


Synthetic MRI with Magnetic Resonance Spin TomogrAphy in Time-Domain (MR-STAT): Results from a Prospective Cross-Sectional Clinical Trial

Jordi P. D. Kleinloog, PhD,^{1,2*}  Stefano Mandija, PhD,^{1,2} Federico D'Agata, PhD,³
 Hongyan Liu, Msc,^{1,2} Oscar van der Heide, Msc,^{1,2} Beyza Koktas, Bsc,¹
 Jan Willem Dankbaar, MD, PhD,² Vera C. Keil, MD, PhD,⁴ Evert-Jan Vonken, MD, PhD,²
 Sarah M. Jacobs, MD,⁵ Cornelis A. T. van den Berg, PhD,^{1,2} Jeroen Hendrikse, MD, PhD,²
 Anja G. van der Kolk, MD, PhD,^{2,6} and Alessandro Sbrizzi, PhD^{1,2}

Background: Magnetic Resonance Spin TomogrAphy in Time-domain (MR-STAT) can reconstruct whole-brain multi-parametric quantitative maps (eg, T_1 , T_2) from a 5-minute MR acquisition. These quantitative maps can be leveraged for synthetization of clinical image contrasts.

Purpose: The objective was to assess image quality and overall diagnostic accuracy of synthetic MR-STAT contrasts compared to conventional contrast-weighted images.

Study Type: Prospective cross-sectional clinical trial.

Population: Fifty participants with a median age of 45 years (range: 21–79 years) consisting of 10 healthy participants and 40 patients with neurological diseases (brain tumor, epilepsy, multiple sclerosis or stroke).

Field Strength/Sequence: 3T/Conventional contrast-weighted imaging (T_1/T_2 weighted, proton density [PD] weighted, and fluid-attenuated inversion recovery [FLAIR]) and a MR-STAT acquisition (2D Cartesian spoiled gradient echo with varying flip angle preceded by a non-selective inversion pulse).

Assessment: Quantitative T_1 , T_2 , and PD maps were computed from the MR-STAT acquisition, from which synthetic contrasts were generated. Three neuroradiologists blinded for image type and disease randomly and independently evaluated synthetic and conventional datasets for image quality and diagnostic accuracy, which was assessed by comparison with the clinically confirmed diagnosis.

Statistical Tests: Image quality and consequent acceptability for diagnostic use was assessed with a McNemar's test (one-sided $\alpha = 0.025$). Wilcoxon signed rank test with a one-sided $\alpha = 0.025$ and a margin of $\Delta = 0.5$ on the 5-level Likert scale was used to assess non-inferiority.

Results: All data sets were similar in acceptability for diagnostic use (≥ 3 Likert-scale) between techniques ($T_1w:P = 0.105$, $PDw:P = 1.000$, $FLAIR:P = 0.564$). However, only the synthetic MR-STAT T_2 weighted images were significantly non-inferior to their conventional counterpart; all other synthetic datasets were inferior ($T_1w:P = 0.260$, $PDw:P = 1.000$, $FLAIR:P = 1.000$). Moreover, true positive/negative rates were similar between techniques (conventional: 88%, MR-STAT: 84%).

View this article online at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/jmri.28425). DOI: 10.1002/jmri.28425

Received Jul 19, 2022, Accepted for publication Aug 23, 2022.

*Address reprint requests to: J.P.D.K., Heidelberglaan 100, 3584 CX Utrecht, The Netherlands.

E-mail: j.p.d.kleinloog@umcutrecht.nl

Contract grant sponsor: Netherlands Organisation for Scientific Research (NWO) Demonstrator Grant 16937.

From the ¹Computational Imaging Group for MR Therapy and Diagnostics, Center for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands; ²Department of Radiotherapy, University Medical Center Utrecht, Utrecht, The Netherlands; ³Department of Neurosciences, University of Turin, Turin, Italy; ⁴Department of Radiology, Amsterdam University Medical Center, Amsterdam, The Netherlands; ⁵Department of Radiology and Nuclear Medicine, University Medical Center Utrecht, Utrecht, The Netherlands; and ⁶Department of Medical Imaging, Radboud University Medical Center, Nijmegen, The Netherlands

Additional supporting information may be found in the online version of this article

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Data Conclusion: MR-STAT is a quantitative technique that may provide radiologists with clinically useful synthetic contrast images within substantially reduced scan time.

Evidence Level: 1

Technical Efficacy: Stage 2

J. MAGN. RESON. IMAGING 2023;57:1451–1461.

A conventional clinical MRI examination consists of several sequences with different image contrast weightings (eg, T_1 and T_2 weighted [T_{1w} and T_{2w}], proton density weighted [PDw], fluid-attenuated inversion recovery [FLAIR]). However, the serial acquisition of these image contrasts results in long examination times. In addition, signal intensity of conventional sequences varies depending on hardware and protocol parameters, which impedes comparison of absolute signal intensity levels between examinations.¹

Recent technological advances have enabled fast multi-parametric quantitative MRI (qMRI) techniques, thereby reducing acquisition time and improving standardization of signal intensities.^{1,2} However, qMRI has not been adopted in clinical settings despite its potential, which include early diagnosis, monitoring of disease progression, and evaluation of therapeutic effects.^{1,2} Also, quantitative tissue maps allow for synthesization of contrast-weighted images,^{3,4} which may provide clinicians with familiar and standardized image contrasts within a substantially reduced examination time.

Several multi-parametric imaging techniques are being developed, such as Multiple-Dynamic Multiple-Echo (MDME)-based Synthetic MR (SyMRI)^{5,6}; Magnetic Resonance Fingerprinting (MRF)^{7–9}; Strategically Acquired Gradient Echo (STAGE)¹⁰; MultiPathway MultiEcho (MPME)¹¹; and Quantitative Transient-state Imaging,^{12–14} some of these already demonstrated the ability of synthesization of contrast-weighted images, which was discussed in a recent review article.¹⁵ Quantified tissue parameters in combination with sequence parameters (eg, TE, TR, and TI) are then used to synthesize the contrast-weighted images using an analytical signal model. Overall, image quality has shown promise, although synthetic FLAIR image quality was inferior to conventional images.¹⁵ The synthesization process for most of these techniques is based on analytical signal models, MPME being the exception since it is based on a data-driven (deep learning) approach. Synthetic MRI technology is approaching a mature stage, as witnessed by SyMRI,^{5,6} which already received FDA and CE approval, highlighting the potential of increased workflow efficiency using these multi-parametric imaging techniques.

