NeuroImage 124 (2016) 157–167

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Comparison of accelerated T1-weighted whole-brain structural-imaging protocols

Pavel Falkovskiy [a,b,c,*], Daniel Brenner [d], Thorsten Feiweier [e], Stephan Kannengiesser [e], Bénédicte Maréchal [a,b,c], Tobias Kober [a,b,c], Alexis Roche [a,b,c], Kaely Thostenson [f], Reto Meuli [b], Denise Reyes [f], Tony Stoecker [d], Matt A. Bernstein [f], Jean-Philippe Thiran [b,c], Gunnar Krueger [b,c,g]

[a] Advanced Clinical Imaging Technology, Siemens Healthcare IM BM PI, Lausanne, Switzerland
[b] Department of Radiology, University Hospital (CHUV), Lausanne, Switzerland
[c] LTS5, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[d] German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
[e] Siemens AG, Healthcare Sector, Erlangen, Germany
[f] Mayo Clinic, Department of Radiology, MN, Rochester, United States
[g] Siemens Medical Solutions USA, Inc., Boston, MA, USA

ARTICLE INFO

Article history:
Received 11 April 2015
Accepted 11 August 2015
Available online 20 August 2015

ABSTRACT

Imaging in neuroscience, clinical research and pharmaceutical trials often employs the 3D magnetisation-prepared rapid gradient-echo (MPRAGE) sequence to obtain structural T1-weighted images with high spatial resolution of the human brain. Typical research and clinical routine MPRAGE protocols with ~1 mm isotropic resolution require data acquisition time in the range of 5–10 min and often use only moderate two-fold acceleration factor for parallel imaging.

Recent advances in MRI hardware and acquisition methodology promise improved leverage of the MR signal and more benign artefact properties in particular when employing increased acceleration factors in clinical routine and research. In this study, we examined four variants of a four-fold-accelerated MPRAGE protocol (2D-GRAPPA, CAIPIRINHA, CAIPIRINHA elliptical, and segmented MPRAGE) and compared clinical readings, basic image quality metrics (SNR, CNR), and automated brain tissue segmentation for morphological assessments of brain structures. The results were benchmarked against a widely-used two-fold-accelerated 3T ADNI MPRAGE protocol that served as reference in this study.

22 healthy subjects (age = 20–44 yrs.) were imaged with all MPRAGE variants in a single session. An experienced reader rated all images of clinically useful image quality. CAIPIRINHA MPRAGE scans were perceived on average to be of identical value for reading as the reference ADNI-2 protocol. SNR and CNR measurements exhibited the theoretically expected performance at the four-fold acceleration.

The results of this study demonstrate that the four-fold accelerated protocols introduce systematic biases in the segmentation results of some brain structures compared to the reference ADNI-2 protocol. Furthermore, results suggest that the increased noise levels in the accelerated protocols play an important role in introducing these biases, at least under the present study conditions.

Introduction

The 3D magnetisation-prepared rapid acquisition gradient-echo (MPRAGE) pulse sequence (Mugler and Brookeman, 1990) with Alzheimer's Disease Neuroimaging Initiative protocol (ADNI-2) (Jack et al., 2010) parameters is a well-established standard for multi-site and longitudinal MRI studies that involve T1-weighted imaging of the human brain. Both in routine clinical and research settings, reduced MRI scan times are desirable for increased patient throughput, improved patient comfort, and better management of patient motion. The 3T imaging protocol of the ADNI employs conventional two-fold parallel imaging (Griswold et al., 2002) to reduce the acquisition time of a whole-brain MPRAGE scan from 9 min in the original ADNI-1 protocol (Jack et al., 2008) to 5 min in ADNI-2. Recently, several strategies have been proposed to further reduce the acquisition time in 3D structural-brain-imaging protocols beyond conventional parallel imaging, e.g. 2D-GRAPPA (Blaimer et al., 2006), CAIPIRINHA (Breuer et al., 2006; Brenner et al., 2014), and segmented MPRAGE (Falkovskiy et al., 2013).

* Corresponding author at: Advanced Clinical Imaging Technology, Siemens Healthcare IM BM PI, Lausanne, Switzerland.
E-mail address: pavel.falkovskiy@epfl.ch (P. Falkovskiy).

http://dx.doi.org/10.1016/j.neuroimage.2015.08.026
1053-8119/

Because of the excellent tissue contrast, the 3D T1-weighted MPRAGE images are appreciated for radiological reading. Moreover, quantitative assessment of brain tissues and the volume of individual brain structures has become an important tool in more research-oriented applications of the MPRAGE (Miller et al., 2009; Camicioli et al., 2009; Jack, 2011; Mills and Tamnes, 2014; Schmitter et al., 2015). For instance, studies aiming at quantification of disease progression, drug efficiency or normal aging apply serial imaging to assess structural changes over time. Since the expected effect sizes in normal aging and disease (e.g. increased atrophy rates) are often subtle over time, it is vital to understand the reliability and reproducibility of the imaging-based quantitative measurements and to understand any inconsistencies that may appear when changing the pulse sequence between longitudinal repeat scans.

The reproducibility of the reference protocol (5-minute ADNI-2) has been studied extensively (Jovicich et al., 2006, 2009, 2013; Wonderlick et al., 2009; Krueger et al., 2012; Kruggel et al., 2010; Morey et al., 2010; Leung et al., 2014; Ching et al., 2015; Reuter et al., 2012). However, to our knowledge, the influence of further accelerations obtained by applying 2D-GRAPPA, CAIPIRINHA or segmented MPRAGE protocols on volumetric brain measurements and clinical readings has not been reported so far.

In this work, we aim to assess the reliability of the data obtained on the same platform across these accelerated protocols both qualitatively and quantitatively. The qualitative analysis is carried out by an experienced observer. The quantitative analysis is performed through assessing the reproducibility of volume measurements with automated brain segmentation algorithm (Schmitter et al., 2015; Roche et al., 2011) contrast-to-noise ratio (CNR) and signal-to-noise ratio (SNR) measurements (Robson et al., 2008).

Theory

Conventional GRAPPA- (Griswold et al., 2002) or SENSE-type accelerations (Pruessmann et al., 1999) are often used in clinical settings. Both methods obtain acceleration by undersampling of the k-space data in one phase-encoding direction (1D). In most cases, only moderate accelerations are employed due to increased aliasing and noise amplification (g-factor penalty) with higher acceleration factors. In order to overcome those limitations partly, and assuming that there is sufficient SNR, it is possible to further generalize the conventional GRAPPA approach to 2D for higher accelerations (Blaimer et al., 2006). A modified undersampling pattern used in 2D-GRAPPA is illustrated in Fig. 1a.

2D-GRAPPA allows for a more effective exploitation of the coil sensitivities of contemporary high-channel-count receive arrays which vary in both phase-encoding directions, resulting in a reduced g-factor compared to 1D-GRAPPA with the same total acceleration. This advantage can be further improved by using the CAIPIRINHA (Breuer et al., 2006; Brenner et al., 2014) approach, which shifts the aliasing pattern by introduction of an additional CAIPIRINHA shift parameter (Fig. 1b).

In addition to parallel imaging techniques, a common approach to reduce scan time is to use elliptical scanning (Bernstein et al., 2001). Elliptical scanning does not sample the corners of k-space and hence reduces the acquisition time.

A segmented MPRAGE approach accelerates the acquisition using a different method which does not rely on coil sensitivity variations in contrast to parallel imaging techniques (Falkovskiy et al., 2013). The standard acquisition scheme is modified to allow acquisition of multiple echoes with bipolar gradient readouts. Between the readouts, phase-encoding blips are inserted so that multiple portions of k-space are sampled per excitation. The echoes are grouped so that the first echoes are acquired in the center of k-space to ensure the intended contrast (Fig. 1c). A phase correction algorithm similar to the one used for EPI acquisitions is applied to remove phase inconsistencies between k-space lines acquired in the different echoes with differing polarity. This correction is computed based on the lines acquired in the center of k-space. Notably, this segmented approach can be introduced in addition to parallel-imaging-based acceleration.

Materials and methods

MR acquisition

All experiments were performed on a standard clinical 3T MRI (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) equipped with a 32-channel head coil array.

