A Standard System Phantom for Magnetic Resonance Imaging

Karl F. Stupic[1], Maureen Ainslie[2], Michael A. Boss[3], Cecil Charles[2], Larry Clarke[4], Andrew M. Dienstfrey[1], Jeffrey L. Evelhoch[5], Paul Finn[6], Zydrunas Gimbutas[1], Jeffrey L. Gunter[7], Derek L. G. Hill[8], Clifford R. Jack[7], Edward F. Jackson[9], Kathryn E. Keenan[1], Guoying Liu[3], Michele Martin[1], Nikki S. Rentz[1], Chun Yuan[10], Stephen E. Russek[1]

[1]National Institute of Standards and Technology, Boulder, CO; [2]Duke University, Durham, NC; [3]American College of Radiology, Philadelphia, PA; [4]National Institutes of Health, Bethesda MD; [5]Merck Research Laboratories, West Point PA; [6]University of California, Los Angeles, CA; [7]Mayo Clinic, Rochester, MN; [8]University College London, London, United Kingdom; [9]University of Wisconsin, Madison, WI; [10]University of Washington, Seattle WA

Running title: A Standard MRI Phantom

Word Count: 7545

Contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

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.

Corresponding author: Karl F. Stupic
National Institute of Standards and Technology
325 Broadway, MC 818.03
Boulder, CO 80305
karl.stupic@nist.gov

ABSTRACT

Purpose: A standard MRI system phantom has been designed and fabricated to assess scanner performance, stability, comparability and assess the accuracy of quantitative relaxation time imaging. The phantom is unique in having traceability to the International System of Units, a high level of precision, long-term stability, and monitoring by a national metrology institute. Here, we describe the phantom design, construction, imaging protocols, and measurement of geometric distortion, resolution, slice profile, signal-to-noise ratio (SNR), proton-spin relaxation times, image uniformity and proton density.

Methods: The system phantom, designed by the ISMRM ad hoc committee on Standards for Quantitative MR, is a 200 mm spherical structure that contains a 57-element fiducial array; two relaxation time arrays; a proton density/SNR array; resolution and slice-profile insets. Standard imaging protocols are presented, which provide rapid assessment of geometric distortion, image uniformity, T_1 and T_2 mapping, image resolution, slice profile, and SNR.

Results: Fiducial array analysis gives assessment of intrinsic geometric distortions, which can vary considerably between scanners and correction techniques. This analysis also measures scanner/coil image uniformity, spatial calibration accuracy, and local volume distortion. An advanced resolution analysis gives both scanner and protocol contributions. SNR analysis gives both temporal and spatial contributions.

Conclusions: A standard system phantom is useful for characterization of scanner performance, monitoring a scanner over time, and to compare different scanners. This type of calibration structure is useful for quality assurance, benchmarking quantitative MRI protocols, and to transition MRI from a qualitative imaging technique to a precise metrology with documented accuracy and uncertainty.

KEYWORDS

Quantitative MRI, phantom, quality assurance, MRI standards

INTRODUCTION

With the rise of quantitative magnetic resonance imaging (MRI) to complement qualitative MRI, it is imperative to assess system stability and understand the comparability of measurements across MRI scanners. One solution is to use a standard imaging phantom, an inanimate object used to characterize or calibrate an imaging system. A standard phantom would be one that is commonly used and accepted by the imaging community; has stable, precisely-defined properties to allow monitoring of scanner performance and accuracy of image-based measurements; is fully documented; has long term monitoring and maintenance; and is traceably connected to the international system of measurements. The need for a standard magnetic resonance imaging system phantom was voiced at a National Institute of Standards and Technology (NIST) workshop entitled "Imaging as a Biomarker: Standards for Change Measurements in Therapy," that was held on September 14-15, 2006 (1). The workshop action items included: (a) Design phantoms that may better characterize the time-related physical performance of imaging systems, and the performance of specific functional and molecular based measurements; (b) Define the physical performance of different imaging platforms required to measure change analysis; and (c) Develop and share open source tools to analyze phantom or simulated data.

The Ad Hoc Committee for Standards in Quantitative Magnetic Resonance (SQMR) was formed, under the auspices of the International Society of Magnetic Resonance in Medicine (ISMRM), to address these action items in the areas that pertained to MRI. The committee developed recommendations for a system phantom (2) that could be used to determine the accuracy, stability and comparability of MRI scanners. The desired measurements include:

1. Radio frequency (RF) field, B_1, non-uniformity
2. Static main magnetic field, B_0, non-uniformity
3. Geometric linearity
4. Gradient amplitude
5. Slice position and profile
6. Image uniformity
7. Resolution (high-contrast detectability)
8. Signal-to-noise ratio (low-contrast detectability)

9. Accuracy and precision of measurement of proton spin relaxation times: T_1 (longitudinal) and T_2 (transverse) and proton density
10. System constancy

Such a system phantom can be used to: (a) track scanner performance at a particular site over time, as well as compare performance with other scanners; (b) determine the accuracy of certain quantitative measurements, such as T_1 mapping, and assist in the development of appropriate imaging protocols to obtain desired accuracy; (c) validate scanner and protocols for participation in clinical trials; and (d) determine the best phantom fabrication techniques for future phantoms, e.g. application-specific phantoms.

Previous successful efforts to create standard phantoms include the American College of Radiology (ACR) MRI phantom (3) and the Alzheimer's Disease Neuroimaging Initiative (ADNI) MRI phantom (4). The ACR magnetic resonance accreditation phantom is a water-filled cylindrical phantom with relaxation times and conductivities in the biological range of interest. The phantom contains a resolution inset, slice profile wedges, a grid to determine geometric distortion, and a contrast array. 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 spherical ADNI phantom is used for measurements of signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), and geometric distortion in scan protocols intended to monitor changes in the brain morphology over time. From 2004 to 2009 the ADNI study acquired a phantom image with each participant study (4). The key results from this phantom plus human-participant study were: 1) scanner calibrations, specifically gradient amplitude, were quite stable with scanner service (minimal drift) and re-calibration introduced changes far exceeding the calibration drift; 2) affine registration of human participant images over time performs as well as the corrections derived from phantom imaging. For ADNI-2 (2010-2017), phantom images were acquired for site certification and after system upgrades. These images were used to ensure that off-line gradient non-linearity corrections based on scanner model could be performed correctly (5). In ADNI-3 (2018-present), all scanners in the study have full three-dimensional (3D) on-board corrections for gradient non-linearity, and sites are requested to use ADNI phantoms for certification if possible. The standard system phantom described here builds off the ACR and ADNI phantoms, addressing the need for a more general

and precise calibration object that has properties traceable to primary standards connected to the International System of Units (SI) (2).

This paper describes the design, construction, accuracy requirements, and measurement protocols of the ISMRM system phantom. The system phantom, designed by the ISMRM SQMR ad hoc committee, is a precise imaging artifact for quantitatively measuring image distortions and validating quantitative mapping. From this design, a 3D model and a materials data base were created, and two prototype phantoms were constructed. In addition to the physical artifact, the phantom package includes the 3D model, a materials data base, suggested measurement protocols, and analysis software. Subsequently, a version of this system phantom has been commercialized, with over 100 copies in existence.

A single phantom cannot provide all the calibration measurements required for assessing the accuracy of MRI scanners and quantitative imaging protocols. However, this precise, well-designed phantom can address many of the important measurements and quality control issues. A universally-used phantom is essential to allow sites to measure system stability and readily compare scanner performance with other scanners. There are several key elements of the MRI system phantom described here that make it unique:

1. SI traceable parameters including geometric dimensions (mm), nuclear magnetic resonance (NMR) parameters such as proton spin relaxation times (s), and material compositions that dictate contrast properties.
2. Dimensionally-stable construction using a rigid precision-joined framework and materials that are both low thermal-expansion and low water-uptake.
3. All components including design, 3D models, imaging protocols, detailed description of construction materials, and imaging analysis routines are in the public domain, https://collaborate.nist.gov/mriphantoms.
4. Long term maintenance/monitoring of stability and accuracy by a national metrology institute.

METHODS

This section gives an overview of the design and construction of prototype MRI system phantoms and the process to provide SI-traceable calibration of the phantom properties. A more comprehensive description of the solutions, chemical components, and materials can be found in

the Supporting Information. The design and construction specifications are meant as a common reference for future commercial versions, not as specific requirements. Subsequent commercial versions may have modifications to improve manufacturability and to reduce cost. The prototype phantoms are modular so that all components may be disassembled and reassembled in a variety of configurations for testing alternative designs.

System Phantom Design

The prototype MRI system phantom, shown in Figure 1, consists of a water-filled spherical polycarbonate shell with a 200 mm inner diameter (ID). The spherical diameter was chosen to mimic a human head and was designed to fit in most head coil assemblies. A spherical design also allows for easy rotation of the phantom. Inside the spherical shell is a framework consisting of five 8.0 mm-thick polyphenylene sulfide (PPS) plates rigidly connected with PPS rods. PPS was chosen since it is a high-performance plastic with low water absorption, low thermal expansion, and good machinability. The PPS plates are annealed for stress relief and then machined to an inspected flatness of less than 0.04 mm. The plates support 57 fiducial spheres, a 14 element $NiCl_2$ array, a 14 element $MnCl_2$ array, a 14 element proton density array, two resolution insets, and wedges for slice profiles. The plates, as shown in Figure 1, contain the following structures: Plate 1, 5 fiducial spheres on bottom; Plate 2, 13 fiducial spheres on top and slice profile wedges on bottom; Plate 3, 26 fiducial spheres on top, proton density array on bottom; Plate 4, coarse resolution inset in the plate, 13 fiducial spheres on top, $MnCl_2$ array on bottom; Plate 5, 5 fiducial spheres on top, serial numbers and phantom ID in the plate, and $NiCl_2$ array and fine resolution inset on bottom. The fine resolution inset can also be mounted on plates 3 or 4.

The phantom coordinate system was chosen to match typical scanner/patient coordinates, with +Z along the scanner bore axis (superior), +X toward the left, and +Y down (posterior), when mounting the phantom in the default orientation. A set of eye decals is included on the phantom shell to allow easy recognition and placement of the phantom within the scanner, as seen in Figure1B. The phantom is placed in the same fashion as a patient laying supine on the imaging bed. The origin of the phantom is defined as the center of the central fiducial sphere as shown in Figure1C. The phantom is designed so that the plates are in the XZ (coronal) plane with left (L), right (R), superior (S), and inferior (I) marked on each plate, along with the plate

identifier. Phantom identification information, along with the serial number, are etched into plate 5, which are easily readable in the prescribed 3D scan, as shown in Figure 1A. In the default orientation, the MR parameter arrays ($NiCl_2$, $MnCl_2$, proton density arrays) are imaged with a coronal scan. Recommended protocols for measuring the arrays and insets on clinical MRI systems are given in the Supporting Information.

