Abstract #05219 ISMRM ABSTRACT

Phantom validation of an Imageless Magnetic Resonance Diagnosis proof-of-concept

Primary: Acquisition & Reconstruction - AI methods 〉 **Secondary:** Physics & Engineering - Low-Field MRI 〉 **Traditional Poster** · 60 min | Acquisitions and Reconstructions · Tuesday, May 12 at 08:20 AM 〉 **Keywords:** SCREENING AI/ML SOFTWARE LOW-FIELD MRI

Alba González-Cebrián[1], **Pablo García-Cristóbal**[1], **Fernando Galve**[1], **José M Algarín**[1], **Viktor Van Der Valk**[2], **Efe Ilıcak**[3], **Marius Staring**[4], **Andrew Webb**[5], **Joseba Alonso**[1]

[1]Institute for Molecular Imaging and Instrumentation (i3M), Consejo Superior de Investigaciones Científicas & Universitat Politècnica de València, Valencia, Spain
[2]Leiden University Medical Center, Leiden, Netherlands
[3]Department of Radiology, Leiden University Medical Center, Leiden, Netherlands
[4]Radiology, Leiden University Medical Center, Leiden, Netherlands
[5]C.J. Gorter MRI Center, Leiden University Medical Center, Netherlands

Presenting Author: *Alba González-Cebrián (algonceb@upv.es)*

Impact

This study provides the first experimental validation of IMRD on real data, supporting feasible clinical decision-making from raw MR signals. In the longer term, these results hold promise towards affordable, fast and deployable MR screening systems fully bypassing image reconstruction.

Synopsis

Motivation: MRI's typical cost and scan times limit its use for large-scale screening. Imageless Magnetic Resonance Diagnosis (IMRD) aims to extract diagnostic information directly from raw MR signals, bypassing image reconstruction.

Goals: To experimentally validate an IMRD sequence and evaluate its quantification ability.

Approach: Agarose-gel phantoms mimicking brain tissues were scanned on a 260 mT system. Sequence parameters were fitted and used to simulate CNN training data.

Results: Imageless scans lasting < 30 s suffice for high-fidelity (95 %) detection of white-matter lesions in phantoms. Lesion quantification is also demonstrated, with a correlation coefficient R^2 = 0,85 with respect to the nominal volumes.

Introduction

MR images remain the gold standard assessment for many imaging-based diagnosis applications, but they are tied to long scan times and costly infrastructure. IMRD redefines MR for diagnostic inference by entirely bypassing image reconstruction and using time-domain MR signals instead.

Our previous in-silico study illustrated IMRD for multiple sclerosis (MS) volume estimation and detection (R^2 > 0,85 and AUC > 0,95)[1]. The current study aims to experimentally validate the IMRD sequence for "MS" quantification with agarose phantoms, evaluating (i) the agreement between simulations and measurements, and (ii) the ability of CNNs trained on synthetic data to quantify MS on real MR signals.

Methods

We scanned four replicates of 7 mL agarose-CuSO4-gelatin phantoms[2] mimicking T1/T2 values of white matter (WM), gray matter (GM), and MS tissues on our 260 mT Dental I LF scanner[3], using an IMRD sequence optimized with BlackBoxOptim[4] to maximize MS contrast. The sequence started with an Inversion Pulse (IR), followed by a train of 40 RF pulses with optimized Flip Angle (FA) and Repetition Times (TR), sampling a single echo mid-TR (at TE = TR/2). The complete scan lasted 26 seconds, providing 40 complex signal points per sample. No gradients (for spatial encoding or active shimming) were used.

For signal model validation, we used WM-only phantoms (7 to 28 mL, Figure 1a) to fit a Bloch-based model, accounting for potential deviations from nominal FAs (quantified by b_{FA} = effective FAs/nominal FAs) and relaxation times. Confidence Intervals (95%) for effective parameters were computed analytically (Jacobian-based) and via bootstrap resampling (nB = 500) in both the signal and residual domains. Confidence bands for the simulated signals were obtained based on parameters' CIs.

