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

Kathryn E. Keenan[1], Maureen Ainslie[2], Alex J. Barker[3], Michael A. Boss[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], Jie Zheng[15]

[1]National Institute of Standards and Technology Boulder, CO; [2]Duke University, Durham, NC; [3]Northwestern University, Evanston, IL; [4]Cincinnati Children's Hospital Medical Center, Cincinnati, OH; [5]University of Michigan, Ann Arbor, MI; [6]National Institutes of Health, Bethesda, MD; [7]Merck Research Laboratories, West Point, PA; [8]University of California, Los Angeles, CA; [9]University of Cooperative Education Dresden, Dresden, Germany; [10]Mayo Clinic, Rochester, MN; [11]King's College, London; [12]University of Wisconsin, Madison, WI; [13]Toshiba Medical Research Institute, Mayfield Village, OH; [14]University of Washington, Seattle, WA; [15]Washington University in St. Louis, St. Louis, MO.

Running title: Quantitative MRI Phantoms Review

Word count: 7863

Certain commercial instruments and software are identified to specify the experimental study adequately. This does not imply endorsement by NIST or that the instruments and software are the best available for the purpose.

Samir D. Sharma is now an employee of Toshiba Medical Research Institute, USA.

Corresponding author: Kathryn E. Keenan
National Institute of Standards and Technology
325 Broadway, MC 818.03
Boulder, CO 80305
kathryn.keenan@nist.gov

ABSTRACT

The magnetic resonance imaging (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 performance of an MRI machine. The International Society of Magnetic Resonance in Medicine (ISMRM) *Ad hoc* Committee on Standards for Quantitative Magnetic Resonance (SQMR) was established in February 2007 to facilitate the expansion of MRI as a mainstream modality for multi-institutional measurements, including, among other things, multi-center trials. The goal of the SQMR was to provide a framework to ensure that quantitative measures derived from magnetic resonance (MR) data are comparable over time, between subjects, between sites, and between vendors. This paper, written by members of the SQMR, 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.

KEYWORDS

Phantom, quantitative, system consistency, quality assurance

INTRODUCTION: NEED FOR QUANTITATIVE PHANTOMS

Over the past two decades, interest in the use of magnetic resonance (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 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. While 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 as well as 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). MR biomarkers are one potential way to address these problems, for example in clinical trials evaluating novel therapeutic agents or to 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 as well as from diamagnetic and/or paramagnetic components arising from the RF coil assembly and sample under study. While 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. While exponential relaxation of the proton magnetization is often observed, multi-exponential or non-exponential 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.*, nuclear magnetic resonance (NMR) spectrometer systems *vs*. 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 (NCI) established the Quantitative Imaging Network (QIN) (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 QIN, 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 SQMR 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 due to 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 'Supplementary Information') or for measurement of a single specific scanner property or for a specific application (see 'Application Specific Phantoms'). 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

MR Quantities to Be Evaluated by the System Phantom

The system phantom should be designed to assess basic system parameters such as 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 3.0 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 Supplementary 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 non-uniformity are primary factors that affect other quantities, such as SNR, slice profile, *etc*.

***B*$_1$ (Transmit) Non-uniformity:** B_1 transmit (B_1^+) non-uniformity 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 T1 mapping model fits to correct the desired *vs.* achieved flip angle (47).

***B*$_1$ (Receive) Non-uniformity:** 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 non-uniformity to be measured from transmit non-uniformity (B_1 map) has been challenged at very high field strengths.

***B*$_0$ Non-uniformity:** 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), non-trivial 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, *e.g.*, proton-density fat fraction estimation, B_0 inhomogeneity causes spatially-dependent phasing of chemical species signals, while for other applications, *e.g.*, EPI-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.

SNR: 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,

post-processing, parameter fitting strategy, *etc*. 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 NEMA and IEC methods (48, 49). This allows for the use of one of the four protocols described in NEMA 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 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 non-uniformities (transmit and receive), gradient linearity, eddy currents, and post-processing. Image uniformity is also a general indicator of the performance of the imaging pipeline.

For measuring image uniformity, NEMA 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 the fill solution has 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 non-uniformity and/or gradient non-linearity. Gradient non-linearity typically leads to geometric distortion and needs to be measured in all three axes. These non-linearities are more pronounced at the edges of the FOV. Manufacturers apply at least a 2D gradient non-linearity correction before image display,

but this correction may not be sufficient for some applications (51) and can degrade image resolution (52).

To assess geometric accuracy for head size volumes, the Alzheimer's Disease Neuroimaging Initiative (ADNI) MagPhan phantom (The Phantom Laboratory, Salem, NY) 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 at odds 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 geometric distortions in order to minimize any measurement variance due to the use of varying analysis software packages (54).

Slice position and profile: B_0 variations, RF amplifier non-linearities, and gradient non-linearity 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, *e.g.*, T_1, T_2, 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, *e.g.*, incomplete signal model.

The recommended system phantom should have 3 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 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.

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.

T_1 Relaxation Time Group: At least 10 compartments in which the T_1 relaxation time varies linearly over the range of 100 ms 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. T_2 Relaxation Time Group: At least 10 compartments in which the T_2 relaxation time varies linearly over the range of 20 ms 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 (PSF), line spread function (LSF) or modulation transfer function (MTF) 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 guidance states "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, *e.g*., 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 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 the items listed.
3. No specific recommendation is provided with respect to spherical *vs.* cylindrical geometry. Given the large variety of MR coils, including multichannel head coils, breast coils, knee coils, *etc*., 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 easily handled 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 one hour. While 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 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 non-uniformity). The software should have well-defined ROIs 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, *e.g.*, if the phantom is dropped, any hazardous materials should be contained such 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, *e.g.*, using ICP 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 and does not need to be known *a priori*. Ideally, it should 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 multi-center 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, *e.g*., region of interest (ROI) 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: requirements may be different for a phantom used to detect whether a system is within manufacturer specification compared to a phantom used for measurements to normalize or alter the data itself.
15. Certain features, such as the high contrast resolution inset, should be compatible with other imaging modalities (CT, PET, ultrasound) to have a single standard, when possible.

CURRENT MRI STANDARDS, PHANTOMS AND QUANTIFICATION EFFORTS

Recognizing the need for standard phantoms, several organizations/initiatives developed MRI phantoms. They include phantoms to: (a) characterize the physical performance of MRI systems for acceptance testing and comparison of different commercial systems performance; (b) characterize time-related changes in the physical performance of imaging systems for specific clinical protocols; and (c) 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 American College of Radiology (ACR) (56, 68, 69), including the ACR MR Accreditation Phantom (Figure S1). Each of these efforts is briefly reviewed in the Supporting Information.

APPLICATION 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, CNR, 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, that 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: minimize large signal voids; use a keyed geometry; enable tight integration with the quality control process; use the phantom as part of site qualification for inclusion in a clinical trial; and complete one phantom scan per human scan in a clinical trial.

