Measuring signal-to-noise ratio in partially parallel imaging MRI

Frank L. Goerner^{a)} and Geoffrey D. Clarke

Department of Radiology, The University of Texas Health Science Center at San Antonio, San Antonio, Texas 78229

(Received 19 April 2011; revised 20 June 2011; accepted for publication 8 July 2011; published 15 August 2011)

Purpose: To assess five different methods of signal-to-noise ratio (SNR) measurement for partially parallel imaging (PPI) acquisitions.

Methods: Measurements were performed on a spherical phantom and three volunteers using a multichannel head coil a clinical 3T MRI system to produce echo planar, fast spin echo, gradient echo, and balanced steady state free precession image acquisitions. Two different PPI acquisitions, generalized autocalibrating partially parallel acquisition algorithm and modified sensitivity encoding with acceleration factors (R) of 2–4, were evaluated and compared to nonaccelerated acquisitions. Five standard SNR measurement techniques were investigated and Bland–Altman analysis was used to determine agreement between the various SNR methods. The estimated g-factor values, associated with each method of SNR calculation and PPI reconstruction method, were also subjected to assessments that considered the effects on SNR due to reconstruction method, phase encoding direction, and R -value.

Results: Only two SNR measurement methods produced g-factors in agreement with theoretical expectations ($g \geq 1$). Bland–Altman tests demonstrated that these two methods also gave the most similar results relative to the other three measurements. R -value was the only factor of the three we considered that showed significant influence on SNR changes.

Conclusions: Non-signal methods used in SNR evaluation do not produce results consistent with expectations in the investigated PPI protocols. Two of the methods studied provided the most accurate and useful results. Of these two methods, it is recommended, when evaluating PPI protocols, the image subtraction method be used for SNR calculations due to its relative accuracy and ease of implementation. © 2011 American Association of Physicists in Medicine.

[DOI: 10.1118/1.3618730]

Key words: MRI, partially parallel imaging, mSENSE, GRAPPA, SNR

I. BACKGROUND

Signal-to-noise ratio (SNR) is an important parameter in obtaining an objective description of image quality in magnetic resonance imaging (MRI). SNR often is used as the primary metric for image quality evaluation and is often referred to as the “currency” of MRI, since SNR can be used to pay for improvements in image quality such as smaller slice thickness or decreased imaging time. The MRI accreditation program, designed by the American College of Radiology (ACR), requires SNR to be measured in order for a facility to become accredited.¹ In the *ACRs MRI Quality Control Manual*¹ an SNR measurement method which uses a single image is suggested, along with a second method that uses two images. The single image method is commonly used by clinical physicists in the field because it is considered to be the easiest and quickest measurement. The two image method is the same as the one described in the Sec. II of this paper called N1.

The ACRs single image method of SNR measurement consists of two measurements: one being the mean signal taken from a region of interest (ROI) in the central portion of the ACR phantom and the standard deviation (SD) from pixels located in the background of the image, i.e., the non-signal

producing region. This second measurement is only accurate when the noise is predictably uncorrelated with signal acquisition time or image pixel coordinates and is free from ghosting artifacts, typically being noise conforming to a Rician distribution.² Conventional approaches used to acquire MR images produce a Rician noise distribution,³ however, with the implementation of partially parallel imaging (PPI) and techniques such as sum-of-squares reconstruction⁴ the noise in MR images does not follow this distribution.^{5–7} PPI involves the use of phased array coils and the acquisition of fewer phase encoding (PE) steps in order to reduce scan time. The acceleration value (or R -value) defines the relative decrease in scan time and reduction of phase encode lines of PPI, compared to conventional MRI. Fewer PE steps would typically result in aliasing; however, data from a calibration scan is used in combination with information taken from each coil element to unfold or reduce aliasing artifact in the final image. This process results in noise being distributed in a spatially varying manner that is roughly described by a parameter called the geometry factor (g-factor).⁶ Since the noise varies spatially and does not follow a normal distribution function, SNR values obtained using the single image method are inaccurate as has been shown previously.^{8–11} The aim of our current study is to validate a method of SNR evaluation for

TABLE I. The imaging protocols that were used to acquire the images used for the SNR analysis in this study.

Sequence	TR (ms)	TE (ms)	No. of slices	Slice thickness (mm)	BW (Hz/pixels)	Matrix	FOV (mm)	Slice gap (mm)	NEX	ETL
EPI	1840	187	5	5	752	256 × 256	220	1	1	256
FLASH	175	4	5	5	240	256 × 256	220	1	1	
Tru-FISP	6.88	3.44	5	5	244	256 × 256	220	1	1	
TSE	1200	76	5	5	122	256 × 256	220	1	1	16

protocols involving PPI methods in MRI that is accurate and relatively easy for physicists to implement in the field, primarily for protocol evaluation but also for coil quality control.^{8,12} However, the current study extends the generality of the method by performing SNR measurements on data acquired using four different types of pulse sequences and two different methods of parallel image reconstruction. This procedure also included evaluating images using four standard methods of SNR determination. These SNR measurement methods were selected for simplicity, and it should be noted that other potentially more accurate methods are available.^{13,14} However, these methods involve strategies that are not practical on a clinical MRI scanner due to time constraints and the difficulty of accessing raw data and sensitivity profiles. The calculated g-factors and SNR measurements were compared to a sophisticated but more elaborate SNR measurement method involving calculation of signal and noise from a series of 100 images, which has been shown to be accurate.^{12,15}