Before synthetic MRI can be adopted into routine clinical practice, its image quality should be sufficient to replace the existing sequences; therefore, clinical trials that compare synthetic image quality with conventional imaging are required. This work describes a comparison study focused on the Magnetic Resonance Spin Tomography in Time-domain

(MR-STAT) acquisition and reconstruction technique.¹⁶ MR-STAT is a quantitative imaging technique for simultaneous T_1 , T_2 and PD mapping, and allows a whole-brain acquisition within 5 minutes.^{17,18} These quantitative parameter maps are obtained by directly fitting a volumetric signal model to transient-state time-domain signal. In this process, quantitative parameter estimation and spatial localization of the signal are performed simultaneously.¹⁷ The fitting is done by numerically solving a (computationally demanding) large-scale nonlinear problem.¹⁸ Advantages of MR-STAT are the flexibility with respect to changes in sequence parameters and a comprehensive signal model which can efficiently deal with any (highly under-sampled) encoding scheme. As a consequence, MR-STAT sequences are typically based on Cartesian acquisitions. This is the most commonly used gradient trajectory in clinical practice, because of, for example, robustness against hardware imperfections and availability on older MR systems. Also, off-resonance effects cause image phase shifts for Cartesian acquisitions that can be easily corrected, but cause blurring for non-Cartesian acquisitions.

The objective of this prospective cross-sectional study was to assess the diagnostic image quality (primary objective) and diagnostic accuracy (secondary objective) of MR-STAT-based synthetic contrast-weighted images compared to conventional contrast-weighted images in healthy participants and patients with four neurological diseases.

Materials and Methods

Study Participants and Design

The study was conducted according to the guidelines described in the Declaration of Helsinki, approved by the Institutional Review Board (NL69544.041.19, METC 19/282). All participants provided written informed consent. The trial was registered in the international clinical trials registry platform with number (<https://trialssearch.who.int/Trial2.aspx?TrialID=NL8437>).

Between October 2019 and October 2021, a total of 50 adults aged 18 years or older were included in this diagnostic accuracy cross-sectional study. Participants had to be aged 18 years or older and able to lie supine in the MRI scanner for 45 minutes. Additional inclusion criteria for healthy participants ($n = 10$) were: 1) no history of neurological disease; and 2) accepting to be informed of any clinically relevant incidental findings. Additional inclusion criteria for patients ($n = 40$) were: 1) a definitive diagnosis of one of the following neurological diseases: primary brain tumor ($n = 11$), epilepsy ($n = 10$), multiple sclerosis (MS; $n = 9$) or ischemic stroke ($n = 10$); 2) a previous clinical MRI examination that clearly showed the characteristics of the particular disease, i.e. a “classical” case of this particular disease.

TABLE 1. Acquisition imaging parameters of the MR-STAT and conventional sequences

Imaging Parameters	MR-STAT	T ₁ w	T ₂ w	PDw	FLAIR
	Spoiled-GRE	SE	TSE	TSE	TSE
FOV (cm)	224 × 224 × 133.5 mm ³				
Spatial resolution	1 × 1 × 3 mm ³				
Gap	1.5 mm				
Slices	30				
TR (msec)	8.9	451	3400	2800	10,000
TI (msec)	-	-	-	-	2800
TE (msec)	4.7	14	80	20	120
Flip angle	Variable	70	90	90	90
TSE factor	-	-	15	14	24
Scan time	5 minutes	11 minutes			

MR-STAT = Magnetic Resonance Spin TomogrAphy in Time-domain; PDw = Proton Density weighted; FLAIR = Fluid Attenuated Inversion Recovery; GRE = Gradient Echo; SE = Spin Echo; TSE = Turbo Spin Echo.

Only four diseases were selected to keep the patient group as homogeneous as possible, while still being able to assess feasibility of the MR-STAT technique for more than just one disease type. These four diseases were chosen because they can have very characteristic findings on MRI. Patient selection was carried out in collaboration with the neurology department of our hospital.

Image Acquisition and Processing

MRI examinations were performed on a 3T Ingenia scanner with a 15 channel receiver head coil (Philips, Best). Conventional (T₁w, T₂w, PDw, and FLAIR) and MR-STAT images were acquired. Sequence parameters are presented in Table 1. The MR-STAT sequence was Cartesian encoded, and consisted of multiple 2D slices with a gradient-spoiled gradient echo and a slowly varying flip angle between 0° and 90° preceded by a non-selective inversion pulse,¹⁸ 30 axial slices with 1 × 1 × 3 mm³ in-plane resolution and a slice gap of 1.5 mm. Acquisition time was approximately 5 minutes. Reconstruction was performed with an alternating direction method of multipliers presented in Liu et al.¹⁹ An automated data-workflow was employed to reconstruct, process, and store the quantitative MRI images. Subsequently, synthetic T₁w, T₂w, PDw, and FLAIR contrast images were generated from analytical signal models²⁰ on voxel-to-voxel basis and thereafter denoised using a denoising autoencoder network.²¹ Figure S1 shows an example of the original and denoised contrast weighted images. Denoising was performed using existing software modified to work as a convolutional neural network with encoder and decoder structure.²² The neural network was trained to reconstruct the conventional image contrasts from the conventional images with random noise added to match the noise level in the synthetic images. The four conventional contrast-weighted data sets were acquired for comparison purposes, using the same spatial parameters as the MR-STAT acquisition and closely

approximating standard clinical brain exams at the time the study commenced (Table 1).

Imaging Assessments

The synthetic MR-STAT and conventional image data sets (n = 400) were randomized and image quality was assessed independently by three radiologists with 6 years, 15 and 16 of experience who were blinded for image type and clinical information (including diagnosis). Assessment was performed in two sittings separated by a 4-week washout period served to reduce recall bias.

PHASE 1: IMAGE QUALITY. Diagnostic quality was scored on a 5-level Likert scale (unacceptable, bad, acceptable, good or excellent, as described in Supplemental Material A)²³ in four sessions separated by at least 2 weeks. Ratings equal or larger than 3 were considered acceptable for diagnostic use. Readers were asked to take into account all slices of a dataset; for instance, if one of the 30 slices had motion artifacts but the other slices were of high diagnostic quality, this dataset would be scored based on the combination of both. In addition to overall quality, anatomical and morphological legibility was evaluated on a binary scale (legible/illegible) in the following features: central sulcus, head of the caudate nucleus, posterior limb of the internal capsule, cerebral peduncle, middle cerebellar peduncle, and cervicomedullary junction. These particular structures were chosen because of their neuroradiological importance, dependence on good image quality, and overlap with a previous study,⁵ facilitating easier comparison with previous synthetic MRI results. Similar to the overall quality score, readers were asked to take the entire dataset (or in this case, all images showing the particular anatomical structure) into account. Readers also recorded whether any of the following artifacts or quality issues were present: low signal-to-noise, motion, fold-over or wrap-around, white pixel or spike noise, phase

encoding, flow, contrast-to-noise, low image resolution, or blurring. To avoid a paired comparison, the dataset scored in each session included one type of contrast from the MR-STAT-based synthetic volumes ($n = 50$) and a different type of contrast from the conventionally acquired protocol ($n = 50$), eg, synthetic T_1w and conventional PDw (total of 100 volumes).