The measurement protocol consisted of five 3D MPRAGE volume acquisitions with protocol parameters similar to the ADNI-2 MPRAGE protocol settings (TR/TI = 2300/900 ms, $\alpha = 9$ deg., BW = 240 Hz/pixel, readout in superior-inferior direction) (Jack et al., 2010) but with 1 mm isotropic resolution at a FoV of $256 \times 240 \times 176$ mm^3: (a) T1w ADNI-2 protocol with 2-fold GRAPPA acceleration (TA = 5:12 min) (Jack et al., 2010) which is used as reference and that we will refer to as reference ADNI-2; (b) 2D-GRAPPA with 4-fold (2×2) undersampling (TA = 2:59 min) (Blaimer et al., 2006); (c) CAIPIRINHA with 4-fold undersampling (2×2 shift 1 in z direction) (TA = 2:59 min) (Brenner et al., 2014); (d) CAIPIRINHA same as *c* but with additional elliptical scanning option, further reducing scan time (TA = 2:40 min) (Brenner et al., 2014); (e) segmented MPRAGE with 4-fold acceleration based on combining conventional parallel imaging (2-fold) and a two-echo segmented acquisition (BW = 480 Hz/pixel, TA = 3:15 min) (Falkovskiy et al., 2013). All 1-D undersampled scans utilized 32 reference lines for reconstruction of the GRAPPA weights, while the 2-D undersampled scans utilized a 32×32 lines reference region.

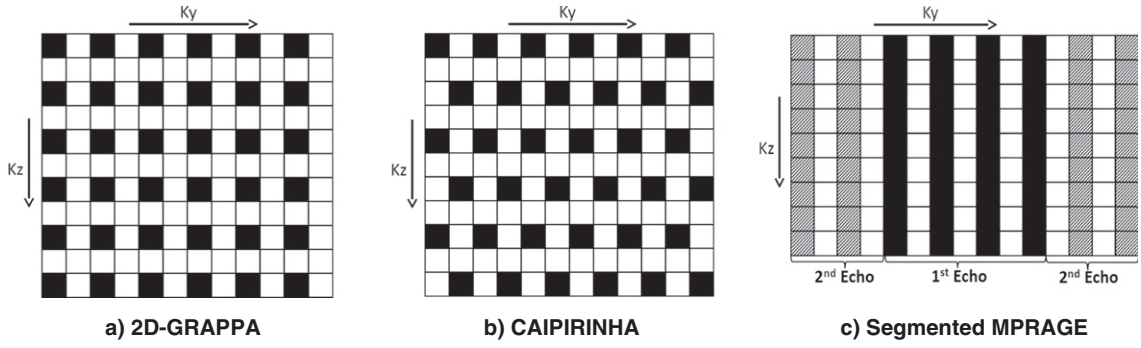

Fig. 1. K-space sampling patterns for the accelerated MPRAGE variants. Solid squares represent acquired lines. Squares with stripes represent lines acquired with a second echo. (a) 2D-GRAPPA acceleration with 2×2 undersampling; (b) CAIPIRINHA with 2×2 undersampling and shift 1; (c) assignment of echoes in segmented MPRAGE.

For the segmented MPRAGE acquisition, the bandwidth was increased to maintain the echo spacing equal to the one of the reference ADNI-2 protocol. With the exception of the reference T1w ADNI-2 measurement, all scans were acquired using in-house prototype sequences. The order of the scans was randomized between sessions.

22 subjects (12 male and 10 female, age 20–44 years, mean 30 ± 6.0 years) were imaged. The subjects were screened prior to enrolment in this study based on a health-assessment questionnaire. Only the subjects in good health and without a history of neurological diseases were considered. The exclusion criteria included: any known active medical conditions; hypertension; excessive smoking; excessive alcohol consumption; drug addiction; diabetes mellitus; history of significant vascular events (i.e., myocardial infarction, stroke or peripheral vascular disease); history of malignant neoplasia of any form; history of cardiac, lung, liver, or kidney failure; active or inadequately treated thyroid disease; active neurological or psychiatric conditions; and history of head trauma with loss of consciousness.

After obtaining written consent, each subject was scanned with the five MPRAGE variants within one session under an IRB-approved protocol. FOV placement was guided by the scanner-integrated AutoAlign feature. Intensity normalization was performed using a scanner software-integrated pre-scan procedure and an adaptive coil combination mode (Walsh et al., 2000) was used to reconstruct the image volumes for all of the protocols used in this study.

Image processing pipeline and analysis

Qualitative

The data were graded by an experienced image analyst from the ADNI MR Core. The analyst was blinded with respect to protocols.

First, each volume was rated with respect to the presence of motion artefacts (none, mild, severe) and overall image quality on the scale of 1–3 based on the presence of blurring/ringing/ghosting or other artefacts (1: good, 2: fair, 3: very low quality). No separation on the nature of artefacts that contributed to overall image quality score was made.

Second, following this ranking, if two or more image volumes were assigned the same image quality grade, image volumes were presented side by side and the observer made a subjective decision which image had higher image quality relative to each-other. In this fashion, relative rankings with-in a session were constructed with a relative scale (1: best, 5: worst). The grade of 1 is assigned to the volume with the best image quality in the session and 5 to the volume with the worst image quality relative to the other scans in the session. Note that it should be considered as a limitation of these qualitative rankings the fact that the observer did not have to justify his choice, and only a single observer was used. These scores were averaged across the sessions, and statistical difference between the protocol variants was tested with the Wilcoxon ranksum test (Wilcoxon, 1945) since we did not expect scores to be normally distributed.

Quantitative

The data was passed through the in-house-developed automatic segmentation framework MorphoBox (Reuter et al., 2012; Roche et al., 2011) to compute the volumes of a number of brain tissues and structures: total intracranial volume (TIV), grey matter (GM), cortical grey matter (cGM), white matter (WM), hippocampus, thalamus, caudate, putamen, pallidum and brain stem. Label masks that define voxels belonging to a particular tissue type or structure were also computed.

In order to assess the consistency of the volumetric observations, intra-class correlation coefficients (ICC) (Shrout and Fleiss, 1979) between all possible combinations of protocol variants were computed. The model used in this study for ICC computation treated subjects as randomly sampled from a larger population and the acquisition protocols employed in this study as fixed parameters.

In order to evaluate a potential systematic bias in the volumetric results, normalized volumetric differences (D) between the reference protocol (V_r) and each variant (V_v)were computed for each structure as:

$$D(V_r, V_v) = \frac{Vv - Vr}{Vr + Vv},$$

where $D(V_r, V_v)$ is in the range $[-1,1]$.

Normalized volumetric differences were averaged across the subjects. All possible choices of the reference protocol (V_r) were examined. The statistical significance of the difference from a zero median in normalized volumetric differences was tested using Wilcoxon signed-rank test (Wilcoxon, 1945) because the differences were not expected to be normally distributed.

Based on the ICCs and normalized volume differences, the most consistent accelerated protocol was selected when compared to the reference ADNI-2 protocol. For this protocol, the normalized volume differences were recomputed using the methodology described in Jovicich et al. (2013) and then compared to the scan–rescan reproducibility study of reference ADNI-2 protocol (Jovicich et al., 2013).

The gold-standard method for SNR computation is the multiple-replica approach that consists of imaging the same object several times. Voxel-wise standard deviations of noise (σ_{noise}) and mean intensities (μ_{signal}) are estimated from those measurements and voxel-wise SNR maps ($\mu_{signal}/\sigma_{noise}$) are constructed. However, the multiple-replica approach is not practical for calculating SNR maps for in-vivo imaging due to patient motion and excessive measurement time. For this reason, pseudo multiple-replica approach (Robson et al., 2008) was used instead, as it just requires knowledge of the noise correlation between coils. It allows the calculation of SNR maps from a single acquisition by mimicking the multiple replica method. First, synthetically generated random noise is scaled and correlated across coils based on the receive coil covariance matrix. The noise covariance matrix was measured with the subject and coil setup unchanged, by acquiring 128 readouts having 512 sampling points each, including oversampling, without an excitation pulse. This correlated and scaled noise is injected into the data at the beginning of the image reconstruction pipeline to produce a stack of images with different noise. The "true" acquisition noise from the subject is still present in this synthetic data and is assumed to be a part of the signal. From this stack of images, the voxel-wise means (μ_{signal}) and standard deviations (σ_{noise}) are computed, and SNR maps ($\mu_{signal}/\sigma_{noise}$) are produced.

The SNR computations were performed for all acquired 3D MPRAGE volumes with 8 pseudo-replicas. To improve the estimation of σ_{noise}, a 2D 5×5 region of neighboring pixels was used to estimate σ_{noise} in a moving-average fashion according to Wiens et al. (2011).

Individual SNR maps were spatially normalized (Roche et al., 2000) to the reference ADNI-2 image volume within the same session. Using the label maps from the reference ADNI-2 datasets, mean SNR values were computed for the following brain tissue types and structures: brain stem (chosen due to its proximity to the center of the head coil; representing the smallest expected SNR), hippocampus, white matter (representing an average of expected SNR), and cortical grey matter (chosen due to its proximity to the coil elements; representing the biggest expected SNR). The differences in SNR of different protocol variants were tested with the Wilcoxon rank-sum test (Wilcoxon, 1945) with Bonferroni correction for multiple comparisons (Dunn, 1961).