It should be noted that a similar study was done by Dietrich *et al.*⁸ There were several differences between the two studies. This study was done using a 3 T MRI magnet as opposed to a 1.5 T MRI magnet. This increased the signal available and subsequently allowed for the use of more *R*-values, *R*-values of 2, 3, and 4 were used rather than just an *R*-value of 2 as in the previous study. The protocols and pulse sequences used in this study also differ from those done previously with the addition of two pulse sequences and the manipulation of an additional echo-planar imaging (EPI) sequence. The Dietrich study also evaluated different filters and large field-of-views (FOVs), where this study did not and instead evaluated the manipulation of the phase encode direction of the readout data. The analysis in the two studies differed slightly as well; this study evaluated the g-factor of each measurement and performed two-way ANOVA analysis to indicate which variables had effects on SNR. Furthermore, the intended audience for this study is a clinical Medical Physicist that wishes to measure SNR in the most accurate and simplistic way possible rather than providing an extremely accurate and detailed method that would be better suited for research protocols.

II. MATERIALS AND METHODS

II.A. Phantom measurements

Images were obtained from a spherical phantom 17.8 cm in diameter (about the average size of a human head) which was filled with 5.45 g NaCl (99.99% pure) 5.29 ml of Magnevist per 1 l distilled water and had a total volume 2415 ml.

This phantom consisted of a #3 soccerball with a double bladder (Balden Series Z Soccer Ball PN S130Z-018). The predicted conductivity and T1 (at 3 T) of the solution is 0.023 S/m and 121 ms, respectively. The phantom was developed by AAPM Task Group #118 on “Parallel imaging in MRI: Technology, applications, and quality control.”¹⁶ Images were acquired using a 12-channel matrix head coil on a 3T MRI system (MAGNETOM Trio, Siemens Medical Systems, Malvern, PA). Images were acquired using EPI, a spoiled gradient-recalled echo which on Siemens scanners is fast low angle shot (FLASH), balanced steady state free precession (bSSFP) on Siemens scanner known as true fast imaging with steady-state precession (Tru-FISP), and turbo spin echo (TSE) pulse sequences. The acquisition parameters for each of the pulse sequences are listed in Table I. All scans used sum of squares coil combined mode with the exception of FLASH and TruFISP, which used adaptive coil combining for nonaccelerated and generalized autocalibrating partially parallel acquisition algorithm (GRAPPA) scans and sum of squares for modified sensitivity encoding (mSENSE). Auto circular polarization was also used for nonaccelerated scans.

The parameters in Table I were not changed throughout image acquisition, except for the second time the EPI sequence was used. The application of PPI to an EPI sequence does not directly shorten scan time but does shorten the echo train length. This allows for a shorter selectable TE and TR time which if implemented not only shortens acquisition time but also reduces susceptibility artifacts. Therefore, to take advantage of PPI implementation, a second set of EPI images were acquired in which the TE and TR were minimized at each *R*-value as shown in Table II.

Each pulse sequence was implemented without parallel imaging and the image data were acquired as a dynamic series of one hundred sequential images. Following this portion of the experiment, the radio frequency (RF) voltage was set to 0 in order to obtain a no signal image (all other parameters remained the same). The various image data acquisition processes were repeated using the partially parallel methods, mSENSE and GRAPPA.^{6,17}

TABLE II. List of changed TE and TR values for second set of EPI pulse sequences where TE and TR times were minimized.

<i>R</i>	TE (ms)	TR (ms)
2	98	956
3	69	666
4	54	515

Additionally, for the FLASH imaging sequence, the amount of reference lines and their relationship to SNR were investigated. For each FLASH sequence, reference lines of 8, 24, 48, and 128 were used for R -values of 2. Reference line values of 12, 24, 48, and 128 were used for R -values of 3 and reference line values of 16, 24, 48, and 128 were used for R -values of 4. This was done to determine whether or not the selectable R -values were actual R -values which would mean the reference lines were not included in the final image. An increase in SNR with increase in the number of reference lines could indicate that the actual R -value is slightly lower than the one selected because the reference lines were included as part of the final image and not just for reconstruction purposes. If this was the case, it was decided that a modified actual R -value would be used to calculate g -factors.

Conventional images and PPI images with mSENSE and GRAPPA reconstruction with selected R -values of 2, 3, and 4 were acquired in the axial orientation with two series, one with the PE orientation anterior–posterior and one in the right–left orientation. With each slice being an image, this resulted in 7070 (5 slices $\times$ 7 R -values $\times$ (100 images + 1 noise image) $\times$ 2PE directions) images per pulse sequence with a total of 35,350 images analyzed for this study. The image analysis was automated by an algorithm developed in-house using a commercial technical computing environment for data visualization and analysis (MATLAB version 7.2, The Mathworks, Inc., Natick, MA). The measurements performed on the images included the five different standard SNR measurements, which are described below.

