should allow for advanced/complete protocol analysis. The requirements may be different for a phantom used to detect whether a system is within manufacturer specifications compared with a phantom used for measurements to normalize or alter the actual data.

15. Certain features, such as the high-contrast resolution inset, should be compatible with other imaging modalities (e.g., CT, positron emission tomography, ultrasound) to have a single standard, when possible.

CURRENT MRI STANDARDS, PHANTOMS, AND QUANTIFICATION EFFORTS

Recognizing the need for standard phantoms, several organizations/initiatives have developed MRI phantoms. They include phantoms to (i) characterize the physical performance of MRI systems for acceptance testing and comparison of different commercial systems performance, (ii) characterize time-related changes in the physical performance of imaging systems for specific clinical protocols, and (iii) develop methods for accreditation of MRI systems for clinical practice. We identified significant efforts by the American Association of Physicists in Medicine (58,59), European Communities Biomedical Engineering Advisory Committee (60–62), Magnetic Resonance National Evaluation Team (63–67), and the ACR (56,68,69), including the ACR MR Accreditation Phantom (Supporting Fig. S1). Each of these efforts is briefly reviewed in the Supporting Information.

APPLICATION OF SPECIFIC PHANTOMS

Application-specific phantoms are developed to evaluate a specific biomarker or to enable quality assurance of a measurement. Unlike phantoms used to characterize fundamental features of MRI systems, such as the proposed system phantom, these phantoms focus on those parameters that are specific to the target application. The purpose of this section is to discuss the objective measurement, the rationale for that measurement, and details of the phantom design for each application. The discussion is limited to phantoms whose descriptions have been disclosed in either publications or publicly available abstracts and proceedings papers.

Structural Brain Imaging Phantom

The ADNI program for structural brain imaging created a phantom that fits within many head coils and is scanned immediately after each patient scan (54). This head-volume phantom was used by three successful study phases: ADNI1, ADNI Grand Opportunity (GO), and ADNI2. The ADNI program has used the phantom to assess more than 350 systems (70). The phantom is used for measurements of SNR, contrast-to-noise ratio, and geometric distortion. The measurements allow correction of patient images with respect to tissue contrast and geometric distortion, as needed for segmentation and for volume measurements.

The ADNI multisite study found several scanner errors, which may have been missed without central monitoring. Errors including misidentification of gradient hardware, disabling of autoshim, and miscalibrated laser alignment light, if undetected would have contributed to imprecision in quantitative metrics at more than a quarter of all ADNI sites (54). In conclusion, the ADNI group's suggestions for best practices include minimizing large signal voids, using a keyed geometry, enabling tight integration with the quality control process, using the phantom as part of site qualification for inclusion in a clinical trial, and completing one phantom scan per human scan in a clinical trial.

Dynamic Contrast-Enhanced MRI Perfusion Phantom

As part of the efforts of the RSNA QIBA, a phantom was designed and produced that could be used to (i) assess bias and variance of signal intensity measurements from T_1 mapping and dynamic contrast-enhanced (DCE) MRI acquisitions across scanners, centers, and time; (ii) assess the effects of parallel imaging and B_1 corrections; (iii)
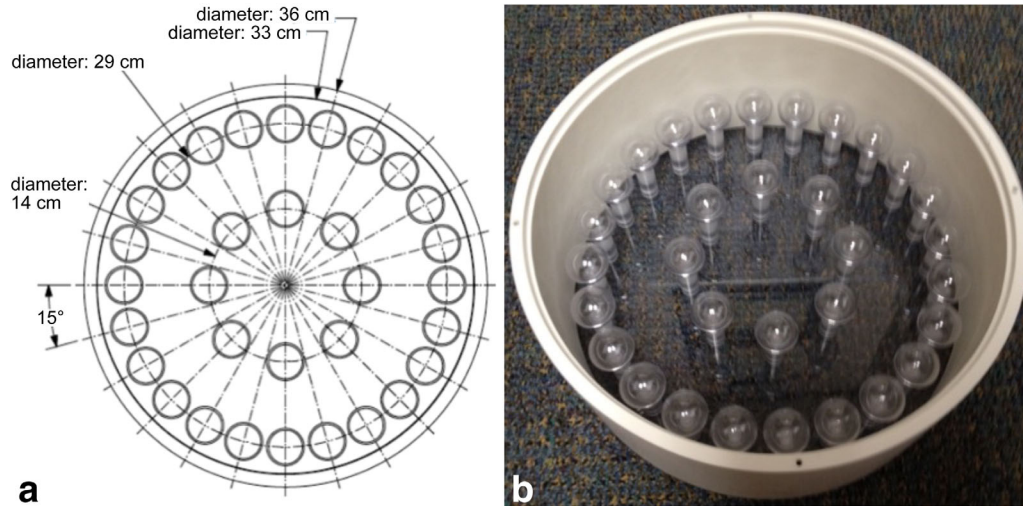
FIG. 1. A diagram (**a**) and an image (**b**) of the QIBA DCE-MRI phantom. The inner set of eight spheres, referred to as the VIF spheres, is distributed around a 14.0-cm-diameter circle. The remaining 24 spheres are uniformly distributed around a 29.0-cm-diameter circle, and consist of three sets of eight "tissue spheres" spaced at 45° increments, with each set having an R_1 range of 0.67 to 7.54 s^{-1}. The lowest R_1 sphere in each set of eight was positioned at 0°, 105°, and 210°, respectively, to produce three virtual rotations of the three sets of tissue spheres, to facilitate the investigation of spatial signal dependencies arising within phased-array coils without the need to physically rotate the phantom between acquisitions.

form one component of a qualification process for imaging centers enrolling in DCE-MRI clinical trials and for ongoing quality control in such studies; and (iv) allow comparisons of T_1 measurements and DCE-MRI data acquired on different scanners and across time, and harmonization of such measures (71). The phantom was critical to the development and implementation of the QIBA DCE-MRI profile (3), which addresses all aspects of a DCE-MRI study, including data acquisition and processing. At the time of development, there were no other phantoms available to assess the contrast response of acquisition sequences across the range of tissue and vascular R_1 values encountered during a DCE-MRI acquisition, particularly at 3 T, while also assessing the effects of nonuniform sensitivity of phased array coils in abdominal imaging.

The RSNA QIBA DCE-MRI phantom (Fig. 1) is a multi-compartment phantom consisting of a 36-cm-diameter, 15-cm-height cylindrical polycarbonate shell containing a set of 32 3.0-cm spheres in a uniform fill solution (71). The spheres are doped with NiCl$_2$ to achieve T_1 values (Table 1) spanning the ranges expected in a vascular input function (VIF) compartment (VIF spheres) and in tissue (tissue spheres) during a typical DCE-MRI study. To appropriately load the RF coil, the phantom is filled with a 30-mM NaCl (Sigma-Aldrich, St. Louis, Missouri, USA) solution in water. The phantom was used for site qualification and requalification in the American College of Radiology Imaging Network 6701 prostate DCE and diffusion-weighted imaging clinical trial (72).

Diffusion Phantoms

Isotropic Diffusion Imaging

Isotropic diffusion imaging is used as a biomarker to identify tumors and track response to treatments (14,15). To have sufficient confidence in diffusion MRI measurements, several research groups performed quality-

assurance testing with phantoms. Laubach et al. used a sucrose solution to alter the apparent diffusion coefficient (ADC) of water molecules (73), whereas Tofts et al. used alkanes to achieve a range of ADC values (74). Delakis et al., using two aqueous test solutions of copper sulfate (CuSO$_4$) and sucrose, developed a quality control protocol to assess the accuracy, precision, and reproducibility of ADC measurement on a clinical MRI system (65). Wang et al. developed an acetone and deuterium oxide (D$_2$O) phantom that demonstrated ADC values in the physiologic range ($0.57 - 3.16 \times 10^{-3}$ mm^2/s) at 0°C without any signal from the D$_2$O solute (75).