The inner frame of the prototype phantom was designed to be very rigid and allow the fiducial-sphere centers to be located within a ± 0.1 mm accuracy. Details of the plate locating mechanism are shown in Supporting Information Figure S1. A hemispherical machined surface on the post mates with a set of inclined plates to precisely locate the plates and eliminate torsional bending. The precision construction of the prototypes comes at an increased cost due to the need for computer-controlled machining. It was deemed necessary for the prototypes to have such precision to allow the study of scanner geometric distortion.

Fiducial Array

The system phantom contains a 3D fiducial array that can be used to assess geometric distortion, image uniformity and B_1 homogeneity. The fiducial array was modeled on the ADNI phantom (4), which has successfully been used to correct geometric distortions and image nonuniformity. The fiducial array spheres all have the same fill solutions with the same T_1 and T_2 properties (T_1 = 407 ms ± 6 ms, T_2 = 347 ms ± 6 ms at 20.0 C, 3.0 T), which means they can be used to assess image homogeneity and calculate B_1 variation.

The fiducial array consists of a set of 57 precision-machined polyvinyl chloride (PVC) 10.0 mm ± 0.1 mm inner-diameter spheres on a 3D cubic lattice, with a lattice spacing of 40 mm ± 0.1 mm. Figure 2 shows a coronal MRI with a subset of the fiducial spheres along with 10 mm diameter circular regions of interest (ROIs) at their prescribed positions, before automated sphere location. Figure 2 shows standard output from the Python-based analysis package, after sphere location, plotting image uniformity and geometric distortion along the *x*,*y*,*z* axes, respectively. These measurements will be discussed in the Results section. The fiducial spheres are numbered 1 through 57, starting at plate 1, increasing from right to left, then inferior to superior. For data presentation purposes, the 57 fiducial array elements are divided into 7 groups: 27 internal spheres on central 3x3x3 grid, 5 spheres at each of the 6 outward faces of the phantom. Each set of fiducials is represented by a different color in the plots in Figure 2.

The spheres are filled with a $CuSO_4$ solution made by mixing 0.802 g $CuSO_4$* $5H_2O$ per liter of deionized water. The Cu^{+2} ion concentration was measured, using an inductively coupled plasma-optical emission spectroscopy and a NIST traceable Cu standard (such as NIST standard reference material 3114) to be 0.18901 ± 0.00055 mg/g or 2.969 ± 0.008 mM. A small amount of blue dye (erioglaucine disodium salt, 0.12 µg/ml) was added to assist in the visual recognition of the fiducial elements. The fiducial spheres are machined in two parts from PVC stock, glued together and attached to the plates using an M8-1.25 threaded stub, as shown in Supporting Information Figure S1C. PVC was chosen because it is inexpensive, can be easily glued, and has a low permeability to water. The spheres are sealed with a tapered polyvinylidene difluoride (PVDF) plug. The fiducial solution was chosen to have short T_1 times to give, for most scans in the protocol, high signal and contrast with respect to the surrounding fill composed of long- T_1 deionized water. High contrast is beneficial for automated location of the fiducial sphere positions.

MR Parameter Arrays ($NiCl_2$, $MnCl_2$, and proton density arrays)

The MR parameter array spheres are 20 mm outer diameter (OD), 15 mm internal diameter (ID) polypropylene spheres that are filled, heat sealed and then epoxy sealed to an M8-1.25 threaded mounting stud (shown in the Supporting Information Figure S1D), similar to those used for the fiducial spheres. Since the spheres are epoxied onto a hemispherical indent, as opposed to being integrally machined with the threaded stud, there is more positional uncertainty in the MR parameter arrays than in the fiducial array. Polypropylene spheres were chosen since they are inexpensive, commercially available, and have one of the lowest water/air permeation rates among all of the plastics tested. However, the OD, ID, and sphericity of the polypropylene spheres are not well controlled: see the micro-computed tomographic image in the Supplemental Information Figure S2D inset. The polypropylene spheres are made by thermally welding together two molded hemispheres, which results in a thicker wall in the weld plane.

The T_1 and T_2 values were modified by doping high-purity water with $NiCl_2$ and $MnCl_2$ in their respective arrays. The proton density values were modified by making solutions of 5 % to 100 % high-purity water with the balance D_2O. Proton density is defined as the concentration of MR-visible protons in a sample/tissue, relative to that in the same volume of water at the same temperature (6). All solutions use American Chemical Society reagent grade, or higher quality,

water. A small quantity of $NiCl_2$ (0.26 mg Ni^{2+}/ml) was added to the proton density solutions to reduce the relaxation times. Green dye (28 µg/mL erioglaucine disodium salt, 28 µg/mL tartrazine) was added to the $NiCl_2$-doped array solutions, red dye (49.6 µg/mL allure red AC) was added to the $MnCl_2$-doped array solutions, and yellow dye (53 µg/mL tartrazine) was added to the proton density solutions to make the arrays easily identifiable by the user. Each array consists of 14 spheres, 10 equally-spaced on a 50 mm radius and 4 spheres internal to the outside ten spheres on a 40 mm grid. The sphere positions, $NiCl_2$, $MnCl_2$, D_2O concentrations, and measured T_1, T_2 values (at 20 C, 1.5 T and 3.0 T) for the $NiCl_2$, $MnCl_2$, and proton density arrays are listed in the Supplemental Information in Tables S1, S2, S3, respectively.

The $NiCl_2$ and $MnCl_2$ arrays were chosen to have a wide range of values to mimic both endogenous tissues and tissue containing standard commercial contrast agents. T_1 and T_2 relaxation times for the arrays, along with typical values for tissue, are shown in Figure 3. The $MnCl_2$-doped array spans a range of relaxation time values much closer to that of tissue; however, it has considerably more temperature and field dependence than the $NiCl_2$-doped array. T_1 and T_2 can be measured on both the $NiCl_2$-doped and $MnCl_2$-doped arrays, although historically, the $NiCl_2$ array was designed to cover a suitable range of T_1 values, while the $MnCl_2$ array was chosen to cover a suitable range of T_2 values.

The $NiCl_2$ doping concentrations were chosen to give a $\sqrt{2}$ progression of T_1 values from approximately 20 ms to 2000 ms at a B_0 field of 1.5 T. $NiCl_2$ was chosen because the relaxivity of paramagnetic spin-1 Ni^{2+} ions on water proton spins has been extensively studied (7,8). T_1 values are relatively insensitive to the magnetic field strength (9) and to temperature over the range of application for this phantom (7). However, the paramagnetic relaxation enhancement due to Ni^{2+} has a complex behavior in the region of application with non-monatomic variation of the relaxivities with field and temperature (9). T_1 and T_2 for the higher concentration $NiCl_2$ solutions, Ni-5 to Ni-14, have a minimum near 22 C at 3 T (as shown in the Supporting Information), close to typical MRI bore temperatures. The presence of a minimum eliminates the large linear variation of T_1 and T_2 with temperature, giving a maximum variation of 4% over the range of expected bore temperatures (16 C to 26 C).

The $MnCl_2$ concentrations are varied from 0.013 mM to 1.704 mM and, as with the $NiCl_2$ array, the concentrations were spaced by factors of $\sqrt{2}$. The variation of T_1 and T_2 is a linear

function of temperature over the expected range of bore temperatures (16 C to 26 C), with a variation at 3 T of 2.7%/C and 1.6%/C for T_1 and T_2 respectively.

Resolution and slice profile insets

The resolution insets are modeled on the ACR guidance (10), and the slice profile wedges are designed in accordance with National Electrical Manufacturers Association (NEMA) MS 5-2018 (11). Photographs and MR images of resolution and slice profile insets are shown in Figure 4.

The phantom contains two resolution insets consisting of 4x4 arrays of holes. A coarse resolution inset is drilled directly into plate 4 with hole diameters of 0.6 mm, 0.7 mm, 0.8 mm, 0.9 mm, and 1.0 mm ± 0.1 mm diameters. The fine resolution inset is contained within a rectangular prism filled with $CuSO_4$ solution to give high contrast and has hole diameters of 0.4 mm, 0.5 mm, 0.6 mm, 0.7 mm, and 0.8 mm ± 0.1 mm diameters. Each hole array is duplicated at a 10° angle, sharing one common hole, to assess resolution along non-principal axes. In total, there are 155 holes in each of the resolution insets (5 hole sizes, 31 per size). The advantage of this resolution inset is that it is easy to interpret visually, compact, easy to manufacture, and there is considerable community experience with it.

The slice profile inset consists of two 10° angle wedges contained within a rectangular prism filled with $CuSO_4$ solution. The wedges are oriented so that one has a positive and one has a negative inclination. The presence of a pair of wedges allows for the slice profile to be corrected for the misalignment of the slice plane relative to the base plane of the wedges.

Safety

For most of the imaging studies, the shell was filled with 4 L of DI water. This was done for three reasons: (1) safety, if the phantom spilled, the outer fill has no toxicity or disposal issues; (2) practicality, the phantom could be easily refilled at any site; and (3) the low RF absorption to insure minimal sample-induced B_1 distortions. The main toxicity issue is due to the $NiCl_2$ solutions. A total of 52 mg Ni^{2+} is included in the phantom. The $NiCl_2$ is contained in the thick-walled polypropylene contrast spheres that are double sealed with both a weld and epoxy. These spheres are very robust and can take large impacts without breakage or leakage.

SI Traceability

A calibration object requires its properties to be precisely measured and traceable to a common set of definitions and units. The measured values must include a well-defined uncertainty. A protocol for traceable spin relaxation measurements for MRI phantoms can be found in NIST SP250-97 (12). Spin relaxation time measurements are SI traceable via a metrology NMR system with a calibrated time base. The NMR time base, a 10 MHz oven-controlled crystal oscillator, is calibrated against the NIST hydrogen maser clock. The metrology NMR system has a maximum of 0.5 ppm time base error. Time base errors, which include jitter of transmit and receive events, give minor contributions to the overall uncertainty in relaxation time measurements.

More important factors that determine the uncertainty in relaxation time measurements are RF power calibration, RF inhomogeneity, B_0 inhomogeneity, and measurement temperature. RF power is calibrated using a nutation experiment on the sample to be calibrated, where the spins are progressively tipped, by varying the RF pulse duration, at increasing angles through at least 2 cycles (720°). Free induction decays are recorded, Fourier transformed to obtain a spectral peak, integrated and plotted against RF pulse duration. The data, which approximate a damped sinusoid, are fit to extract the pulse duration required for the desired tip angle and to extract an RF field inhomogeneity in the system. The RF field inhomogeneity is minimized during calibration by precisely centering a 10 mm long test sample within a 14 mm long RF homogenous zone. This field inhomogeneity and uncertainty in the RF power calibration are fed into Monte Carlo Bloch simulations, along with all other uncertainties, to calculate the uncertainties in the measured relaxation times.