Dynamic Contrast Enhanced MRI Perfusion Phantom

As part of the efforts of the Radiological Society of North America (RSNA) Quantitative Imaging Biomarkers Alliance® (QIBA), a phantom was designed and produced that could be used to: (1) assess bias and variance of signal intensity measurements from T_1-mapping and dynamic contrast enhanced (DCE) MRI acquisitions across scanners, centers, and time; (2) assess the effects of parallel imaging and B_1 corrections; (3) 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 (4) 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 3.0 T, while also assessing the effects of non-uniform sensitivity of phased array coils in abdominal imaging.

The RSNA QIBA DCE-MRI phantom (Figure 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 VIF compartment (VIF spheres) and in tissue (tissue spheres) during a typical DCE-MRI study. To appropriately load the radiofrequency coil, the phantom is filled with a 30 mM NaCl (Sigma-Aldrich, USA) solution in water. The phantom was used for site qualification and re-qualification in the ACRIN 6701 prostate DCE and DWI 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), while 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_2O) phantom that demonstrated ADC values in the physiologic range ($0.57 – 3.16 \times 10^{-3}$ mm^2/s) at 0 ºC without any signal from the D_2O 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 non-diffusivity phantom, which contains a highly viscous, very large mono-disperse polymer with very low ADC, for diffusion MRI experiments (76). The sucrose 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 non-diffusivity phantom is used for testing eddy current-induced image distortions.

Diffusion is a thermally-driven process and 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 ºC) to give a stable water proton ADC of 1.1 x 10^{-3} mm^2/s. With this phantom, large errors in the ADC were observed when measured off magnet isocenter, due to nonlinear gradients (78). Boss *et al*.(44) demonstrated an improved isotropic diffusion phantom (Figure 2), developed by coordinated efforts of NCI, RSNA QIBA, and NIST, that 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 ms to 1450 ms, and T_2 spans 126 ms 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 (DTI) (83). The phantom provided accurate geometric information over the scanning volume for echo planar imaging (EPI) based functional MRI and DTI 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 experimental DW-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 DTI and magnetic resonance 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 (PC-MRI) is used to assess hemodynamics in cardiovascular blood flow for a range of clinical applications, for example, 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 PC-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 PC-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 PC-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 multi-site 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 impact on the accuracy of spatially and temporally integrated phase contrast flow measurements. The most comprehensive multi-site effort to investigate phase offset errors was undertaken by members of the European Society of Cardiology (EuroCMR) 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 L -15 L tanks of aqueous gelatin solution, which were doped with 5 mmol/L of gadolinium-diethylenetriamine pentaacetic acid (Gd-DTPA) to facilitate the measurement of small background phase offsets. Gelatin eliminated phase differences due to 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 (3 experiments per scanner) were found to exceed the quality threshold. A follow-up study in nine 1.5 T 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).

While 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 ISMRM in the Assessment of Methodology of Phase Mapping for Flow Measurement (AMPMFM) 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 field of view 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 (due to 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 (such as: head height and tubing length, compliance, resistance, etc.) 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 is 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 employed 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 (UCSF/NIST) Breast phantom design with flexible outer shell easily fits into 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 ms - 1400 ms and 1500 ms - 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, since 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.

PDFF phantoms have been used to analyze the accuracy of MR-based PDFF techniques for liver fat quantification (Figure 3) (115). This phantom is composed of separate vials with approximately 40 mL volume for each, consisting of a different PDFF, typically in the range of 0% - 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. PDFF phantoms have also been used to assess the accuracy and reproducibility of PDFF measurements across different sites, vendors, and field strengths (116, 117).

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 (QC) standards, standardized analysis, and correction of collected images, if necessary.

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

MRI 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 non-uniformity, geometric nonlinearity, gradient amplitude, image uniformity, SNR, transmitter and receiver gain, receiver bandwidth, image ghosting, and eddy currents, using the methods described in IEC 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 QC standard for evaluation and acceptance of a new scanner, benchmarks for comparisons to 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 CNR required by the analytical technique.

One approach uses the strategy of ensuring specific pulse sequence parameters are identical, to the degree possible. The seeming 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 (contrast, artifact propagation, *etc*.). Further complications exist even within 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. 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. 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 post-processing 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 if 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 impacts 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 datasets that were acquired by multiple sites. However, such methods are often developed on datasets 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 datasets 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 impact on analysis should be determined and decision made 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.

CONCLUSION

Quantitative MRI enables non-invasive measurements of biomarkers pertinent to clinical trials and diagnostic tests. This paper, prepared by the ISMRM SQMR, 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 3.0 T. The system phantom prototype was constructed (125) and later commercialized; both the prototype and commercial phantoms were used by the ISMRM SQMR for studies of T1 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, efficacy and efficiency of clinical trials and is a critical step towards the full potential of MR biomarkers.

Any quantitative MRI phantom data needs 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 needs to 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 (MR-PET). 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 NCI QIN. Research developments will be enabled by the system phantom, for example, 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.

TABLES

Table 1: Theoretical sphere R_1 (=1/T_1) values at 3.0 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

The R_1 values were chosen to mimic the range of values typically encountered in a DCE-MRI study for both the vascular input function (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 $(mM \cdot s)^{-1}$ at 3.0 T.

FIGURE CAPTIONS

Figure 1

A diagram (A) and an image (B) of the QIBA DCE-MRI Phantom. The inner set of 8 spheres, referred to as the vascular input function (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 3 sets of 8 "tissue spheres" spaced at 45° increments, with each set having an R_1 range of 0.67 – 7.54 s^{-1}. The lowest R_1 sphere in each set of 8 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.

Figure 2

Ice water diffusion phantom (A) with an array of PVP solutions to obtain ADC values from 0.1 x 10^{-3} to 1.1 x 10^{-3} to 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.

Figure 3

(A) Vial consisting of a predetermined proton-density fat fraction (PDFF). (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 %.

Figure S1

ACR Large Phantom (images courtesy of J.M. Specialty Parts).

REFERENCES

1. Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, Oates JA, Peck CC, Schooley RT, Spilker BA, Woodcock J, Zeger SL. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89-95.

2. Kessler LG, Barnhart HX, Buckler AJ, Choudhury KR, Kondratovich MV, Toledano A, Guimaraes AR, Filice R, Zhang Z, Sullivan DC, Group QTW. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. Statistical methods in medical research. 2014. doi: 10.1177/0962280214537333. PubMed PMID: 24919826.

3. Raunig DL, McShane LM, Pennello G, Gatsonis C, Carson PL, Voyvodic JT, Wahl RL, Kurland BF, Schwarz AJ, Gonen M, Zahlmann G, Kondratovich M, O'Donnell K, Petrick N, Cole PE, Garra B, Sullivan DC, Group QTPW. Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment. Statistical methods in medical research. 2014. doi: 10.1177/0962280214537344. PubMed PMID: 24919831.

