Data Acquisition Working Group Challenge: Assessment of interplatform variability of T_1 quantification methods used for DCE-MRI in a multicenter phantom study

Bane O[1], Hectors S[1], Wagner M[1], Arlinghaus L[2], Aryal M[3], Boss M[4], Cao Y[3], Chenevert T[5], Fennessy F[6], Huang W[7], Hylton N[8], Kalpathy-Cramer J[9], Keenan K[4], Malyarenko D[5], Mulkern R[6], Newitt D[8], Wilmes L[8], Yen Y F[9], Yankeelov T[10], Taouli B[1]

[1]Icahn School of Medicine at Mount Sinai, [2]Vanderbilt University, [3]University of Michigan Radiation Oncology (2), [4]National Institute of Standards and Technology, [5]University of Michigan Radiology (1), [6]Brigham and Women's Hospital, [7]Oregon Health and Science University, [8]University of California San Francisco, [9]Massachusetts General Hospital, [10]University of Texas-Austin

INTRODUCTION

T_1-weighted dynamic contrast-enhanced MRI (DCE-MRI) is used to quantify perfusion and flow in tumors or other pathologies. DCE-MRI captures the signal change in time with the intravenous injection of a gadolinium (Gd)-based contrast agent by acquiring T_1-weighted images before, during and after injection of contrast agent at a high temporal resolution. Pharmacokinetic models are then applied to the contrast uptake curves to estimate parameters such as flow and vascular permeability. The precision of pharmacokinetic parameters is highly dependent on the conversion of T_1-weighted signal to Gd concentration, and thus on the baseline T_1 value of the tissue of interest.

Objectives:

1. To measure interplatform variability in T_1 quantification in a multicenter NCI QIN study by testing common inversion-recovery spin-echo (IR-SE) and variable flip angle (VFA) protocols using a dedicated T_1 phantom
2. To determine the precision of several T_1 mapping methods currently used by participating centers in a phantom with known reference T_1 values
3. To determine the feasibility of a harmonized T_1 mapping protocol across platforms and centers.

METHODS

Sites

- Preliminary survey identified 8 QIN sites that acquire T_1 values for DCE-MRI study rather than using literature values. Organs of interest, vendor and used T_1 sequences are listed in Table 1 for each site.

Table 1. Center-specific T1 mapping protocols

Site #	Site	Organ	Scanner #	Scanner	Sequences
1	BWH	Prostate	1	GE 3T Discovery w750	VTR
2	MGH	Brain	2,3	Siemens 3T Skyra/Tim Trio	VFA
3	MSinai	Liver, prostate	4,9	Siemens 1.5 Aera/3T Skyra	VFA, Look-Locker
4	OHSU	Breast, extremity	5	Siemens 3T Tim Trio	Proton density
5	UCSF	Breast	10	GE 1.5T HDx	VFA
6	UMich (1)	Brain	6	Philips 3T Ingenia	VTR/ progressive saturation
7	UMich (2)	Brain	7	Siemens 3T Skyra	VFA
8	Vanderbilt	Breast	8	Philips 3T Achieva	VFA

Phantom (Figure 1)

- T_1 phantom produced by the National Institute of Standards and Technology
- Fourteen spherical vials containing deionized water doped with varying concentration of T_1 shortening $NiCl_2$; T1 values measured by NMR spectroscopy.

BWH: Brigham and Women's Hospital; MGH: Massachusetts General Hospital; MSinai: Icahn School of Medicine at Mount Sinai; UCSF: University of California San Francisco; UMich: University of Michigan; OHSU: Oregon Health and Science University.

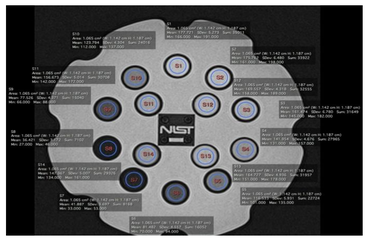

Figure 1. Central plate of T_1 NIST phantom on coronal IR-SE image (TI=150 ms), showing positioning of the 14 spheres of $NiCl_2$ solution

Table 2. NMR T1 values (ms).