The Cancer Research UK Clinical MR Group and Royal Marsden Hospital uses the sucrose diffusivity phantom (containing a solution with known water diffusivity) and nondiffusivity phantom, which contains a highly viscous, very large mono-disperse polymer with very low ADC, for diffusion MRI experiments (76). The sucrose

Table 1
Theoretical Sphere R_1 ($= 1/T_1$) Values at 3 T and Corresponding NiCl$_2$ Concentration

	VIF spheres		Tissue spheres	
Sphere	R_1 (s^{-1})	[NiCl$_2$] (mg/L)	R_1 (s^{-1})	[NiCl$_2$] (mg/L)
1	0.75	87.1	0.67	69.68
2	2.63	479.03	0.94	127.40
3	6.56	1302.09	1.33	209.03
4	11.56	2347.24	1.89	324.48
5	17.56	3601.42	2.67	487.74
6	24.56	5064.64	3.77	718.63
7	32.56	6736.88	5.33	1045.15
8	41.56	8618.16	7.54	1506.93

Note: The R_1 values were chosen to mimic the range of values typically encountered in a DCE-MRI study for both the VIF and tissue compartments. To achieve these relaxation rates, the corresponding concentrations of NiCl$_2$ are provided, assuming a water-relaxation rate of 0.33 s^{-1} and NiCl$_2$ relaxivity of 0.62 (mM·s)$^{-1}$ at 3 T.
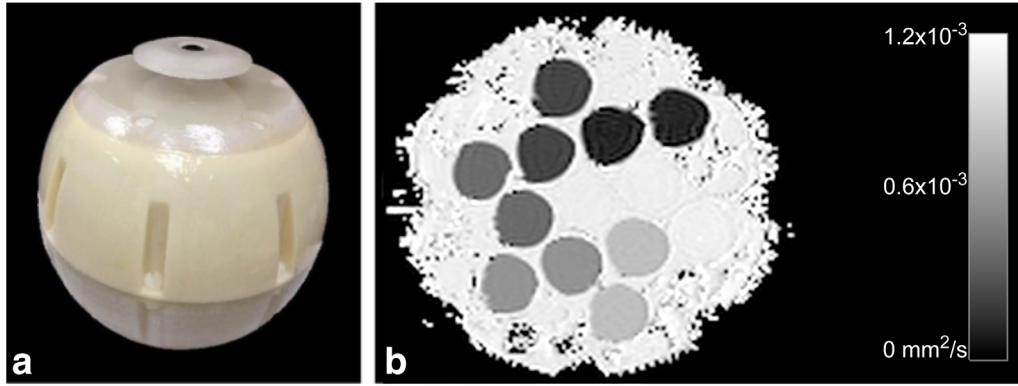
FIG. 2. Ice-water diffusion phantom (**a**) with an array of PVP solutions to obtain ADC values from 0.1×10^{-3} to 1.1×10^{-3} mm^2/s at 0°C (**b**). The phantom has a spherical geometry with an outer diameter of 194 mm, designed to fit in existing multichannel head coils. Thirteen high-density polyethylene vials (31.5 mm outer diameter and approximately 68 mm tall) contain PVP solutions ranging from 0 to 50% PVP by mass fraction in water, arranged in two concentric circles, with a central vial filled with deionized water. These vials are in one plane of the phantom; to characterize all three imaging planes, the phantom must be physically rotated.

phantom allows detection of drift in the measured ADC on the same MRI system over time, and how well ADC values calculated from a single magnetic field gradient agree across three orthogonal gradient directions. The nondiffusivity phantom is used for testing eddy current–induced image distortions.

Diffusion is a thermally driven process and is highly sensitive to temperature variation; the ADC of pure water changes approximately 3%/K near room temperature. To obtain an accurate reference standard, an ADC phantom must have accurately controlled or measured temperature. Padhani et al. recommended the use of ice water in a phantom to eliminate thermal variability (18), leading to the efforts of Chenevert et al. (77) and Malyarenko et al. (78) to develop an ice water phantom (0°C) to give a stable water proton ADC of 1.1×10^{-3} mm^2/s. With this phantom, large errors in the ADC were observed when measured off magnet isocenter, as a result of nonlinear gradients (78). Boss et al. (44) demonstrated an improved isotropic diffusion phantom (Fig. 2), developed by the coordinated efforts of the National Cancer Institute, RSNA QIBA, and the National Institute of Standards and Technology, which incorporates a variable ADC array using aqueous solutions of polyvinylpyrrolidone (PVP) (79). The phantom has a modular polycarbonate shell that can be disassembled to fill the phantom with an ice-water bath to accurately control temperature. Users do find the temperature control of this phantom challenging and would prefer to measure the phantom temperature rather than set up an ice bath. Additionally, the T_1 and T_2 values of the PVP material do not span the full physiologic range at 0°C: T_1 ranges from 157 to 1450 ms, and T_2 spans 126 to 1040 ms for the PVP solutions at 1.5 T. At 0°C, the ADC values of PVP do not span the physiologic range; however, at higher temperatures (e.g., 37°C), the ADC values of PVP cover the full range of isotropic diffusion in the human body.

Diffusion Tensor MR Imaging

Anisotropic diffusion imaging characterizes the path of water molecule diffusion and is used to characterize brain injury (16,17). Several different approaches were used to mimic the anisotropic diffusion of water seen in the brain (80–82). An acrylic water-filled phantom with a grid structure was used to evaluate geometric distortions in functional MRI and diffusion tensor imaging (83). The phantom provided accurate geometric information over the scanning volume for echo-planar imaging–based functional MRI and diffusion tensor imaging of the human brain. The results suggest this phantom can reveal geometric distortions not easily detected by standard MRI phantoms. In another study, separate water phantom calibration experiments were conducted to accurately determine and correct eddy current–induced image distortions for in vivo diffusion anisotropy (84). Further investigations examined the practicalities of using separate phantom calibration data to correct high b-value diffusion tensor imaging measurements by investigating the stability of these distortion parameters, and hence the eddy currents, with time (85). Rayon fibers were used to mimic axonal bundles, crossing at 90° to validate q-ball imaging (86). Fieremans et al. used high molecular-weight polyethylene fibers, packed together tightly in heat-shrink tubing, to compare the experimental diffusion-weighted MRI and NMR results with Monte Carlo simulations of the apparent diffusion coefficient, fractional anisotropy, and kurtosis (82). Anisotropic diffusion and elastic properties of the brain were mimicked with Spandex fibers in a polyvinylalcohol hydrogel to simultaneously evaluate diffusion tensor imaging and MR elastography in a single reference object (87). These fiber phantoms generate a porous structure that mimics the hindered space seen in axonal bundles; however, they do not also mimic the restricted water space. Glass capillaries have been used to simulate a restricted space for water diffusion, allowing comparison of different reconstruction techniques (88,89); these phantoms lack the hindered water space and cannot easily create fiber crossings. Hollow polypropylene fibers have been incorporated into a phantom, allowing for fiber crossing and changes in packing density to create both a hindered and restricted water space (90).

Flow Phantom

Phase-contrast MRI is used to assess hemodynamics in cardiovascular blood flow for a range of clinical applications, such as assessment of pulmonary to systemic flow shunting (91,92), measurement of peak velocity to assess valvular disease (93,94), and the assessment of pressure gradient through stenosis in arteries (95,96). In all cases, guideline-driven quantitative thresholds exist to inform the need for therapy or intervention. The use of guideline-driven thresholds underlines a need for accuracy and repeatability to be assessed at a system level for phase-contrast MRI blood-flow measurements.