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(6):1774-81. doi: 10.2214/ajr.184.6.01841774. PubMed PMID: 15908529.

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. Magnetic resonance imaging. 1997;15(3):323-34. PubMed PMID: 9201680.

6. Mayr NA, Taoka T, Yuh WT, Denning LM, Zhen WK, Paulino AC, Gaston RC, Sorosky JI, Meeks SL, Walker JL, Mannel RS, Buatti JM. 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(1):14-22. PubMed PMID: 11777618.

7. Lemieux L, Hagemann G, Krakow K, Woermann FG. Fast, accurate, and reproducible automatic segmentation of the brain in T1-weighted volume MRI data. Magnetic resonance in medicine. 1999;42(1):127-35. PubMed PMID: 10398958.

8. Jack CR, Jr., Shiung MM, Gunter JL, O'Brien PC, Weigand SD, Knopman DS, Boeve BF, Ivnik RJ, Smith GE, Cha RH, Tangalos EG, Petersen RC. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology. 2004;62(4):591-600. PubMed PMID: 14981176; PMCID: PMC2730165.

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(5):464-72. PubMed PMID: 10847596.

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(3):279-85. PubMed PMID: 10588575; PMCID: PMC2733348.

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(1):253-8. doi: 10.1073/pnas.0135058100. PubMed PMID: 12506194; PMCID: PMC140943.

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(13):4637-42. doi: 10.1073/pnas.0308627101. PubMed PMID: 15070770; PMCID: PMC384799.

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(11):1868-81. doi: 10.1002/hbm.21153. PubMed PMID: 21259382; PMCID: PMC3208821.

14. Moffat BA, Chenevert TL, Lawrence TS, Meyer CR, Johnson TD, Dong Q, Tsien C, Mukherji S, Quint DJ, Gebarski SS, Robertson PL, Junck LR, Rehemtulla A, Ross BD. Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. Proc Natl Acad Sci U S A. 2005;102(15):5524-9. doi: 10.1073/pnas.0501532102. PubMed PMID: 15805192; PMCID: PMC555936.

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(3):272-9. doi: 10.1002/nbm.1320. PubMed PMID: 19009568.

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 American journal of neuroradiology. 2004;25(3):370-6. PubMed PMID: 15037457.

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(8):643-50. doi: 10.1212/WNL.0b013e3181d0ccdd. PubMed PMID: 20089939; PMCID: PMC2830922.

18. Padhani AR, Liu G, Koh DM, Chenevert TL, Thoeny HC, Takahara T, Dzik-Jurasz A, Ross BD, Van Cauteren M, Collins D, Hammoud DA, Rustin GJ, Taouli B, Choyke PL. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia. 2009;11(2):102-25. PubMed PMID: 19186405; PMCID: 2631136.

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(4):498-505. PubMed PMID: 10953180.

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(22):10855-60. doi: 10.1158/0008-5472.CAN-06-2564. PubMed PMID: 17108122.

21. Reference Values for Arterial Stiffness C. 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(19):2338-50. doi: 10.1093/eurheartj/ehq165. PubMed PMID: 20530030; PMCID: PMC2948201.

22. Ances BM, Sisti D, Vaida F, Liang CL, Leontiev O, Perthen JE, Buxton RB, Benson D, Smith DM, Little SJ, Richman DD, Moore DJ, Ellis RJ, group H. Resting cerebral blood flow: a potential biomarker of the effects of HIV in the brain. Neurology. 2009;73(9):702-8. doi: 10.1212/WNL.0b013e3181b59a97. PubMed PMID: 19720977; PMCID: PMC2734291.

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. Journal of magnetic

resonance imaging : JMRI. 2012;35(5):1162-8. doi: 10.1002/jmri.22856. PubMed PMID: 22271330.

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(3):767-75. doi: 10.1148/radiol.13121360. PubMed PMID: 23382293.

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(2):507-11. doi: 10.1148/radiol.2372040539. PubMed PMID: 16244259.

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(4):542-7. doi: 10.1111/j.1651-2227.2007.00186.x. PubMed PMID: 17306008.

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. Journal of magnetic resonance imaging : JMRI. 2009;29(6):1332-9. doi: 10.1002/jmri.21751. PubMed PMID: 19472390; PMCID: PMC2689318.

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(11):1423-9. doi: 10.1016/j.acra.2005.07.008. PubMed PMID: 16253854.

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(2):W279-85. doi: 10.2214/AJR.10.5350. PubMed PMID: 21785054.

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(7):792-800. doi: 10.1002/nbm.3303. PubMed PMID: 25943246.

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(10):1222-9. doi: 10.1002/nbm.3177. PubMed PMID: 25125325; PMCID: PMC4491353.

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. Journal of magnetic resonance imaging : JMRI. 2013;38(6):1462-71. doi: 10.1002/jmri.24117. PubMed PMID: 23559437; PMCID: PMC3775924.

33. Loomba R, Wolfson T, Ang B, Hooker J, Behling C, Peterson M, Valasek M, Lin G, Brenner D, Gamst A, Ehman R, Sirlin C. Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: a prospective study. Hepatology. 2014;60(6):1920-8. doi: 10.1002/hep.27362. PubMed PMID: 25103310; PMCID: PMC4245360.

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. Clinical gastroenterology and hepatology : the official clinical practice journal of the American

Gastroenterological Association. 2007;5(10):1207-13 e2. doi: 10.1016/j.cgh.2007.06.012. PubMed PMID: 17916548; PMCID: 2276978.

35. Clarke L, Sriram RD. Imaging as a Biomarker: Standards for Change Measurements in Therapy Workshop Summary. 2006.

36. Woodcock J, Woosley R. The FDA Critical Path Initiative and Its Influence on New Drug Development. Annual Review of Medicine. 2008;59(1):1-12. doi: doi:10.1146/annurev.med.59.090506.155819. PubMed PMID: 18186700.

37. Glaser KJ, Manduca A, Ehman RL. Review of MR elastography applications and recent developments. Journal of magnetic resonance imaging : JMRI. 2012;36(4):757-74. doi: 10.1002/jmri.23597. PubMed PMID: 22987755; PMCID: 3462370.

38. Padhani AR. Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. Journal of magnetic resonance imaging : JMRI. 2002;16(4):407-22. doi: 10.1002/jmri.10176. PubMed PMID: 12353256.

39. Huisman TA, Boltshauser E, Martin E, Nadal D. Diffusion tensor imaging in progressive multifocal leukoencephalopathy: early predictor for demyelination? AJNR American journal of neuroradiology. 2005;26(8):2153-6. PubMed PMID: 16155175.

40. Quantitative Imaging Biomarkers Alliance: Radiological Society of North America; 2016 [cited 2016 August 24, 2016]. Available from: http://www.rsna.org/qiba/.

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(4):195-7. PubMed PMID: 19956378; PMCID: PMC2781068.