SNR efficiency can be defined as: $\mathrm{SNR}/\sqrt{\mathrm{TA}}$, where TA is total acquisition time. SNR efficiency was computed for all protocol variants and differences were tested with the Wilcoxon rank-sum test (Wilcoxon, 1945) with Bonferroni correction for multiple comparisons (Dunn, 1961).

In order to assess the contrast-to-noise ratio (CNR), data were spatially normalized (Roche et al., 2000) to the reference ADNI-2 image volume within session. In addition to intensity normalization performed using a scanner software-integrated pre-scan procedure, a separate B1

bias field correction was applied to all of the datasets (Van Leemput et al., 1999). This additional bias field correction was shown to be very robust in 3T settings. CSF, WM and GM were masked using the label masks from the reference ADNI-2 image volumes.

Two CNR values were computed as follows:

$$CNR_{\text{csf-gm}} = \frac{\left(\mu_{csf} - \mu_{gm}\right)^2}{\sigma^2{}_{csf} + \sigma^2{}_{gm}},$$

$$CNR_{\text{gm-wm}} = \frac{\left(\mu_{gm} - \mu_{wm}\right)^2}{\sigma^2{}_{gm} + \sigma^2{}_{wm}}$$

where μ_{csf}, μ_{gm}, μ_{wm} are mean intensities and $\sigma^2{}_{csf}$, $\sigma^2{}_{gm}$, $\sigma^2{}_{wm}$ are image intensity variances within CSF, GM and WM volumes. The differences in CNR of different protocol variants were tested with the Wilcoxon rank-sum test (Wilcoxon, 1945) with Bonferroni correction for multiple comparisons (Dunn, 1961).

CNR efficiency can be defined as: $CNR/\sqrt{TA}$, where TA is total acquisition time. CNR efficiency was computed for all of the protocols used in this study and differences were tested with the Wilcoxon rank-sum test (Wilcoxon, 1945) with Bonferroni correction for multiple comparisons (Dunn, 1961).

It is expected that the SNR will be decreased in accelerated acquisitions when compared to the reference ADNI-2 protocols, at a first-glance approximation, proportionally to the square-root of the acceleration factor. The decrease in SNR can be attributed to the lower total acquisition time. A further factor is the reconstruction-induced noise amplification mediated by the g-factor. Furthermore, the different protocol variants can induce a different signal modulation in k-space.

In terms of GM-WM and CSF-GM contrast, the inversion recovery curve with the segmented MPRAGE acquisition is mapped along the partition-encoding axis, and undersampling is performed along the phase-encoding axis. Therefore, the resulting contrast is not expected to change compared to the standard MPRAGE. In 2D-GRAPPA and CAIPIRINHA acquisitions, the inversion recovery curve is mapped mostly along the partition-encoding, and the changes in contrast are consequently expected to be minimal (Brenner et al., 2014; Busse et al., 2008). We expect that the changes in CNR using accelerated acquisitions are driven by changes in noise (due to different sampling durations between the reference and the four-fold accelerated scans) rather than by changes in contrast. Therefore, the last part of our image processing pipeline only addresses the noise-related effects on the results obtained with MorphoBox.

To study the noise dependency of the segmentation results, the following numerical experiment was designed: scaled noise was added to the raw data of the reference ADNI-2 protocols to mimic the SNR performance of the accelerated protocols. Based on the SNR measurements within-session, the average SNR value was calculated between 4-fold accelerated protocols. The level of noise added to ADNI-2 data was set to match the average SNR performance of 4-fold accelerated protocols on a per session basis. These synthetic data were reconstructed and then passed on to the automated segmentation framework. This procedure was repeated 32 times for each subject to produce 32 synthetic image volumes and segmentation results per subject.

The subsequently obtained volumetric results from synthetic data of the different brain structures, and tissue types were pooled from all the subjects. ICCs between synthetic volumetric results were used as a measure of consistency of data. The model used to assess the consistency of the synthetic noise experiment considered subjects to be randomly sampled from a larger population of subjects and 32 repetitions with additional noise to be judges that are sampled from a bigger population of judges (Shrout and Fleiss, 1979). Normalized volumetric differences between synthetic data and the reference ADNI-2 protocol were calculated to assess whether there is a systematic bias in the measurements. The statistical significance of the difference in normalized volumetric differences was tested using the Wilcoxon signed-rank test (Wilcoxon, 1945).

Results

Each row in Fig. 2 demonstrates images from the same subject (three shown in total) obtained within one session using the reference ADNI-2 protocol and the 3-minute accelerated protocol variants. The

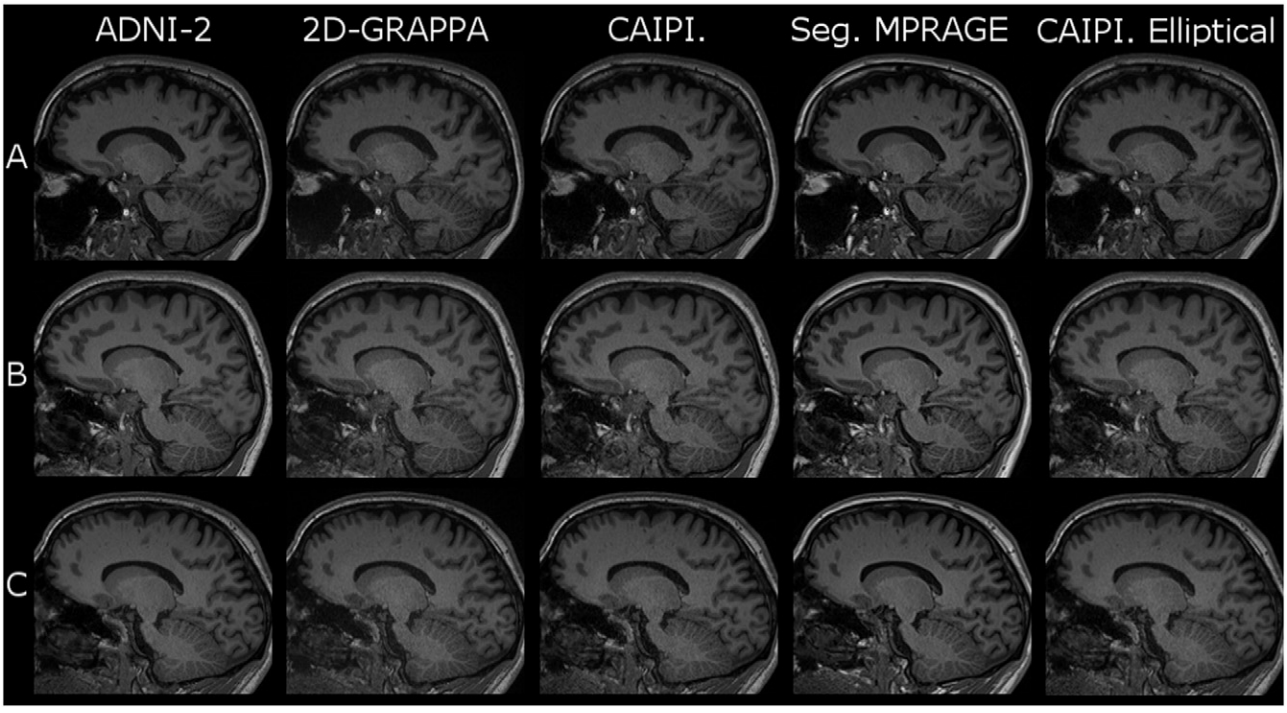

Fig. 2. Sample images showing representative sagittal views of all protocol variants for 3 subjects (A, B, C). Please note the increased noise level of the accelerated protocols that is most visible in the brain stem region of the subjects.

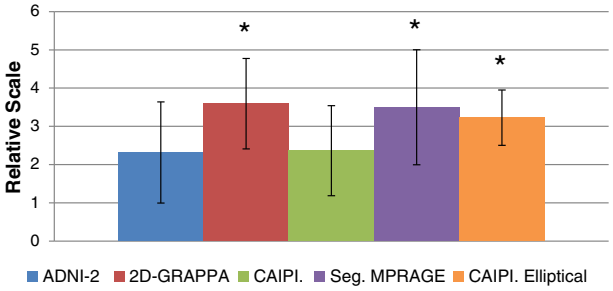

Fig. 3. Observer rankings. Relative scale (1, best to 5, worst). * indicates significant difference from the ADNI-2 protocol.

windowing was set to the same level for all shown images. In the brain stem region, there is a visible increase of noise in accelerated protocols when compared to the reference ADNI-2 protocol.