Numerous efforts have constructed flow phantoms to mimic vascular territories and disease conditions (97–99). These single-site "in-house" studies have reported the accuracy and precision of phase-contrast MRI to measure case- and site-specific regional velocity, bulk flow rates (velocity integrated over a region of interest), and net flow (temporally and spatially integrated velocity). Computational fluid dynamics, particle image velocimetry, and bulk flow transducers are typically used to validate the phase-contrast MRI flow-field measurements. However, no literature exists on a proven, robust, "dynamic fluid" phantom that sufficiently addresses all challenges associated with creating a reliable and reproducible fluid flow field for multisite use with proven test-retest stability.

To date, the most extensive studies have used static tissue phantoms. Static tissue phantoms can be used to study phase offset errors, an error postulated to have a large effect on the accuracy of spatially and temporally integrated phase-contrast flow measurements. The most comprehensive multisite effort to investigate phase-offset errors was undertaken by members of the European Society of Cardiology Working Group. Concerned that background phase offsets were a cause of flow-measurement inaccuracies in commercial MRI systems, the group designed a 10-site, 3-manufacturer, 12-system study (all 1.5 T) (100). The phantoms consisted of 10- to 15-L tanks of aqueous gelatin solution, which were doped with 5 mmol/L of gadolinium-diethylenetriamine pentaacetic acid to facilitate the measurement of small background phase offsets. Gelatin eliminated phase differences as a result of convection and motion-induced fluid currents. By assuming a worst-case error from spatial and temporal integration, an offset of 0.6 cm/s was chosen as a quality threshold (given the potential to cause a 10% error in a pulmonary to systemic shunt measurement). Noting that this was a worst-case scenario, 35 of the 36 uncorrected exams (three experiments per scanner) were found to exceed the quality threshold. A follow-up study in nine 1.5T scanners (involving three different manufacturers) used the same phantoms to understand the correlation of exam protocol parameters on phase-offset errors (101). No generic protocol was found to generate acceptable offset values across all scanners (using a 0.6-cm/s quality metric). Both studies recommended post hoc corrections to improve accuracy of the measurements, although no universally accepted algorithm was recommended. Multisite temporal stability of background offsets was also examined with this phantom design (102).

Although significant efforts established the importance of phase offsets with static phantoms, a need exists for a robust, dynamic phantom to replicate spatially and temporally varying velocities across a large range of magnitudes. In single-center, in-house studies, dynamic fluid phantoms were used to replicate pulsatile flow (103), stenosis geometries (104), and other patient-specific geometries (105). The most comprehensive multisite effort to date was initiated in 1999 by the Flow and Motion Study Group of the International Society for Magnetic Resonance in Medicine in the Assessment of Methodology of Phase Mapping for Flow Measurement trial. A preliminary two-site report was published in 2005, which detailed the design criteria and plans for a dynamic-flow phantom capable of mimicking various vessels and FOV configurations (106). However, there were no spatially and/or temporally resolved flow measurements with this flow phantom. The design of a robust dynamic fluid-filled phantom is challenging to implement across multiple sites for a multitude of reasons. Without significant effort (and cost), pump systems and control hardware must reside outside of the scan room (as a result of electromagnetic noise and ferromagnetic components). This means fluid tubing must be routed through waveguides to the control room, and plumbing connections must be repeatedly disassembled and reassembled, thereby risking joint failure, introduction of air bubbles, or catastrophic leakage in the scan room. Furthermore, the inlet conditions, position, and configuration of the assembly (e.g., head height and tubing length, compliance, resistance) will vary according to waveguide location and exam room layout. For these reasons, a dynamic fluid-filled phantom may suffer in terms of reliability and repeatability. The most promising alternative is a rotating gelatin disk phantom, whereby a large range of known velocities can be measured using a priori knowledge of the angular rotation and measurement position in relation to the axis of rotation. Challenges associated with fluid motion, leakage, and presence of air bubbles are mitigated. A few studies reported the use of such phantoms to test velocity and phase-contrast measurement methods (107–109). This configuration compromises the ability to evaluate the effects associated with fluid flow in vessels and boundary interactions, including partial volume artifacts or the presence of turbulence, with robustness and cost effectiveness.

Breast Phantom

Breast MRI with quantitative methods is increasingly used for breast cancer diagnosis, staging, and monitoring. For these quantitative applications, it is important to understand and mitigate the sources of variability, such as fat suppression, variations in the left and right sides of the coil, and B_0 inhomogeneity across the large image volume. To address these issues, breast phantoms have been created for quality control (110), well-mixed fat and fibroglandular tissue (111), and DCE-MRI in the breast (112).

The University of California San Francisco and National Institute of Standards and Technology breast phantom design with flexible outer shell easily fits into
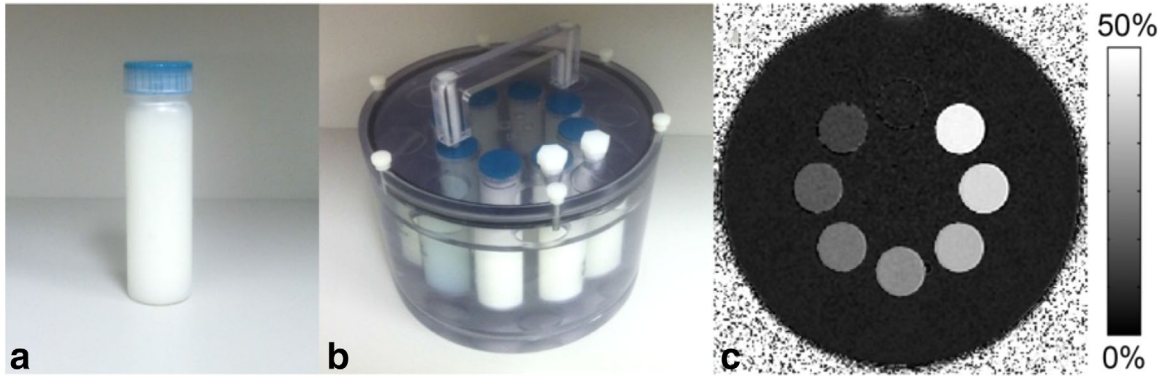
FIG. 3. **a**: Vial consisting of a PDFF mimic. **b**: Multiple vials can be scanned simultaneously by placing them in a phantom holder that is filled with deionized water (128). **c**: The PDFF for each of the vials can be estimated using MRI. In this example, the PDFF values are (beginning at 12 o'clock position and moving counterclockwise) 0, 5, 10, 15, 20, 30, 40, and 50%.

different coils and is useful for clinical breast imaging techniques (43,113). The phantom was tested using the sequences of a particular breast imaging clinical trial. The fibroglandular mimic exhibited target T_1 values of 1300 to 1400 ms and 1500 to 1850 ms on 1.5 T and 3 T clinical systems, respectively. Fat was suppressed using standard techniques, and PVP solutions mimicked the range of ADC values from malignant tumors to normal breast tissue (43). The phantom does not include any dynamic components for DCE-MRI as other designs have (112). Additionally, the T_1 and T_2 values of the PVP are not physiologic for breast fibroglandular or tumor tissue. One challenge of the two-phantom design is it requires twice the scan time to assess both sides of the coil. It is important to assess both sides of the coil, as one study found geometric distortion between the right and left coil sides of multiple platforms when using echo-planar imaging diffusion techniques (113). A chest cavity model may need to be added to properly replicate the B_1 homogeneity challenges in breast imaging.

Proton-Density Fat Fraction Phantom

Quantification of fat in the body has many important applications in the liver, heart, and pancreas, as well as in skeletal muscle. Proton-density fat fraction (PDFF) is currently regarded as the most practical and meaningful MR-based biomarker of tissue fat concentration (114). The PDFF represents the ratio of MR-visible fat protons to the total number of MR-visible water and fat protons.