PHASE 2: DIAGNOSTIC ACCURACY. Diagnostic accuracy was assessed in two sessions separated by 4 weeks. In each session, each radiologist received 50 datasets in a randomized order. Each dataset consisted of the four conventional or synthetically generated contrast-weighted images for a given participant (Supplemental Material B). Diagnostic accuracy in 50 participants was assessed (for both the conventional and MR-STAT technique) by three radiologists resulting in 300 cases in total. Radiologists reported the level of confidence (not likely [10%], possibly [50%], likely [75%], matching [90%]) of the following pathological subtypes adapted

from Osborn et al²⁴: 1) traumatic, complex, indeterminate, or other condition or injury; 2) congenital malformation; 3) ischemic or hemorrhagic stroke, subarachnoid hemorrhage/aneurysm; 4) vascular malformation; 5) neoplasm/primary neoplastic cysts; 6) infectious/demyelinating disease; or 7) metabolic/degenerative disorders. The scores of the radiologists were compared to the confirmed diagnosis of the patient which was provided by the treating neurologist at inclusion; this confirmed diagnosis was the result of an extensive clinical workup well before being included in this study. Based on the level of confidence and the confirmed diagnosis, observer scores per participant were divided into four levels: 0 = false positive (healthy volunteers) and false negative (patients); 1 = uncertain about a lesion (at least one pathological subtype was detected with 50% confidence); 2 = at least one lesion observed ($\geq 75\%$ confidence) but incorrect pathological subtype classification; 3 = correct classification of the lesion with at least 75% confidence for a true positive, or a true negative classification of a healthy volunteer.

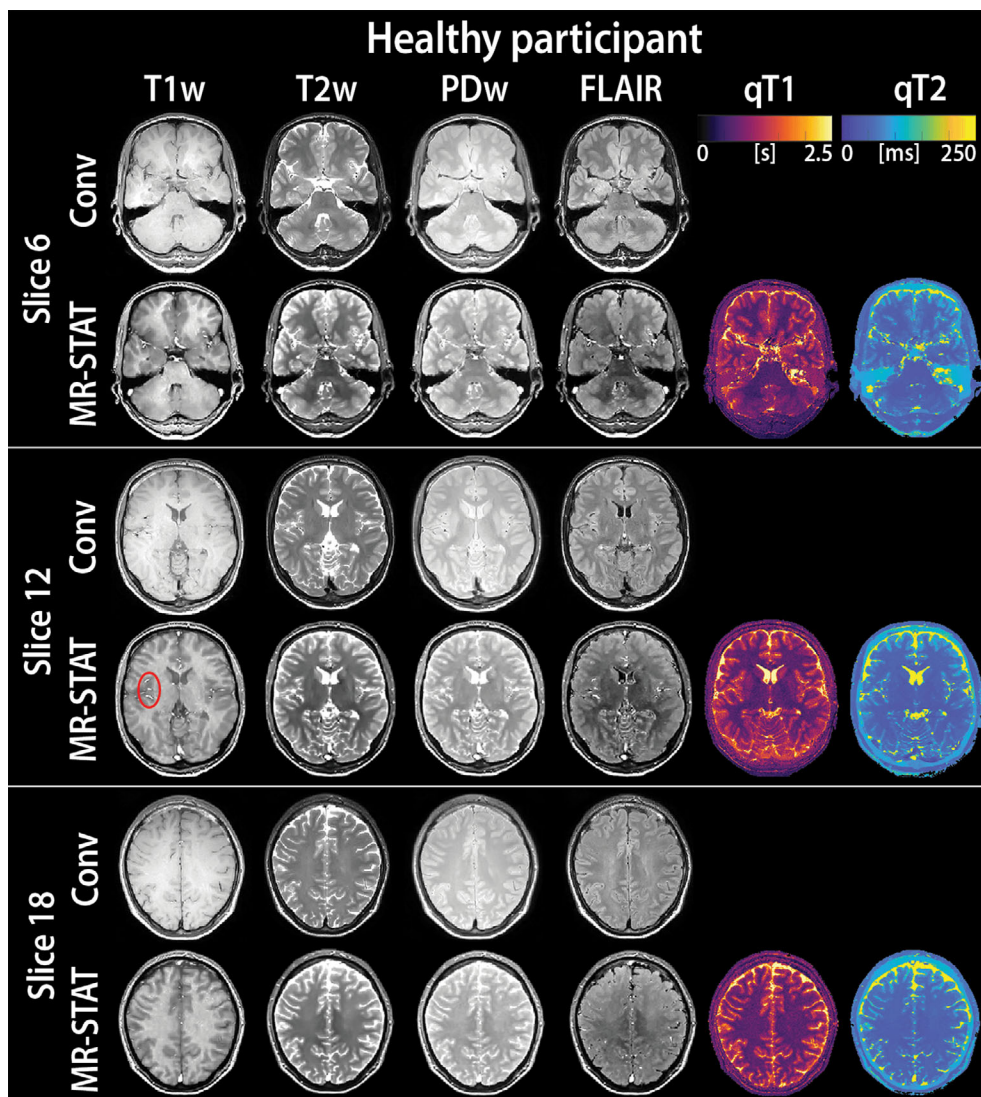


FIGURE 1: Comparison of T_1 weighted (T_1w), T_2 weighted (T_2w), proton density weighted, and fluid-attenuated inversion recovery images obtained from conventional MR and synthetically generated from quantitative (q) T_1 and (q) T_2 maps obtained with Magnetic Resonance Spin Tomography in Time-Domain. Different axial slices near the skull base (Slice 6), halfway through the lateral ventricles (Slice 12), and high in the brain (Slice 18) of one representative healthy participant.

Statistical Analysis

Statistical analysis was performed using SPSS (IBM Corp., IBM SPSS Statistics, V26, Armonk, NY), R-studio version 1.4.1717 and R version 4.1.2. A one-sided α of 0.025 was considered to be statistically significant. Multiple comparisons were corrected using the Bonferroni method (P -value \times number of tests). Results are presented as percentage per contrast and overall, which were calculated as $(n/N\%)$, where n is the count and N is the total reads per patient group ($N = 50$ per contrast and 200 overall).