B_0 inhomogeneity in the NMR metrology system is minimized by performing an automated shimming procedure on the sample to be calibrated. The inhomogeneous line width (the full-width-half-maximum of the Fourier transform of the free induction decay) is required to be no more than 10 Hz larger than the homogenous line width. Monte Carlo Bloch simulations (12) indicate that this level of B_0 inhomogeneity gives a negligible contribution to the overall uncertainty in relaxation times with the pulse sequences used (inversion recovery sequence with composite 180° pulses for T_1 and Carr-Purcell-Meiboom-Gill (CPMG) sequence with 2.0 ms refocusing times for T_2).

Temperature measurement and control are a major source of uncertainty both in the primary calibrations and in phantom-based MRI scanner assessment. Phantom components can

have several percent change in relaxation time per °C, as shown in the Supporting Information. For primary calibration in the NMR metrology system, sample capillaries are immersed in a fluorinated heat transfer solution within a 5 mm diameter NMR tube. The fluorinated solution also serves as a susceptibility matching solution. A fiber optic thermometer is positioned 15 mm away from the sample and is calibrated in the same NMR tube in a water bath against two NIST-calibrated platinum resistance thermometers.

Geometric accuracy is determined at manufacture time using computer controlled machining techniques. Post-assembly geometric accuracy can be checked with computed tomography techniques using SI traceable geometric calibrations, although this remains to be done. Here, the geometric distortion is reported relative to the 3D model values, with uncertainty given by the manufacturing tolerances.

Composition of the MR parameter arrays is verified using inductively coupled plasma-mass or optical spectroscopy techniques calibrated with traceable NIST reference standards. A reference library of all solutions, flame sealed in capillaries, is kept to establish stability over time. A test of the stability of proton relaxation times of an archived sample of Ni-12 solution at 3 T and 20.0 C, for example, gave T_1 = 44.53 ms ± 0.01 ms, T_2 = 31.86 ms ± 0.03 ms in May 2015 and T_1 = 44.67 ms ± 0.003 ms, T_2 = 31.97 ms ± 0.026 ms in June 2019, a difference of 0.3% for T_1 and 0.3% for T_2. Here, the reported errors are the standard deviation of 3 consecutive measurements at each time point and is a measure of system stability and is not the uncertainty in the values, which must incorporate all systematic errors as described in NIST SP250-97. Uncertainties will depend on the material under test, the precise pulse sequence used, the geometry of the sample, and on the values of the relaxation times. The uncertainty for proton spin relaxation time calibrations, defined as the range within which there is greater than 96 % probability of the true value, are given in the NIST calibration certificates for each material and are on the order of 1.5 %.

RESULTS

Geometric Distortion, Image Uniformity, and B_1 Homogeneity

Geometric distortion is assessed in a manner similar to that described in Ref. (4). A 3D gradient echo image, with short echo time, of the whole phantom is obtained with isotropic voxels with typical voxel dimension of 1.0 mm. This sequence can be used to assess the intrinsic geometric

accuracy of the scanner, whereas sequences with longer echo times will be more sensitive to distortions due to the patient/phantom. The apparent locations of the 10.0 mm fiducial spheres are determined by cross-correlation maximums between a fiducial sphere mask and the image (discussed in the Supporting Information). We compare this approach with alternative methods for determining apparent sphere positions in the Supporting Information. The sphere center can be determined to < 0.1 mm, which is comparable to the accuracy in the construction. Determination of the center of the fiducial spheres is sensitive to defects, such as small air bubbles in the solution. These defects can be detected manually or numerically, and the defective sphere data discarded. After locating the apparent sphere centers, a similarity transformation (an affine transformation consisting of a translation, rotation, and isotropic scale factor) which minimizes the distances between the prescribed and apparent locations is determined by a generalized Procrustes algorithm. The resulting rotation and translation define the map from phantom to scanner coordinates. The scale factor provides a measure of overall dimensional rescaling introduced by the scanner. We define geometric distortion as the difference between the apparent position $\vec{R}_a$ and the prescribed position $\vec{R}_p$ of the sphere centers after the similarity transformation has been applied to the later, $\delta\vec{R} = \vec{R}_a - \vec{R}_p$. An example geometric distortion analysis for an image without corrections for gradient non-uniformity, is shown in Figure 2. A large (~2 mm) *x*-component of geometric distortion is seen on the right face, with a negative sign, in Figure 2 since the apparent position is farther along the negative *x*-axis than the prescribed position. Similarly, there is a large *x*-component of geometric distortion on the left face, with a positive sign, since the apparent position is farther along the positive x-axis. Examples of geometric distortion with non-uniform gradient corrections applied are given in the Supporting Information and show the need for the phantom geometric accuracy to be < 0.1 mm. Local volumetric distortion can be assessed by comparing cross correlation volumes.

We assess the uniformity of image intensity as part of this same analyses by calculating the integrated intensity of the 57 fiducial spheres. Figure 2 shows image uniformity obtained from a head coil, which shows signal loss in the inferior or chin region due to low receive coil sensitivity. The image uniformity analysis can also be done as part of other protocols, such as the of T_1, T_2 protocols.

Finally, B_1 mapping can be done with the system phantom fiducial array using flip angle mapping techniques (13-15). The fiducial spheres, as well as the surrounding background fill

liquid (DI water), provide regions with identical properties to allow identification of RF non-uniformities. The phantom, in its default configuration with a background solution of DI water, will not substantially contribute to RF absorption. To assess the effect of dielectric resonances, the fill solution can be swapped out with a low dielectric constant fluorocarbon fluid. We do not perform this analysis here.

T_1 Measurement

Figure 5 shows automated T_1 analysis of the $NiCl_2$ array using the Python-based analysis package. Figure 5A shows an image of the $NiCl_2$ array using the inversion-recovery (IR) sequence protocol described in the Supporting Information with recommended 10.0 mm circular ROIs centered on the $NiCl_2$ array spheres. The suggested T_1-IR protocol requires 10 inversion times, TI, ranging from 50 ms to 3000 ms and long TR = 4500 ms. Note, the actual inversion recovery protocol may vary slightly among the different vendors due to details of the scanner design and programmable control. Figure 5B shows a typical fit to the data using a standard IR model (16-18), appropriate when TR >> T_1, and the inversion and detection flip angles are precisely 180° and 90°. The average signal in each ROI is given by $S(TI) = \mathrm{A}\left|1 - (1+\delta)e^{-\frac{TI}{T_1}}\right|$, where A is the signal amplitude when the magnetization is allowed to fully relax back to its equilibrium state, and δ is the inversion efficiency (16,18,19), the signal amplitude immediately following inversion normalized to A. Here, only magnitude data is used since the phase data is often not available and can be hard to interpret. This is in contrast with the NMR IR reference protocol, which uses the real part of the complex signal for analysis. Figure 5C shows the results of the T_1 analysis along with the deviation relative to the NMR reference values given in Table 1. The deviation is defined as $100(T_{1m} - T_{1r})/T_{1r}$, where T_{1m}, T_{1r} are the measured and reference values, respectively. The large deviation for sphere Ni-14, which has the shortest T_1, results from having only 3 data points where the signal is changing before full recovery. This is reflected in the large error bar, which is the standard error in the non-linear least squares fit. The analysis was done both by 1) fitting the average of the ROI signal to the model and 2) fitting each voxel independently to the model, making a T_1 map, and then averaging. Both methods agree within the standard errors.

More information is obtained from these fits than simply the relaxation time. The inversion efficiency, δ, is an important indicator of B_1 and flip angle uniformity. As seen in Figure 5D, the inversion efficiency is close to 0.85, whereas the ideal value is 1.0. Typical values obtained from NMR data on the same solutions, with a more homogenous RF field geometry and using composite inversion pulses, are between 0.97 and 1.0. A Monte Carlo study of Bloch simulations indicate that there is a correlation between the deviation of the inversion efficiency from its ideal value and the error in the T_1 measurements (12).

IR sequences, while fairly accurate in measuring T_1, require long acquisition times. Variable flip angle (VFA) sequences, with shorter TRs and smaller flip angles are considerably faster with less RF deposition. VFA protocols typically use 3D gradient echo sequences (see Supporting Information) and vary the flip angle α. Ideally, the signal is given by $S(\alpha) = S_{90}\sin(\alpha)\frac{1-e^{-\frac{TR}{T_1}}}{1-e^{-\frac{TR}{T_1}}\cos(\alpha)}$ (17,18). Figure 6A shows a coronal image of the $NiCl_2$ array using the VFA protocol described in the Supporting Information, with TR = 5.37 ms and TE = 1.49 ms. Figure 6B shows a series of sagittal images for a series of flip angles varying from 2° to 30°. The signal in each ROI has a maximum at the Ernst angle, $\cos(\alpha_E) = e^{-\frac{TR}{T_1}}$, which for ROI 5 is near 10° and ROI 2 is near 5°. Figure 6C shows the mean signal for ROIs 2, 5, 8, 11 along with fits to the model listed above and the residuals. As seen by the non-Gaussian distribution of the residuals, there are systematic deviations from the simple model listed above due, in large part, to the inability to achieve the prescribed flip angle at each location in the phantom. Figure 6D shows the deviations in the measured T_1 values from the NMR reference values for three consecutive 3 mm slices through the array. The slice-to-slice variation in the measured T_1 values is correlated to scanner-induced intensity modulations, which can be seen as the vertical slice-to-slice modulations seen in Figure 6B. The error bars in Figure 6C are the standard error in the fits, which are large for the longer T_1 ROIs. The large standard errors on the fits, along with a systematic variation in the residuals, indicate that the model, which assumes an accurately prescribed flip angle that is the same for all regions in the phantom, is limiting measurement accuracy. More advanced models and protocols exist to improve on VFA T_1 measurements(20); however, here we are focusing on using the system phantom and simple protocols to identify and characterize scanner non-idealities.

T_2 Measurement

Figure 7 shows T_2 measurements of the $MnCl_2$ array using the Python-based analysis package. Figure 7A shows an image of the $MnCl_2$ array using the spin echo (SE) sequence listed in the Supporting Information with TE varying from 10 ms to 320 ms by 10 ms and TR= 5000 ms. Figure 7B shows a typical fit to the data using a simple exponential model: $S(TE\ ) = S_0 e^{-TE/T_2}$, where $S(TE)$is average signal in each ROI and TE is the spin echo time. Figure 7C shows the T_2 obtained from the fits along with the deviation relative to the NMR reference values given in Table 2. The error bars in Figure 7C are the standard errors derived from the non-linear least squares fitting procedure. For the ROIs with short T_2 times, there are not many points used in the fit, and the standard error is large.

Proton Density and SNR Measurement

A coronal spin echo image of the PD array is shown in Figure 8A with the PD ROIs shown in yellow. Substantial background nonuniformity exists; the signal changes by a factor of ~6 going from the periphery to the center along the superior/inferior direction. Figure 8B shows the average signal in each PD ROI and the measured proton density, given by average signal normalized to the local background and the 100% water sphere, plotted versus prescribed proton density. The average signal in the PD ROIs shows a clear nonlinearity due to signal loss towards the periphery of the phantom. The normalized signal shows the expected linearity of the signal with respect to proton density. The maximum error in the measured proton density is 8.3 %, and the standard deviation of the errors is 2.8 %.