42. Nordstrom RJ. The Quantitative Imaging Network in Precision Medicine. Tomography. 2016;2(4):239-41.

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. Journal of magnetic resonance imaging : JMRI. 2016;44(3):610-9. doi: 10.1002/jmri.25214. PubMed PMID: 26949897; PMCID: PMC4983524.

44. Boss M, Chenevert T, Waterton J, Morris D, Ragheb H, Jackson A, deSouza N, Collins D, van Beers B, Garteiser P, Doblas S, Persigehl T, Hedderich D, Martin A, Mukherjee P, Keenan K, Russek S, Jackson E, Zahlmann G. Thermally-Stabilized Isotropic Diffusion Phantom for Multisite Assessment of Apparent Diffusion Coefficient Reproducibilty. Medical Physics. 2014;41(6):464-. doi: 10.1118/1.4889298. PubMed PMID: 218CC116-698A-4B5C-969A-ED3928D410D9.

45. Friedman L, Glover GH. Report on a multicenter fMRI quality assurance protocol. Journal of magnetic resonance imaging : JMRI. 2006;23(6):827-39. doi: 10.1002/jmri.20583. PubMed PMID: 16649196.

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. Magnetic resonance in medicine. 2006;55(3):566-74. doi: 10.1002/mrm.20791. PubMed PMID: 16450365.

47. Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, Pike GB. On the accuracy of T1 mapping: searching for common ground. Magnetic resonance in medicine. 2015;73(2):514-22. doi: 10.1002/mrm.25135. PubMed PMID: 24578189.

48. NEMA Standards Publication MS 1-2008 R2014 Determination of Signal-to-Noise Ratio SNR in Diagnostic MRI. National Electrical Manufacturers Association; 2014.

49. IEC. Magnetic Resonance Equipment for Medical Imaging - Part 1: Determination of Essential Image Quality Parameters. 62464-1. International Electrotechnical Commission; 2007.

50. NEMA Standards Publication MS 3-2008 R2014 Image Uniformity. Determination of Image Uniformity in Diagnostic Magnetic Resonance Images 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. Journal of magnetic resonance imaging : JMRI. 2015;42(4):908-19. doi: 10.1002/jmri.24883. PubMed PMID: 25758543.

52. Tao S, Trzasko JD, Shu Y, Huston J, 3rd, Bernstein MA. Integrated image reconstruction and gradient nonlinearity correction. Magnetic resonance in medicine. 2015;74(4):1019-31. doi: 10.1002/mrm.25487. PubMed PMID: 25298258; PMCID: PMC4390402.

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. International Society of Magnetic Resonance in Medicine; 2007; Berlin, Germany.

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(6):2193-205. doi: 10.1118/1.3116776. PubMed PMID: 19610308; PMCID: PMC2754942.

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: American College of Radiology; 2015.

56. MRI Accreditation Program Requirements. American College of Radiology, 2016 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. Magnetic resonance in medicine. 2013;70(5):1293-305. doi: 10.1002/mrm.24589. PubMed PMID: 23325564.

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. 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: Report of AAPM nuclear magnetic resonance Task Group No. 1. MEDICAL PHYSICS. 1990;17.

60. Firbank MJ, Harrison RM, Williams ED, Coulthard A. Quality assurance for MRI: practical experience. The British journal of radiology. 2000;73(868):376-83. doi: 10.1259/bjr.73.868.10844863. PubMed PMID: 10844863.

61. Lerski RA, de Certaines JD. Performance assessment and quality control in MRI by Eurospin test objects and protocols. Magnetic resonance imaging. 1993;11(6):817-33. PubMed PMID: 8371637.

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. Magnetic resonance imaging. 1988;6(2):201-14. PubMed PMID: 3374294.

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 in Magnetic Resonance (Magnetic Resonance Engineering). 2002;15(1):111-6.

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. Physics in medicine and biology. 2004;49(8):1409-22. PubMed PMID: 15152682.

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. Journal of Magnetic Resonance Imaging. 2001;13(2):288-93.

67. Williams J. MagNET's MRI Test Objects 2005 [cited 2006]. Available from: http://magnet-mri.org/products/testobjects/index.htm.

68. Marquez D. Personal Communication ed2016.

69. Price R, Allison J, Clarke G, Dennis M, Hendrick E, Keener C, Masten J, Nessaiver M, Och J, Reeve D. 2015 MRI Quality Control Manual: American College of Radiology; 2015.

70. Gunter JL. Personal Communication ed2016.

71. Bosca R, Ashton E, Zahlmann G, Jackson E, editors. RSNA Quantitative Imaging Biomarker Alliance (QIBA) DCE-MRI Phantom: Goal, Design, and Initial Results. Radiological Society of North America 2012 Scientific Assembly and Annual Meeting; 2012; Chicago, IL.

72. Repeatability Assessment of Quantitative DCE-MRI and DWI: A Multicenter Study of Functional Imaging Standardization in the Prostate [March 20, 2017]. Available from: https://www.acrin.org/6701_protocol.aspx.

73. Laubach HJ, Jakob PM, Loevblad KO, Baird AE, Bovo MP, Edelman RR, Warach S. A phantom for diffusion-weighted imaging of acute stroke. Journal of magnetic resonance imaging : JMRI. 1998;8(6):1349-54. PubMed PMID: 9848751.

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. Magnetic resonance in medicine. 2000;43(3):368-74. PubMed PMID: 10725879.

75. Wang X, Reeder SB, Hernando D. An Acetone Based Phantom for Quantitative Diffusion Magnetic Resonance Imaging. Journal of Magnetic Resonance Imaging. Accepted 2017.

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. Journal of magnetic resonance imaging : JMRI. 2011;34(4):983-7. doi: 10.1002/jmri.22363. PubMed PMID: 21928310; PMCID: PMC3177164.

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. Journal of Magnetic Resonance Imaging. 2013;37(5):1238-46. doi: 10.1002/jmri.23825.

79. Pierpaoli C, Sarlls J, Nevo U, Basser PJ, Horkay F. Polyvinylpyrrolidone (PVP) water solutions as isotropic phantoms for diffusion MRI studies. Proc Intl Soc Mag Reson Med. 2009; 17:1414.

80. Pullens P, Roebroeck A, Goebel R. Ground truth hardware phantoms for validation of diffusion-weighted MRI applications. Journal of magnetic resonance imaging : JMRI. 2010;32(2):482-8. doi: 10.1002/jmri.22243. PubMed PMID: 20677281.

81. Hubbard PL, Zhou FL, Eichhorn SJ, Parker GJ. Biomimetic phantom for the validation of diffusion magnetic resonance imaging. Magnetic resonance in medicine. 2015;73(1):299-305. doi: 10.1002/mrm.25107. PubMed PMID: 24469863.

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(2):189-99. doi: 10.1016/j.jmr.2007.10.014. PubMed PMID: 18023218.

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. Magnetic resonance in medicine. 2007;57(4):754-63. doi: 10.1002/mrm.21218. PubMed PMID: 17390354.