Nonparametric tests were performed as the data distribution was non-normal, which was inferred using a Shapiro–Wilkinson test. The median across observers was used to determine whether contrasts were scored differently between the synthetic and conventional technique (contrasts \times technique) using the nonparametric analysis of longitudinal data (nparLD) R-package.²⁵ Accordingly, differences between techniques in acceptability for diagnostic use (≥ 3 on the Likert scale) was assessed per contrast using a McNemar's test (one-sided $\alpha = 0.025$). Non-inferiority of synthetic image quality was assessed by means of the Wilcoxon signed rank test with a one-sided $\alpha = 0.025$ and a margin of $\Delta = 0.5$ on the 5-level Likert scale.²⁶ The margin of 0.5 was previously substantiated by Tanenbaum et al in a related study.⁵ Anatomical and morphological legibility by anatomic region, artifact prevalence, and diagnostic accuracy were reported with descriptive statistics. Overall diagnostic accuracy between techniques was described as well as differences per patient group.

Results

Diagnostic Image Quality

Different slices of the four contrast-weighted images from both conventional and MR-STAT methods from a healthy participant are shown in Fig. 1. The median quality ratings across observers significantly differed per contrast between techniques. Differences between techniques were non-significant in acceptability for diagnostic use (T_{1w} : $P = 0.105$, PDw: $P = 1.000$, FLAIR: $P = 0.564$) (Table 2). Note that no McNemar's test could be performed for the T_{2w} data set as all images were acceptable for diagnostic use. Only the T_{2w} synthetic MR-STAT contrast was significantly non-inferior with a median $\Delta = 0$, while the T_{1w} (median $\Delta = -0.5$, $P = 0.260$), PDw (median $\Delta = -0.5$, $P = 1.000$) and FLAIR (median $\Delta = -1$, $P = 1.000$) synthetic MR-STAT contrasts received a lower rating (Table 2). The most common quality issue in the unacceptable T_{1w} images was a difference in image contrast between slices. For synthetic FLAIR images, the most common issue was that lesions appeared hypointense, while they appeared hyperintense on conventional FLAIR images (Fig. 2). Example images of the quantitative maps (T_1 , T_2 and PD) and conventional and synthetic contrast weighted images

TABLE 2. Image-quality ratings by contrast type and overall for conventional and synthetic images^a

Diagnostic Quality ^b	T_{1w}		T_{2w}		PDw		FLAIR		Overall	
	Conv	MR-STAT	Conv	MR-STAT	Conv	MR-STAT	Conv	MR-STAT	Conv	MR-STAT
Acceptable for diagnostic use (3, 4, 5)	98%	86%	100%	100%	100%	98%	98%	92%	99%	94%
Excellent (5)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Good (4)	30%	0%	90%	82%	86%	20%	90%	16%	74%	30%
Acceptable (3)	68%	86%	10%	18%	14%	78%	8%	76%	25%	65%
Unacceptable for diagnostic use (2) ^c	2%	14%	0%	0%	0%	2%	2%	8%	1%	6%
Non-inferiority median Δ (P -value) ^d	-0.5 (0.325)		0.0 (<0.001)		-0.5 (1.000)		-1.0 (1.000)		-	
P -value difference diagnostic use ^e	0.105		-		1.000		0.564		-	

Conv = conventional MR imaging (control); MR-STAT = synthetic contrasts from quantitative Magnetic Resonance Spin Tomography in Time-domain (MR-STAT) images; T_{1w} = T_1 weighted; T_{2w} = T_2 weighted, PDw = proton density weighted; FLAIR = Fluid Attenuated Inversion Recovery.

^aAll data are shown as n ($n/N\%$), where n is the count and N is the total reads per category ($N = 50$ per category and 200 overall).

^bMedian across observers on a five-point Likert scale.

^cNo image had a score of 1.

^dNon-inferiority Wilcoxon signed rank test per contrast with a margin of $\Delta = 0.5$, as there was a significant contrast \times technique interaction tested with nonparametric analysis of longitudinal data.

^eMcNemar's test. P -values were adjusted for multiple comparisons with Bonferroni correction.