II.B. SNR measurement methods

Each of the SNR measurement procedures are described in order, below. Placement of the subsequent ROIs can be seen in Fig. 1, an EPI sequence taken at mSENSE $R = 3$. For three of the methods listed, noise is collected from the background of the image and subjected to the correction factor 0.66. This factor is used because noise from the background portion of an MR image theoretically follows a Rician distribution.^{3,18} However, practical implementations that include the use of sum of squares and other vendor specific techni-

FIG. 1. ROI placement in phantom for SNR analysis. White areas indicate measurement areas. The signal ROI also indicates where measurements for both signal and noise were taken for the M100 method. Method N1 is the subtraction method. Method N2 is the no signal method, while ACR and N4 are ACR method and the background method that NEMA recommends.

ques such as Fermi filters, edge enhancement algorithms, and gradient distortion correction will cause the noise distribution to be modified. For this reason, methods using background portions of an image are inadequate to determine SNR. In the work described below, the factor 0.66 is still used in cases where the image background is used for noise calculation with this caveat in mind.

1. NEMA method 1 (N1):^{19–21} This method utilized two images acquired with the same imaging parameters within 5 min of each other.

Signal: To calculate the signal, the two images were added and divided by two to get an average image. Following this process, the average pixel value from a circular ROI that encompassed 75% of the phantom volume was used as the signal. This ROI was drawn automatically by an image analysis function in MATLAB.

Noise: To calculate noise, the two images were subtracted to produce a third image. An ROI of the same geometry, which was used in the above signal calculation, was drawn in the subtraction image. The SD from this ROI was taken as the image noise where $\bar{R}$ is the average of an ROI, $R(i, j)$ is an individual pixel within the ROI, and n and m are pixel numbers in the row and column directions,

$$SD = \left[\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} (R(i, j) - \bar{R})^2}{\sum_{i=1}^n (m_i) - 1} \right]^{1/2}. \quad (1)$$

The SD was then corrected based on the fact that the noise within the ROI follows an uncorrelated Gaussian distribution. Histograms of the images were obtained and shown to follow a Gaussian distribution. This calculation was accomplished by dividing the SD by $\sqrt{2} = 1.41$ as is standard with propagation of error when two images are used.²²

2. NEMA method 2 (N2):²⁰ This method used a single image to find signal and the non-signal image to find the noise.

Signal: The average of an ROI encompassing 75% of the signal producing volume of one of the acquired signal producing images was used as the signal.

Noise: The noise computation was performed using data from the non-signal-producing image. On the MRI system used in this study, the signal was eliminated by setting the transmit RF coil voltage to 0 V. In other scanners, images without signal may be obtained by setting the flip angle to 0°. The standard deviation of the ROI with the same geometry used to compute the signal was then measured on the non-signal image and divided by 0.66.

3. NEMA method 4 (N4):²⁰ For this method, a single signal-producing image was acquired and the following analysis was applied.

Signal: The signal used was the average of an ROI encompassing 75% of the signal-producing region of the image.

Noise: Four rectangular ROIs adding to a total of 1024 pixels were drawn outside of the phantom and away from any artifacts in the background of the image. The noise ROIs in both the frequency-encoding and phase-encoding directions were well removed from the signal-producing phantom. This

requirement placed four 16×16 pixel noise ROIs in the corners of the image resulting in 1024 pixels per image.

The SDs of the noise ROIs were averaged to obtain the mean noise and divided by the 0.66 Rician distribution correction factor to obtain total image noise.

4. ACR method (ACR):¹ This method was measured in essentially the same manner as the N4 method, described in Sec. III. The only difference between is that only 50 pixels were used for the noise ROIs in the ACR method.

5. 100 image method (M100):⁸ In this method of SNR measurement, one hundred identical images were acquired in rapid succession. If the pixel position, $r = (x, y, z)$, then $S(r, k)$ is the image intensity at image number k , slice z , and pixel position (x, y) . To calculate the SNR on a pixel-by-pixel basis, the pixel intensity (I) was measured at every position r for all 100 images to produce mean pixel intensity,

$$\bar{I}(r) = \frac{\sum_{k=1}^{100} I(r, k)}{100}. \quad (2)$$

The pixel-by-pixel signal information was used to find the noise by calculating the standard deviation (σ) of the signal for each pixel over the entire series of images,

$$\sigma(r) = \sqrt{\frac{1}{100-1} \sum_{k=1}^{100} (I(r, k) - \bar{I}(r))^2}. \quad (3)$$

Thus, the signal-to-noise ratio on a pixel-by-pixel basis (Ψ) was defined as

$$\Psi(r) = \frac{\bar{I}(r)}{\sigma(r)}. \quad (4)$$

In order to compare the SNR as defined by Dietrich *et al.*⁸ to the SNR obtained from the other methods described in this study, it is necessary to determine an SNR value that characterizes the entire image. To obtain the overall image signal, all the pixels for an ROI that are located in the 75% signal region (75% SR) were averaged,

$$S(\text{ROI}) = \bar{I}(r, k)_{75\% \text{SR}}. \quad (5)$$

The noise was found by using the same 75%SR to determine the standard deviation of the values calculated in the pixel-by-pixel method,

$$\text{SD}(\text{ROI}) = \bar{\sigma}_{(75\% \text{SR})}. \quad (6)$$

To determine the SNR for the entire image, the ratio of these values is simply calculated,

$$\text{SNR}(\text{ROI}) = \frac{S(\text{ROI})}{\text{SD}(\text{ROI})}. \quad (7)$$

This metric gives a single number for each image that is readily comparable to the values obtained from the previously described methods.