We later used effective sequence parameters derived from these fits to re-simulate BrainWeb data[5], adding Gaussian noise based on our measurements. These synthetic signals (935 slices) were used to train 1D CNNs predicting MS volume. 1D CNNs consisted of a single convolutional layer with a kernel size 40, spanning the receptive field to all TRs to extract 100 features, fed to a fully connected layer producing the outcome (Figure 4a). Simulation-trained CNNs were evaluated on 100 slices from two unseen brains and fine-tuned with phantom data. One set of mixtures with 7 mL of WM and GM and varying volumes of MS (from 0 to 7 mL, Figure 1b), was used for model fine-tuning, freezing the convolutional layer, and another set was held out for model evaluation.

Results

Figure 1(c-d) presents parametric and non-parametric 95 % CIs for all WM datasets (R^2>0,98; χ^2≈1). Figure 2 displays simulated signal confidence bands alongside the corresponding experimental measurements (R^2≥0,97). Figure 3 shows the average residuals across the 40 TRs for each dataset. Table 1 lists the performance metrics obtained on test sets when the training data were simulated with nominal and with effective sequence parameters. Figures 4c and 4d plot the MS volume ($\hat{vol}_{MS}$) predicted by CNNs against the true MS volumes for both in silico (R^2≥0,96) and phantom (R^2=0,85) datasets, and their corresponding prediction errors, respectively.

Discussion

Parameter estimates followed similar trends across experiments (Figure 1c-d) and confidence bands of signals simulated with effective parameters overlapped with measured signals (R^2≥0,97, Figure 2), confirming proper signal modelling. Small deviations from nominal FA and T1/T2 values might originate from a position effect related to less homogeneous B_0 or B_1 regions, discrepancies with T1/T2 calibration protocols or progressive phantom property degradation. Systematically higher residuals between measurements and simulations at TRs 19 and 20 (Figure 3) also suggest an unexpected signal behaviour, but do not compromise the validity of the simulation model.

Table 1 reports CNNs' performance for MS quantification and detection on test sets, using either nominal or effective parameters for training data simulation. The promising relation between predicted ($\hat{vol}_{MS}$) and true MS (R^2 = 0,85, Figure 4c), and the detection of MS in 95% of samples by $\hat{vol}_{MS} > 0$ (TPR, Table 1) indicate valid informative patterns learned from simulations. The increased scatter across phantom predictions compared to those in simulations (R^2≥0,96, Figures 4c and 4d), might be reduced by increasing phantoms' variability and the number of samples.

Conclusion

This work experimentally validates IMRD on ultra-fast, gradient-free, real low-field MR measurements. The agreement between simulations and measurements (R^2≥0,97), together with MS volume quantification (R^2 = 0,85) and positive detection in 95% of samples, offer an intermediate quantitative validation of MR-answered questions using minimal hardware and acquisition complexity. Future work will extend validation to heterogeneous and in-vivo datasets.

Acknowledgements

This work was supported by European Innovation Council (NextMRI [101136407]); Programa Fondo Europeo de Desarrollo Regional (FEDER) Comunitat Valenciana 2021-2027, (Views INNVA1/2023/30); PROMETEO 2022 (CIPROM/2021/003).

References

1. González-Cebrián A, García-Cristóbal P, Galve F, Ilıcak E, Valk VVD, Staring M, et al. An imageless magnetic resonance framework for fast and cost-effective decision-making. Progress in Nuclear Magnetic Resonance Spectroscopy. noviembre de 2025;150-151:101574.

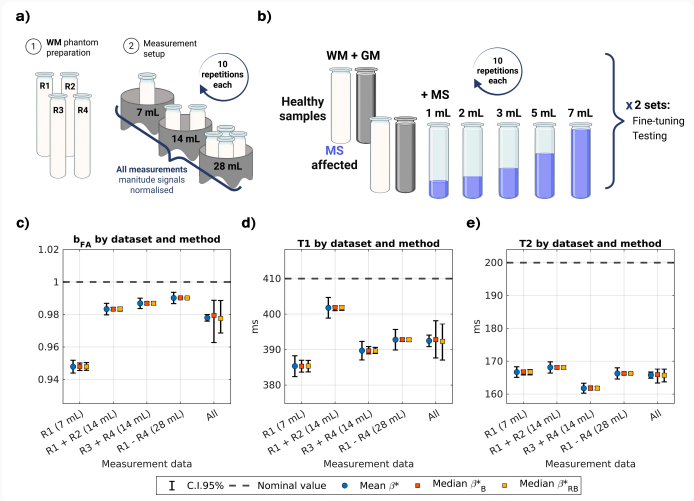
2. Jordanova KV, Fraenza CC, Martin MN, Tian Y, Shen S, Vaughn CE, et al. Paramagnetic salt and agarose recipes for phantoms with desired T1 and T2 values for low-field MRI. NMR in Biomedicine. enero de 2025;38(1):e5281.