In addition to measuring proton density, this array is used to measure SNR using methods similar to those described in Ref. (18). The noise is measured by taking the difference of two identical scans, $\Delta S = S_2 - S_1$, taken sequentially (Figure 8C), as discussed in Method 1 in Ref. (21). The SNR is calculated for each PD ROI by taking the average signal $S_{av} = \frac{1}{2}(S_{2av} + S_{1av})$ and dividing by the noise $\mathrm{N} = SD/\sqrt{2}$, where SD is the standard deviation of the difference signal ΔS within the ROIs. The SNR is plotted versus proton density in Figure 8D. Also shown is the average signal in each ROI normalized to the standard deviation within each ROI. This is a measure of the ratio of the average signal to the nonuniformity within each region, only part of which comes from noise.

Two additional SNR ROIs are shown in Figure 8 in blue. SNR1 ROI is meant to more closely mimic the geometry suggested in Ref. (21). However, the larger ROIs are more susceptible to scanner and phantom temporal drift, which gives rise to the low spatial frequency structure in Figure 8C. SNR2 ROI is meant for Method 4 analysis discussed in Ref. (21) . This method was judged unsuitable since the background signal outside the phantom is anomalously low (less than 1 count) and does not reflect the noise in the image.

Resolution

Resolution is determined either using the manual method suggested by ACR (10) or an automated computer algorithm that compares the MR image to synthesized images with different resolution distortions. Figure 9A shows an image of the resolution inset taken with a voxel size of 0.35 mm and a slice thickness of 4.0 mm. As per ACR guidelines, visual inspection of the image yields a resolution of 0.5 mm. Figure 9B shows a synthetic image generated by analytically computing the Fourier transform of the resolution-inset disk array, calculating a finite k-space image, and converting into a real-space image. Comparison of the synthetic image to the measured image shows additional blurring beyond that expected from the point spread function due to finite k-space sampling. Applying an additional Gaussian point spread function can improve the match between the synthetic and real image (Figure 9), with the difference between the images minimized using a point spread function with a width $2\sigma = 0.075$ mm, as seen in Figure 9D. The resolution is then reported as a finite k-space sample limited resolution of 0.35 mm and a machine specific Gaussian broadening of 0.075 mm.

Slice profile measurement

Slice profile measurements, shown in Figure 10, are done in accordance with NEMA MS-5-2018 (11): the edge response function is calculated from the imaging data by placing a 4 mm x 50 mm rectangular ROI over each wedge, differentiating the data with respect to image position, and then multiplying the image position by tan(α) to determine the slice profile. $\alpha = 10° + \theta$ is the corrected angle, where $\theta = 1/2\sin^{-1}((w_2 - w_1)\sin(2 * 10°)/(w_2 + w_1))$ measures the angle of the image plane with respect to the base of the wedges (see Figure 10C inset), and w_1, w_2 are the projected thicknesses of the two wedges. The slice thickness, t_{sl}, is then determined by measuring the full width at half maximum of the corrected slice profile. The suggested

prescribed slice thicknesses are 3 mm and 5 mm, with the maximum being set by the wedge height of 10 mm.

The slice thickness may also be measured by fitting the edge response function with a piecewise linear function, extracting the projected slice thicknesses, and then multiplying by the tangent of the wedge angle: $t_{sl} = \frac{1}{2}(w_2 + w_1) \times tan(10°)$. Averaging both ramp orientations corrects, in lowest order, for errors due to the misorientation of the image slice plane relative to the base of the wedges. The latter method has the advantage that it is easily automated and has a well-defined uncertainty.

DISCUSSION

The system phantom described here, or modified versions of it, have been used for several studies including repeatability of MR fingerprinting (22), assessing changes occurring during scanner upgrades (23), and assessing variation in multi-site T_1 measurements (24). Many other uses are ongoing, including monitoring of scanners for quality assurance and homogenizing scan protocols across multi-vendor scanners. The phantom was designed to be used in several modes, starting from simple visual inspection for identification of geometric distortion and resolution, to semi-automated analysis to determine relaxation times and SNR, and finally for custom analysis to validate advanced quantitative protocols. While application specific phantoms may provide better structures for particular measurements, e.g. the ADNI phantom for geometric distortion or a uniform agar phantom for image homogeneity, the system phantom is meant for characterization of a large number of scanner parameters and to have sufficient accuracy, generality, and complexity to assess the accuracy of many types of advanced hardware and protocols. The standard protocols provided with the system phantom are not intended to be the best protocols for accurate measurement of geometric distortion, proton spin relaxation times, resolution, and SNR. Rather, they are meant as a common method to test the ideality of the scanner and to provide a comparison for other quantitative protocols that may be faster or more accurate.

Here we have shown how the phantom can be used for scanner diagnostics assessing geometric distortion, image homogeneity, resolution and slice profile. The resulting data can provide both parameters to benchmark scanner performance and insight on what is determining these parameters. For example, resolution analysis can separate the pulse-sequence component

from the intrinsic scanner component of the point spread function. The SNR protocol can separate temporal and spatial noise components, which contribute in different ways to SNR. Finally, the MR-parameter arrays provide a very broad range of properties to allow validation of both standards and advanced quantitative protocols.

One specific challenge we encountered is that T_2 is not a rigorously defined quantity (25). $1/T_2$ is the component of the total transverse spin dephasing rate, $1/T_2^*$, that is considered to be intrinsic to the sample, with the extrinsic component, $1/T_2'$, being eliminated by using a spin echo: $\frac{1}{T_2^*} = \frac{1}{T_2'} + \frac{1}{T_2}$. There is ambiguity on what dephasing components are considered extrinsic and intrinsic, and the components being eliminated by the refocusing pulses will depend on the details of the refocusing pulses, such as refocusing times and phases of the RF pulses. Hence, T_2 measured by a CPMG sequence is different than that measured by a simple SE sequence. CPMG sequences, with shorter refocusing times, will rephase the spin packets more effectively than a SE sequence, which use longer TEs. Given ideal MRI/NMR systems, measured T_2 for CPMG sequences should always be longer than T_2 for SE sequences. CPMG sequences also effectively null out the longitudinal magnetization, which if present along with B_0 and B_1 inhomogeneities, can lead to stimulated echoes. Stimulated echoes, which can be present in SE sequences, add long-time signal and can increase the measured T_2. This may cause the measured T_2 for SE sequences to be longer than measured T_2 for CPMG reference sequences. The variability in T_2 measurements have been extensively studied in the literature (26-28). Here, we note that simple SE sequences, which may appear to be similar, can have very different spin manipulations and signal production. Using a standard phantom can identify these differences, which can be important for both conventional qualitative interpretation as well as quantitative analysis.

From the feedback of users of the prototype phantoms, several limitations and potential improvements have been identified. During the time since this phantom was conceived, head coil geometries are becoming more complex, and the size is decreasing. Newer head coils have minimum dimension less than 195 mm, which is smaller than the current system phantom design. There is a trade-off in the system phantom between sampling large volumes to assess uniformity of measurements and the ability to fit into newer, smaller coils. A smaller phantom shell has been investigated, and a smaller version of the system phantom will be needed for certain coil sets. Another concern is the cost of the phantom. Tradeoffs can be made, such as

relaxing the 0.1 mm geometric accuracy, to reduce costs to make the phantoms commercially viable.

The contrast array components are arranged in an orderly fashion to allow easier interpretation by users. More information may potentially be obtained by random placement of the contrast elements, given that their properties are highly determined and errors due to location may be better extracted using computer algorithms. There could be benefit as well in reducing the number of unique spheres in the contrast array to provide replicate spheres. The prototype phantoms presented here are easily reconfigurable to test these modifications.

The contrast and fiducial arrays use plastic components that have sharp signal/-to-no-signal interfaces that cause Gibb's ringing artifacts, as can be seen in Figure 6A. These artifacts can confound SNR and other measurements. While these artifacts can be accounted for in certain measurements, for instance they can be included in the kernel for the cross correlations determining geometric distortion, it would be desirable to minimize these artifacts by making the plastic shells thinner or by allowing them to have some signal.

Dielectric resonances can give rise to image distortion, particularly for large water phantoms at high fields(29). These effects still need to be studied for the system phantom. However, the outer fill is easily changeable, without affecting the measurement protocols. Fills with different conductivities and dielectric properties can easily be explored.

Temperature control and monitoring remain a critical issue, especially when trying to verify the accuracy of relaxation time measurements to better than 2 %. While the system phantom is shipped with a thermometer to record temperature, the bore temperatures are often different from the ambient scanner room temperature or even the control room, where the temperature is typically measured. There is no guarantee that the phantom temperature is the same as the recorded temperature. Further, the temperature information is not easily encoded in the DICOM files. This establishes the need for in-situ NMR thermometers, and an MRI-readable liquid crystal thermometer has been developed for MRI phantoms (30,31). Further, the system phantom would greatly benefit from the use of materials that are less sensitive to variations in temperature and magnetic field strength. Most paramagnetic salts such as $MnCl_2$ and $CuSO_4$ have considerable temperature (7) and field dependence (see Figure 3) in their relaxivity values r_1, r_2 ($mM^{-1}s^{-1}$). $NiCl_2$, while having less temperature and field dependence, has safety/toxicity issues.

The materials used in the system phantom were chosen for their stability and simple properties, e.g. exhibiting mono-exponential spin relation. As shown in the Supporting Information, a reference library sample, sealed in a glass capillary, has 4-year T_1 and T_2 stability better than 0.1%. Assessing stability of the phantom components is more challenging, since sampling the solutions for SI-traceable NMR measurements would be destructive. The plastic spheres in the phantoms have some permeability to water and may slowly absorb paramagnetic salts, leading to long-term drift in T_1 and T_2. A sampling protocol of the MR-property spheres, stored under conditions similar to those contained within the phantom, is required to assess long term stability and is in progress.

There is demand for more biomimetic materials that match multiple tissue properties, such as the apparent water diffusion coefficient as well as proton density and proton spin relaxation times. It is challenging to obtain complex biomimetic materials that are stable and also have properties that can be rigorously quantified. However, this work is underway, as illustrated by the stable fat mimic developed for the commercial version of the breast phantom developed by NIST and the University of California San Francisco (32).

While the current implementation of the MRI system phantom can be improved and will never be suitable for all calibration needs, it is a first attempt at developing and implementing a rigorous calibration structure with traceable properties, community-driven open construction details and evaluation protocols, and long-term monitoring and assessment with state-of-the-art metrology tools.

ACKNOWLEDGEMENTS

We thank Michael Snow and William Hollander at High Precision Devices for the computer aided design and manufacturing guidance. We thank Elizabeth Mirowski from Verellium and Joshua Levy from Phantom Labs for useful discussions on phantom design and implementation. We thank all of the ISMRM SQMR committee members, past and present, for their advice and support.