II.C. Phantom statistical analysis

The g-factor is defined as the degradation in the SNR due to the parallel imaging process,⁶

$$\frac{\text{SNR}_{R=1}}{\text{SNR}_R} = g\sqrt{R}. \quad (8)$$

Here $\text{SNR}_{R=1}$ is the baseline SNR found with non-PPI protocol, SNR_R is SNR of the PPI protocol and R is the acceleration factor. In this study, we calculated the g-factors for the images produced by each of the measurement methods used and each of the protocols imaged. For R -value, the selected value indicated was not used because it accounts for only the reduction in the amount of phase encode acquisitions but not for the use of reference lines required in the images. This dependence was determined from the FLASH image data which was acquired multiple times using different numbers of reference lines. In each case, the SNR increased with the number of reference lines indicating that SNR was dependent on the number of reference lines which modify the actual R -value. This was true for all pulse sequence with the exception of EPI scans, for which the reference scan is taken prior to image acquisition and only used for coil calibration purposes.²³ Thus, R -values used to calculate g-factors were adjusted based on the number of reference lines used in the reconstruction process. For each pulse sequence, with the exception of TSE, 24 reference lines were used. The TSE sequence used reference lines of 34, 40, and 44, respectively, for R -values of 2, 3, and 4. Adjusted R -values were 1.68, 2.34, and 2.91 for FLASH and TRUFisp. They were 1.58, 2.04, and 2.37 for TSE and remained 2, 3, and 4 for EPI scans. Since the parallel imaging process cannot improve SNR, the g-factor should be equal to or greater than 1.⁵ For this reason, any measurement method, which resulted in g-factors less than 1, was deemed to be inaccurate. Additionally, a two-way ANOVA was performed on the g-factor measurements to determine the degree to which the R -value and PE direction influenced the results.

In addition to the g-factor evaluation, the SNR values calculated using the N1, N2, ACR, and N4 SNR measurement methods were compared for agreement with the SNR measurements made using the M100 method. First the deviations between M100 and each method were calculated as follows:

$$\% \text{Diff} = 100 \cdot \frac{|\text{SNR}_x - \text{SNR}_{\text{M100}}|}{\text{SNR}_{\text{M100}}}, \quad (9)$$

where SNR_x is the SNR of the method being compared to M100. Additionally, the Bland–Altman method^{24,25} was used to display graphically the agreement between the SNR values, determined using the various SNR measurement methods.

II.D. In-vivo measurements and analysis

Three healthy volunteers were used to evaluate four of the methods *in-vivo*. All four of the pulse sequences mentioned previously were utilized. The volunteers were positioned in the same 3T MRI system and scanned with the same 12 channel head coil as the phantom mentioned previously. Scan parameters were held constant with a few exceptions: 15 slices were used instead of 5, which also changed the TR times for the EPI and TSE pulse sequences. Additionally, the data for the EPI sequence was acquired with both TR and TE held constant, based on the values used for

TABLE III. Different TR values were used with various data acquisitions acquired using the echo-planar pulse sequence in order to evaluate improvements in overall image quality.

Sequence	TR (ms)	TE (ms)
TSE	3170	76
EPI	5510	187
EPI $R = 2$	2870	98
EPI $R = 3$	2440	69
EPI $R = 4$	1550	54

the non-PPI images, and the TR and TE minimized as the R -values were increased. Table III displays changes in TR for the EPI and TSE sequences.

The M100 method of SNR measurement was not evaluated in the *in-vivo* study since the study time would have been several hours for each of the subjects. To evaluate the other methods, two slices per subject were selected; the slices were selected based on the availability of a relatively uniform area of white matter located centrally in the brain. Two square ROIs were placed on each image slice over the largest region of white matter in the central area of the brain axially on both the left and right sides. These ROIs were used to obtain the signal for all methods and used for the noise measurements in methods N1 and N2. Figure 2 displays, where ROIs were placed in one of the subjects, position of the ROIs in the other two subjects varied slightly due to individual patient anatomy. The noise ROIs for methods N4 and ACR are the same as described for phantom measurement. Since the M100 method was not utilized, the g-factor was used to determine if any of the measurements adhered to the theory that g-factor > 1 . A two-way ANOVA was performed to assess the effects of R -value and reconstruction method on g-factor and SNR.

III. RESULTS

III.A. Phantom results

Following a linear least squares fit of the SNR calculation, a g-factor > 1 should be obtained.⁶ The average g-factor for

FIG. 2. Image of a non-PPI FLASH sequence from one of the volunteers used in this study. The left image is one slice (slice 8) superior to the right image (slice 7). The white ROIs indicate where signal measurements for all SNR measurement methods were taken and noise measurements for N1 and N2 were taken in this specific subject. The ROIs varied slightly according to individual patient anatomy.