Observer ratings

All raw images before noise injection were rated by the observer as diagnostically useful with high image quality (76 image volumes) or mild image quality issues (34 image volumes). Segmented MPRAGE acquisitions exhibited, in some cases, mild ringing artefacts largely arising from hyper-intense signals due to abundant fat in the neck. No systematic pattern of image quality issues was reported by the observer. Overall, qualitative observations yield that all protocols provide clinically useful image quality.

The mean observer rankings (1, best to 5, worst) are shown in Fig. 3. Note that the 3-minute CAIPIRINHA MPRAGE scans are perceived on average of identical value for radiological reading as the reference 5-minute ADNI-2 scans despite a subtle, but visible noise degradation. However, the other protocols appear statistically different ($p < 0.05$), as assessed by the Wilcoxon rank-sum test.

Volumetric measurements

Visually, the segmentation results from all protocols used in this study display a high degree of similarity (Fig. 4). The segmentation errors are at the single voxel level, and therefore hard to visually notice. Visual inspection suggests that most of these differences occur in regions affected by partial-volume effects. All segmentation results were visually inspected for large segmentation errors and no gross segmentation errors have been observed in this study. It is important to note that MorphoBox was run in a fully automated fashion and no manual editing was applied at any stage of the segmentation process.

Intra-class correlation analysis shows highly consistent volumetric measurements of accelerated protocols when compared to the reference ADNI-2 protocol (Table 1).

Fig. 5 illustrates the intra-class correlation coefficients between all protocols used in this study. Each column represents an ICC between two protocol variants and each row represents a structure. The CSF volumes in segmented MPRAGE acquisitions exhibit smaller correlation coefficients relative to the other protocol combinations. Pallidum intra-class correlation coefficients are systematically smaller than the correlation coefficients of other structures.

Fig. 6 summarizes the outcome of the analysis of normalized volumetric differences when taking computed volumes from the reference ADNI-2 protocol as a reference volume (V_r). The label "ADNI-2 + NOISE" corresponds to a synthetic dataset generated through addition of noise to match the SNR of accelerated protocols. There is a strong indication that the increase in noise levels introduces a systematic bias to the computed volumes of some structures when compared to the reference volume of the ADNI-2 protocol. Most apparent, a statistically significant increase in the white matter volumes in the accelerated acquisitions and a trend towards decreased cortical grey matter volumes were observed in this investigation.

Fig. 7 summarizes the outcome of the normalized volumetric differences analysis for all possible choices of the reference volume (V_r). Each column represents the outcome of computing the normalized volume difference between two protocol variants used in this study and averaging it across the subjects. The results indicate that the segmented MPRAGE acquisitions introduce a change to estimated CSF volumes when compared to other accelerated protocol variants.

Signal-to-noise ratio (SNR)

Fig. 8 demonstrates sagittal SNR maps computed with the pseudo multiple-replica method (Robson et al., 2008) for a single subject within one session. Theoretically, a decrease in SNR proportional to $\frac{1}{\sqrt{2}}$ (reducing the acquisition time from 5 min from 3 min) is expected in the accelerated acquisitions when compared to the reference ADNI-2 protocol. Obtained SNR maps demonstrate the expected decrease of SNR in the accelerated protocols.

The results of the SNR analysis performed within different brain substructures are summarized in Fig. 9. Mean SNR values of 50 (white matter), 29 (cortical grey matter), 22 (brain stem) and 20 (hippocampus structures) were observed, using the conventional ADNI-2 protocol. There is a statistically significant drop in SNR between ADNI-2 and four-fold-accelerated protocols. SNR dropped consistently by around 34%, 36%, 31%, and 34% for the accelerated variants: 2D GRAPPA, CAIPIRINHA, segmented MPRAGE and CAIPIRINHA with elliptical scanning. White matter SNR is significantly different between 2D-GARPPA and segmented MPRAGE acquisitions. The rest of the four-fold-accelerated protocols are not significantly different from each other.

In terms of SNR efficiency, Fig. 9 illustrates performance of the protocols used in this study and significant differences. With the conventional ADNI-2 protocol, average SNR efficiency values of 2.8 (white matter), 1.7 (cortical grey matter), 1.2 (brain stem) and 1.1 (hippocampus structures) were observed. Noteworthy is a decrease in SNR efficiency of 2D-GRAPPA, CAIPIRINHA, and segmented MPRAGE acquisitions in white matter, cortical grey matter and hippocampus regions when compared to the ADNI-2 protocol. CAIPIRINHA with elliptical scanning was not statistically different from the ADNI-2 protocol in terms of SNR efficiency.

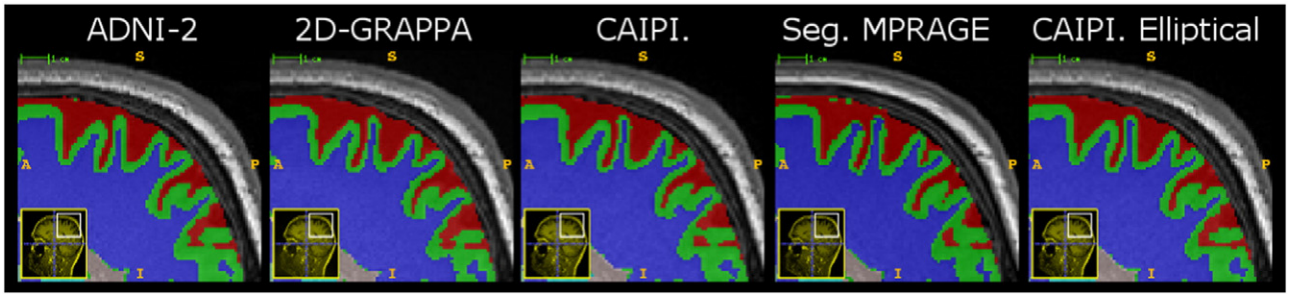

Fig. 4. Segmentation results from a single subject within one session. Blue: white matter; Green: grey matter; Red: CSF.

Table 1
Intra-class correlation coefficients (95% confidence interval shown in brackets).

	ADNI-2 vs 2D-GRAPPA	ADNI-2 vs CAIPI	ADNI-2 vs Seg. MPRAGE	ADNI-2 vs CAIPI. Elliptical
TIV	1.00	1.00	0.99 [0.98–1.00]	1.00
Cortical grey matter	0.96 [0.90–0.98]	0.97 [0.93–0.99]	0.93 [0.84–0.97]	0.97 [0.92–0.99]
White matter	0.99 [0.98–1.00]	0.99 [0.97–1.00]	0.99 [0.97–1.00]	1.00 [0.99–1.00]
Hippocampus	0.94 [0.85–0.97]	0.95 [0.88–0.98]	0.94 [0.87–0.98]	0.94 [0.86–0.97]
Thalamus	0.98 [0.94–0.99]	0.98 [0.96–0.99]	0.97 [0.93–0.99]	0.97 [0.94–0.99]
Putamen	0.96 [0.90–0.98]	0.97 [0.94–0.99]	0.95 [0.89–0.98]	0.95 [0.88–0.98]
Caudate	0.98 [0.95–0.99]	0.99 [0.97–0.99]	0.98 [0.94–0.99]	0.98 [0.94–0.99]
Pallidum	0.89 [0.76–0.95]	0.94 [0.86–0.97]	0.88 [0.74–0.95]	0.85 [0.68–0.94]

In the brain stem region, only 2D-GRAPPA and CAIPIRINHA appeared to be statistically different. On average, SNR efficiency decreased by 12%, 15%, and 13% for 2D-GRAPPA, CAIPIRINHA, and segmented MPRAGE acquisitions when compared to ADNI-2 protocol.

Contrast-to-noise ratio (CNR)

The results of the CNR assessment of all protocols used in this study are shown in Fig. 10. The greater the value of CNR, the better is the separation between the intensity distributions of the tissues, which simplifies the segmentation problem. Analogously to the SNR measurements, contrast to noise is decreased in accelerated acquisitions. There is a statistically significant decrease in CNR between ADNI-2 and four-fold-accelerated protocols. On average, CNR in the conventional ADNI-2 protocol was found to be 4.47 and 7.16 for CSF-GM and GM-WM, respectively, which significantly decreased on average by 29%, 26%, 38% and 28% in the 2D-GRAPPA, CAIPIRINHA, segmented MPRAGE, and elliptical CAIPIRINHA approaches. Both CSF-GM and GM-WM CNR appear to be significantly different between segmented MPRAGE and other four-fold-accelerated variants. The rest of the four-fold-accelerated protocols are not significantly different from each other.