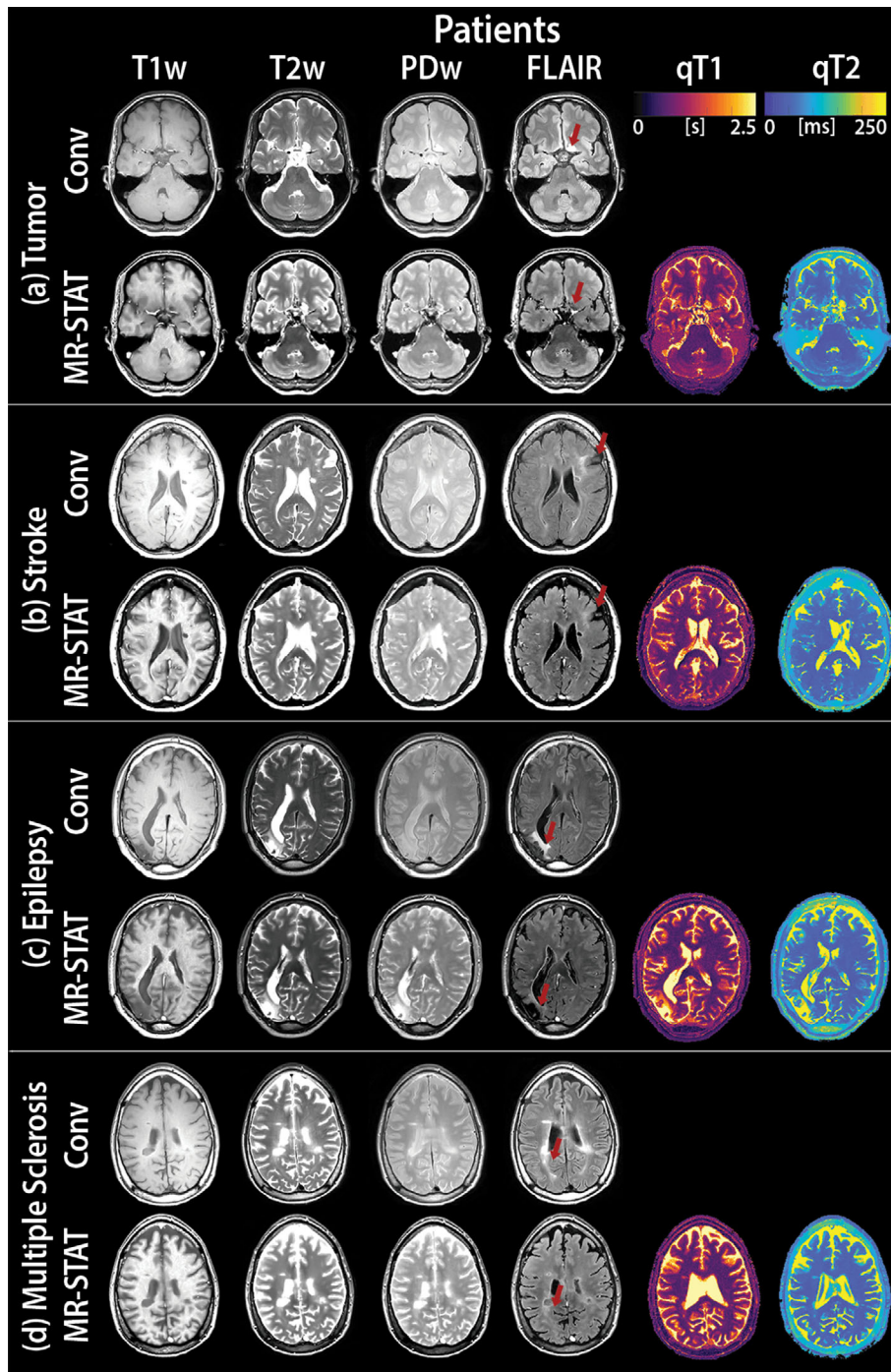


FIGURE 2: Comparison of T_1 weighted (T_{1w}), T_2 weighted (T_{2w}), proton density weighted, and fluid-attenuated inversion recovery (FLAIR) contrast images obtained from conventional MR and synthetically generated from quantitative (q) T_1 and (q) T_2 maps obtained with Magnetic Resonance Spin TomogrAphy in Time-Domain. Axial slices of interest from patients within the following groups: (a) tumor, (b) stroke, (c) epilepsy, or (d) multiple sclerosis. The lesions are indicated with a red arrow in the FLAIR contrasts.

including all 30 slices from a healthy participant and a patient of each group are in a publicly available repository (<https://gitlab.com/asbrizzi/mr-stat-synthetic-images>).

Legibility of Anatomic/Morphologic Features

All anatomic/morphologic features were legible in $\geq 96\%$ of participants for all data sets in both techniques, except for the

head of caudate nucleus (64%) and posterior limb of internal capsule (30%) on the conventional T_{1w} (Table 3).

Artifact Occurrence and Characterization

Artifact occurrence and characterization for each of the eight image contrast data sets are specified in Table 4. Overall, more motion artifacts were observed in conventional images (MR-STAT: 0% vs. Conventional: 21%), primarily in the

TABLE 3. Legibility of anatomic/morphologic features by contrast type and overall for conventional and synthetic images^a

Legibility of Feature ^b	T ₁ w		T ₂ w		PDw		FLAIR		Overall	
	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)
Central sulcus	100	100	100	100	100	100	98	100	100	100
Cerebral peduncle	100	100	100	100	100	100	98	100	100	100
Cervicomedullary junction	100	100	100	100	98	100	98	100	99	100
Head of caudate nucleus	64	100	100	100	100	100	98	100	91	100
Middle cerebellar peduncle	100	100	100	100	100	100	100	100	100	100
Posterior limb of internal capsule	30	100	100	98	100	100	98	96	82	99

Conv = conventional MR imaging (control); MR-STAT = synthetic contrasts from quantitative Magnetic Resonance Spin Tomography in Time-domain (MR-STAT) images; T₁w = T₁ weighted; T₂w = T₂ weighted; PDw = proton density weighted; FLAIR = Fluid Attenuated Inversion Recovery.

^aAll data are shown as n (n/N%), where n is the count and N is the total reads per category (N = 50 per category and 200 overall).

^bMedian across observers.

FLAIR images (50%). White pixels/spike noise were only observed in synthetic MR-STAT T₁w data sets (92%). Flow artifacts were observed in both techniques, but in 94% of all synthetic MR-STAT data sets compared to 31% in conventional data sets, specifically in the conventional T₁w (92%) and FLAIR (28%). An example of the white pixels/spike noise and flow can be observed in the synthetic T₁w (Fig. 1). Overall, more synthetic MR-STAT images were blurred (MR-STAT: 74% vs. conventional: 52%), only the synthetic FLAIR did not suffer from blurring. Additionally, susceptibility artifacts were more common in synthetic MR-STAT images (MR-STAT: 43% vs. conventional: 28%). Synthetic MR-STAT T₁w images did not contain susceptibility artifacts compared to 66% incidence in conventional T₁w images. In contrast, more susceptibility artifacts were noted in synthetic T₂w (MR-STAT: 54%, vs conventional: 2%), PDw (MR-STAT: 100% vs. conventional: 36%) and FLAIR (MR-STAT: 16% vs. conventional: 8%) contrasts.

Diagnostic Accuracy

Diagnostic accuracy is shown per level and patient group in Fig. 3. Overall, radiologists observed the correct pathology (level 3) more often in the conventional (77%) compared to synthetic MR-STAT (65%) datasets. The 12% difference between techniques was accounted to the following patient groups: tumor $\Delta = -3$ percentage point (pp), stroke $\Delta = -2$ pp, epilepsy $\Delta = -3$ pp, MS $\Delta = -4$ pp. However, the lesion in the MR-

STAT datasets was still observed but incorrectly classified (level 2; MR-STAT: 19% vs. conventional: 11%). In other words, true positive/negative rates were similar at 88% and 84% for the conventional and synthetic MR-STAT images, respectively. Radiologists were uncertain about a lesion (level 1) in 6% of MR-STAT datasets and 5% of conventional datasets. False negative (level 0) occurred in 10% of MR-STAT datasets and 6% of conventional datasets, while a lesion was incorrectly observed by one radiologist in a healthy participant in a conventional dataset.

Discussion

MR-STAT is a multi-parametric quantitative imaging technique that can be leveraged to synthesize multiple contrast-weighted images from one single 5-minute MR acquisition. In this clinical study, diagnostic image quality and accuracy of these MR-STAT-based synthetic images were compared to conventional contrast-weighted images. Image quality was similar between both techniques in acceptability for diagnostic use, while only the synthetic MR-STAT T₂w images were rated as non-inferior to the conventional images. In addition, more artifacts were identified (i.e., white pixels/spike noise, flow, blurring, and susceptibility) in the synthetic data sets, whereas motion artifacts did not occur. Strikingly, brain lesions on synthetic FLAIR images oftentimes appeared hypointense while being hyperintense on conventional FLAIR images; this might explain—at least in part—why the diagnostic accuracy was higher using the conventional images.