3. Algarín JM, Díaz-Caballero E, Borreguero J, Galve F, Grau-Ruiz D, Rigla JP, et al. Simultaneous imaging of hard and soft biological tissues in a low-field dental MRI scanner. Sci Rep. 8 de diciembre de 2020;10(1):21470.

4. Feldt R. BlackBoxOptim.jl Retrieved from https://github.com/robertfeldt/BlackBoxOptim.jl. 2024.

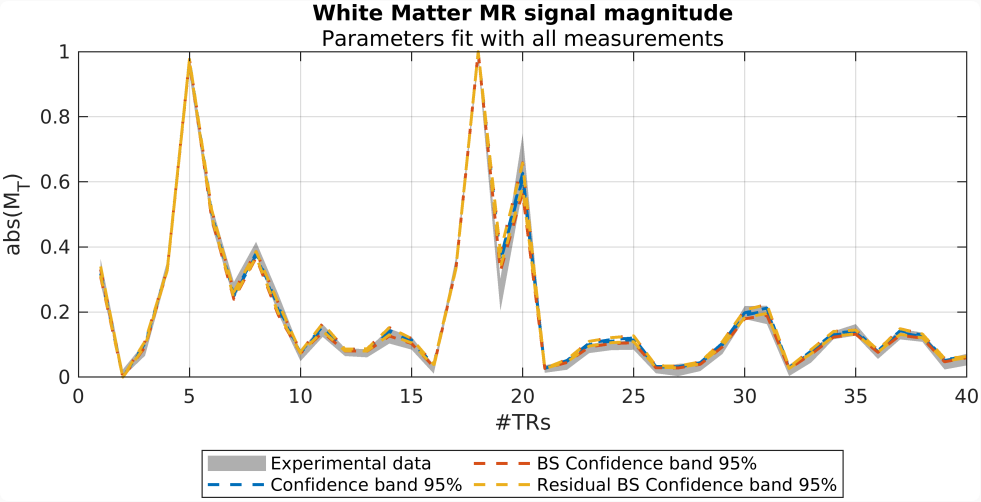
5. Aubert-Broche B, Griffin M, Pike GB, Evans AC, Collins DL. Twenty new digital brain phantoms for creation of validation image data bases. IEEE transactions on medical imaging. 2006;25(11):1410-6.