Fig. 10 illustrates CNR efficiency of the protocols used in this study. CNR efficiency in the conventional ADNI-2 protocol was found to be 0.25 and 0.41 for CSF-GM and GM-WM. Segmented MPRAGE appeared to be statistically different from the rest of the protocols in terms of CSF-GM and GM-WM CNR efficiency. 2D-GRAPPA is statistically different from the ADNI-2 protocol in terms of GM-WM CNR efficiency.

Noise effects

In order to understand whether changes in the noise or signal properties cause the observed changes in SNR and CNR, Fig. 11 illustrates the mean changes in the distance between the peaks of intensity distributions of the tissues and the sum of variances of those tissues as extracted by the MorphoBox. A statistically significant increase in sums of variances of $\sigma^2_{csf}+\sigma^2_{gm}$ and $\sigma^2_{wm}+\sigma^2_{gm}$ between ADNI-2 and CAIPIRINHA, segmented MPRAGE, and elliptical CAIPIRINHA acquisitions is observed. There is a statistically significant increase in $\sigma^2_{wm}+\sigma^2_{gm}$ sums of variances between ADNI-2 and 2D-GRAPPA acquisitions. Both sums of variances, $\sigma^2_{csf}+\sigma^2_{gm}$ and $\sigma^2_{wm}+\sigma^2_{gm}$ appear to be significantly different between segmented MPRAGE and other four-fold-accelerated variants. Changes in the distance between the peaks of the intensity distributions were not significantly different.

Numerical experiment with addition of synthetic noise

For each subject, 32 volumes of synthetic data were computed by adding synthetic noise to the reference ADNI-2 raw data to reduce the SNR by approximately 35%, i.e. to match the SNR of the four-fold accelerated scans.

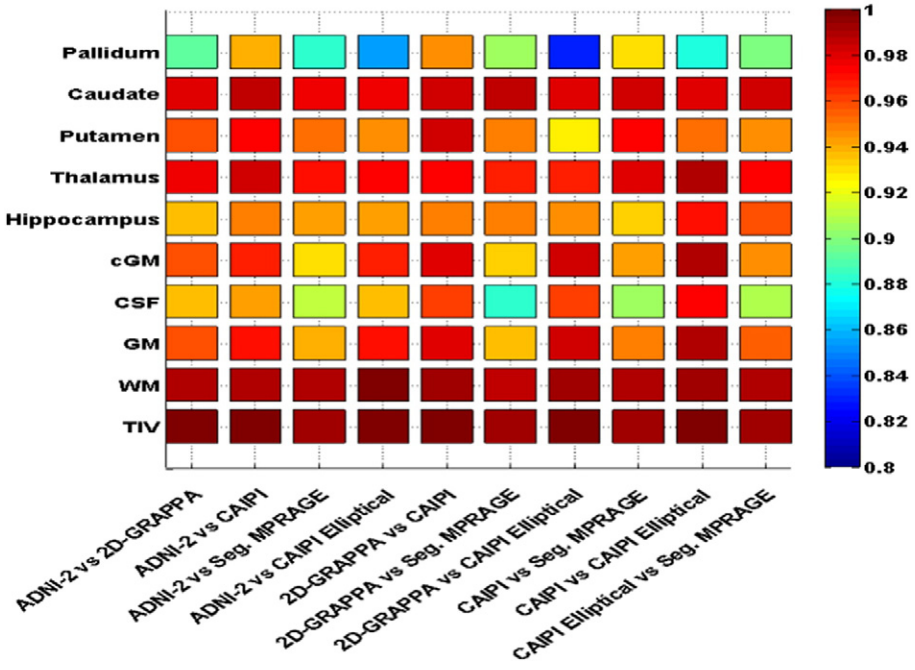

Fig. 5. Intra-class correlation coefficients for different brain structures between all possible protocol combinations.

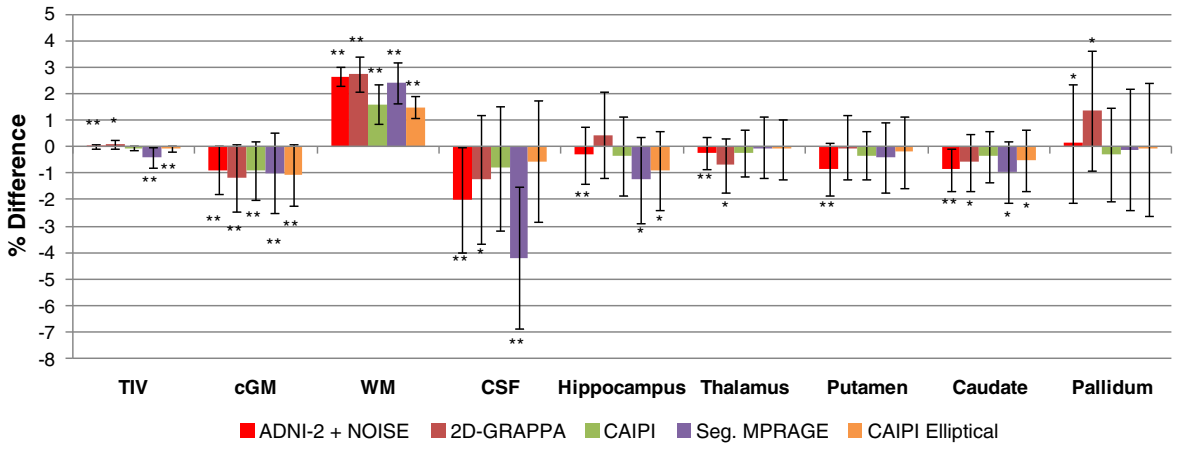

Fig. 6. Volumetric percent difference with ADNI-2 protocol used as reference scan (V_r). * indicates difference from 0% median at the 5% significance level. *$p \sim 10^{-2}$; **$p < 10^{-4}$.

Table 2 illustrates ICCs between the segmentation results of synthetic image volumes. For all examined structures, ICC values greater than 0.93 were observed. The estimated ICC of pallidum volumes exhibited the smallest correlation coefficient of 0.93 when compared to other structures.

Introduction of artificial 35% noise amplification to the reference ADNI-2 data introduced a bias in the volumetric data similar to the bias level observed with the accelerated acquisitions (Fig. 6). A statistically significant increase in the measurements of white matter volumes and a trend towards decreased cortical grey matter volumes are evident when compared to the reference ADNI-2 volumes.

Discussion

In longitudinal studies that use morphometric assessments of brain tissues and structures, the choice of the imaging protocol can potentially influence qualitative readings and degrade reproducibility of serial automated brain segmentations. The main objective of this study was to investigate the impact that accelerated protocols (2D-GRAPPA, CAIPIRINHA, segmented MPRAGE, and CAIPIRINHA with elliptical scanning) have on the observer's qualitative readings and on an automated brain segmentation procedure.

Qualitative analysis demonstrated that all of the data obtained using the accelerated 3-minute protocols have clinical value, even if artefacts are present and despite the changes in noise distribution. This observation can be explained by the fact that radiologists are used to "read through" artefacts and indicates that, in these settings (3T with a 32-channel head coil), the "gold-standard" reference ADNI-2 protocol provides an SNR value that is sufficiently high. It is important to note that the radiological readings of the CAIPIRINHA scans were perceived to provide on average the same diagnostic image quality as the reference ADNI-2 scans despite a 36% reduction in SNR. This may further indicate that the images obtained with the conventional ADNI-2 protocol using a 32-channel coil operate in an SNR regime and/or have the artefact-to-noise ratio that provides an image quality beyond the minimum routine reading requirements. Also, the shorter scan time can reduce motion sensitivity and may lead on average in the longer scans to some subtle degradation of the overall image quality.