84. Bastin ME. Correction of eddy current-induced artefacts in diffusion tensor imaging using iterative cross-correlation. Magnetic resonance imaging. 1999;17(7):1011-24. PubMed PMID: 10463652.

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. Magnetic resonance imaging. 2000;18(6):681-7. PubMed PMID: 10930777.

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(1457):881-91. doi: 10.1098/rstb.2005.1650. PubMed PMID: 16087433; PMCID: PMC1854933.

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. Journal of magnetic resonance imaging : JMRI. 2013;37(1):217-26. doi: 10.1002/jmri.23797. PubMed PMID: 22987805.

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. doi: 10.3389/fnins.2016.00407. PubMed PMID: 27656122; PMCID: PMC5013034.

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(1):128-35. doi: 10.1016/j.jmr.2010.10.014. PubMed PMID: 21084204; PMCID: PMC3021618.

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(44):29960-7. doi: 10.1021/acsami.6b09809. PubMed PMID: 27723307.

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(2):453-8. doi: 10.1148/radiology.211.2.r99ma43453. PubMed PMID: 10228528.

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(7):1892-7. doi: 10.1007/s00330-006-0559-9. PubMed PMID: 17225131.

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(1):53-61. doi: 10.1148/radiol.09091437. PubMed PMID: 20308444.

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(2):344-50. doi: 10.2214/AJR.10.5941. PubMed PMID: 22268176.

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(1):266-73. doi: 10.1148/radiol.11101175. PubMed PMID: 21813739; PMCID: PMC3184231.

96. Deng Z, Fan Z, Lee SE, Nguyen C, Xie Y, Pang J, Bi X, Yang Q, Choi BW, Kim JS, Berman D, Chang HJ, Li D. Noninvasive measurement of pressure gradient across a coronary stenosis using phase contrast (PC)-MRI: A feasibility study. Magnetic resonance in medicine. 2016. doi: 10.1002/mrm.26579. PubMed PMID: 28019028.

97. Frayne R, Steinman DA, Ethier CR, Rutt BK. Accuracy of MR phase contrast velocity measurements for unsteady flow. Journal of magnetic resonance imaging : JMRI. 1995;5(4):428-31. PubMed PMID: 7549205.

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(2 Pt 1):415-25. doi: 10.1118/1.597141. PubMed PMID: 8497234.

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(6):565-70. PubMed PMID: 1813750.

100. Gatehouse PD, Rolf MP, Graves MJ, Hofman MB, Totman J, Werner B, Quest RA, Liu Y, von Spiczak J, Dieringer M, Firmin DN, van Rossum A, Lombardi M, Schwitter J, Schulz-Menger J, Kilner PJ. 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. doi: 10.1186/1532-429X-12-5. PubMed PMID: 20074359; PMCID: PMC2818657.

101. Rolf MP, Hofman MB, Gatehouse PD, Markenroth-Bloch K, Heymans MW, Ebbers T, Graves MJ, Totman JJ, Werner B, van Rossum AC, Kilner PJ, Heethaar RM. Sequence optimization to reduce velocity offsets in cardiovascular magnetic resonance volume flow quantification--a multi-vendor study. J Cardiovasc Magn Reson. 2011;13:18. doi: 10.1186/1532-429X-13-18. PubMed PMID: 21388521; PMCID: PMC3065419.

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. doi: 10.1186/1532-429X-14-72. PubMed PMID: 23083397; PMCID: PMC3514147.

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(3):788-800. doi: 10.1007/s10439-009-9854-3. PubMed PMID: 19953319; PMCID: PMC2872988.

104. Casas B, Lantz J, Dyverfeldt P, Ebbers T. 4D Flow MRI-based pressure loss estimation in stenotic flows: Evaluation using numerical simulations. Magnetic resonance in medicine. 2016;75(4):1808-21. doi: 10.1002/mrm.25772. PubMed PMID: 26016805.

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. Magnetic resonance in medicine. 2008;59(3):535-46. doi: 10.1002/mrm.21331. PubMed PMID: 18306406.

106. Summers PE, Holdsworth DW, Nikolov HN, Rutt BK, Drangova M. Multisite trial of MR flow measurement: phantom and protocol design. Journal of magnetic resonance imaging : JMRI. 2005;21(5):620-31. doi: 10.1002/jmri.20311. PubMed PMID: 15834913.

107. Nordell B, Stahlberg F, Ericsson A, Ranta C. A rotating phantom for the study of flow effects in MR imaging. Magnetic resonance imaging. 1988;6(6):695-705. PubMed PMID: 2850436.

108. Durand EP, Jolivet O, Itti E, Tasu JP, Bittoun J. Precision of magnetic resonance velocity and acceleration measurements: theoretical issues and phantom experiments. Journal of magnetic resonance imaging : JMRI. 2001;13(3):445-51. PubMed PMID: 11241820.

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(6):663-71. Epub April 30, 2013. doi: 10.1177/0284185113478005.

110. Tuong B, Gardiner I. Development of a novel breast MRI phantom for quality control. AJR Am J Roentgenol. 2013;201(3):W511-5. doi: 10.2214/AJR.12.9571. PubMed PMID: 23971483.

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(2):743-53. doi: 10.1118/1.3533899. PubMed PMID: 21452712; PMCID: PMC3033878.

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(10):5601-11. doi: 10.1118/1.3633911. PubMed PMID: 21992378; PMCID: PMC3195376.

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. Journal of magnetic resonance imaging : JMRI. 2016;44(4):846-55. doi: 10.1002/jmri.25237. PubMed PMID: 27008431; PMCID: PMC5098898.

114. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: A standardized MR-based biomarker of tissue fat concentration. Journal of Magnetic Resonance Imaging. 2012;36(5):1011-4.

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. Journal of Magnetic Resonance Imaging. 2009;30(5):1215-22.

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. Journal of Magnetic Resonance Imaging. 2013;37(6):1359-70.

117. Hernando D, Sharma SD, Aliyari Ghasabeh M, Alvis BD, Arora SS, Hamilton G, Pan L, Shaffer JM, Sofue K, Szeverenyi NM, Welch EB, Yuan Q, Bashir MR, Kamel IR, Rice MJ, Sirlin CB, Yokoo T, Reeder SB. Multisite, multivendor validation of the accuracy and reproducibility of proton-density fat-fraction quantification at 1.5T and 3T using a fat-water phantom. Magnetic resonance in medicine. 2017;77(4):1516-24. doi: 10.1002/mrm.26228. PubMed PMID: 27080068; PMCID: PMC4835219.

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(3):767-75. doi: 10.1148/radiol.10100708. PubMed PMID: 21248233; PMCID: PMC3042638.

119. Horng DE, Hernando D, Hines CD, Reeder SB. Comparison of R2* correction methods for accurate fat quantification in fatty liver. Journal of magnetic resonance imaging : JMRI. 2013;37(2):414-22. doi: 10.1002/jmri.23835. PubMed PMID: 23165934; PMCID: PMC3557786.