The PDFF phantoms have been used to analyze the accuracy of MR-based PDFF techniques for liver fat quantification (Fig. 3) (115). This phantom consists of separate vials with approximately 40-mL volume for each, and a different PDFF, typically in the range of 0 to 50% to reflect clinically relevant liver fat fractions. The vials store a gel mixture of peanut oil and deionized water, together with minute concentrations of additional substances to ensure mixture of the oil and water and to prevent spoiling. The PDFF phantoms have also been used to assess the accuracy and reproducibility of PDFF measurements across different sites, vendors, and field strengths (116,117).

A further development to a PDFF phantom includes the need for it to reflect physiologically relevant relaxation (i.e., R_2^*) rates. It is important that the R_2^* values for both the fat component and water component remain similar to one another, which has been measured in vivo (118,119).

IMPLEMENTATION OF QUANTITATIVE MRI

For quantitative MRI to be widely adopted, a framework is required to ensure that the quantitative measures are comparable over time, between subjects, between scanner sites, and between manufacturers. The proposed system phantom and application-specific phantoms are one part of the framework, along with standardized protocols and data collection. With an established framework, quantitative MRI can be used to assess outcomes in clinical trials and for clinical diagnostics. In particular, clinical trials and clinical use of quantitative MRI can benefit from the use of a phantom in many ways, including protocol development, selection of RF coils, training of technologists, quality-control standards, standardized analysis, and correction of collected images, if necessary.

This paper discusses the design process, requirements, and recommendations for a phantom to assess the performance and stability of an MRI system. In addition, we reviewed application-specific phantoms that are designed to evaluate the performance of a particular technique. In this final section, we discuss the implementation of quantitative MRI.

Magnetic Resonance Imaging System Constancy

System constancy data should be tracked regularly at all MRI systems and especially those doing quantitative measurements. The described system phantom enables assessment of scanner performance over time (stability or constancy) for many parameters. For example, the system phantom can be used to track B_1 and B_0 nonuniformity, geometric nonlinearity, gradient amplitude, image uniformity, SNR, transmitter and receiver gain, receiver bandwidth, image ghosting, and eddy currents, using the methods described in International Electrotechnical Commission 62464-1 (49). For any stability parameter, prospective criteria should be developed to generate a

service call for "out of specification" results. Friedman and Glover present the advantages of a quality-control standard for evaluation and acceptance of a new scanner, benchmarks for comparisons with other MRI centers, monitoring system constancy through hardware and software upgrades, and planning of multicenter studies (45).

Standard Protocols

A clinical trial requires protocol standardization across participating sites to ensure that conclusions can be drawn from the data. The challenge is to minimize differences in effective acquisition parameters across sites for multicenter studies and across time for multicenter and single-center studies. This is confounded by varying hardware and software configurations within and across manufacturers' platforms. The implication for protocol standardization is that the actual scan protocol may be slightly different across manufacturers to get the same contrast-to-noise ratio required by the analytical technique.

One approach uses the strategy of ensuring that specific pulse-sequence parameters are identical, to the degree possible. The advantage of this approach is that it is easy to implement. The disadvantage is that subtle variations in implementation, which are not always known outside of the manufacturer, can cause significant changes in image appearance (e.g., contrast, artifact propagation). Further complications exist even within the manufacturers' platforms over time. Nevertheless, this should be a starting point for protocol definition and standardization.

Clinical use of quantitative MRI requires harmonization, and these efforts are led in part by RSNA QIBA. Clinical trials can build on the work of QIBA to create standard protocols for each technique. The QIBA profiles provide claim statements for quantitative imaging biomarkers within a specified clinical context. These claim statements indicate the reproducibility of the quantitative measurement as determined by existing literature and biomarker-specific groundwork projects. The QIBA profiles provide a list of requisite activities and associated actors to meet the claim statements, as well as assessment procedures to ensure proper quality assurance. These activities can include subject selection and preparation, image acquisition and reconstruction, analysis, and interpretation. Profiles undergo a strict vetting procedure within QIBA, and then pass through the stages of public comment, consensus, technical confirmation, claim confirmation, and clinical confirmation, as the profile is adopted and thoroughly tested in the clinical environment. Physical phantoms and virtual phantoms (digital reference objects) are essential in QIBA profiles.

An appropriate phantom can be used to refine the "identical" protocols, such that the images obtained across platforms are equivalent. This process includes removing any postprocessing steps, which may not be readily apparent to the user. Once a standard protocol is established, all protocols should be provided in an electronic fashion to the sites, if possible, to minimize entry errors at the console. Each site should provide images of an appropriate phantom from the standard protocol as a qualification step to be included in a clinical trial.

Standardized Training of Technologists

The "gatekeeper" for image quality is the local technologist (radiographer, technician). It is recommended that uniform training be provided for all sites and that such training consider variations in hardware and software platforms. Such training can be accomplished at a group meeting, individual training visits to the site, and/or by video instructions that provide specific details of the study (e.g., positioning criteria). The phantom should be used to provide hands-on training, and the phantom images will allow the coordinating center to determine whether a site is ready to be included in a clinical trial.

Specific Image Analysis Procedures

A well-designed phantom is useful for developing analysis methods, evaluating how system-to-system error affects results, and correction of images for uniform analysis/error reduction.

Large-scale clinical studies of MR images often require the application of quantitative image analysis methods on data sets that were acquired by multiple sites. However, such methods are often developed on data sets from a single MR system vendor and/or scanner model. To evaluate the variability across different MRI systems, the methods should be tested and validated on data sets from multiple scanners with different properties using the same standard protocols.

All metadata (e.g., header information) received from clinical trial sites should be checked for protocol adherence. At the beginning of the study, acceptable deviations should be determined (ideally by evaluation of such deviations on the analytical procedure), and the ranges documented in the project manual(s). If a hardware or software upgrade generates parameters outside of the prospective criteria, the effect on analysis should be determined and decision made about whether to drop the site or accept the protocol deviation. If the deviation is acceptable, the initial range perhaps should have been broader. Ideally, all such parameter restrictions should be prospectively determined and based on the actual outcome measure. When feasible, "electronic protocols" should be centrally distributed to avoid errors associated with users translating information from written protocols into scanners.

For example, Chenevert et al. used an ice-water phantom to compare measurement of ADC across systems, including multiple manufacturers and platforms (120). The images generated by one of the scanners appeared to have image intensity scaling that was not accounted for by most quantitative image analysis tools. Incorrect image scaling leads to measurement bias, and the scaling of images must be accounted for in the image analysis routine.

Future Implications

The use of MR-based measurements as biomarkers is a driver for developing a framework for quantitative MRI

adoption, and clinical applications will also benefit considerably from these developments. Of interest, a decade ago, quantitative imaging was seen as the future by the leaders of the radiology community, as reflected by the statement "the RSNA remains committed to helping to transform Radiology from a qualitative to quantitative science" (121). With the advent of methods such as magnetic resonance fingerprinting (122) and compressed sensing (123,124), quantitative MRI can now be performed in a clinically appropriate timeline; however, quality control with a quantitative MRI phantom is necessary to ensure the accuracy and precision of results. Such sequences and corresponding reconstruction methods develop behavior that significantly differs from that of "classic" MRI methods, and a system phantom can provide a way to rigorously characterize the behavior of these methods when standard image quality metrics like SNR are no longer valid. Looking forward, a comprehensive system phantom along with MR imaging acquisition (i.e., pulse sequence), reconstruction, and analysis software and quality assurance recommendations could be an accreditation program for quantitative MRI, similar to the ACR MR Accreditation Program currently in place for qualitative MRI.