Figures and Tables



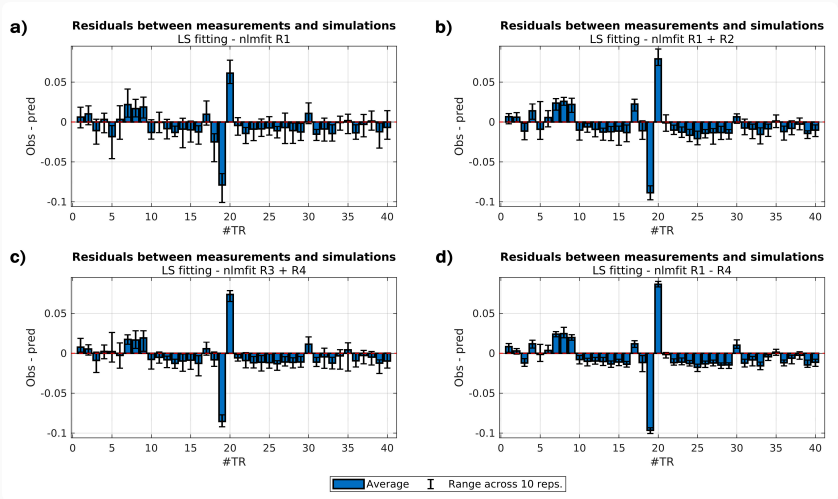
Scan for high-resolution version

Figure 1: *Phantom data, with (a) WM volumes and their joint analysis ("All") and (b) WM and GM combined with six MS volumes (0–7 mL) measured with 10 repetitions per condition, using separate train and test phantom sets; (c) b_{FA}, (d) T1 and (e) T2 fitted with nonlinear model (β*), bootstrap ($β^*_B$), and residual bootstrap ($β^*_{RB}$), estimating CIs analytically (β*), or via percentiles ($β^*_B$, $β^*_{RB}$; nB = 500). WM fittings showed high R^2 (>0,98) and χ^2 values below or close to 1 and below critical values.*



Scan for high-resolution version

Figure 2: *Confidence bands on the signal domain derived by simulating the expected signal with the confidence interval extremes estimated with different methods using all data. Goodness-of-fit coefficients for models fitted using all measurements were of R^2 = 0,988 (using classical non-linear fit on all samples, blue dashed lines); R^2 = 0.988 (average R^2 values across bootstrap samples); and R^2 = 0.975 (average R^2 values across residual bootstrap samples).*



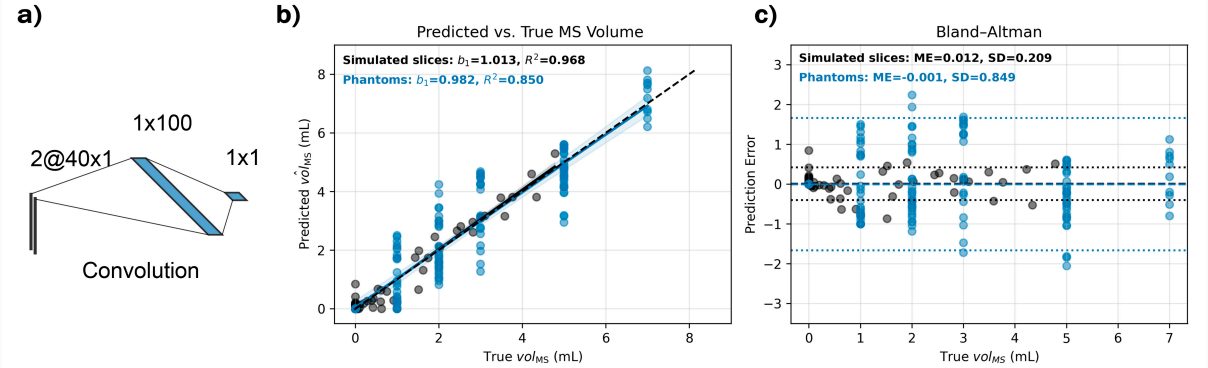
Scan for high-resolution version

Figure 3: *Residuals by TR between the measured and the simulated values using the model fitted with each one of the WM datasets, having ten measurement repetitions in ach. Bars and error bars indicate the average and range of residuals across ten measurements, respectively.*

Dataset	R^2	Slope (m; $CI_{95\%}$)	TPR	FPR	TNR	FNR
Slices nominal	0.98	0.98 [0.96; 1.02]	0.85	0.20	0.80	0.15
Slices effective	0.97	0.95 [0.88; 1.01]	0.82	0.20	0.80	0.18
Phantoms nominal	0.85	0.95 [0.88; 1.01]	0.95	0.00	1.00	0.05
Phantoms effective	0.85	0.98 [0.13; 1.05]	0.95	0.00	1.00	0.05

Scan for high-resolution version

Figure 4: *Results on test sets for simulated signals from in silico slices and for measured ones from test-tube phantoms. The threshold for detection was$\hat{vol}_{MS}>0$. Regression ($\hat{vol}_{MS} = b_0+b_1\cdot vol_{MS}$) metrics are the R^2 and the slope, with non-significant intercept terms (p-val> 0.05). Classification metrics are the True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR) and False Negative Rate (FNR).*



Scan for high-resolution version

Figure 5: *(a) CNN architecture used for MS quantification; (b) Predicted vs. true MS volumes with regression metrics (simulated R^2 = 0.96, phantoms R^2 = 0.85), and (c) Bland-Altman plot indicating bias and dispersion between predicted and true MS volumes, with normal distribution limits as dashed lines.*