120. Chenevert TL, Malyarenko DI, Newitt D, Li X, Jayatilake M, Tudorica A, Fedorov A, Kikinis R, Liu TT, Muzi M, Oborski MJ, Laymon CM, Li X, Thomas Y, Jayashree KC, Mountz

JM, Kinahan PE, Rubin DL, Fennessy F, Huang W, Hylton N, Ross BD. Errors in Quantitative Image Analysis due to Platform-Dependent Image Scaling. Transl Oncol. 2014;7(1):65-71. PubMed PMID: 24772209; PMCID: PMC3998685.

121. Becker GJ. 2007 Report of the Chairman of the Board of Directors 2007 [cited 2015 July 24, 2005]. Available from: https://rsna.org/2007_Report_of_the_Chairman_of_the_Board_of_Directors.aspx.

122. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, Griswold MA. Magnetic resonance fingerprinting. Nature. 2013;495(7440):187-92. doi: 10.1038/nature11971. PubMed PMID: 23486058; PMCID: PMC3602925.

123. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic resonance in medicine. 2007;58(6):1182-95. doi: 10.1002/mrm.21391. PubMed PMID: 17969013.

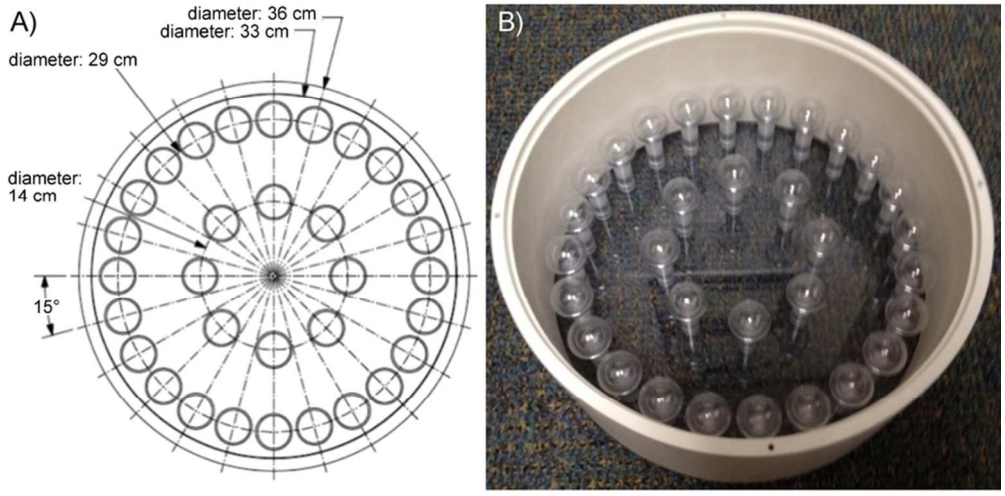
124. Doneva M, Bornert P, Eggers H, Stehning C, Senegas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. Magnetic resonance in medicine. 2010;64(4):1114-20. doi: 10.1002/mrm.22483. PubMed PMID: 20564599.

125. Russek S, Boss M, Jackson E, Jennings D, Evelhoch J, Gunter J, Sorensen A, editors. Characterization of NIST/ISMRM MRI System Phantom. International Society for Magnetic Resonsance in Medicine; 2012; Melbourne, Australia.

126. Keenan K, Stupic K, Boss M, Russek S, Chenevert T, Prasad P, Reddick W, Cecil K, Zheng J, Hu P, Jackson E, editors. Multi-site, multi-vendor comparison of T1 measurement using NIST/ISMRM system phantom. International Society for Magnetic Resonance in Medicine; 2016; Singapore.

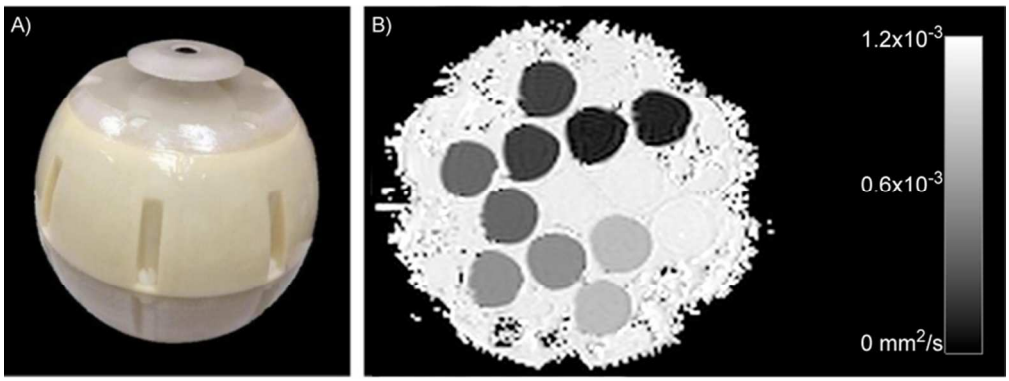
127. Keenan K, Jackson E, Boss M, Kown S, Jennings D, Russek S, editors. ISMRM/NIST System Phantom: T1 Measurements on Multiple MRI Systems. International Society for Magnetic Resonance in Medicine; 2013; Salt Lake City, UT.

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



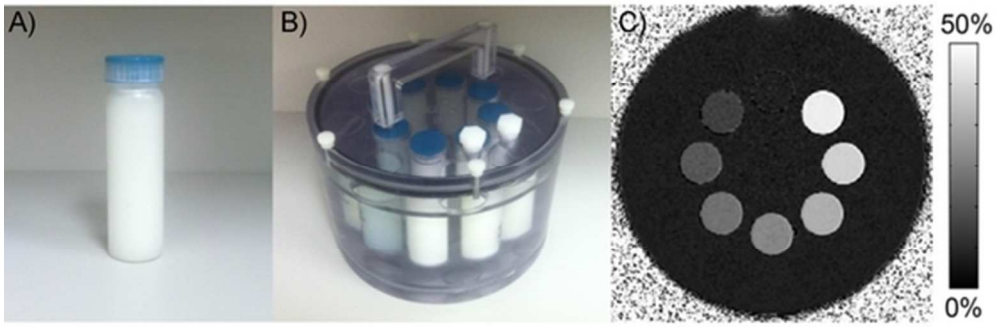
A diagram (A) and an image (B) of the QIBA DCE-MRI Phantom. The inner set of 8 spheres, referred to as the vascular input function (VIF) spheres, is distributed around a 14.0 cm diameter circle. The remaining 24 spheres are uniformly distributed around a 29.0 cm diameter circle, and consist of 3 sets of 8 "tissue spheres" spaced at 45 increments, with each set having an R1 range of 0.67 – 7.54 s-1. The lowest R1 sphere in each set of 8 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.