CONCLUSIONS

Quantitative MRI enables noninvasive measurements of biomarkers pertinent to clinical trials and diagnostic tests. This paper, prepared by the International Society for Magnetic Resonance in Medicine, Ad Hoc committee on Standards for Quantitative Magnetic Resonance, describes the need for phantoms, previous standardization attempts, an overview of available phantoms, and the desired features of a system phantom for quantitative MRI. The system phantom is designed to be used for quality control (assessing system constancy) and with the intentions of comparing results across manufacturer systems, hardware and software, across time, and across physiologic ranges at 1.5 T and 3 T. The system phantom prototype was constructed (125) and later commercialized; both the prototype and commercial phantoms were used by the International Society for Magnetic Resonance in Medicine Ad Hoc committee on Standards for Quantitative Magnetic Resonance for studies of T_1 variation (126,127). The full manuscript describing the system phantom is in preparation. The improved accuracy and reproducibility of quantitative results through use of a system phantom should increase statistical power, patient safety, and efficacy and efficiency of clinical trials, and is a critical step toward the full potential of MR biomarkers.

Any quantitative MRI phantom data need to be easy to analyze, to enable adoption by many different scanner locations and users. This is true both of a general system phantom and application-specific phantoms. It is therefore important that any phantom designed to characterize MRI performance meet certain requirements to be amenable for quantitative analysis. To allow regular quality control, a technologist should be able to position and image the phantom and import the images to the software package, and the analysis software should generate

a report with the system status. Another benefit of the system phantom is that it can be used for comparative studies of processing strategies, such as those available from vendors or research groups for quantitative MRI signal models.

In this paper, we reviewed the application-specific phantoms developed for certain quantitative MRI techniques. Additional application-specific phantoms are still needed, such as for musculoskeletal techniques and the combination of MRI and positron emission tomography. The components and materials research for the system phantom can be used to develop application-specific phantoms, especially given the proposed modular structure.

MRI system stability is required for implementation of quantitative MRI, especially to enable biomarkers for diagnostic use. A standardized MR system phantom will support the efforts of the quantitative MRI community, including RSNA QIBA and the National Cancer Institute's Quantitative Imaging Network. Research developments will be enabled by the system phantom, such as acquisition and modeling for relaxometry. A standard system phantom, with SI-traceable components that will be monitored for long-term stability by a national metrology institute, will further facilitate the use of MRI measurements as a biomarker. Most importantly, to support clinical use of quantitative MRI, such a phantom must be adopted by the user community and equipment manufacturers for regular use.

REFERENCES

1. Atkinson AJ, Colburn WA, DeGruttola VG, et al. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69:89–95.
2. Kessler LG, Barnhart HX, Buckler AJ, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. Stat Methods Med Res 2015;24:9–26.
3. Raunig DL, McShane LM, Pennello G, et al. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. Stat Methods Med Res 2015;24:27–67.
4. Partridge SC, Gibbs JE, Lu Y, Esserman LJ, Tripathy D, Wolverton DS, Rugo HS, Hwang ES, Ewing CA, Hylton NM. MRI measurements of breast tumor volume predict response to neoadjuvant chemotherapy and recurrence-free survival. AJR Am J Roentgenol 2005;184:1774-1781.
5. Vaidyanathan M, Clarke LP, Hall LO, Heidtman C, Velthuizen R, Gosche K, Phuphanich S, Wagner H, Greenberg H, Silbiger ML. Monitoring brain tumor response to therapy using MRI segmentation. Magn Reson Imaging 1997;15:323–334.
6. Mayr NA, Taoka T, Yuh WT, et al. Method and timing of tumor volume measurement for outcome prediction in cervical cancer using magnetic resonance imaging. Int J Radiat Oncol Biol Phys 2002;52:14–22.
7. Lemieux L, Hagemann G, Krakow K, Woermann FG. Fast, accurate, and reproducible automatic segmentation of the brain in T1-weighted volume MRI data. Magn Reson Med 1999;42:127–135.
8. Jack CR Jr, Shiung MM, Gunter JL, et al. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology 2004;62:591–600.
9. Resnick SM, Goldszal AF, Davatzikos C, Golski S, Kraut MA, Metter EJ, Bryan RN, Zonderman AB. One-year age changes in MRI brain volumes in older adults. Cereb Cortex 2000;10:464–472.
10. Schuff N, Amend DL, Knowlton R, Norman D, Fein G, Weiner MW. Age-related metabolite changes and volume loss in the hippocampus by magnetic resonance spectroscopy and imaging. Neurobiol Aging 1999;20:279–285.
11. Greicius MD, Krasnow B, Reiss AL, Menon V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci U S A 2003;100:253–258.

12. Greicius MD, Srivastava G, Reiss AL, Menon V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci U S A 2004;101: 4637–4642.

13. Wu X, Li R, Fleisher AS, Reiman EM, Guan X, Zhang Y, Chen K, Yao L. Altered default mode network connectivity in Alzheimer's disease—a resting functional MRI and Bayesian network study. Hum Brain Mapp 2011;32:1868–1881.

14. Moffat BA, Chenevert TL, Lawrence TS, et al. Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. Proc Natl Acad Sci U S A 2005;102:5524–5529.

15. Liimatainen T, Hakumaki JM, Kauppinen RA, Ala-Korpela M. Monitoring of gliomas in vivo by diffusion MRI and (1)H MRS during gene therapy-induced apoptosis: interrelationships between water diffusion and mobile lipids. NMR Biomed 2009;22:272–279.

16. Huisman TA, Schwamm LH, Schaefer PW, Koroshetz WJ, Shetty-Alva N, Ozsunar Y, Wu O, Sorensen AG. Diffusion tensor imaging as potential biomarker of white matter injury in diffuse axonal injury. AJNR Am J Neuroradiol 2004;25:370–376.

17. Mayer AR, Ling J, Mannell MV, Gasparovic C, Phillips JP, Doezema D, Reichard R, Yeo RA. A prospective diffusion tensor imaging study in mild traumatic brain injury. Neurology 2010;74:643–650.

18. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia 2009;11:102–125.

19. Parsons MW, Li T, Barber PA, Yang Q, Darby DG, Desmond PM, Gerraty RP, Tress BM, Davis SM. Combined (1)H MR spectroscopy and diffusion-weighted MRI improves the prediction of stroke outcome. Neurology 2000;55:498–505.

20. Golman K, Zandt RI, Lerche M, Pehrson R, Ardenkjaer-Larsen JH. Metabolic imaging by hyperpolarized 13C magnetic resonance imaging for in vivo tumor diagnosis. Cancer Res 2006;66:10855–10860.

21. Reference Values for Arterial Stiffness Collaboration. Determinants of pulse wave velocity in healthy people and in the presence of cardiovascular risk factors: 'establishing normal and reference values'. Eur Heart J 2010;31:2338–2350.

22. Ances BM, Sisti D, Vaida F, et al. Resting cerebral blood flow: a potential biomarker of the effects of HIV in the brain. Neurology 2009;73:702–708.

23. Markl M, Wallis W, Strecker C, Gladstone BP, Vach W, Harloff A. Analysis of pulse wave velocity in the thoracic aorta by flow-sensitive four-dimensional MRI: reproducibility and correlation with characteristics in patients with aortic atherosclerosis. J Magn Reson Imaging 2012;35:1162–1168.

24. Idilman IS, Aniktar H, Idilman R, Kabacam G, Savas B, Elhan A, Celik A, Bahar K, Karcaaltincaba M. Hepatic steatosis: quantification by proton density fat fraction with MR imaging versus liver biopsy. Radiology 2013;267:767–775.

25. Qayyum A, Goh JS, Kakar S, Yeh BM, Merriman RB, Coakley FV. Accuracy of liver fat quantification at MR imaging: comparison of out-of-phase gradient-echo and fat-saturated fast spin-echo techniques—initial experience. Radiology 2005;237:507–511.