TABLE 4. Artifact identification and image quality issues characterization by contrast type and overall for conventional and synthetic images^a

Artifact Type ^b	T _{1w}		T _{2w}		PDw		FLAIR		Overall	
	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)	Conv (%)	MR-STAT (%)
Signal-encoding artifacts										
Motion	16	0	12	0	6	0	50	0	21	0
Infolding or wraparound	0	0	0	0	0	0	0	0	0	0
White pixels/spike noise	0	92	0	0	0	0	0	0	0	23
Flow	92	82	2	96	2	100	28	98	31	94
Phase-encoding	2	0	2	8	0	6	10	0	4	4
Quality issues										
Low resolution	0	2	0	0	0	2	0	0	0	1
Low SNR	0	0	0	0	0	0	0	0	0	0
Poor CNR	0	4	0	0	2	0	0	0	1	1
Blurring	8	100	60	96	36	100	0	0	26	74
Susceptibility	66	0	2	54	36	100	8	16	28	43

Conv = conventional MR imaging (control); MR-STAT = synthetic contrasts from quantitative MR-STAT images; Conv = conventional MR imaging (control); MR-STAT = synthetic contrasts from quantitative Magnetic Resonance Spin Tomography in Time-domain (MR-STAT) images; T_{1w} = T₁ weighted; T_{2w} = T₂ weighted; PDw = proton density weighted; FLAIR = Fluid Attenuated Inversion Recovery; SNR = signal-to-noise ratio; CNR = contrast-to-noise ratio.

^aAll data are shown n/N%, where n is the count and N is the total reads per category (N = 50 per category and 200 overall).

^bMedian across observers.

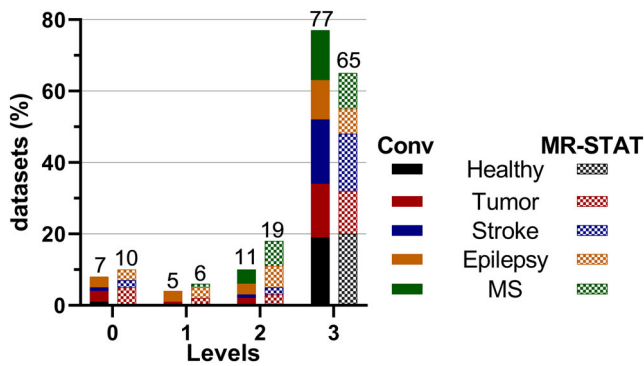


FIGURE 3: Diagnostic accuracy of conventional and synthetic datasets (Magnetic Resonance Spin Tomography in Time-Domain) consisting of a T_1 weighted (T_1w), T_2 weighted (T_2w), proton density weighted, and fluid-attenuated inversion recovery volume. Each dataset (10 healthy volunteers and 40 patients with neurological diseases; tumor, stroke, epilepsy and multiple sclerosis) was assessed by three radiologists, resulting in 150 cases per technique. The scores of the radiologists were compared to the reference diagnosis as determined in the hospital based on all patient information and additional scans if available. The accuracy per participant was divided into four levels: 0 = false positive (healthy volunteers) and false negative (patients); 1 = uncertain about a lesion (at least one pathological subtype was detected with 50% confidence); 2 = at least one lesion observed ($\geq 75\%$ confidence), but incorrect pathological subtype classification; 3 = correct classification of the lesion with at least 75% confidence for a true positive, or true negative classification of a healthy volunteer.

Image Quality

Previous studies have already assessed the diagnostic image quality of synthetically generated contrast-weighted images from quantitative parameter maps^{5,7,27,28} and several issues were observed particularly in FLAIR images.¹⁵ Tanenbaum et al reported a similar image quality of seven synthetically generated contrast-weighted data sets from an MDME-based acquisition compared to conventional images when the median across all contrasts was used.⁵ Therefore, differences per contrast may have been overlooked, whereas they did note that FLAIR images received lower quality scores. Two other studies^{27,28} using SyMRI diagnostic techniques⁶ reported a lower diagnostic image quality of synthetic T_1w , T_2w , and FLAIR. An additional study using an MRF-based synthetic procedure⁷ showed issues with the synthetic FLAIR images. In contrast to our results, synthetic FLAIR images from the study by Betts et al²⁸ were still inferior when results were dichotomized: only 82% were of acceptable diagnostic quality compared with 94% of the conventional FLAIR images. In general, MDME-derived synthetic FLAIR images were rated inferior due to CSF pulsation artifacts and hyperintense artifacts at tissue-CSF interface,^{5,15,27,29} which could lead to misdiagnosis.¹⁵

Analogously to above mentioned literature, also in our study the synthetization of FLAIR images revealed to be challenging. In particular, normally hyperintense brain lesions that appeared hypointense and were therefore challenging to distinguish on synthetic MR-STAT FLAIR images could have influenced the

diagnostic confidence of our observers. For MDME and MRF, partial volume effects were the proposed cause of the FLAIR issues, which could also apply to the FLAIR images synthesized from MR-STAT acquisition. In addition, possible magnetization transfer-related effects, which were not considered, may also be one of the underlying causes of these issues.

Alternative techniques, such as deep learning-based synthetization via quantitative maps,^{30,31} were proposed to try to overcome issues of model-based synthetic image quality. Improved FLAIR images were already generated from a MDME-based acquisition using a deep neural network with conventional FLAIR images as targets.³² After the analysis of this study, we also used deep learning based on the same dataset to optimize the image quality of synthetic contrast-weighted images using a physics-informed model, which led to promising results, although thorough assessment still needs to be performed.^{30–33} Synthetization directly from the acquired data also circumvents the computational reconstruction to multi-parametric maps for quicker visualization.

Artifact Occurrence and Characterization

Contrast-weighted images from both techniques were affected by artifacts, although different types were identified. The relatively short MR-STAT acquisition time (10 sec per slice) resulted in no motion artifacts for this technique, while half of the conventional FLAIR images suffered from motion artifacts. Motion artifacts in conventional FLAIR images may be caused by the longer acquisition time. Albeit overall less motion artifacts ($\Delta = 8$ pp) were observed in MDME-based synthetic contrast compared to conventional serial acquisition, still 21% had motion artifacts.⁵ Additionally, the single MR-STAT acquisition inherently prevents misregistration between contrasts, whereas interscan subject motion during conventional serial acquisitions hampers comparison between acquisitions. On the other hand, flow effects are intrinsic to the 2D gradient echo MR-STAT acquisition technique, which also leads to white pixels/spike noise in the synthetic MR-STAT T_1w contrast. Flow artifacts and white pixels/spike noise were also more common in synthetic MDME-based contrast weighted images.⁵ In addition, blurring of the synthetic MR-STAT images occurred during the synthetization procedure as a result of de-noising, which is a common problem.³⁴

Diagnostic Accuracy

The radiologists could more accurately identify the correct pathology on the conventional datasets, although in most incongruent cases the lesion was observed but not correctly classified on MR-STAT datasets. These results were consistent with the study from Tanenbaum et al,⁵ whereas two other studies reported similar diagnostic accuracy of synthetic compared to conventional images.^{27,35} However, in the study by Granberg et al³⁵ radiologists only differentiated MS

patients and healthy controls, while radiologists were given only three options (MS, ischemia, or non-specific) in the other study.²⁷ In our study, we increased the uncertainty for the classification of pathologies by increasing the diagnostic classification options to reduce bias in the readers. In a study by Ryu et al, MDME-derived synthetic images could be adequately used in daily practice to replace conventional contrasts.²⁹ This suggests that radiologic diagnosis was not affected by degradation of synthetic FLAIR images.