78x39mm (300 x 300 DPI)



Ice water diffusion phantom (A) with an array of PVP solutions to obtain ADC values from 0.1 x 10-3 to 1.1 x 10-3 to mm2/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.

76x28mm (300 x 300 DPI)



(A) Vial consisting of a predetermined proton-density fat fraction (PDFF). (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 %.

50x16mm (300 x 300 DPI)

SUPPORTING MATERIAL

CURRENT MRI STANDARDS, PHANTOMS AND QUANTIFICATION EFFORTS

Recognizing the need for standard phantoms, several organizations/initiatives developed MRI phantoms. They include phantoms to: (a) characterize the physical performance of MRI systems for acceptance testing and comparison of different commercial systems performance; (b) characterize time-related changes in the physical performance of imaging systems for specific clinical protocols; and (c) develop methods for accreditation of MRI systems for clinical practice.

American Association of Physicists in Medicine (AAPM)

Since the early 1980s, the American Association of Physicists in Medicine (AAPM) produced multiple MRI task group reports that provide consensus recommendations on how to design phantoms to characterize MRI system physical performance, such as resonance frequency, signal-to-noise ratio (SNR), image uniformity, spatial linearity, spatial resolution, slice thickness, slice position/separation, and phase related image artifacts. The AAPM published a report (AAPM Report No. 28) in 1990 on “Quality Assurance Methods and Phantoms for Magnetic Resonance Imaging” to describe a set of standard acceptance testing procedures for emerging commercial clinical MRI systems (1). The report also provided recommendations for acceptable MR phantom materials, phantom designs, and analysis procedures. An updated task group report was published in 2010 as Acceptance Testing and Quality Assurance Procedures for Magnetic Resonance Imaging Facilities (AAPM Report 100) (2). The primary goal of these efforts was to implement quality assurance methods in the routine clinical practice setting. These measurements, however, are focused on the overall physical performance of the MRI scanner as part of a routine quality assurance program. In addition, there is a need to assess the performance of specific protocols for MR biomarkers over time.

European Communities Biomedical Engineering Advisory Committee (COMAC-BME)

The European research community implemented a project referred to as the EuroSpin Test Objects, which were designed as part of the work of European Concerted Action "Tissue Characterization by MRS and MRI" (part of COMAC-BME). It expended considerable effort in the field of MRI quality control. Their original design consisted of five test objects (T0l-T05) and was tested for quality control from acceptance testing to routine quality assurance in the late 1980s (3). They were manufactured and distributed commercially by Diagnostic Sonar Ltd

(Livingston, Scotland). However, there are limitations associated with the test set, and generally detailed measurements must be made through a combination of test object studies and some adjustments to the mode of operation of the scanner. Subsequently in the early 1990s, modifications were made to the test objects, which enhanced functionality (4).

More recently, Firbank et al. described their experience with the EuroSpin phantoms for quality assurance (QA) in MRI, based on a comprehensive assessment of QA parameters undertaken on clinical MRI scanners (5). SNR and image uniformity were measured daily. Slice thickness and position, geometric distortion, image resolution and image ghosting were assessed monthly. They found that apart from some drift of the RF amplifier voltage, all measurements were within acceptable error limits and were stable over the course of one year. They stated that the SNR, geometric distortion and RF amplifier voltage were simple to determine and could be measured in less than 15 minutes by the scanner operator, using the scanner software. They recommended weekly in-house recording of these parameters for clinical MRI scanners that are frequently used. In addition, comprehensive QA routines were discussed for systems used for quantitative measurements.

Magnetic Resonance National Evaluation Team (MagNET)

The MagNET program in the United Kingdom (UK) performed image quality evaluations from 1988 until 2007. They were an independent evaluation program supported by the Medicines and Healthcare products Regulatory Agency (MHRA), an executive agency of the UK Department of Health. Using the methods described by COMAC-BME and the National Electrical Manufacturers Association (NEMA) and their own unique methods, the MagNET team conducted testing at manufacturing facilities and in hospitals to advise the UK on MRI purchasing. The results from each test were published as reports from the UK Medical Devices agency, and manufacturers were given the opportunity to include comments in the published report. They also provided acceptance testing on installation at a hospital (6). The MagNET program used test objects similar to the EuroSpin test objects, including gadolinium doped agarose gels with known T_1 and T_2 values (T_1 range: ~300 ms - 1600 ms; T_2 range ~50 ms - 400 ms) (7). The test objects were available for purchase (8). The MagNET team published several papers comparing MRI scanners, including an investigation of acoustic noise (9). In addition, the group published guidance on quality control protocols, including diffusion MRI (10). Much of

the program's work is summarized in an intersystem comparison paper focusing on SNR, geometric distortion, slice width and imaging speed (6).

American College of Radiology (ACR)

The American College of Radiology (ACR) has a long history of accreditation for diagnostic imaging and radiation oncology facilities, going back as far as 1963. The ACR MRI accreditation program was initiated in the late 1980s and was formally introduced in the mid-1990s. The ACR released its initial MRI Quality Control manual in 2001 and revised manuals in 2004 and 2015 (11). The program was designed to provide recommended procedures for evaluating equipment performance and a formal recommendation for quality control measures (at least weekly quality control assessments and annual system performance testing assessments). The program requires submission of phantom images, multiple sets of clinical images, copies of most recent annual system performance test results, and site information, including continuing medical education information for radiologists, technologists, and physicists. Information on the program can be obtained at http://www.acr.org/Quality-Safety/Accreditation/MRI.

An ACR MRI phantom was designed by the ACR and is manufactured by J.M. Specialty Parts (San Diego, CA, USA). The phantom is cylindrically shaped, weighs 6.4 kg, and has an inner diameter of 190 mm, outer diameter of 204 mm and length of 165 mm. It is filled with a solution containing 10 mM nickel chloride ($NiCl_2$) and 15 mM sodium chloride (NaCl) to give relaxation times and conductivities in the biological range of interest, and it fits in many common head coils. The cost for the large phantom currently is over one thousand US dollars (12). The ACR phantom is not designed specifically for quantitative imaging biomarker measurements. It was designed to generally assess geometric accuracy, section thickness, high contrast spatial resolution, low contrast object detectability, position accuracy, and signal uniformity and ghosting. The assessment of geometric accuracy does provide minimal support for quantitative MRI measurements (e.g. assessment of tumor volume), but the geometric accuracy components are only in one plane. The phantom and recommended imaging protocol have the advantages of low cost and rapid scan time. The phantom was primarily designed as a method to assist site accreditation performed by the ACR and as a mechanism to improve quality control by repeated imaging on a weekly basis. The phantom is an integral component of the most commonly used MRI accreditation program in the United States, and thus, it is undoubtedly the most widely available phantom with over 13,000 copies sold since 1997 (13).