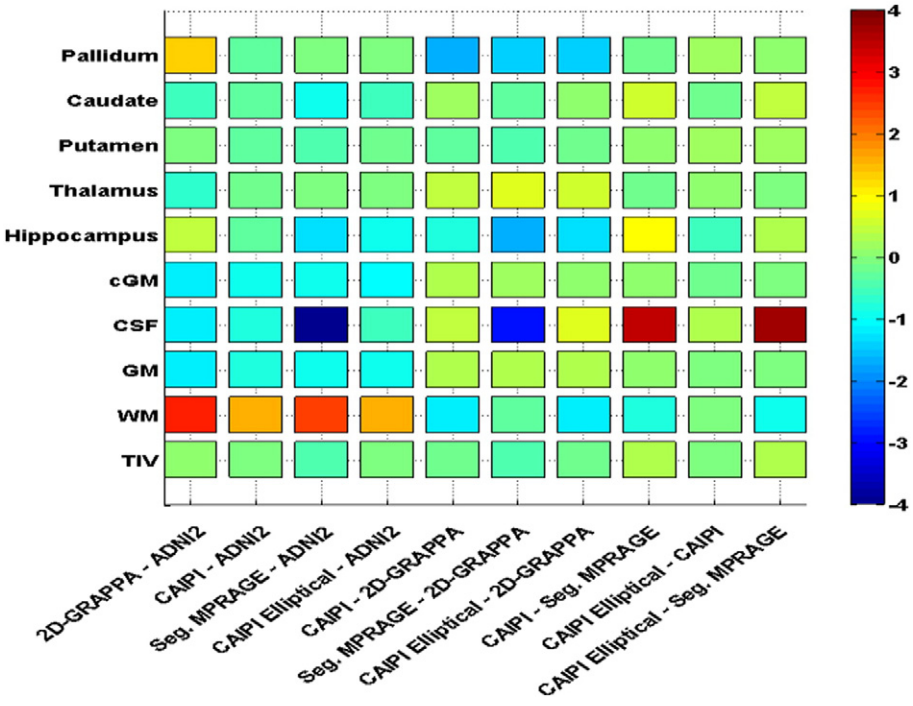

Fig. 7. Volumetric percent difference with all possible choices of reference protocol (V_r).

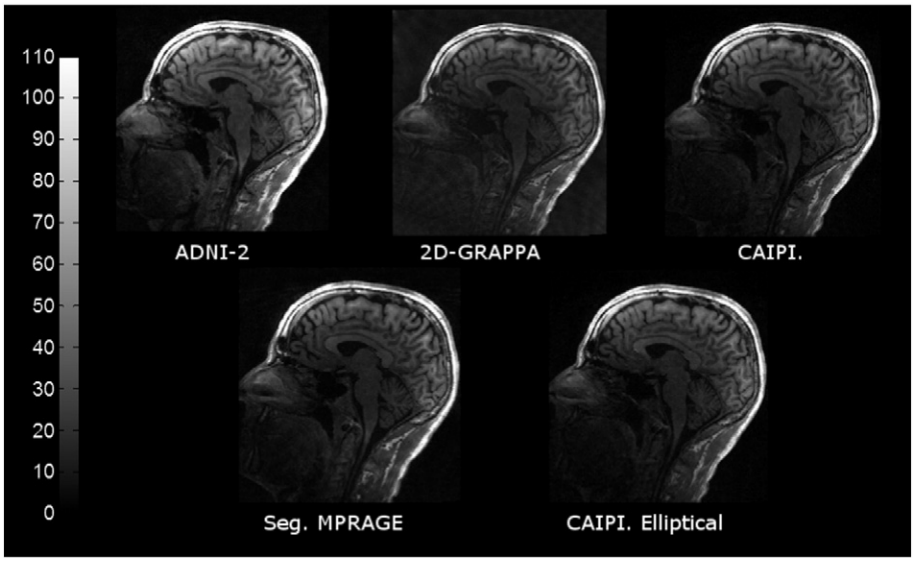

Fig. 8. Voxel-wise SNR maps obtained using the pseudo multiple-replica method for a single subject within one session.

However, when volumetric data acquired with different protocols are pooled in longitudinal studies, one has to be cautious of potentially introducing a bias to any quantitative analysis. In our analysis, we found such a bias of the order of 2% in white matter and of the order of 1% in cortical grey matter volumes when using the accelerated protocols.

If we look at the specific structures, for instance, we found that the hippocampus volume assessment differs most when comparing the reference ADNI-2 standard with the segmented MPRAGE and CAIPIRINHA with elliptical-scanning acquisition protocols. Pallidum volumes were most affected by 2D-GRAPPA and caudate volumes were most affected by segmented MPRAGE when compared to the reference ADNI-2 protocol.

Overall, the most consistent accelerated acquisition scheme for morphological analysis when compared to the reference ADNI-2 protocol is CAIPIRINHA, since the systematic changes in volumes that were observed are the smallest and a strong correlation between volumes is present. However, this MPRAGE variant can only be applied in the situations when multi-channel coils are used and if there is a sufficient variation in coil sensitivity profiles.

For small structures, the observed changes in volumes segmented by MorphoBox were consistent with the results presented in several reproducibility studies (Jovicich et al., 2013; Kruggel et al., 2010; Morey et al., 2010; Reuter et al., 2012) comparing repeat scans with an identical protocol. In particular, if we recomputed our normalized volumetric differences between the CAIPIRINHA and the reference ADNI-2 protocol using the methodology described in Jovicich et al. (2013) we observed: hippocampus 2.65 +/− 1.45%; caudate 1.57 +/− 1.25%; pallidum 1.45 +/− 1.35%; putamen 1.63 +/− 1.03%; and thalamus 1.35 +/− 1.16%. Those values are in agreement with values reported in a reproducibility study (Jovicich et al., 2013) processed with the FreeSurfer (Dale et al., 1999; Fischl et al., 1999) software package (version 5.1.0): hippocampus 3.26 +/− 0.93%; caudate 2.57 +/− 0.36%; pallidum 7.44 +/− 1.95%; putamen 4.61 +/− 0.88%; and thalamus 4.97 +/− 1.29%. Slightly smaller absolute normalized volumetric differences in the current study can be attributed to the use of a single scanner platform, not repositioning the subject within the session, time span between scan and rescan, hydration level changes, different degree of acceleration, different types of accelerations, different age group, smaller subject number, and may also be attributed to the differences in the segmentation software.

ICCs between the reference ADNI-2 protocol and the respective accelerated variants were also similar to previously reported values in reproducibility studies (Jovicich et al., 2013; Wonderlick et al., 2009; Morey et al., 2010), comparing repeat scans with an identical protocol. In particular, the correlations of volumes of small structures from CAIPIRINHA and reference ADNI-2 protocols are comparable to the values reported in Jovicich et al. (2009) processed with FreeSurfer (Dale et al., 1999; Fischl et al., 1999) (version 4.0.1): hippocampus 0.95 [0.88–0.98] (this work) vs 0.989 [0.976–0.997] (Wonderlick et al., 2009), caudate 0.98 [0.95–0.99] vs 0.994 [0.988–0.998], pallidum 0.88 [0.74–0.95] vs 0.706 [0.445–0.897], putamen 0.97 [0.94–0.99] vs 0.971 [0.939–0.991], thalamus 0.98 [0.96–0.99] vs 0.984 [0.965–0.995]; numbers in brackets represent the 95% confidence intervals. The differences in ICCs are small and can be attributed to the bigger subject number and

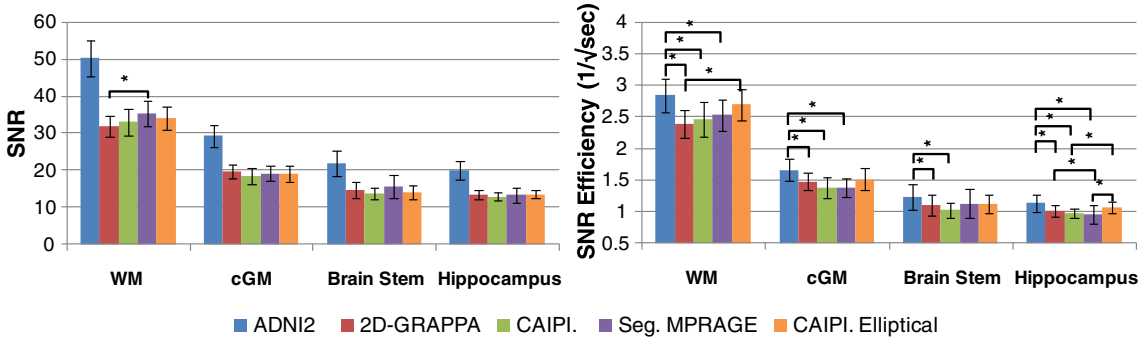

Fig. 9. SNR measurement for different brain structures and protocols. Error bars show the standard deviations across the subjects. * indicates significant difference between two protocols.

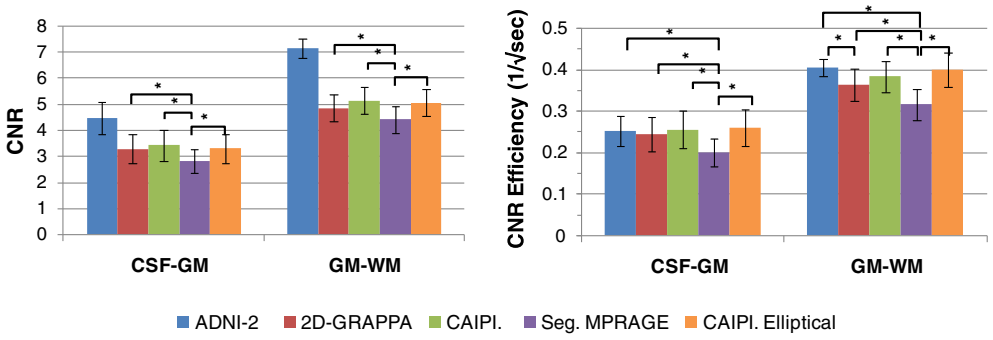

Fig. 10. CNR averaged across subjects. Error bars indicate the standard deviations across the subjects. * indicates significant difference between protocols.

slightly different protocol parameters (TR/TI/TE/flip angle) in the current study as well as different scanner platform, not repositioning the subject within the session, time span between scan and rescan, hydration level changes, different degree of acceleration and different segmentation software.