	1.5T	3T
S1	2033	1989
S2	1489	1454
S3	1012	984.1
S4	730.8	706
S5	514.1	496.7
S6	367.9	351.5
S7	260.1	247.13
S8	184.6	175.3
S9	132.7	125.9
S10	92.7	89
S11	65.4	62.7
S12	46.32	44.53
S13	32.45	30.84
S14	22.859	21.719

Table 3. Standardized acquisition parameters of IR-SE and VFA

	IR-SE	VFA
Orientation	Coronal	Coronal
Flip angle	180	2, 5, 10, 15, 20, 25, 30
Echo time (ms)	9	2
Repetition time (ms)	5000	12
Inversion times (ms)	24, 50, 75, 100, 125, 150, 250, 500, 750, 1000, 2000, 3000	n/a
Field of view (mm^2)	200x200	200x200
Number of slices	1	16
Slice thickness (mm)	5-6	5-6
Matrix	256x256	256x256
Echo train length	5-6	n/a
Number of averages	1	3
Acquisition time (min)	45	13

Acquisition parameters

- Acquisition performed at room temperature, temperature monitored before and after acquisition
- Duplicate (test-retest) measurements
- Phantom scanned by each center using coil normally employed during DCE-MRI studies
- T_1 acquisitions using center-specific acquisitions and standardized common IR-SE and VFA sequences (Table 3)

Data analysis

- ROI placement in each vial on a single slice by one observer in OsiriX
- T_1 fitting of ROI curves with sequence-specific custom-written Matlab scripts
- Calculation of accuracy and test-retest precision errors:

$$\text{Accuracy (\%)} = 100 * (T_{1\ protocol} - T_{1\ NMR}) / T_{1\ NMR}$$

$$\text{Test-Retest Precision (\%)} = 100 * (T_{1\ test} - T_{1\ retest}) / \text{Mean}(T_{1\ test}, T_{1\ retest})$$

Statistical analysis

- Interplatform variability assessment by coefficient of variation (CV)
- Test-retest repeatability by calculation of CV and Bland-Altman for common sequences (IR-SE and VFA) at each platform
- Agreement of of IR-SE and NMR T_1 values assessed by Lin's concordance correlation
- Identification of independent predictors of accuracy and test-retest precision of standard VFA and site-specific sequences by general linear mixed models with fixed effects vial, scanner, vendor, site, field strength for the common VFA protocol, and additional protocol and method for site-specific sequences.

RESULTS

Interplatform variability (Table 4)

- High CV between platforms for VFA (up to 46%), significantly higher than for IR-SE

Test-retest repeatability (Table 5)

- IR-SE showed better repeatability (CV range 0.17-6.4%) compared to VFA (range 1.17-18.03%)

Agreement of IR-SE and NMR T1 values

- Strong overall agreement between NMR T1 and IR-SE T1 (Fig.2), Lin's concordance correlations rc >0.99, p<10^-6
- Deviation from unity line at large T1 values

General Linear Mixed Model

Common VFA protocol

- **Field strength** was identified as a significant independent predictor of **accuracy,** with values less accurate at 3T (Fig 3a)
- **Scanner** was identified as a significant independent predictor **of test-retest** precision (Fig 3b)

Site-specific protocols

- **Site** was identified as a significant independent predictor of **accuracy (Fig 3c)**
- **Protocol** was identified as a significant independent predictor of **test-retest precision** (Fig 3d)

Table 4. Interplatform coefficient of variation (CV, %) for T_1 values measured in each phantom sphere at 3T, with the IR-SE and VFA protocols.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
IR-SE	1.67	1.31	1.05	1.51	10.59	0.99	4.2	1.65	2.07	2.75	3.76	14.58	3.69	4.88
VFA	17.80	18.67	18.47	19.83	20.26	17.70	22.35	22.26	20.24	18.40	16.57	17.72	28.21	45.53

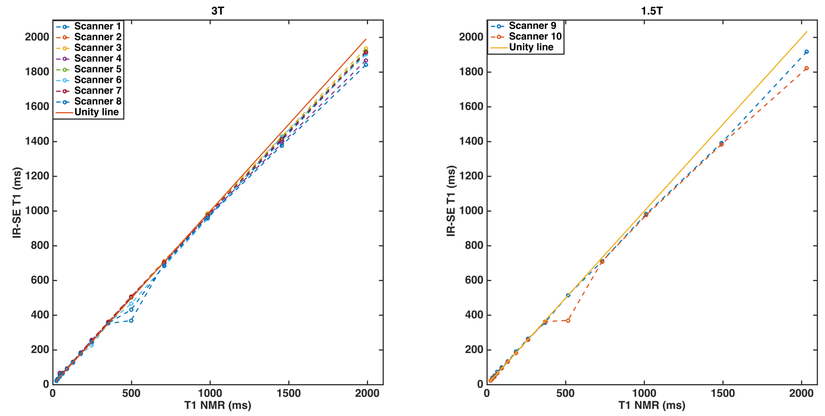

Figure 2. IR-SE and NMR T1 measurements at 3T and 1.5T.