FIGURES

FIGURE 1 A, Schematic of system phantom showing fiducial array (grey/brown), $NiCl_2$ array (green), $MnCl_2$ array (red), proton density array (yellow), and resolution and slice profile insets. Plate 5 contains MR visible engravings with the serial number and phantom type. All plates have orientational marks and plate 3 contains a grid for manual estimation of geometric distortion. B, Top view of phantom showing eye decals. C, Sagittal slice in the Y-Z plane showing the MR parameter arrays, fiducial elements, phantom origin, and the resolution inset.

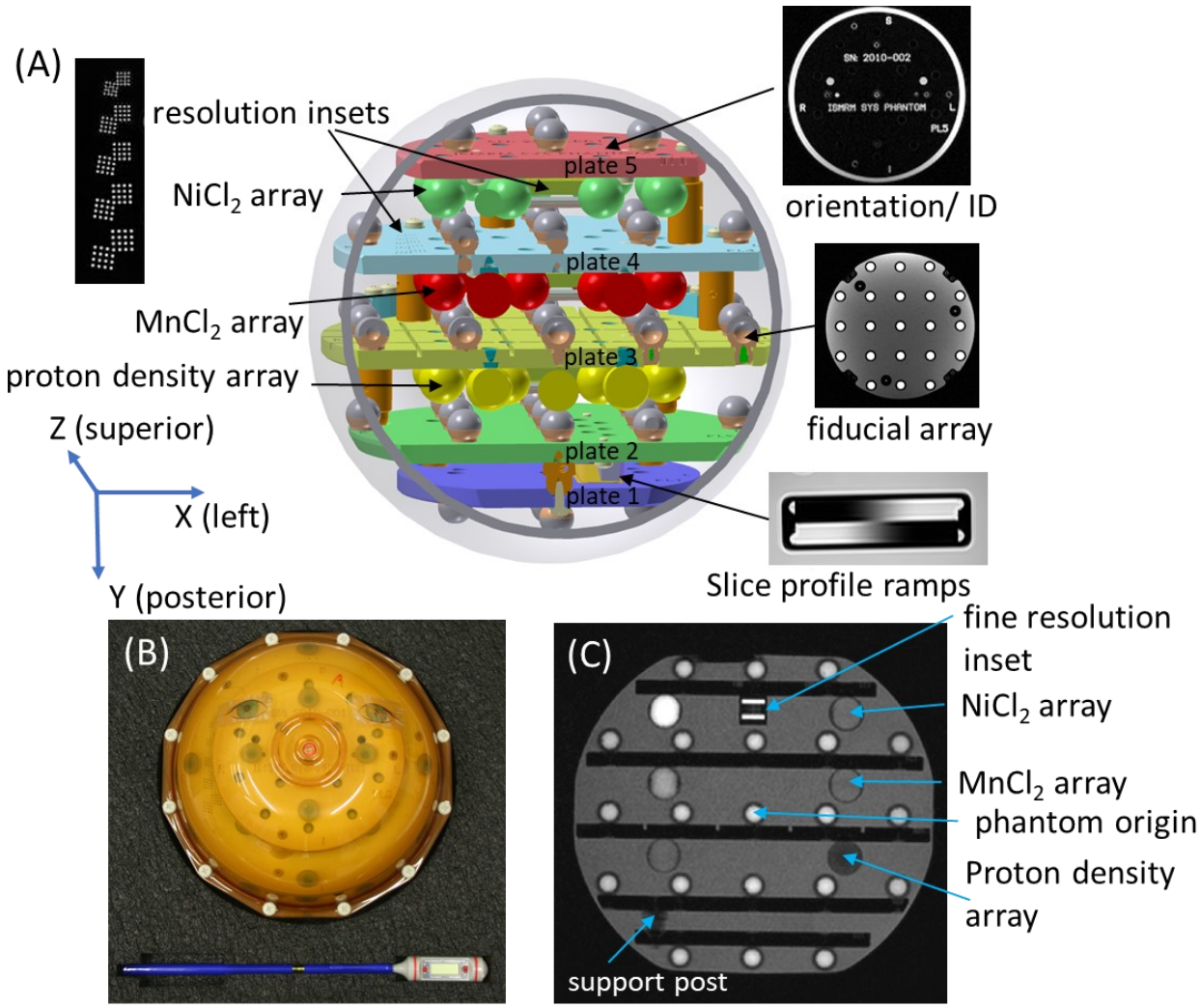


FIGURE 2 A, Coronal 3D gradient echo image of the fiducial array along with 10.0 mm circular ROIs located on prescribed 40 mm grid, which allows for visual identification of geometric distortion. S, I, L, R refer to superior, inferior, left, right directions, respectively. The scan is at 3 T using a body transmit coil and a head receive coil with 0.97 mm isotropic voxels, TR = 6.3 ms, TE =1.89 ms. The nonuniform gradient corrections are turned off. B, Normalized integrated intensity of all 57 spheres obtained after the sphere location procedure. C,D,E, show geometric distortion along the x, y, z directions respectively. Geometric distortion, $\delta\vec{R} = \vec{R}_a - \vec{R}_p$, is defined as the distance between the apparent position $\vec{R}_a$ and the prescribed position $\vec{R}_p$, of the sphere centers after an affine transformation has accounted for rotation, translation, and overall scale correction.

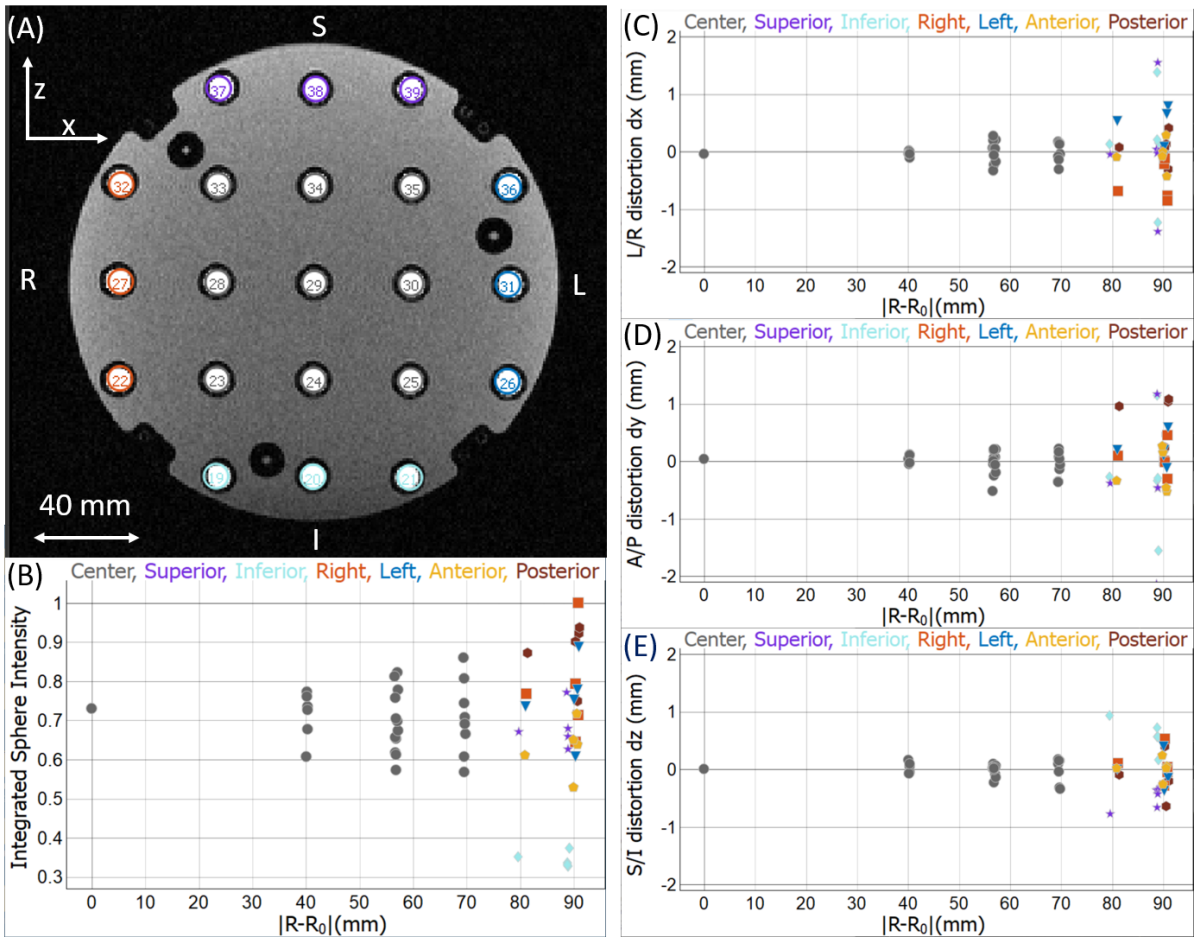


FIGURE 3 Overview of T_1 and T_2 properties of MR parameter arrays at 20.0 °C. 1.5 T data are plotted as circles, 3.0 T data as squares. Also plotted in orange, for reference, are values for various human tissues at 1.5 T and 3.0 T at 37 °C, taken from the online IT'IS data base(28).

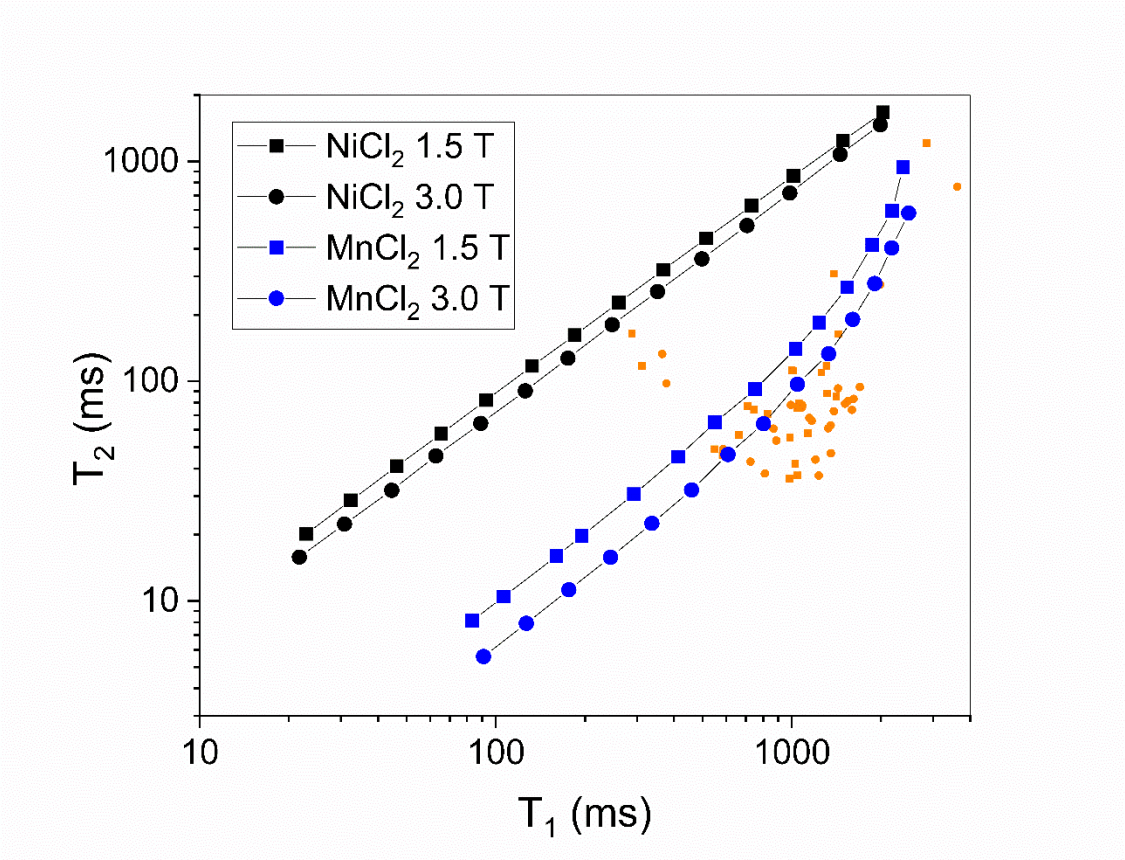


FIGURE 4 A,B Photo and MRI of resolution inset. C,D Photo and MRI of slice profile inset.