26. Pacifico L, Celestre M, Anania C, Paolantonio P, Chiesa C, Laghi A. MRI and ultrasound for hepatic fat quantification:relationships to clinical and metabolic characteristics of pediatric nonalcoholic fatty liver disease. Acta Paediatr 2007;96:542–547.

27. Reeder SB, Robson PM, Yu H, Shimakawa A, Hines CD, McKenzie CA, Brittain JH. Quantification of hepatic steatosis with MRI: the effects of accurate fat spectral modeling. J Magn Reson Imaging 2009; 29:1332–1339.

28. Mentore K, Froh DK, de Lange EE, Brookeman JR, Paget-Brown AO, Altes TA. Hyperpolarized HHe 3 MRI of the lung in cystic fibrosis: assessment at baseline and after bronchodilator and airway clearance treatment. Acad Radiol 2005;12:1423–1429.

29. Ohno Y, Koyama H, Yoshikawa T, Matsumoto K, Takahashi M, Van Cauteren M, Sugimura K. T2* measurements of 3-T MRI with ultrashort TEs: capabilities of pulmonary function assessment and clinical stage classification in smokers. AJR Am J Roentgenol 2011;197: W279–W285.

30. Babourina-Brooks B, Simpson R, Arvanitis TN, Machin G, Peet AC, Davies NP. MRS thermometry calibration at 3 T: effects of protein, ionic concentration and magnetic field strength. NMR Biomed 2015; 28:792–800.

31. Babourina-Brooks B, Wilson M, Arvanitis TN, Peet AC, Davies NP. MRS water resonance frequency in childhood brain tumours: a novel potential biomarker of temperature and tumour environment. NMR Biomed 2014;27:1222–1229.

32. Rieke V, Instrella R, Rosenberg J, Grissom W, Werner B, Martin E, Pauly KB. Comparison of temperature processing methods for monitoring focused ultrasound ablation in the brain. J Magn Reson Imaging 2013;38:1462–1471.

33. Loomba R, Wolfson T, Ang B, et al. Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: a prospective study. Hepatology 2014;60:1920–1928.

34. Yin M, Talwalkar JA, Glaser KJ, Manduca A, Grimm RC, Rossman PJ, Fidler JL, Ehman RL. Assessment of hepatic fibrosis with magnetic resonance elastography. Clin Gastroenterol Hepatol 2007;5:1207–1213.e2.

35. Clarke L, Sriram RD. Imaging as a biomarker: standards for change measurements in therapy workshop summary. Acad Radiol 2008;15: 501–530.

36. Woodcock J, Woosley R. The FDA critical path initiative and its influence on new drug development. Ann Rev Med 2008;59:1–12.

37. Glaser KJ, Manduca A, Ehman RL. Review of MR elastography applications and recent developments. J Magn Reson Imaging 2012;36: 757–774.

38. Padhani AR. Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. J Magn Reson Imaging 2002;16: 407–422.

39. Huisman TA, Boltshauser E, Martin E, Nadal D. Diffusion tensor imaging in progressive multifocal leukoencephalopathy: early predictor for demyelination? AJNR Am J Neuroradiol 2005;26:2153–2156.

40. Radiological Society of North America. Quantitative Imaging Biomarkers Alliance. http://www.rsna.org/qiba/. Accessed August 24, 2016.

41. Clarke LP, Croft BS, Nordstrom R, Zhang H, Kelloff G, Tatum J. Quantitative imaging for evaluation of response to cancer therapy. Transl Oncol 2009;2:195–197.

42. Nordstrom RJ. The quantitative imaging network in precision medicine. Tomography 2016;2:239–241.

43. Keenan KE, Wilmes LJ, Aliu SO, Newitt DC, Jones EF, Boss MA, Stupic KF, Russek SE, Hylton NM. Design of a breast phantom for quantitative MRI. J Magn Reson Imaging 2016;44:610–619.

44. Boss M, Chenevert T, Waterton J, et al. Thermally-stabilized isotropic diffusion phantom for multisite assessment of apparent diffusion coefficient reproducibilty. Med Phys 2014;41:464.

45. Friedman L, Glover GH. Report on a multicenter fMRI quality assurance protocol. J Magn Reson Imaging 2006;23:827–839.

46. Cheng HL, Wright GA. Rapid high-resolution T(1) mapping by variable flip angles: accurate and precise measurements in the presence of radiofrequency field inhomogeneity. Magn Reson Med 2006;55: 566–574.

47. Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, Pike GB. On the accuracy of T1 mapping: searching for common ground. Magn Reson Med 2015;73:514–22.

48. National Electrical Manufacturers Association (NEMA). Determination of signal-to-noise ratio SNR in diagnostic MRI. In: NEMA standards publication MS 1-2008 R2014; 2014.

49. International Electrotechnical Commission (IEC). Magnetic resonance equipment for medical imaging. 1: Determination of essential image quality parameters. In: IEC standards 62464-1; 2007.

50. National Electrical Manufacturers Association (NEMA). Determination of image uniformity in diagnostic magnetic resonance images. In: NEMA standards publication MS 3-2008 R2014 image uniformity; 2014.

51. Newitt DC, Tan ET, Wilmes LJ, Chenevert TL, Kornak J, Marinelli L, Hylton N. Gradient nonlinearity correction to improve apparent diffusion coefficient accuracy and standardization in the American College of Radiology Imaging Network 6698 breast cancer trial. J Magn Reson Imaging 2015;42:908–919.

52. Tao S, Trzasko JD, Shu Y, Huston J 3rd, Bernstein MA. Integrated image reconstruction and gradient nonlinearity correction. Magn Reson Med 2015;74:1019–1031.

53. Gunter JL, Britson PJ, Felmlee JP, Ward CP, Schuff N, Weiner M, Levy J, Jack CR, editors. The ADNI phantom and analysis algorithm: a new and accurate tool to measure scanner performance. Berlin, Germany: International Society of Magnetic Resonance in Medicine; 2007.

54. Gunter JL, Bernstein MA, Borowski BJ, Ward CP, Britson PJ, Felmlee JP, Schuff N, Weiner M, Jack CR. Measurement of MRI scanner performance with the ADNI phantom. Med Phys 2009;36:2193–205.

55. Price RR, Allison J, Clarke GD, Dennis M, Hendrick E, Keener C, Masten J, Nessaiver M, Och J, Reeve D. 2015 MRI quality control manual. Reston, VA: American College of Radiology; 2015.

56. American College of Radiology. MRI accreditation program requirements. 2016. http://www.acraccreditation.org/~/media/ACRAccreditation/Documents/MRI/Requirements.pdf. Accessed June 23, 2016.

57. Xu D, Maier JK, King KF, Collick BD, Wu G, Peters RD, Hinks RS. Prospective and retrospective high order eddy current mitigation for diffusion weighted echo planar imaging. Magn Reson Med 2013;70: 1293–1305.

58. Jackson EF, Bronskill MJ, Drost DJ, Och J, Pooley RA, Sobol WT, Clarke GD. Acceptance testing and quality assurance procedures for magnetic resonance imaging facilities. Alexandria, VA: American Association of Physicists in Medicine; 2010.

59. Price RR, Axel L, Morgan T, Newman R, Perman W, Schneiders N, Selikson M, Wood ML, Thomas SR. Quality assurance methods and phantoms for magnetic resonance imaging. Alexandria, VA: American Association of Physicists in Medicine; 1990. p 17.

60. Firbank MJ, Harrison RM, Williams ED, Coulthard A. Quality assurance for MRI: practical experience. Brit J Radiol 2000;73:376–383.

61. Lerski RA, de Certaines JD. Performance assessment and quality control in MRI by Eurospin test objects and protocols. Magn Reson Imaging 1993;11:817–833.