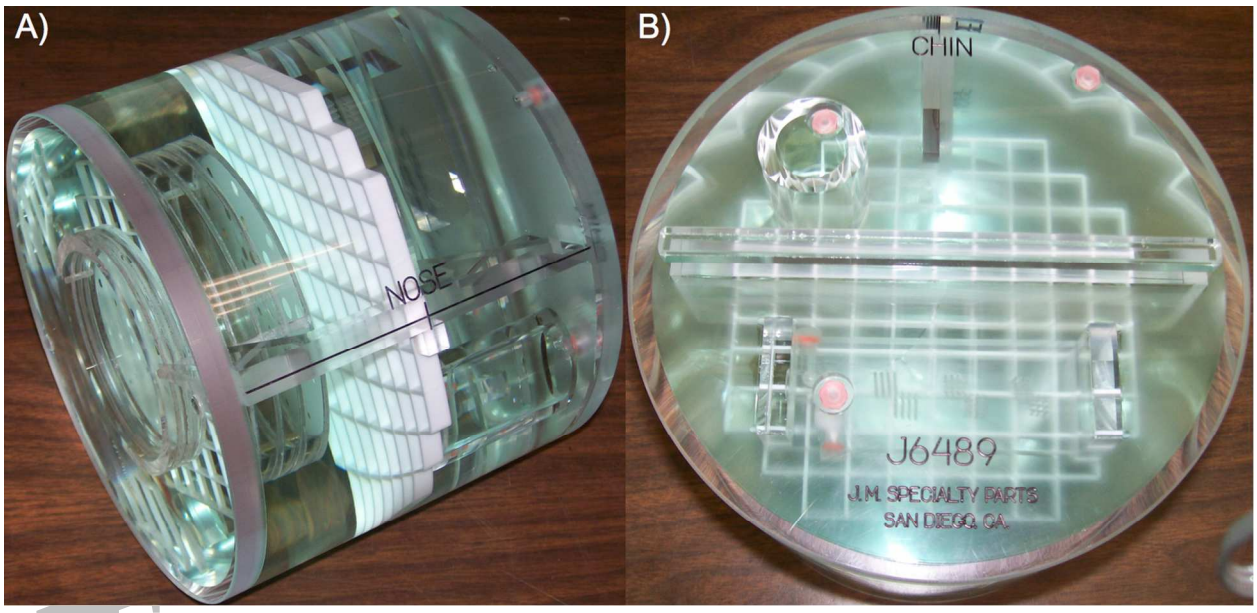


Figure S1: ACR large phantom (images courtesy of J.M. Specialty Parts).

ACRONYMS

AAPM	American Association of Physicists in Medicine
ACR	American College of Radiology
ADNI	Alzheimer's Disease Neuroimaging Initiative
COMAC-BME	European Communities Biomedical Engineering Advisory Committee
ISMRM	International Society of Magnetic Resonance in Medicine
MagNET	Magnetic Resonance National Evaluation Team
MedPAC	Medicare Payment Advisory Commission
MHRA	Medicines and Healthcare products Regulatory Agency
NCI	National Cancer Institute
NEMA	National Electrical Manufacturers Association
NIST	National Institute of Standards and Technology
QIBA	Quantitative Imaging Biomarker Alliance
QIN	Quantitative Imaging Network
RSNA	Radiological Society of North America
SQMR	Standards for Quantitative Magnetic Resonance
UCSF	University of California San Francisco

GLOSSARY

ACCURACY: The accuracy of an analytical procedure expresses the closeness of agreement between the value, which is accepted either as a conventional true value or an accepted reference value and the value found.

ANALYTICAL PROCEDURE: The image analytical procedure refers to the way of performing the analysis. It should describe in detail the steps necessary to perform each analytical test. This may include but is not limited to: the sample, the reference standard and the reagents preparations, use of the apparatus, generation of the calibration curve, use of the formulae for the calculation, etc.

LINEARITY: The linearity of an analytical procedure is its ability (within a given range) to obtain test results, which are directly proportional to the concentration (amount) of analyte in the sample.

PRECISION: The precision of an analytical procedure expresses the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Precision may be considered at three levels: repeatability, intermediate precision and reproducibility. Precision should be investigated using homogeneous, authentic samples. However, if it is not possible to obtain a homogeneous sample it may be investigated using artificially prepared samples or a sample solution. The precision of an analytical procedure is usually expressed as the variance, standard deviation or coefficient of variation of a series of measurements.

- **Repeatability**: Repeatability expresses the precision under the same operating conditions over a short interval of time. Repeatability is also termed intra-analytical precision
- **Intermediate precision**: Intermediate precision expresses within laboratories variations: different days different analysts, different equipment, etc.
- **Reproducibility**: Reproducibility expresses the precision between laboratories (collaborative studies usually applied to standardization of methodology).

PHANTOM:

- A biomimetic or non-biomimetic object that can be imaged using a specific modality (e.g. MRI) and generates a signal (image) that can be analyzed against a gold standard value. An example of a spatial fidelity phantom for MRI is described in section 3.2.
- A digitally synthesized image of known composition that can be analyzed against the gold standard values.

RANGE: The range of an analytical procedure is the interval between the upper and lower concentration (amounts) of analyte in the sample (including these concentrations) for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and linearity.

ROBUSTNESS: The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage.

REFERENCES

1. 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: Report of AAPM nuclear magnetic resonance Task Group No. 1. MEDICAL PHYSICS. 1990;17.

2. 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. 2010.

3. 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. Magnetic resonance imaging. 1988;6(2):201-14. PubMed PMID: 3374294.

4. Lerski RA, de Certaines JD. Performance assessment and quality control in MRI by Eurospin test objects and protocols. Magnetic resonance imaging. 1993;11(6):817-33. PubMed PMID: 8371637.

5. Firbank MJ, Harrison RM, Williams ED, Coulthard A. Quality assurance for MRI: practical experience. The British journal of radiology. 2000;73(868):376-83. doi: 10.1259/bjr.73.868.10844863. PubMed PMID: 10844863.

6. De Wilde JP, Price D, Curran J, Williams J, Kitney RI. Standardization of Performance Evaluation in MRI: 13 Years' Experience of Intersystem Comparison. Concepts in Magnetic Resonance (Magnetic Resonance Engineering). 2002;15(1):111-6.

7. 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.

8. Williams J. MagNET's MRI Test Objects 2005 [cited 2006]. Available from: http://magnet-mri.org/products/testobjects/index.htm.

9. 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. Journal of Magnetic Resonance Imaging. 2001;13(2):288-93.

10. Delakis I, Moore EM, Leach MO, De Wilde JP. Developing a quality control protocol for diffusion imaging on a clinical MRI system. Physics in medicine and biology. 2004;49(8):1409-22. PubMed PMID: 15152682.
11. Price R, Allison J, Clarke G, Dennis M, Hendrick E, Keener C, Masten J, Nessaiver M, Och J, Reeve D. 2015 MRI Quality Control Manual: American College of Radiology; 2015.
12. MRI Accreditation Program Requirements. American College of Radiology, 2016 June 23, 2016.
13. Marquez D. Personal Communication ed2016.