Considering other possible protocol combinations, the most interchangeable protocol combinations are CAIPIRINHA/2D-GRAPPA and CAIPIRINHA/CAIPIRINHA with elliptical scanning. This conclusion is based on the observed high ICC values and small changes in volumes relative to the other protocol combinations (Fig. 5, Fig. 7). There is a strong indication that segmented MPRAGE acquisitions introduce a change to CSF volumes. These differences can be attributed to the reduction in susceptibility-related distortions due to the increased bandwidth. Qualitatively, it can be observed that there is a better delineation of CSF in segmented MPRAGE acquisitions. We assume that this MPRAGE variant exhibits similar properties as the multi-echo MPRAGE variant MEMPRAGE (van der Kouwe et al., 2008).

The CNR and SNR analysis performed in this study depends on segmentation quality of MorphoBox and the accuracy of the registration algorithm. An alternative option would be to perform brain tissue segmentations manually. However, manual segmentations would be vulnerable to human error and are very time consuming. Another alternative option would be to not fully segment brain tissues but rather select small ROIs within them. However, using this method for SNR and CNR analysis will not necessarily capture the noise amplification (g-factor).

In order to assess the segmentation errors made by MorphoBox, it would be ideal to have the ground truth segmentations. However, no brain segmentation method, be it automatic or manual, is error-free. Experimentally, MorphoBox and FreeSurfer (Dale et al., 1999; Fischl et al., 1999) segmentation tools have been compared on the ADNI (Jack et al., 2008) database in terms of disease detection accuracy and showed that MorphoBox and FreeSurfer achieve similar performance levels (Schmitter et al., 2015). In the current study, no gross segmentation errors were observed in all of the acquired image volumes and the differences were at the single voxel level. Visually, most of these differences occurred in regions affected by partial volume effects. Potential segmentation errors can be a consequence of partial volume effects since they make the segmentation problem intrinsically ambiguous and CNR\SNR changes can amplify this ambiguity. An indication of this effect is given by the order of magnitude of the observed differences in volumes between the protocol variants, suggesting that only a small number of voxels are affected.

The SNR analysis yielded approximately a $\frac{1}{\sqrt{2}}$ decrease in SNR when moving from the accelerated reference ADNI-2 protocol to further accelerated variants. This result is expected from theoretical considerations when assuming that noise is proportional to $\sqrt{\text{sampling time}}$ (Macovski, 1996). In 2D-GRAPPA, CAIPIRINHA and CAIPIRINHA with elliptical-scanning acquisitions, the reduced sampling time, and thus the decrease in SNR, are attributed to undersampling of k-space. The small difference to the theoretically expected value indicates a negligible additional g-factor penalty. In segmented MPRAGE, the decrease in SNR is due to the increased bandwidth.

The SNR efficiency of 2D-GRAPPA, CAIPIRINHA, and segmented MPRAGE acquisitions decreased compared to the ADNI-2 protocol. This can be attributed to prolonged fraction of scan time as a result of the acquisition of reference lines in these protocol variants. CAIPIRINHA with elliptical scanning was not statistically different from the ADNI-2 protocol due to reduced scan time and increased SNR because of the use of elliptical scanning.

Systematic changes in the CNR were observed in the accelerated acquisitions. The analysis of the means and variances of intensities of the accelerated protocols revealed that there is a systematic change in the noise figure compared to the reference protocol. This change in variance is largely explained by the observed changes in SNR of the accelerated acquisitions.

It is important to note that the changes in SNR and CNR can also potentially influence the scan–rescan reproducibility of the four-fold-accelerated protocols in terms of segmentation results. Even though, a

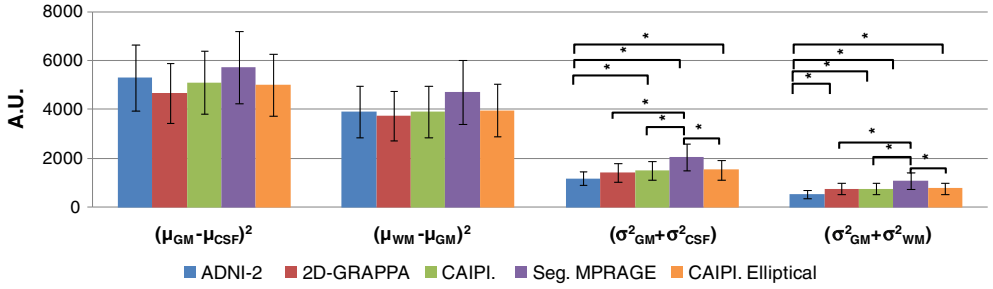

Fig. 11. Mean distance between the peaks of the intensity distributions and their variance for the different protocol variants. * indicates significant difference between protocols.

Table 2
Intra-class correlation coefficients (95% confidence interval shown in brackets) for the numerical experiment with addition of noise to the reference ADNI-2 scan.

TIV	1.00	Thalamus	0.99 [0.99–1.00]
Cortical Grey Matter	0.98 [0.97–0.99]	Putamen	0.97 [0.95–0.99]
White Matter	1.00	Caudate	0.99 [0.99–1.00]
Hippocampus	0.99 [0.98–0.99]	Pallidum	0.93 [0.88–0.97]

recent study (Krueger et al., 2012) demonstrated negligible impact on the test-retest reproducibility within identical protocols, with our current experimental design, we are unable to fully address scan–rescan reproducibility of the four-fold-accelerated protocols and further research is needed. Therefore caution must be exercised when using the four-fold-accelerated protocols in research or clinical settings for volumetric analysis.

In the numerical experiment that added synthetic noise to the reference ADNI-2 data, there is a high degree of consistency between the segmentation results of synthetic data within a subject. This conclusion is based on the high ICC values for all considered structures. In terms of volumetric results, the synthetically noise-matched ADNI-2 scans and the accelerated protocols exhibited similar bias when compared to the reference ADNI-2 protocol. Although the results suggest that noise properties affect volume bias, it should be noted that the exact relationship between noise and volume biases and the impact of other sources of biases cannot be disentangled by the current study design.

Overall, our analysis of the influence of noise on the segmentation results suggests that, in contrast to other reproducibility studies, inconsistencies between scans can be at least partly explained by the noise, CNR, and SNR—values which are measureable in dedicated settings. In other words, we hypothesize that it may be possible to correct for those systematic changes in the volumetric assessments of brain structures based on the knowledge of the noise level in the image. This statement is based on our experimental observations of volumetric measurements with high ICC values between protocol variants and the presence of a systematic bias.

However, further investigation is needed to generalize these findings, i.e. to determine if it is possible to extrapolate the findings to different coil setup, scanner field-strength, acquisition time, resolution, scan orientation and other parameters that may result in different levels of partial volume contamination and different SNR\CNR levels.

Conclusion

In summary, there are three main practical results of this study. First, accelerations of 3D structural brain scans beyond the routinely used acceleration factor of two have a measureable impact on some image analysis metrics (CNR, SNR, and noise). However, obtained images provide at least very similar information for qualitative readings in this 3T setting. Second, our analysis suggests that using or combining data from different variants of MPRAGE protocols should be done with caution. This statement is based on the findings of a number of quantitative image analysis metrics including SNR, CNR and volumetric assessments that all showed differences between the variants. This holds especially true for small brain structures that are subject to higher partial-volume effects than larger structures. Third, our results suggest that volumetric biases (at least under the given study conditions and subject to limitations mentioned in discussion) are largely affected by the noise properties of the images.

Overall, one has to carefully consider the exact use case of the accelerated protocols. In some situations, the benefits of using such accelerated protocols may potentially outweigh the drawbacks. For example, the management of patient motion is expected to improve through the use of accelerated protocols. In our study, it was demonstrated that accelerated protocols may be used in routine clinical readings. An indicator of this statement is given by the results of the qualitative readings that find the reference ADNI-2 and the CAIPIRINHA scans to have equal image quality based on qualitative assessment. We provide an error estimate on the volumetric results, when accelerated protocols are used and compared to a "standardized" ADNI-2 protocol. Therefore, if quantitative volumetric assessment is of interest, a careful consideration must be given to the effect size of changes in volumes for the structure of interest when using data from accelerated protocols. The future direction of this study will be to investigate scaling procedures as a feasible way to correct for inconsistencies in the accelerated protocols.