Future Perspectives

It would be of interest how the accuracy and precision of the quantitative maps affect the reliability and reproducibility of the synthetic contrast weighted images. Additionally, quantitative confidence maps could be reconstructed and related to image quality. The quantitative maps could also be improved by optimizing the variable flip angle scheme of the MR-STAT sequence by reducing the Cramer-Rao-based predictions of noise standard deviation values on quantitative maps.³⁶ Additionally, a 3D MR-STAT sequence could be developed to achieve higher signal-to-noise ratio and a 1 mm³ isotropic resolution,³⁷ and T₂ sensitivity could be increased by small quadratic radiofrequency phase increments.³⁸ Computation times could also be further reduced by using a GPU and Julia programming language for the reconstruction of the parameter maps.³⁹ Moreover, multi-parametric quantitative maps acquired with MR-STAT could provide new opportunities such as earlier detection of disease and monitoring lesion progression. Furthermore, the quantitative maps are expected to enhance data standardization; which is required for deep learning algorithms aimed at disease classification or lesion segmentation.² Finally, the application of MR-STAT for diagnostic decision making is still in development and could provide new and more sensitive biomarkers.⁴⁰ For this purpose, MR-STAT diffusion imaging and postcontrast sequences could also be investigated.

Limitations

The MR-STAT technique is characterized by shorter acquisition, but longer reconstruction times compared to conventional sequences. The quantitative maps are reconstructed offline, which currently takes about 2 minutes per slice leading to about 1 hour of processing time. While this time lag does not necessarily hamper the diagnostic performance, it would be desirable to reduce the computation time to just a few minutes, thereby facilitating the wide adoption of MR-STAT in the clinical workflow. We note that immediate image quality assessment can still be provided to the technicians by presenting the fast Fourier transformed images: these images reveal accurate anatomical information and thus anomalies in the data acquisition process that can be easily detected.

In addition, although a higher image contrast between gray and white matter could be achieved in the synthetic MR-STAT T_{1w} images, some of these images were unacceptable for diagnostic use due to an image contrast difference between slices, which was already present in the raw data.

Transmit RF field (B₁) variation was not taken into account in the reconstruction model, because potential B₁-induced biases in the T₂ quantification were negligible when generating the synthetic contrast weighted images.

We should also note that the current study was not designed to provide a confident and accurate diagnosis, as only T_{1w}, T_{2w}, FLAIR and PDw were supplied. For accurate diagnosis, additional information is required such as patient information, contrast-enhanced and diffusion imaging, which was not available in the scoring sessions. In addition, assessment of diagnostic quality would require a larger sample size and plethora of neurological diseases, which was beyond the scope of the current study. Moreover, the spatial resolution used in this study was approximating the one used in standard clinical exams of non-specialized hospitals when designing the study (2018). However, since then we note that spatial parameters and in particular slice thickness of conventional scans have decreased. Nowadays, 3D acquisitions are often used. Other study design limitations include the small sample size and only three experienced neuroradiologists rated the images; for clinical translation interpretation of general radiologists would be of additional value.

Conclusion

The synthetic MR-STAT images obtained from a single 5-minute scan time sequence were overall acceptable for diagnostic use, although the perceived image quality was rated lower—except for T_{2w} images—and in some cases the signal intensity of lesions differed in FLAIR images. These results highlight the potential of MR-STAT to provide clinicians with clinically useful image contrasts within substantially reduced acquisition time, in addition to multi-parametric quantitative maps.

Acknowledgments

We would like to thank the neuroradiologists for assessing the MR images, Carlo Lucci from the University Medical Center Utrecht for supporting the data acquisition, and Shanta Kalaykhan-Sewradj for her administrative work. Additionally, we would like to thank all involved research technicians and participants for their contribution to the study.

References

1. Seiler A, Nöth U, Hok P, et al. Multiparametric quantitative MRI in neurological diseases. *Front Neurol* 2021;12:640239.
2. Gulani V, Seiberlich N. Quantitative MRI: Rationale and challenges. In: Seiberlich N, Gulani V, Calamante F, et al., editors. *Advances in*