TABLE IV. Average g-factors $\pm$ standard deviation for mSENSE reconstruction methods listed by selected R -value and the type of SNR measurement.

R	mSENSE				
	M100	N1	N2	N4	ACR
2	1.04 ± 0.12	1.13 ± 0.19	0.81 ± 0.05	0.9 ± 0.09	0.89 ± 0.07
3	1.24 ± 0.03	1.4 ± 0.13	0.83 ± 0.06	0.9 ± 0.08	0.9 ± 0.08
4	1.92 ± 0.04	2.28 ± 0.35	0.88 ± 0.11	0.93 ± 0.13	0.92 ± 0.09

each method across all protocols is listed in Tables IV and V. The EPI sequences that were taken with minimum TE and TR values were not used in this portion of the study since additional SNR changes due to reduction of TE and TR preclude direct assessment of the g-factor. The only methods that consistently had g-factors greater than unity in any of the protocols investigated were methods M100 (100 image method) and N1 (image subtraction method). The mean g-factor for method M100 was 1.40 ± 0.46 for mSENSE reconstruction and 1.3 ± 0.39 with GRAPPA reconstruction. The g-factor for method M100 ranged from 0.83 to 2.75. Method N1 had a mean g-factor of 1.61 ± 0.62 for mSENSE reconstruction and 1.48 ± 0.50 for GRAPPA. The g-factors measured for method N1 ranged from 0.96 to 3.22. The M100 method had g-factors less than 1 at $R = 2$ for each pulse sequence and at $R = 3$ for the bSSFP sequence. The N1 method had g-factors less than 1 for the EPI sequence at $R = 2$ also in the bSSFP protocol with GRAPPA reconstruction at $R = 2$. The only other SNR measurement method with any g-factors greater than 1 occurred with the N4 and ACR methods using the FLASH sequence when mSENSE reconstruction was used. Additionally, based on previously recorded g-factors that are clinically useful, our g-factors for methods N1 and M100 were deemed to be reasonable.²⁶

Data from 5 slices were measured over 14 protocols in 4 different sequences in addition to data from an EPI sequence with varied TE and TR. A total of 350 SNR measurements were performed for each method. Of the 350 N1 measurements, 276 of them were within 10% of the M100 method. For method N2, 21 of the 350 measurements were within 10% of the M100 results, and 20 of these were from non-PPI protocols. For method N4, 24 of the measurements were within 10% of the M100 method and 14 of these were from non-PPI protocols. The rest of the measurements were all at R -values of 2. For the ACR method, 20 of the measurements were within 10% of the M100 method, 16 of these measurements were from non-PPI protocols, and the rest were from protocols with R -values of 2.

TABLE V. Average g-factors $\pm$ standard deviation for GRAPPA reconstruction methods listed by selected R -value and the type of SNR measurement.

R	GRAPPA				
	M100	N1	N2	N4	ACR
2	0.99 ± 0.01	1.08 ± 0.02	0.58 ± 0	0.59 ± 0.02	0.58 ± 0.04
3	1.15 ± 0.01	1.31 ± 0.02	0.44 ± 0	0.4 ± 0.01	0.38 ± 0.03
4	1.76 ± 0.03	2.06 ± 0.03	0.35 ± 0	0.38 ± 0.01	0.35 ± 0.02

FIG. 3. Bland–Altman analysis of agreement of difference method and 100 image method.

The percent differences in SNR between the M100 method and each of the other methods were also determined. The N1 method had a range 0% to 54%, the N2 method had a range of 0.4% to 646%, the N4 method had a range of 0.4% to 992.1%, and the ACR method had a range of 0.2% to 1192.0%. Averaged over all 350 measurements, the N1 method resulted in an average difference of $7.4\% \pm 8.3\%$, the N2 method had an average of $171.6\% \pm 161.4\%$, the N4 method produced an average of $187.8\% \pm 198.8\%$, and the ACR method resulted in an average of $220.2\% \pm 235.9\%$.

Bland–Altman analysis was performed to determine the agreement between the data obtained using the M100 method and each of the additional SNR measurement methods. Each analysis included every measurement that was made for each slice in each pulse sequence, the reconstruction method, the R -value, and the phase encoding direction. The results are displayed in Figs. 3–6. The N1 SNR measurement method showed the greatest agreement with the M100 measurement method. The bias between these two methods was -4.33 and the 95% confidence intervals (limits of agreement) were -30.17 to 21.51 . The data acquired with the N2 method and the data from the M100 method had a bias of -76.73 with the 95% confidence intervals falling between -254.3 and 100.9 . The analysis of the N4 method data compared with the M100 data had a bias of -129.9 and a 95% confidence interval between -433.5 and 173.7 . Finally, analysis of the data from the ACR method and the

FIG. 4. Bland–Altman analysis of agreement of ACR method and 100 image method.

FIG. 5. Bland–Altman analysis of agreement of no signal method and 100 image method.

M100 method revealed the poorest agreement with the largest bias of -257.9 and with confidence intervals between -808.8 and 292.9 .