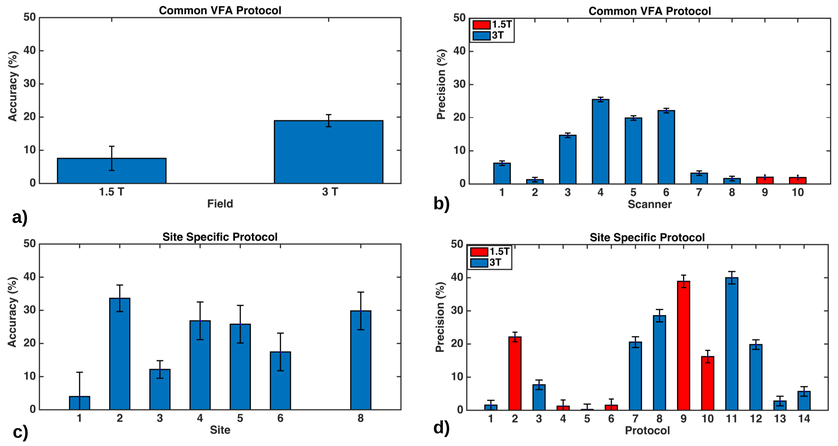

Figure 3. Accuracy (a,c) and test-retest precision (b,d) of common and site specific protocols, with their independent predictors. Data is presented as least square means ± standard error. Smaller numbers represent better accuracy/precision. Protocols: 1=UMich1 Brain VTR, 2=UCSF Breast VFA 1.5T, 3=Vanderbilt Breast VFA 3T, 4= MSinai Liver Look-Locker 1.5T, 5=MSinai Liver Look-Locker 3T, 6=MSinai Liver VFA 1.5T, 7=MSinai Liver VFA 3T, 8=MSinai Prostate VFA 3T, 9=MSinai Prostate VFA 1 1.5T, 10=MSinai Prostate VFA 2 1.5T, 11=BWH Prostate VTR, 12= OHSU Sarcoma PD, 13=MGH Brain Skyra, 14=MGH Brain Trio; Site 7, UMich2, did not provide site-specific data.

Table 5. Test-retest repeatability at participating sites using the common imaging protocols.

		IR-SE			VFA		
		Mean CV (%)	Bias (%)	BA-LA (%)	Mean CV (%)	Bias (%)	BA-LA (%)
1.5T	MSinai Siemens Aera	0.3	-0.29	(-1.54,0.96)	1.45	-2.05	(-3.87,0.233)
	UCSF GE HDx	0.86	1.2	(-2.5, 4.46)	1.38	1.95	(-0.3, 4.22)
3.0T	BWH GE Discovery	0.65	-0.57	(-4.78,3.64)	4.15	-3.6	(-17.95,10.62)
	MGH Siemens Skyra	0.22	-0.18	(-1.17,0.81)	0.93	1.21	(-1.38,3.8)
	MGH Siemens Trio	0.17	-0.12	(-0.7,0.477)	10.39	14.7	(9.3,20.1)
	MSinai Siemens Skyra	0.31	-0.13	(-1.17,0.91)	18.03	-25.5	(-28.4,-22.6)
	OHSU Siemens Trio	0.25	-0.16	(-0.87,0.55)	14.21	19.92	(14.94,24.9)
	UMich1 Philips Ingenia	1.2	0.06	(-4.86,4.98)	15.66	-22.15	(-30.5,-13.8)
	UMich2 Siemens Skyra	6.4	-8.04	(-24.94,8.86)	2.31	-3.27	(-4.82,-1.73)
	Vanderbilt Philips Achieva	0.58	0.35	(-3.24,3.94)	1.17	-1.49	(-4.28,1.29)

CONCLUSIONS

We observed high interplatform variability in T_1 values for the common VFA protocol, among spheres and in test-retest scans. Although there was very strong agreement between IR-SE and NMR values, the deviation from unity at large T1 values precluded the use of IR-SE as internal "gold standard" for each scanner. The general linear mixed model analysis showed less accuracy and precision at 3T with the VFA protocol. Among site-specific protocol, accuracy depended on how well the protocol was optimized for its specific application. Precision for site-specific protocols was lower at 3T than at 1.5T for VFA protocols.