- magnetic resonance technology and applications*, Vol 1. Cambridge: Academic Press; 2020. xxxvii-li.
3. Callaghan MF, Mohammadi S, Weiskopf N. Synthetic quantitative MRI through relaxometry modelling. *NMR Biomed* 2016;29(12):1729-1738.
 4. Iglesias JE, Dinov I, Singh J, Tong G, Tu Z. Synthetic MRI signal standardization: Application to multi-atlas analysis. *Med Image Comput Comput Assist Interv* 2010;13:81-88.
 5. Tanenbaum LN, Tsiouris AJ, Johnson AN, et al. Synthetic MRI for clinical neuroimaging: Results of the magnetic resonance image compilation (MAGiC) prospective, multicenter, multireader trial. *AJNR Am J Neuroradiol* 2017;38:1103-1110.
 6. Gonçalves FG, Serai SD, Zuccoli G. Synthetic brain MRI: Review of current concepts and future directions. *Top Magn Reson Imaging* 2018; 27(6):387-393.
 7. Deshmane A, McGivney D, Badve C, et al. Accurate synthetic FLAIR images using partial volume corrected MR fingerprinting. *Proc Intl Soc Mag Reson Med* 2016;24:1909.
 8. Ma D, Gulani V, Seiberlich N, et al. Magnetic resonance fingerprinting. *Nature* 2013;495(7440):187-192.
 9. Panda A, Mehta BB, Coppo S, et al. Magnetic resonance fingerprinting-an overview. *Curr Opin Biomed Eng* 2017;3:56-66.
 10. Haacke EM, Chen Y, Utraiainen D, et al. Strategically acquired gradient echo (STAGE) imaging, part iii: Technical advances and clinical applications of a rapid multi-contrast multi-parametric brain imaging method. *Magn Reson Imaging* 2020;65:15-26.
 11. Cheng CC, Preiswerk F, Hoge WS, Kuo TH, Madore B. Multipathway multi-echo (MPME) imaging: All main MR parameters mapped based on a single 3D scan. *Magn Reson Med* 2019;81(3):1699-1713.
 12. Gómez PA, Molina-Romero M, Buonincontri G, Menzel MI, Menze BH. Designing contrasts for rapid, simultaneous parameter quantification and flow visualization with quantitative transient-state imaging. *Sci Rep* 2019;9(1):8468.
 13. Gómez PA, Cencini M, Golbabaee M, et al. Rapid three-dimensional multiparametric MRI with quantitative transient-state imaging. *Sci Rep* 2020;10(1):13769.
 14. Pirkl CM, Nunez-Gonzalez L, Kofler F, et al. Accelerated 3D whole-brain T1, T2, and proton density mapping: Feasibility for clinical glioma MR imaging. *Neuroradiology* 2021;63(11):1831-1851.
 15. Ji S, Yang D, Lee J, Choi SH, Kim H, Kang KM. Synthetic MRI: Technologies and applications in neuroradiology. *J Magn Reson Imaging* 2022;55(4):1013-1025.
 16. Mandija S, D'Agata F, Liu H, et al. A five-minute multi-parametric high-resolution whole-brain MR-STAT exam: First results from a clinical trial. *Proc Intl Soc Mag Reson Med* 2020;28:0558.
 17. Sbrizzi A, Heide OVD, Cloos M, et al. Fast quantitative MRI as a nonlinear tomography problem. *Magn Reson Imaging* 2018;46:56-63.
 18. van der Heide O, Sbrizzi A, Luijten PR, van den Berg CAT. High-resolution in vivo MR-STAT using a matrix-free and parallelized reconstruction algorithm. *NMR Biomed* 2020;33(4):e4251.
 19. Liu H, Heide OV, Mandija S, Berg CATV, Sbrizzi A. Acceleration strategies for MR-STAT: Achieving high-resolution reconstructions on a desktop pc within 3 minutes. *IEEE Trans Med Imaging* 2022;1. <https://doi.org/10.1109/TMI.2022.3168436>
 20. Warntjes JBM, Leinhard OD, West J, Lundberg P. Rapid magnetic resonance quantification on the brain: Optimization for clinical usage. *Magn Reson Med* 2008;60(2):320-329.
 21. Goodfellow I, Bengio Y, Courville A. *Deep learning*. Cambridge: MIT press; Vol 775; 2016.
 22. Vincent P, Larochelle H, Bengio Y, Manzagol P-A. *Extracting and composing robust features with denoising autoencoders. Proceedings of the 25th international conference on Machine learning*. Helsinki, Finland: Association for Computing Machinery; 2008. p 1096-1103.
 23. Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932;22:155.
 24. Banerjee AK. Diagnostic imaging: Brain. 2(nd) edition. *Br J Radiol* 2010;83(989):450-451.
 25. Noguchi K, Gel YR, Brunner E, Konietzschke F. Nparld: An R software package for the nonparametric analysis of longitudinal data in factorial experiments. *J Stat Softw* 2012;50(12):1-23.
 26. Mascha EJ, Sessler DI. Equivalence and noninferiority testing in regression models and repeated-measures designs. *Anesth Analg* 2011; 112(3):678-687.
 27. Blystad I, Warntjes J, Smedby O, Landtblom AM, Lundberg P, Larsson EM. Synthetic MRI of the brain in a clinical setting. *Acta Radiol* 2012;53(10):1158-1163.
 28. Betts AM, Leach JL, Jones BV, Zhang B, Serai S. Brain imaging with synthetic MR in children: Clinical quality assessment. *Neuroradiology* 2016;58(10):1017-1026.
 29. Ryu KH, Baek HJ, Moon JI, et al. Initial clinical experience of synthetic MRI as a routine neuroimaging protocol in daily practice: A single-center study. *J Neuroradiol* 2020;47(2):151-160.
 30. Virtue P, Tamir JI, Doneva M, Yu SX, Lustig M. Learning contrast synthesis from MR fingerprinting. *Proc Intl Soc Mag Reson Med* 2018;26: 0676.
 31. Moya-Sáez E, Peña-Nogales Ó, Luis-García RD, Alberola-López C. A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data. *Comput Methods Programs Biomed* 2021;210:106371.
 32. Hagiwara A, Otsuka Y, Hori M, et al. Improving the quality of synthetic FLAIR images with deep learning using a conditional generative adversarial network for pixel-by-pixel image translation. *Am J Neuroradiol* 2019;40(2):224-230.
 33. Jacobs L, Mandija S, Hongyan L, van den Berg CAT, Sbrizzi A, Maspero M. Generalizable synthetic multi-contrast MRI generation using physics-informed convolutional networks. *Proc Intl Soc Mag Reson Med* 2022;30:2850.
 34. Tian C, Fei L, Zheng W, Xu Y, Zuo W, Lin C-W. Deep learning on image denoising: An overview. *Neural Netw* 2020;131:251-275.
 35. Granberg T, Uppman M, Hashim F, et al. Clinical feasibility of synthetic MRI in multiple sclerosis: A diagnostic and volumetric validation study. *Am J Neuroradiol* 2016;37(6):1023-1029.
 36. Fuderer M, van der Heide O, Liu H, van den Berg CAT, Sbrizzi A. Non-steady-state sequences for multi-parametric MRI need to be evaluated in the context of gradient-encoding. *Proc Intl Soc Mag Reson Med* 2022;30:2786.
 37. Liu H, van der Heide O, Fuderer M, van den Berg CAT, Sbrizzi A. 3D MR-STAT: Towards a fast multi-parametric protocol with increased snr. *Proc Intl Soc Mag Reson Med* 2022;30:1348.
 38. Liu H, Bruijnen T, van der Heide O, Fuderer M, van den Berg CAT, Sbrizzi A. Increasing the T2 sensitivity of MR-STAT sequences by small quadratic RF phase increments. *Proc Intl Soc Mag Reson Med* 2022;30: 0625.
 39. van der Heide O, Sbrizzi A, van den Berg CAT. Faster bloch simulations and MR-STAT reconstructions on GPU using the Julia programming language. *Proc Intl Soc Mag Reson Med* 2021;29:3063.
 40. Weiskopf N, Edwards LJ, Helms G, Mohammadi S, Kirilina E. Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. *Nat Rev Phys* 2021;3(8):570-588.