Two separate two-way ANOVA analyses were also performed on the data obtained from the M100 method and the N1 method data to determine how much of an effect the R -value and reconstruction method had on the measured SNR and how much the direction of phase-encoding and the R -value influenced the SNR measurement. The two-way ANOVA for the M100 data revealed that the R -value had a 99.6% effect on the total variance ($p < 0.0001$) which is interpreted as being extremely significant relative to the SNR. The parallel imaging reconstruction method (mSENSE vs GRAPPA) had a 0.27% effect on the variance of SNR which was not significant. The phase-encoding direction had a 0.18% effect on SNR ($p < 0.005$) which was statistically significant but of little importance. For the data from the N1 method, the two-way ANOVA revealed that the R -value accounted for 99.7% of the variance ($p < 0.0001$). The results were similar to the data from the M100 method, with regards to the effect of reconstruction method which accounted for 0.21% of the variance in SNR ($p = \text{NS}$). The phase-encoding direction accounted for 0.16% of the variance in SNR ($p < 0.05$).

III.B. In-vivo results

The average g-factor for each method across all protocols is listed in Tables VI and VII. The EPI sequences acquired

FIG. 6. Bland–Altman analysis of agreement of NEMA single method and 100 image method.

TABLE VI. The average g-factors $\pm$ standard deviations measured in three volunteers across four pulse sequences and two slices for mSENSE based reconstruction methods listed by selected R -value and SNR measurement method.

R	mSENSE			
	N1	N2	N4	ACR
2	1.06 ± 0.15	0.15 ± 0.85	0.85 ± 0.24	0.24 ± 0.82
3	1.42 ± 0.53	0.53 ± 0.87	0.87 ± 0.14	0.14 ± 0.86
4	1.84 ± 0.8	0.8 ± 0.86	0.86 ± 0.18	0.18 ± 0.82

with minimum TE and TR values were not used in this portion of the study since the g-factor cannot be calculated because of additional SNR changes with changes in TE and TR. The only method that produced g-factors greater than unity in any of the protocols investigated was method N1 (image subtraction method). The mean g-factor for method N1 was 1.44 ± 0.39 for mSENSE reconstruction and 1.34 ± 0.43 with GRAPPA reconstruction. The g-factor for method N1 data ranged from 0.67 to 2.81. The N1 method data had no g-factors less than 1 for the TSE sequence. For the bSSFP pulse sequence, the g-factors were less than 1 in two of the three volunteers for GRAPPA with $R=2$ and $R=3$ and for all mSENSE $R=2$ measurements. The FLASH sequence had no g-factors less than 1 with a minimum g-factor of 0.105. The EPI sequence only had g-factors greater than 1 for mSENSE reconstructed data acquired with $R=2$.

IV. DISCUSSION

SNR is a valuable metric that allows for the objective measurement of image quality in MRI protocols. Higher SNR typically indicates an image of greater clinical utility and images with low SNR often appear grainy. Thus, it is important to select an accurate measurement method for SNR that also can be implemented with relative ease. The present study was designed to evaluate several different methods of simple SNR measurements for accuracy. The results demonstrate that the M100 and N1 methods for SNR measurement are in agreement and the most accurate of the five methods tested. However, the N1 method is more straightforward to implement.

IV.A. Phantom data

The results of g-factor analysis from the phantom data indicated that the M100 and N1 methods were the two meth-

TABLE VII. This table lists the average g-factors $\pm$ standard deviations measured in three volunteers across four pulse sequences and two slices for mSENSE based reconstruction methods listed by selected R -value and SNR measurement method.

R	GRAPPA			
	N1	N2	N4	ACR
2	1.02 ± 0.15	0.15 ± 0.64	0.64 ± 0.15	0.15 ± 0.62
3	1.18 ± 0.29	0.29 ± 0.52	0.52 ± 0.21	0.21 ± 0.49
4	1.83 ± 0.79	0.79 ± 0.46	0.46 ± 0.24	0.24 ± 0.46

ods that conformed most closely to the theory that g increases with R from a base value of unity in protocols where PPI methods are implemented. The N1 method showed average g-factors larger than or equal to unity in all cases and the M100 method did so as well except in the case of GRAPPA $R=2$, with the smallest average g-factor being 0.99. All the other methods evaluated resulted in average g-factors below 0.92 in every case and below 0.59 in the case of GRAPPA reconstruction. However, there are a number of factors which require a decrease in SNR when PPI methods are used, which limits the g-factor to values greater than or equal to unity.

First, PPI methods necessarily collect fewer data points in order to increase imaging speed, thus SNR is always lower with PPI methods than with standard FT-MRI methods. The noise characteristics of PPI also suffer due to the mutual interactions between coil elements of a multi-channel RF coil, which can result in different coil elements sharing noise simultaneously. The degree of noise interaction between coil elements can be mathematically characterized by the noise correlation matrix Ψ .¹² Finally, in the reconstruction process, PPI algorithms may include steps, such as masking, filtering, and fitting that distort the noise characteristics of the raw image data, which can lead to mischaracterization of the SNR by the N2, N4, and ACR methods evaluated in the current study.^{27,28} Thus, the use of PPI protocols result in a different noise distribution and subsequently invalidates the non-signal methods of noise evaluation and thus those methods for the purpose of SNR assessment. An additional reason for error occurs in the case of method N2. For accurate reconstruction, the sensitivity profile should not be non-zero for the noise image. Due to the fact that both mSENSE and GRAPPA are autocalibrated, the sensitivity profile of a non-signal image will always be non-zero and cannot be changed without modifying the reconstruction algorithm and obtaining raw data. Dietrech *et al.*⁸ previously noted this circumstance and thus this method was excluded from their evaluation. However, in the current study, this method was included despite this knowledge and with no sensitivity profile correction in order to demonstrate its inaccuracy empirically.