62. Lerski RA, McRobbie DW, Straughan K, Walker PM, de Certaines JD, Bernard AM. Multi-center trial with protocols and prototype test objects for the assessment of MRI equipment. EEC Concerted Research Project. Magn Reson Imaging 1988;6:201–214.

63. De Wilde JP, Price D, Curran J, Williams J, Kitney RI. Standardization of performance evaluation in MRI: 13 years' experience of intersystem comparison. Concepts Magn Reson 2002;15:111–116.

64. De Wilde JP, Price D, Papadaki AM, Curran J, Kitney RI. Siemens Magnetom Symphony Quantum gradient system 1.5T MR imaging system: technical and user evaluation. MDA 01014. 2001.

65. Delakis I, Moore EM, Leach MO, De Wilde JP. Developing a quality control protocol for diffusion imaging on a clinical MRI system. Phys Med Biol 2004;49:1409–1422.

66. Price DL, De Wilde JP, Papadaki AM, Curran JS, Kitney RI. Investigation of acoustic noise on 15 MRI scanners from 0.2 T to 3 T. J Magn Reson Imaging 2001;13:288–293.

67. Williams J. MagNET's MRI test objects. http://magnet-mri.org/products/testobjects/index.htm. 2005. Accessed June 16, 2006.

68. Marquez D. Personal communication. 2016.

69. Price R, Allison J, Clarke G, Dennis M, Hendrick E, Keener C, Masten J, Nessaiver M, Och J, Reeve D. MRI quality control manual. Reston, VA: American College of Radiology; 2015.

70. Gunter JL. Personal communication. 2016.

71. Bosca R, Ashton E, Zahlmann G, Jackson E, editors. RSNA Quantitative Imaging Biomarker Alliance (QIBA) DCE-MRI phantom: goal, design, and initial results. In Proceedings of the 98th Scientific Assembly and Annual Meeting of RSNA, Chicago, Illinois, USA, 2012.

72. Repeatability assessment of quantitative DCE-MRI and DWI: a multicenter study of functional imaging standardization in the prostate. https://www.acrin.org/6701_protocol.aspx. Accessed March 20, 2017.

73. Laubach HJ, Jakob PM, Loevblad KO, Baird AE, Bovo MP, Edelman RR, Warach S. A phantom for diffusion-weighted imaging of acute stroke. J Magn Reson Imaging 1998;8:1349–1354.

74. Tofts PS, Lloyd D, Clark CA, Barker GJ, Parker GJ, McConville P, Baldock C, Pope JM. Test liquids for quantitative MRI measurements of self-diffusion coefficient in vivo. Magn Reson Med 2000;43:368–374.

75. Wang X, Reeder SB, Hernando D. An acetone based phantom for quantitative diffusion magnetic resonance imaging. J Magn Reson Imaging 2017;46:1683–1692.

76. Collins D, Koh DM. IMI-QuiConCePT conference call. 2013.

77. Chenevert TL, Galban CJ, Ivancevic MK, Rohrer SE, Londy FJ, Kwee TC, Meyer CR, Johnson TD, Rehemtulla A, Ross BD. Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies. J Magn Reson Imaging 2011;34:983–987.

78. Malyarenko D, Galbán CJ, Londy FJ, Meyer CR, Johnson TD, Rehemtulla A, Ross BD, Chenevert TL. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. J Magn Reson Imaging 2013;37:1238–1246.

79. Pierpaoli C, Sarlls J, Nevo U, Basser PJ, Horkay F. Polyvinylpyrrolidone (PVP) water solutions as isotropic phantoms for diffusion MRI studies. In Proceedings of the 17th Annual Meeting of ISMRM, Honolulu, Hawaii, USA, 2009. p. 1414.

80. Pullens P, Roebroeck A, Goebel R. Ground truth hardware phantoms for validation of diffusion-weighted MRI applications. J Magn Reson Imaging 2010;32:482–488.

81. Hubbard PL, Zhou FL, Eichhorn SJ, Parker GJ. Biomimetic phantom for the validation of diffusion magnetic resonance imaging. Magn Reson Med 2015;73:299–305.

82. Fieremans E, De Deene Y, Delputte S, Ozdemir MS, D'Asseler Y, Vlassenbroeck J, Deblaere K, Achten E, Lemahieu I. Simulation and experimental verification of the diffusion in an anisotropic fiber phantom. J Magn Reson. 2008;190:189–199.

83. Mattila S, Renvall V, Hiltunen J, Kirven D, Sepponen R, Hari R, Tarkiainen A. Phantom-based evaluation of geometric distortions in functional magnetic resonance and diffusion tensor imaging. Magn Reson Med 2007;57:754–763.

84. Bastin ME. Correction of eddy current-induced artefacts in diffusion tensor imaging using iterative cross-correlation. Magn Reson Imaging 1999;17:1011–1024.

85. Bastin ME, Armitage PA. On the use of water phantom images to calibrate and correct eddy current induced artefacts in MR diffusion tensor imaging. Magn Reson Imaging 2000;18:681–687.

86. Perrin M, Poupon C, Rieul B, Leroux P, Constantinesco A, Mangin JF, Lebihan D. Validation of q-ball imaging with a diffusion fibre-crossing phantom on a clinical scanner. Philos Trans R Soc Lond B Biol Sci 2005;360:881–891.

87. Qin EC, Sinkus R, Geng G, Cheng S, Green M, Rae CD, Bilston LE. Combining MR elastography and diffusion tensor imaging for the assessment of anisotropic mechanical properties: a phantom study. J Magn Reson Imaging 2013;37:217–226.

88. Lichenstein SD, Bishop JH, Verstynen TD, Yeh FC. Diffusion capillary phantom vs. human data: outcomes for reconstruction methods depend on evaluation medium. Front Neurosci 2016;10:407.

89. Komlosh ME, Ozarslan E, Lizak MJ, Horkay F, Schram V, Shemesh N, Cohen Y, Basser PJ. Pore diameter mapping using double pulsed-field gradient MRI and its validation using a novel glass capillary array phantom. J Magn Reson 2011;208:128–135.

90. Guise C, Fernandes MM, Nobrega JM, Pathak S, Schneider W, Fangueiro R. Hollow polypropylene yarns as a biomimetic brain phantom for the validation of high-definition fiber tractography imaging. ACS Appl Mater Interfaces 2016;8:29960–29967.

91. Arheden H, Holmqvist C, Thilen U, Hanseus K, Bjorkhem G, Pahlm O, Laurin S, Stahlberg F. Left-to-right cardiac shunts: comparison of measurements obtained with MR velocity mapping and with radionuclide angiography. Radiology 1999;211:453–458.

92. Ley S, Mereles D, Puderbach M, Gruenig E, Schock H, Eichinger M, Ley-Zaporozhan J, Fink C, Kauczor HU. Value of MR phase-contrast flow measurements for functional assessment of pulmonary arterial hypertension. Eur Radiol 2007;17:1892–1897.

93. Hope MD, Hope TA, Meadows AK, Ordovas KG, Urbania TH, Alley MT, Higgins CB. Bicuspid aortic valve: four-dimensional MR evaluation of ascending aortic systolic flow patterns. Radiology 2010;255: 53–61.

94. Hodnett PA, Glielmi CB, Davarpanah AH, Scanlon TG, Ward E, Collins JD, Weale PJ, Carr JC. Inline directionally independent peak velocity evaluation reduces error in peak antegrade velocity estimation in patients referred for cardiac valvular assessment. AJR Am J Roentgenol 2012;198:344–350.

95. Bley TA, Johnson KM, Francois CJ, Reeder SB, Schiebler ML, B RL, Consigny D, Grist TM, Wieben O. Noninvasive assessment of transstenotic pressure gradients in porcine renal artery stenoses by using vastly undersampled phase-contrast MR angiography. Radiology 2011;261:266–273.