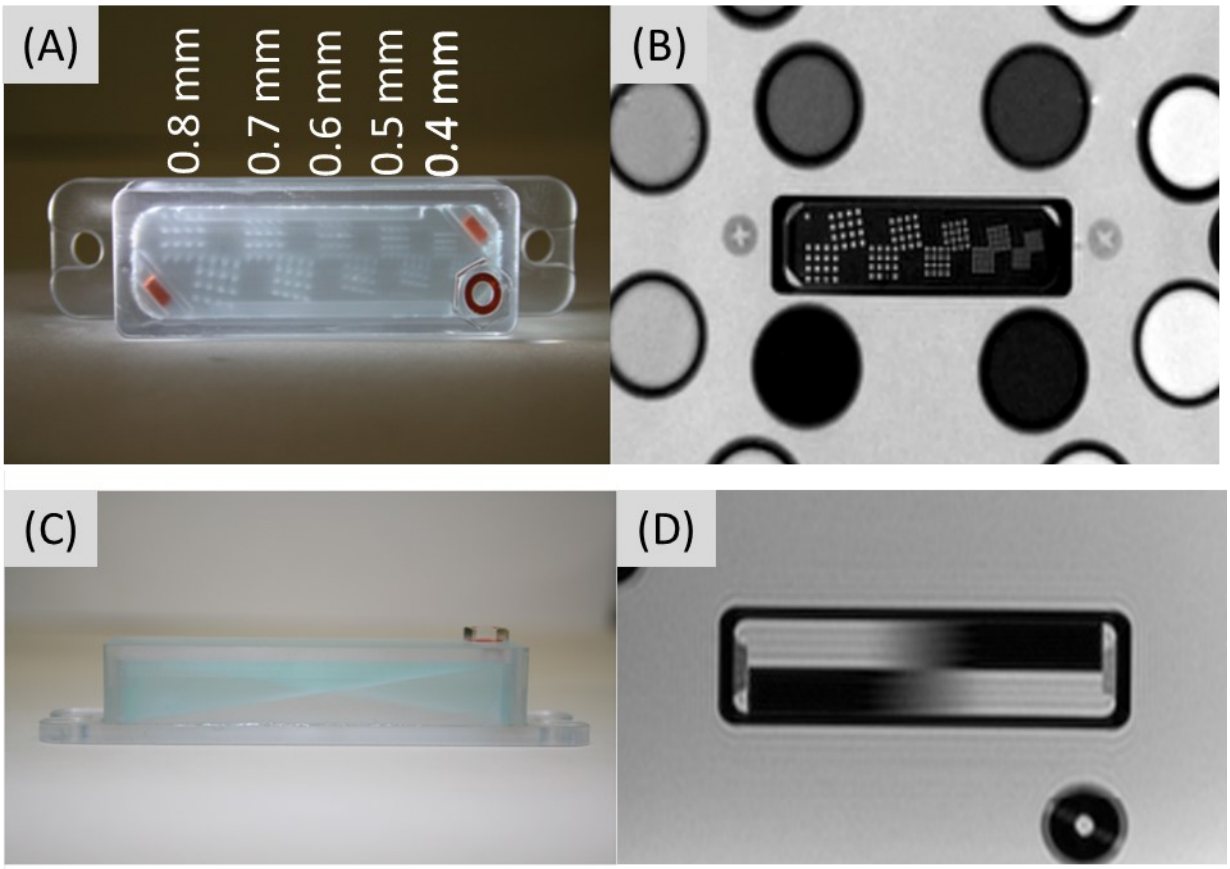


FIGURE 5 A, Coronal image of the $NiCl_2$ array at 3 T ~20 °C using an IR protocol. B, average signal in ROIs 1,5,9,13 in the $NiCl_2$ array, plotted as a function of inversion time, TI, and fits to inversion recovery model. C, automated analysis of T_1 for each of the 14 $NiCl_2$ array elements, with the deviation from the NMR reference measurements. The analysis was done both by 1) fitting the average of the ROI signal to the model and 2) fitting each voxel independently to the model, making a T_1 map, and then averaging. The deviations are shown in blue and red squares, respectively. The error bars on the deviations are the standard errors in the nonlinear least squares fits and represent the quality of the model and the fit. D, The inversion efficiency, δ determined from the fits. The line at δ = 1 represents perfect inversion, and the error bars are the standard errors in the parameter determined from the nonlinear least squares fits. Unobserved error bars are contained within the symbol.

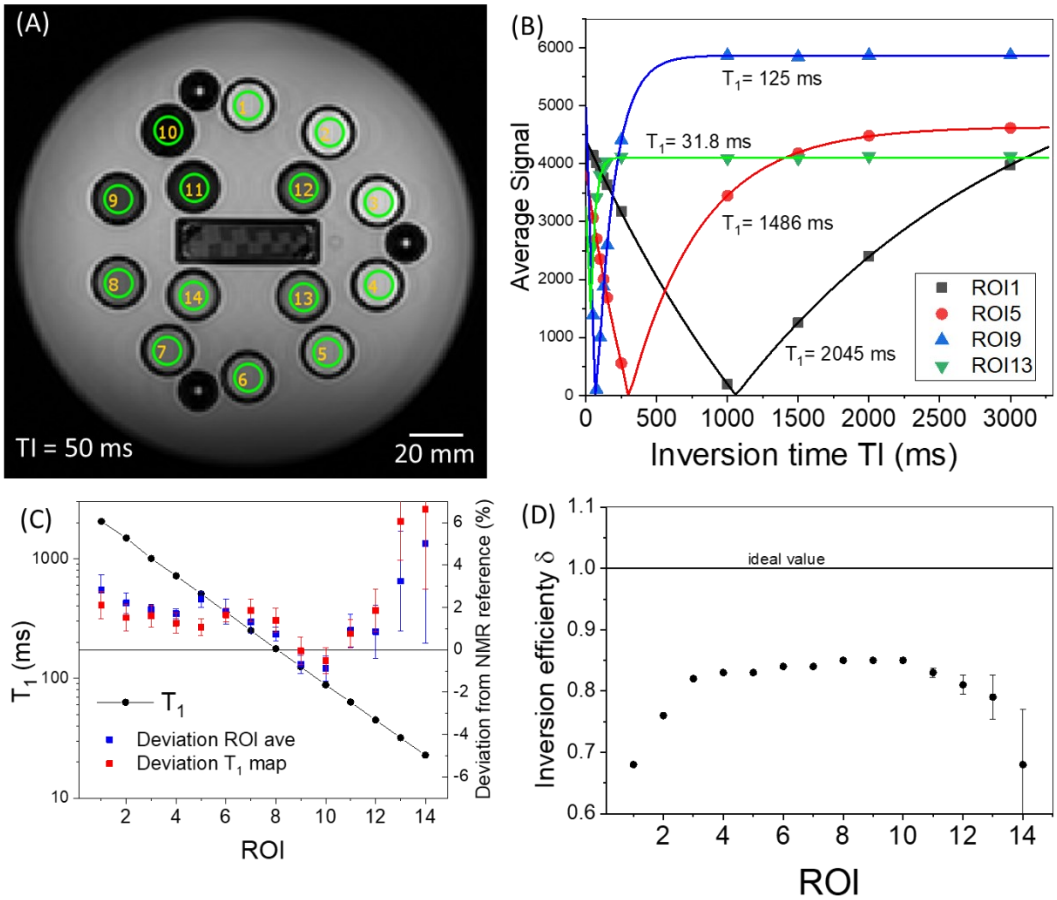


FIGURE 6 A, Coronal image of the $NiCl_2$ array at 3 T ~20 °C using a 3D gradient echo VFA sequence, B, Average signal and fit to VFA model for ROIs 2,5,8,11. C, sagittal images showing sphere 2 and sphere 5 for different flip angles. ROI 5 has a maximum signal near 10°, while ROI 2 has a maximum signal near 5°. D, Deviation of measured T_1 for each of the 14 ROIs from the NMR reference measurements.