G-factor analysis demonstrated the mutual agreement of methods N1 and M100 as well as their general agreement with PPI theory. Of the 350 measurements taken with N1 and M100, 276 of them ($\sim 80\%$) were within 10% of each other. Also Bland–Altman analysis demonstrated that the N1 data most closely agrees with the M100 measurements. In contrast, results from the other three methods of SNR measurement average more than 100% difference from the M100 method with the minimum average difference found in the N2 method at 140.3%. The M100 method, as indicated in Sec. III, had a range of g-factors including some that were less than unity. This occurred in 18% of g-factor calculations and 4% of g-factor calculations were more than one standard deviation below unity. These results can be attributed to the fact that the definition of the g-factor in Eq. (8) is not completely accurate when using nonlinear reconstruction methods.¹³

The two-way ANOVA test results indicated that the primary contributions of degradation in SNR are due to a change in R -value rather than a change in reconstruction

protocol. Additionally a change in phase-encoding direction had minimal effect on the measured SNR values with the coil used in this study, so the *in-vivo* study measurements were not performed in different phase-encoding directions. Additional data could be taken in different slice orientations with different PE directions which could potentially change SNR results but in typical clinical settings with head coil use the axial orientation or orthogonal orientation with an acute angle from axial is used. Thus other slice orientations were not investigated.

IV.B. *In-vivo* data

The *in-vivo* measurements further validated the concept that determining noise in regions of low signal is a poor predictor of noise in PPI data. The M100 method was not evaluated *in-vivo* since it would require more than 10 h to produce 15 slices 100 times for each protocol and each sequence. Additionally the M100 method would be subject to physiological noise and motion from the subjects. The N1 method is also subject to noise and motion from the subject since it uses a set of two dynamic images. The other four measurements were evaluated and the results obtained were similar to the phantom findings. The g-factors again were less than unity for each of the three non-signal based methods and ranged on average from 0.15 to 0.87. For the N1 method, the g-factors on average were greater than unity in every case. Additionally all methods of measurement aside from N1 showed an increase in SNR with an increase in R-value and the implementation of PPI which is the opposite of what theory dictates.

All the work, described in the current project, was carried out without detailed knowledge of the MRI system's manufacturer's proprietary PPI algorithms and relied on descriptions published in the scientific literature. Thus a limitation of the current study is that the conclusions drawn may have limited generalizability in regards to MRI systems that use other specific PPI post-processing algorithms, such as three-dimensional PPI. In addition, these data were acquired with a specific RF coil and similar data acquired with multi-channel RF coils of different configurations may not produce exactly the same results. For instance, the relative effects on SNR due to the direction which is used for phase-encoding may be very dependent on RF coil geometry. Furthermore, the relationship between SNR and g-factor presented in Eq. (8) assumes a linear reconstruction method and all but two of the sequences studied used a sum-of-squares reconstruction method, which is nonlinear. This would suggest that the g-factor calculations are subject to error from artifacts and image scaling,¹³ however, the calculation of g-factor in the manner described clearly indicates changes in SNR that occur from R-value to R-value. With g-factors < 1, SNR can actually increase with increasing R-value (depending on R-value and exact g-factor) which is a violation of theory as well. Thus the g-factors were useful in eliminating some methods from further analysis but were not exact. Nevertheless, the overall conclusions of this study are likely applicable to the PPI imaging methods currently found on most clinical MRI scanners.

V. CONCLUSION

This study demonstrated how methods of SNR measurement that determine the noise in regions of little or no signal produce noise statistics that lead to inaccurate SNR values. One of these inaccurate metrics is currently recommended by ACR for SNR measurement. The results also suggest that the N1 method, or image subtraction method, offers the most accurate and straight-forward approach of SNR determination currently available. However, the results also indicate that both the N1 method and M100 methods have their limitations in determining SNR accurately. The g-factors for these methods should always be greater than unity and this was not the case in all measurements.

Currently it is the recommendation of the authors that the N1 method is implemented when SNR measurements are required in a protocol using PPI. Future investigations should be aimed at determining specifically where these methods fail so that improved methods of measuring SNR can be identified for use where PPI or other reduced-sampling MRI acquisition methods are utilized.

ACKNOWLEDGMENTS

We express our appreciation to Dr. Jason Stafford (The University of Texas M.D. Anderson Cancer Center) for his contributions in review prior to article submission. Frank Goerner was supported by the National Institute of Biomedical Imaging and Bioengineering of NIH Training Grant T-32 EB000817-04.

^{a)} Author to whom correspondence should be addressed. Electronic mail: fgoerner@yahoo.com

¹ American College of Radiology, *Magnetic Resonance Imaging Quality Control Manual* (American College of Radiology, Reston, VA, 2004).

² R. M. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Med. Phys.* **12**, 232–233 (1985).