96. Deng Z, Fan Z, Lee SE, et al. Noninvasive measurement of pressure gradient across a coronary stenosis using phase contrast (PC)-MRI: a feasibility study. Magn Reson Med 2016;18:218.

97. Frayne R, Steinman DA, Ethier CR, Rutt BK. Accuracy of MR phase contrast velocity measurements for unsteady flow. J Magn Reson Imaging 1995;5:428–431.

98. Frayne R, Gowman LM, Rickey DW, Holdsworth DW, Picot PA, Drangova M, Chu KC, Caldwell CB, Fenster A, Rutt BK. A geometrically accurate vascular phantom for comparative studies of x-ray,

ultrasound, and magnetic resonance vascular imaging: construction and geometrical verification. Med Phys 1993;20:415–425.

99. Holdsworth DW, Rickey DW, Drangova M, Miller DJ, Fenster A. Computer-controlled positive displacement pump for physiological flow simulation. Med Biol Eng Comput 1991;29:565–570.

100. Gatehouse PD, Rolf MP, Graves MJ, et al. Flow measurement by cardiovascular magnetic resonance: a multi-centre multi-vendor study of background phase offset errors that can compromise the accuracy of derived regurgitant or shunt flow measurements. J Cardiovasc Magn Reson 2010;12:5.

101. Rolf MP, Hofman MB, Gatehouse PD, et al. Sequence optimization to reduce velocity offsets in cardiovascular magnetic resonance volume flow quantification—a multi-vendor study. J Cardiovasc Magn Reson 2011;13:18.

102. Gatehouse PD, Rolf MP, Bloch KM, Graves MJ, Kilner PJ, Firmin DN, Hofman MB. A multi-center inter-manufacturer study of the temporal stability of phase-contrast velocity mapping background offset errors. J Cardiovasc Magn Reson 2012;14:72.

103. Barker AJ, Lanning C, Shandas R. Quantification of hemodynamic wall shear stress in patients with bicuspid aortic valve using phase-contrast MRI. Ann Biomed Eng 2010;38:788–800.

104. Casas B, Lantz J, Dyverfeldt P, Ebbers T. 4D flow MRI-based pressure loss estimation in stenotic flows: evaluation using numerical simulations. Magn Reson Med 2016;75:1808–1821.

105. Canstein C, Cachot P, Faust A, Stalder AF, Bock J, Frydrychowicz A, Kuffer J, Hennig J, Markl M. 3D MR flow analysis in realistic rapid-prototyping model systems of the thoracic aorta: comparison with in vivo data and computational fluid dynamics in identical vessel geometries. Magn Reson Med 2008;59:535–546.

106. Summers PE, Holdsworth DW, Nikolov HN, Rutt BK, Drangova M. Multisite trial of MR flow measurement: phantom and protocol design. J Magn Reson Imaging 2005;21:620–631.

107. Nordell B, Stahlberg F, Ericsson A, Ranta C. A rotating phantom for the study of flow effects in MR imaging. Magn Reson Imaging 1988;6:695–705.

108. Durand EP, Jolivet O, Itti E, Tasu JP, Bittoun J. Precision of magnetic resonance velocity and acceleration measurements: theoretical issues and phantom experiments. J Magn Reson Imaging 2001;13:445–451.

109. Nilsson A, Bloch KM, Töger J, Heiberg E, Ståhlberg F. Accuracy of four-dimensional phase-contrast velocity mapping for blood flow visualizations: a phantom study. Acta Radiologica 2013;54:663–671.

110. Tuong B, Gardiner I. Development of a novel breast MRI phantom for quality control. AJR Am J Roentgenol 2013;201:W511–W515.

111. Freed M, de Zwart JA, Loud JT, El Khouli RH, Myers KJ, Greene MH, Duyn JH, Badano A. An anthropomorphic phantom for quantitative evaluation of breast MRI. Med Phys 2011;38:743–753.

112. Freed M, de Zwart JA, Hariharan P, Myers MR, Badano A. Development and characterization of a dynamic lesion phantom for the quantitative evaluation of dynamic contrast-enhanced MRI. Med Phys 2011;38:5601–5611.

113. Keenan KE, Peskin AP, Wilmes LJ, Aliu SO, Jones EF, Li W, Kornak J, Newitt DC, Hylton NM. Variability and bias assessment in breast ADC measurement across multiple systems. J Magn Reson Imaging 2016;44:846–855.

114. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging 2012;36:1011–1014.

115. Hines CD, Yu H, Shimakawa A, McKenzie CA, Brittain JH, Reeder SB. T1 independent, T2* corrected MRI with accurate spectral modeling for quantification of fat: validation in a fat-water-SPIO phantom. J Magn Reson Imaging 2009;30:1215–1222.

116. Mashhood A, Railkar R, Yokoo T, Levin Y, Clark L, Fox-Bosetti S, Middleton MS, Riek J, Kauh E, Dardzinski BJ. Reproducibility of hepatic fat fraction measurement by magnetic resonance imaging. J Magn Reson Imaging 2013;37:1359–1370.

117. Hernando D, Sharma SD, Aliyari Ghasabeh M, et al. Multisite, multivendor validation of the accuracy and reproducibility of proton-density fat-fraction quantification at 1.5T and 3T using a fat-water phantom. Magn Reson Med 2017;77:1516–1524.

118. Meisamy S, Hines CD, Hamilton G, Sirlin CB, McKenzie CA, Yu H, Brittain JH, Reeder SB. Quantification of hepatic steatosis with T1-independent, T2-corrected MR imaging with spectral modeling of fat: blinded comparison with MR spectroscopy. Radiology 2011;258:767–775.

119. Horng DE, Hernando D, Hines CD, Reeder SB. Comparison of R2* correction methods for accurate fat quantification in fatty liver. J Magn Reson Imaging 2013;37:414–422.

120. Chenevert TL, Malyarenko DI, Newitt D, et al. Errors in quantitative image analysis due to platform-dependent image scaling. Transl Oncol 2014;7:65–71.

121. Becker GJ. 2007 Report of the Chairman of the Board of Directors. https://rsna.org/2007_Report_of_the_Chairman_of_the_Board_of_Directors.aspx. 2007. Accessed July 24, 2015.

122. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, Griswold MA. Magnetic resonance fingerprinting. Nature 2013;495:187–192.

123. Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Reson Med 2007;58:1182–1195.

124. Doneva M, Bornert P, Eggers H, Stehning C, Senegas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. Magn Reson Med 2010;64:1114–1120.

125. Russek S, Boss M, Jackson E, Jennings D, Evelhoch J, Gunter J, Sorensen A, editors. Characterization of NIST/ISMRM MRI system phantom. In Proceedings of the 20th Annual Meeting of ISMRM, Melbourne, Australia, 2012. p. 2456.

126. Keenan K, Stupic K, Boss M, et al., editors. Multi-site, multi-vendor comparison of T1 measurement using NIST/ISMRM system phantom. In Proceedings of the 24th Annual Meeting of ISMRM, Singapore, 2016. p. 3290.

127. Keenan K, Jackson E, Boss M, Kown S, Jennings D, Russek S, editors. ISMRM/NIST system phantom: T1 measurements on multiple MRI systems. In Proceedings of the 21st Annual Meeting of ISMRM, Salt Lake City, Utah, USA, 2013. p. 4338.

128. Bosca RJ. Methodological development of a multi-parametric quantitative imaging biomarker framework for assessing treatment response with MRI. Houston, Texas, USA: The University of Texas Graduate School of Biomedical Sciences at Houston; 2014.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Fig. S1. American College of Radiology large phantom (images courtesy of J.M. Specialty Parts).