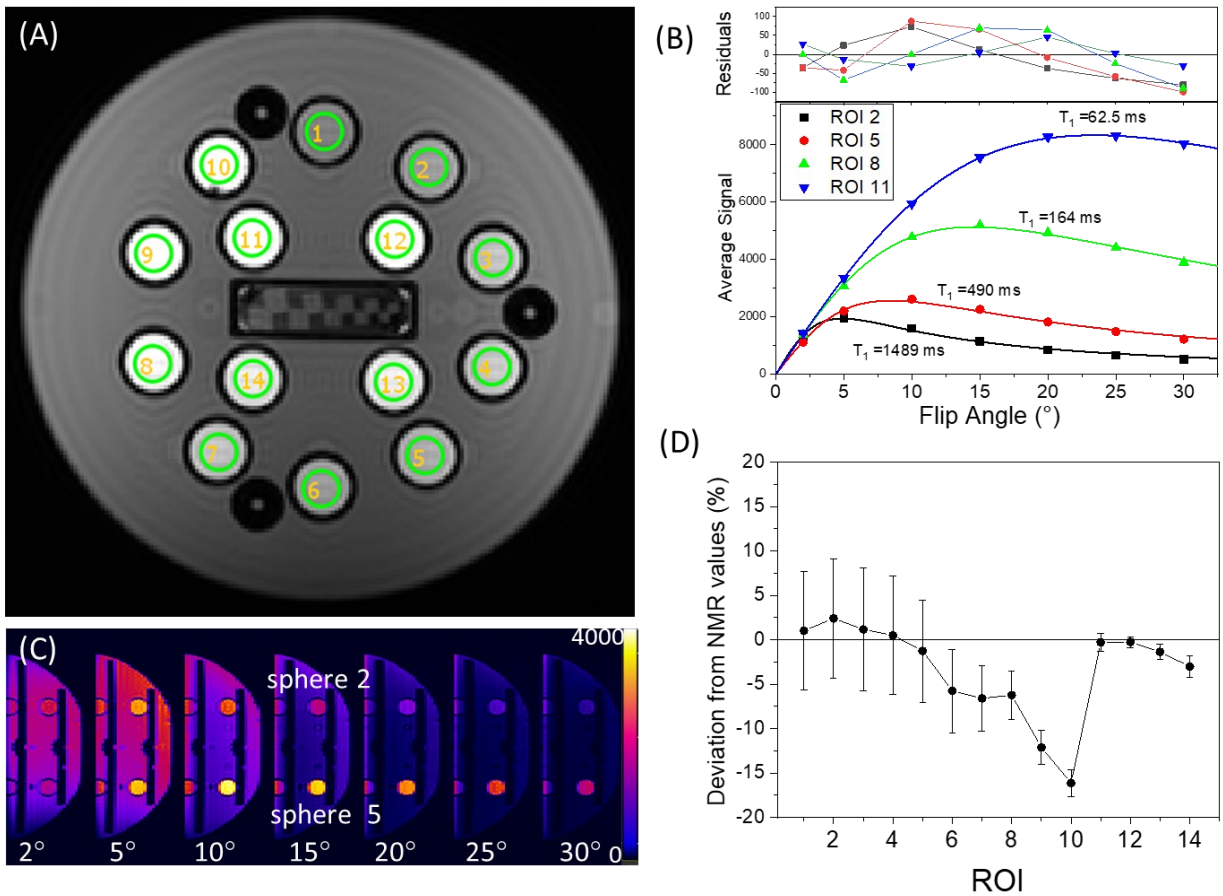


FIGURE 7 A, Coronal image of the $MnCl_2$ array at 3 T ~20 °C using a spin-echo sequence with TR = 5000 ms, B, typical fit to a spin echo protocol and C, automated analysis of T_2 for each of the 14 array elements, with the deviation from the NMR reference measurements. The TE = 10 ms image was not used due to anomalously low signal, and points below the noise floor, the orange line in B, were not included in the fits.

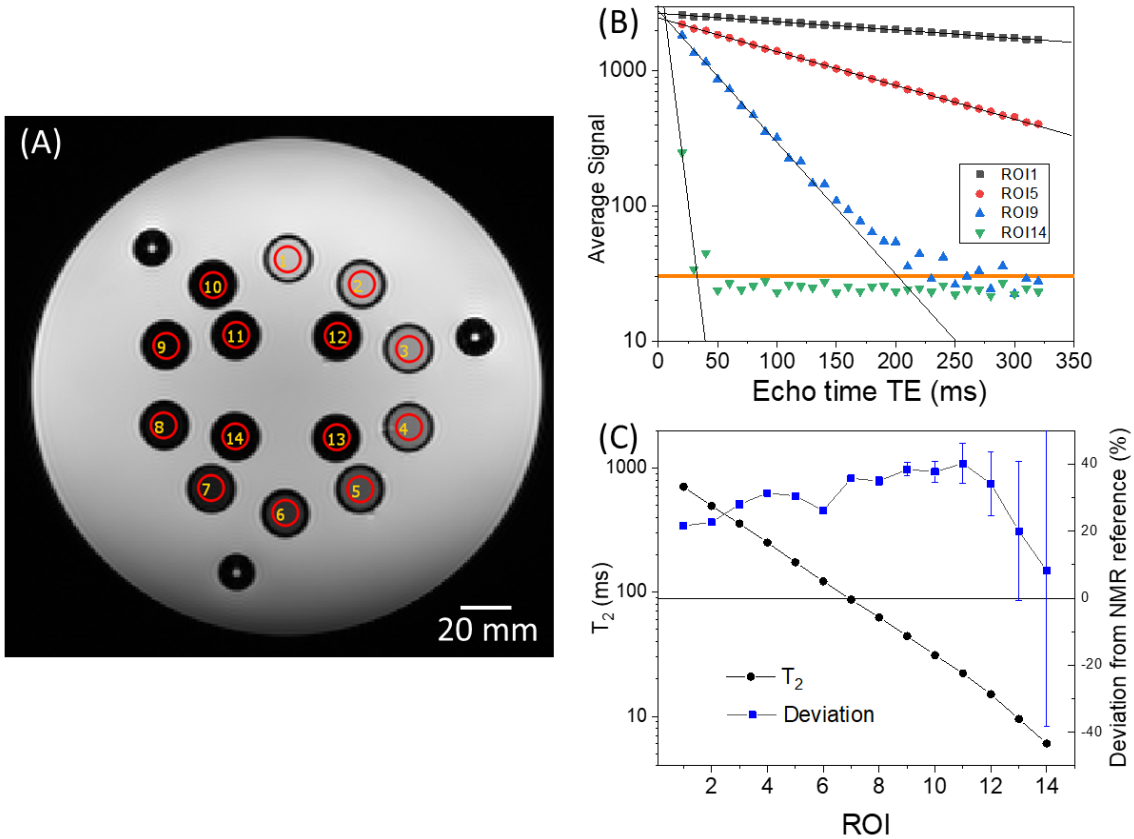


FIGURE 8 A, Coronal image of proton density array using a spin-echo sequence with TR = 5000 ms, TE = 10 ms. Two SNR ROIs, in addition to the 14 yellow proton density ROIs, are shown in blue. B, Average intensity of PD ROIs and measured PD as a function of prescribed proton density. C, Difference of two identical images and D, plot of the SNR as a function of proton density. Also shown in D, is the ratio of the signal to the standard deviation within each ROI.

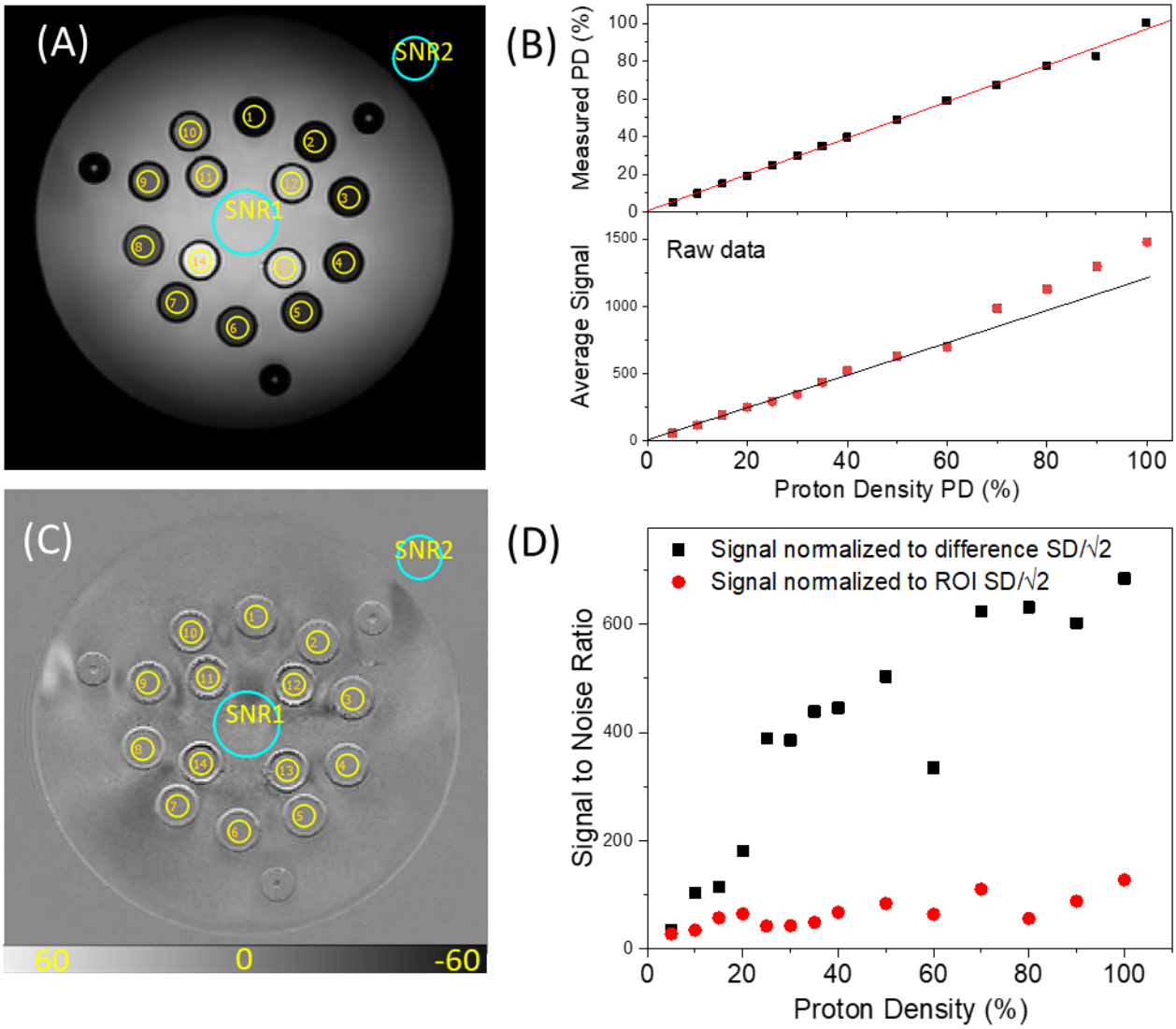


FIGURE 9 A, MRI of resolution inset taken with a voxel size of 0.35 mm and slice thickness of 4.0 mm. Here the resolution, according to ACR protocol is 0.5 mm since all the arrays with hole diameters ≥ 0.5 mm can be resolved, while the 0.4 mm hole size arrays cannot. B, Synthetic image of hole array with broadening only from finite k-space sampling. C, Synthetic image with additional blurring from a Gaussian point spread function of width $2\sigma = 0.075$ mm. D, The magnitude of the difference of the measured and synthetic image, using an L2 norm, as a function of the width of the point spread function, showing a minimum in the difference at a blurring of 0.075 mm.