³ H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magn. Reson. Med.* **34**, 910–914 (1995).

⁴ P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller, "The NMR phased array," *Magn. Reson. Med.* **16**, 192–225 (1990).

⁵ D. K. Sodickson and W. J. Manning, "Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays," *Magn. Reson. Med.* **38**, 591–603 (1997).

⁶ K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.* **42**, 952–962 (1999).

⁷ C. D. Constantinides, E. Atalar, and E. R. McVeigh, "Signal-to-noise measurements in magnitude images from NMR phased arrays," *Magn. Reson. Med.* **38**, 852–857 (1997).

⁸ O. Dietrich, J. G. Raya, S. B. Reeder, M. F. Reiser, and S. O. Schoenberg, "Measurement of signal-to-noise ratios in MR images: Influence of multi-channel coils, parallel imaging, and reconstruction filters," *J. Magn. Reson. Imaging* **26**, 375–385 (2007).

⁹ A. Ogura, A. Miyai, F. Maeda, H. Fukutake, and R. Kikumoto, "Accuracy of signal-to-noise ratio measurement method for magnetic resonance images," *Nippon Hoshasen Gijutsu Gakkai Zasshi* **59**, 508–513 (2003).

¹⁰ A. Ogura, T. Miyati, M. Kobayashi, H. Imai, K. Shimizu, T. Tsuchihashi, T. Doi, and Y. Machida, "Method of SNR determination using clinical images," *Nippon Hoshasen Gijutsu Gakkai Zasshi* **63**, 1099–1104 (2007).

¹¹ H. Imai, T. Miyati, A. Ogura, T. Doi, T. Tsuchihashi, Y. Machida, M. Kobayashi, K. Shimizu, and Y. Kitou, "Signal-to-noise ratio measurement in parallel MRI with subtraction mapping and consecutive methods," *Nippon Hoshasen Gijutsu Gakkai Zasshi* **64**, 930–936 (2008).

¹² S. B. Reeder, B. J. Wintersperger, O. Dietrich, T. Lanz, A. Greiser, M. F. Reiser, G. M. Glazer, and S. O. Schoenberg, "Practical approaches to the evaluation of signal-to-noise ratio performance with parallel imaging:

- Application with cardiac imaging and a 32-channel cardiac coil," *Magn. Reson. Med.* **54**, 748–754 (2005).
- ¹³P. M. Robson, A. K. Grant, A. J. Madhuranthakam, R. Lattanzi, D. K. Sodickson, and C. A. McKenzie, "Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions," *Magn. Reson. Med.* **60**, 895–907 (2008).
 - ¹⁴F. A. Breuer, S. A. R. Kannengiesser, M. Blaimer, N. Seiberlich, P. M. Jakob, and M. A. Griswold, "General formulation for quantitative G-factor calculation in GRAPPA reconstructions," *Magn. Reson. Med.* **62**, 739–746 (2009).
 - ¹⁵P. Kellman and E. R. McVeigh, "Image reconstruction in SNR units: A general method for SNR measurement," *Magn. Reson. Med.* **54**, 1439–1447 (2005).
 - ¹⁶AAPM TG #118 (Personal Correspondence, 2009).
 - ¹⁷M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.* **47**, 1202–1210 (2002).
 - ¹⁸S. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.* **24**, 46–156 (1945).
 - ¹⁹B. Murphy, P. Carson, J. Ellis, Y. Zhang, R. Hyde, and T. Chenevert, "Signal-to-noise measures for magnetic resonance imagers," *Magn. Reson. Imag.* **11**, 425–428 (1993).
 - ²⁰National Electrical Manufacturers Association, *Determination of Signal-to-Noise Ratio (SNR) in Diagnostic Magnetic Resonance Imaging*, NEMA Standards Publication MS1-2008 (2008).
 - ²¹M. J. Firbank, A. Coulthard, R. M. Harrison, and E. D. Williams, "A comparison of two methods for measuring the signal to noise ratio on MR images," *Phys. Med. Biol.* **44**, N261–N264 (1999).
 - ²²G. F. Knoll, *Radiation Detection and Measurement*, 3rd ed. (John Wiley & Sons, Inc., Hoboken, NJ, xxxx).
 - ²³Applications Guide Operating Instructions MAGNETOM TRIO, syngo MR 2003T (2002).
 - ²⁴J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet* **1**, 307–310 (1986).
 - ²⁵J. M. Bland and D. G. Altman, "Comparing methods of measurement: Why plotting difference against standard method is misleading," *Lancet* **346**, 1085–1087 (1995).
 - ²⁶M. A. Ohliger and D. K. Sodickson, "An introduction to coil array design for parallel MRI," *NMR Biomed.* **19**, 300–315 (2006).
 - ²⁷K. Pruessmann, "Sensitivity-encoded magnetic resonance imaging," Ph.D. dissertation, Swiss Federal Institute of Technology Zurich, 2000.
 - ²⁸D. J. Larkman, P. G. Batchelor, D. Atkinson, D. Rueckert, and J. V. Hajnal, "Beyond the g-factor limit in sensitivity encoding using joint histogram entropy," *Magn. Reson. Med.* **55**, 153–160 (2006).