References

Bernstein, M.A., Fain, S.B., Riederer, S.J., 2001. Effect of windowing and zero-filled reconstruction of MRI data on spatial resolution and acquisition strategy. J. Magn. Reson. Imaging 14 (3), 270–280.

Blaimer, M., Breuer, F.A., Mueller, M., Seiberlich, N., Ebel, D., Heidemann, R.M., Griswold, M.A., Jakob, P.M., 2006. 2D-GRAPPA-operator for faster 3D parallel MRI. Magn. Reson. Med. 56 (6), 1359–1364 (Dec.).

Brenner, D., Stirnberg, R., Pracht, E.D., Stöcker, T., 2014. Two-dimensional accelerated MP-RAGE imaging with flexible linear reordering. MAGMA 27 (5), 455–462 (Oct.).

Breuer, F.A., Blaimer, M., Mueller, M.F., Seiberlich, N., Heidemann, R.M., Griswold, M.A., Jakob, P.M., 2006. Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA). Magn. Reson. Med. 55 (3), 549–556 (Mar.).

Busse, R.F., Brau, A.C.S., Vu, A., Michelich, C.R., Bayram, E., Kijowski, R., Reeder, S.B., Rowley, H.A., 2008. Effects of refocusing flip angle modulation and view ordering in 3D fast spin echo. Magn. Reson. Med. 60 (3), 640–649 (Sep.).

Camicioli, R., Gee, M., Bouchard, T.P., Fisher, N.J., Hanstock, C.C., Emery, D.J., Martin, W.R.W., 2009. Voxel-based morphometry reveals extra-nigral atrophy patterns associated with dopamine refractory cognitive and motor impairment in parkinsonism. Parkinsonism Relat. Disord. 15 (3), 187–195 (Mar.).

Ching, C.R.K., Hua, X., Hibar, D.P., Ward, C.P., Gunter, J.L., Bernstein, M.A., Jack, C.R., Weiner, M.W., Thompson, P.M., 2015. Does MRI scan acceleration affect power to track brain change? Neurobiol. Aging 36 (Suppl. 1), S167–S177 (Jan.).

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. NeuroImage 9 (2), 179–194.

Dunn, O.J., 1961. Multiple comparisons among means. J. Am. Stat. Assoc. 56 (293), 52–64.

Falkovskiy, P., Kober, T., Reyes, D., Steinert, K., Seeger, M., Bernstein, M., Krueger, G., 2013. Segmented multi-echo MPRAGE acquisition for accelerated T1-weighted brain imaging. Proc. Intl. Soc. Mag. Reson. Med. 21

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. NeuroImage 9 (2), 195–207.

Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A., 2002. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn. Reson. Med. 47 (6), 1202–1210 (Jun.).

Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L.G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27 (4), 685–691 (Apr.).

Jack, C.R., Bernstein, M.A., Borowski, B.J., Gunter, J.L., Fox, N.C., Thompson, P.M., Schuff, N., Krueger, G., Killiany, R.J., Decarli, C.S., Dale, A.M., Carmichael, O.W., Tosun, D., Weiner, M.W., 2010. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. Alzheimers Dement. 6 (3), 212–220 (May).

Jack, C.R., 2011. Alliance for aging research AD biomarkers work group: structural MRI. Neurobiol. Aging 32 (Suppl. 1), S48–S57 (Dec.).

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., Macfall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage 30 (2), 436–443 (Apr.).

Jovicich, J., Czanner, S., Han, X., Salat, D., van der Kouwe, A., Quinn, B., Pacheco, J., Albert, M., Killiany, R., Blacker, D., Maguire, P., Rosas, D., Makris, N., Gollub, R., Dale, A., Dickerson, B.C., Fischl, B., 2009. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. NeuroImage 46 (1), 177–192 (May).

Jovicich, J., Marizzoni, M., Sala-Llonch, R., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Nobili, F., Hensch, T., Tränkner, A., Schönknecht, P., Leroy, M., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.-P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Blin, O., Frisoni, G.B., 2013. Brain morphometry reproducibility in multi-center 3T MRI studies: a comparison of cross-sectional and longitudinal segmentations. NeuroImage 83, 472–484 (Dec.).

Krueger, G., Granziera, C., Jack, C.R., Gunter, J.L., Littmann, A., Mortamet, B., Kannengiesser, S., Sorensen, A.G., Ward, C.P., Reyes, D.A., Britson, P.J., Fischer, H., Bernstein, M.A., 2012. Effects of MRI scan acceleration on brain volume measurement consistency. J. Magn. Reson. Imaging 36 (5), 1234–1240 (Nov.).

Kruggel, F., Turner, J., Muftuler, L.T., 2010. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. NeuroImage 49 (3), 2123–2133 (Feb.).

Leung, K.K., Malone, I.M., Ourselin, S., Gunter, J.L., Bernstein, M. a, Thompson, P.M., Jack, C.R., Weiner, M.W., Fox, N.C., 2014. Effects of changing from non-accelerated to

accelerated MRI for follow-up in brain atrophy measurement. Neuroimage 107C, 46–53 (Dec.).
Macovski, A., 1996. Noise in MRI. Magn. Reson. Med. 36 (3), 494–497.
Miller, M.I., Priebe, C.E., Qiu, A., Fischl, B., Kolasny, A., Brown, T., Park, Y., Ratnanather, J.T., Busa, E., Jovicich, J., Yu, P., Dickerson, B.C., Buckner, R.L., 2009. Collaborative computational anatomy: an MRI morphometry study of the human brain via diffeomorphic metric mapping. Hum. Brain Mapp. 30 (7), 2132–2141 (Jul.).
Mills, K.L., Tamnes, C.K., 2014. Methods and considerations for longitudinal structural brain imaging analysis across development. Dev. Cogn. Neurosci. 9C, 172–190 (Jul.).
Morey, R.A., Selgrade, E.S., Wagner, H.R., Huettel, S.A., Wang, L., McCarthy, G., 2010. Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. Hum. Brain Mapp. 31 (11), 1751–1762 (Nov.).
Mugler, J.P., Brookeman, J.R., 1990. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). Magn. Reson. Med. 15 (1), 152–157.
Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999. SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42 (5), 952–962.
Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 61 (4), 1402–1418 (Jul.).
Robson, P.M., Grant, A.K., Madhuranthakam, A.J., Lattanzi, R., Sodickson, D.K., McKenzie, C. a, 2008. Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions. Magn. Reson. Med. 60 (4), 895–907 (Oct.).
Roche, A., Malandain, G., Ayache, N., 2000. Unifying maximum likelihood approaches in medical image registration. Int. J. Imaging Syst. Technol. 11 (1), 71–80.
Roche, A., Ribes, D., Bach-Cuadra, M., Krüger, G., 2011. On the convergence of EM-like algorithms for image segmentation using Markov random fields. Med. Image Anal. 15 (6), 830–839.
Schmitter, D., Roche, A., Maréchal, B., Ribes, D., Abdulkadir, A., Bach-Cuadra, M., Daducci, A., Granziera, C., Klöppel, S., Maeder, P., Meuli, R., Krueger, G., 2015. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. NeuroImage. Clin. 7, 7–17 (Jan.).
Shrout, P., Fleiss, J., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86 (2), 420–428.
van der Kouwe, A.J.W., Benner, T., Salat, D.H., Fischl, B., 2008. Brain morphometry with multiecho MPRAGE. NeuroImage 40 (2), 559–569 (Apr.).
Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 1999. Automated model-based tissue classification of MR images of the brain. IEEE Trans. Med. Imaging 18 (10), 897–908.
Walsh, D.O., Gmitro, A.F., Marcellin, M.W., 2000. Adaptive reconstruction of phased array MR imagery. Magn. Reson. Med. 43 (5), 682–690.
Wiens, C.N., Kisch, S.J., Willig-Onwuachi, J.D., McKenzie, C.A., 2011. Computationally rapid method of estimating signal-to-noise ratio for phased array image reconstructions. Magn. Reson. Med. 66 (4), 1192–1197 (Oct.).
Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biom. Bull. 1 (6), 80–83.
Wonderlick, J.S., Ziegler, D.A., Hosseini-Varnamkhasti, P., Locascio, J.J., Bakkour, A., van der Kouwe, A., Triantafyllou, C., Corkin, S., Dickerson, B.C., 2009. Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging. NeuroImage 44 (4), 1324–1333 (Feb.).