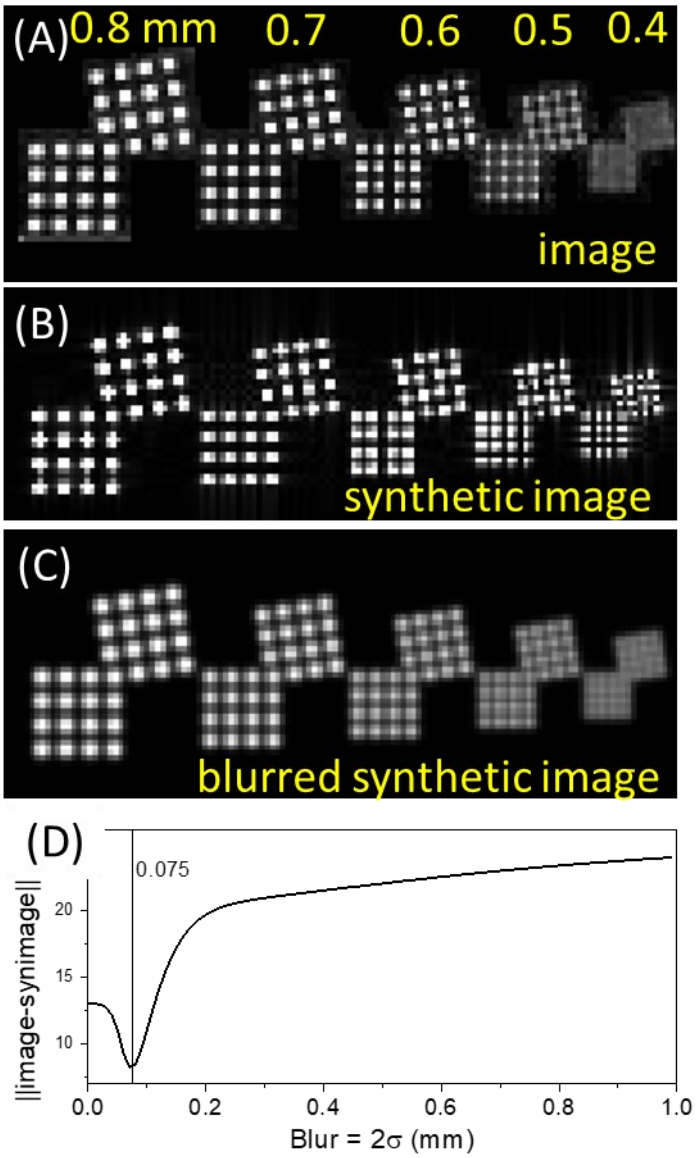
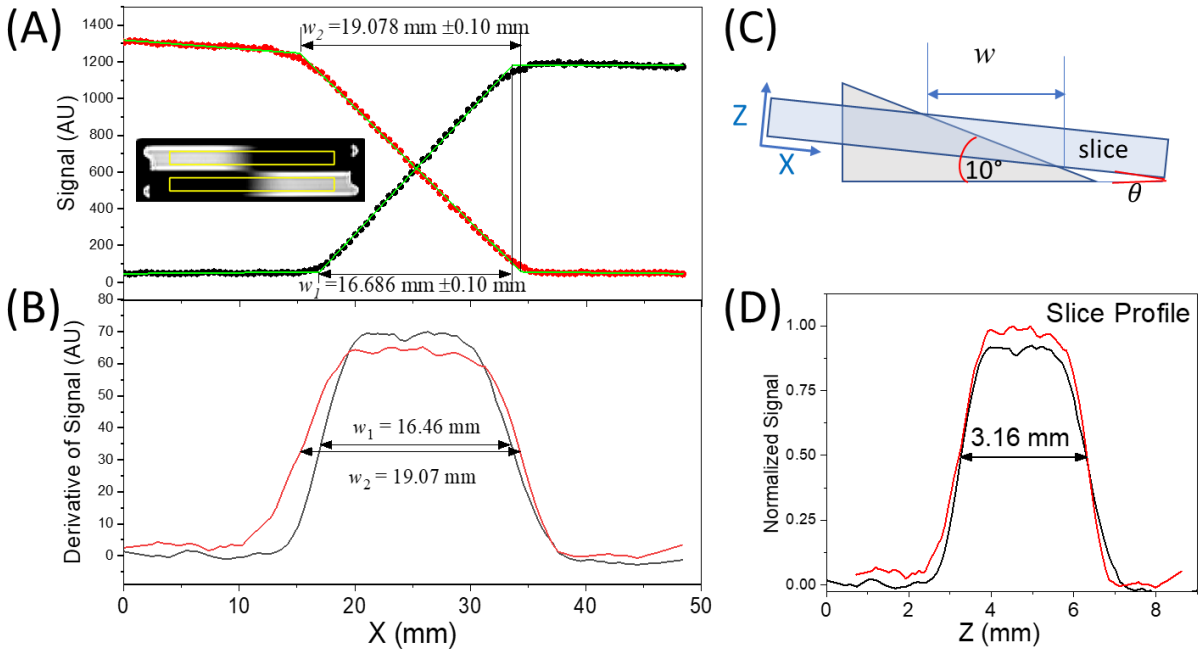


FIGURE 10 Slice profile analysis from an MR image with prescribed 3.0 mm slice thickness. A, B, The MR signal and its derivative, for rectangular ROIs centered on the two wedges (inset in A). C, D The measurement geometry and the slice profile, corrected for angular misalignment θ. The measured slice thickness is 3.15 mm using the NEMA MS 5-2018 protocol (11) and 3.16 mm ± 0.2 mm using the automated protocol described in the text. The misalignment of the slice profile inset and the scan plane is $\theta = 0.72°$.



REFERENCES

1. Clarke LP, Sriram RD, Schilling LB. Imaging as a Biomarker: Standards for Change Measurements in Therapy workshop summary. J Academic Radiology 2008;15 4:501-530.
2. Keenan Kathryn E, Ainslie M, Barker Alex J, Boss Michael A, Cecil Kim M, Charles C, Chenevert Thomas L, Clarke L, Evelhoch Jeffrey L, Finn P, Gembris D, Gunter Jeffrey L, Hill Derek LG, Jack Clifford R, Jackson Edward F, Liu G, Russek Stephen E, Sharma Samir D, Steckner M, Stupic Karl F, Trzasko Joshua D, Yuan C, Zheng J. Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom. Magnetic Resonance in Medicine 2017;79(1):48-61.
3. 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.
4. 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):2193-2205.
5. ADNI2. Alzheimer's Disease Neuroimaging Initiative 3T MRI Technical Procedures Manual. 2011.
6. Tofts Paul S. PD: Proton Density of Tissue Water. Quantitative MRI of the Brain 2004.
7. Kraft KA, Fatouros PP, Clarke GD, Kishore PRS. An MRI phantom material for quantitative relaxometry. Magnetic Resonance in Medicine 1987;5(6):555-562.
8. Kowalewski J, Maler L. Nuclear spin relaxation in liquids : theory, experiments, and applications. New York: Taylor & Francis; 2006. 426 pages p.
9. Kowalewski J, Egorov A, Kruk D, Laaksonen A, Nikkhou Aski S, Parigi G, Westlund PO. Extensive NMRD studies of Ni(II) salt solutions in water and water-glycerol mixtures. J Magn Reson 2008;195(1):103-111.
10. ACR. Phantom Test Guidance for Use of the Large MRI Phantom for the ACR MRI Accreditation Program. American College of Radiology; 2018.

11. NEMA. Standards Publication MS 5-2018 Slice Thickness. Determination of Slice Thickness in Diagnostic Magnetic Resonance Imaging. 2018.
12. Boss MA, Dienstfrey AM, Gimbutas Z, Keenan KE, Splett JD, Stupic KF, Russek SE. Magnetic Resonance Imaging Biomarker Calibration Service: Proton Spin Relaxation Times. Volume Special Publication (NIST SP) - 250-97: NIST; 2018.
13. Lutti A, Hutton C, Finsterbusch J, Helms G, Weiskopf N. Optimization and validation of methods for mapping of the radiofrequency transmit field at 3T. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 2010;64(1):229-238.
14. Stollberger R WP, McKinnon G, Justich E, Ebner F. . Rf-field mapping in vivo. In: Proceedings of the 7th Annual Meeting of SMRM, San Francisco, CA, USA 1988:106.
15. van Schie JJ, Lavini C, van Vliet LJ, Vos FM. Feasibility of a fast method for B1-inhomogeneity correction for FSPGR sequences. Magn Reson Imaging 2015;33(3):312-318.
16. Kingsley PB. Methods of measuring spin-lattice (T1) relaxation times: An annotated bibliography. Concepts in Magnetic Resonance 1999;11(4):243-276.
17. Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3 T? Magn Reson Imaging 2017;35:69-80.
18. Cercignani M, Dowell NG, Tofts P. Quantitative MRI of the brain : principles of physical measurement. Boca Raton: CRC Press, Taylor & Francis Group; 2018. xvii, 342 pages p.
19. Rakow-Penner R, Daniel B, Yu H, Sawyer-Glover A, Glover GH. Relaxation times of breast tissue at 1.5T and 3T measured using IDEAL. Journal of magnetic resonance imaging : JMRI 2006;23(1):87-91.
20. Taylor AJ, Salerno M, Dharmakumar R, Jerosch-Herold M. T1 Mapping: Basic Techniques and Clinical Applications. JACC: Cardiovascular Imaging 2016;9(1):67-81.
21. NEMA Standards Publication MS 1-2008 R2014 Determination of Signal-to-Noise Ratio SNR in Diagnostic MRI. National Electrical Manufacturers Association; 2014.
22. Jiang Y, Ma D, Keenan KE, Stupic KF, Gulani V, Griswold MA. Repeatability of magnetic resonance fingerprinting T1 and T2 estimates assessed using the ISMRM/NIST MRI system phantom. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 2016.

23. Keenan KE, Gimbutas Z, Dienstfrey A, Stupic KF. Assessing effects of scanner upgrades for clinical studies. Journal of Magnetic Resonance Imaging 2019;0(0).
24. Bane O, Hectors SJ, Wagner M, Arlinghaus LL, Aryal MP, Cao Y, Chenevert TL, Fennessy F, Huang W, Hylton NM, Kalpathy-Cramer J, Keenan KE, Malyarenko DI, Mulkern RV, Newitt DC, Russek SE, Stupic KF, Tudorica A, Wilmes LJ, Yankeelov TE, Yen Y-F, Boss MA, Taouli B. Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI: Results from a multicenter phantom study. Magnetic Resonance in Medicine 2018;79(5):2564-2575.
25. Abragam A. The principles of nuclear magnetism: Clarendon Press; 1983.
26. Vold RL, Vold RR, Simon HE. Errors in measurements of transverse relaxation rates. Journal of Magnetic Resonance (1969) 1973;11(3):283-298.
27. Majumdar S, C Orphanoudakis S, Gmitro A, O'Donnell M, C Gore J. Errors in the measurements of T2 using multiple-echo MRI techniques. II. Effects of static field inhomogeneity. 1986;3:562-574.
28. Majumdar S, C Orphanoudakis S, Gmitro A, O'Donnell M, C Gore J. Error in the measurements of T2 using multiple-echo MRI techniques. I. Effects of radiofrequency pulse imperfections1986. 397-417 p.
29. Collins CM, Liu W, Schreiber W, Yang QX, Smith MB. Central brightening due to constructive interference with, without, and despite dielectric resonance. Journal of magnetic resonance imaging : JMRI 2005;21(2):192-196.
30. Mirowski E, Snow M, Keenan K; Liquid Crystal Thermometer for MRI. US patent US20190271750A1. 2019.
31. Keenan KE, Stupic KF, Russek SE, Mirowski E. MRI-visible liquid crystal thermometer. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 2020;n/a(n/a).
32. 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 2016;44(3):610-619.