Pergamon

Magnetic Resonance Imaging, Vol. 12, No. 7, pp. 1089-1097, 1994

0730-725X/94 $6.00 + .00

0730-725X(94)00043-3

● *Original Contribution*

MRI SCANNER VARIABILITY STUDIES USING A SEMI-AUTOMATED ANALYSIS SYSTEM

ROSEMARY J. HYDE, JAMES H. ELLIS, EDWARD A. GARDNER, YANTIAN ZHANG, AND PAUL L. CARSON

Ann Arbor Department of Veterans Affairs Medical Center, University of Michigan Medical Center, Department of Radiology, Kresge III, Ann Arbor, MI 48109-0553, USA

Due to the unique design of the Parallel Rod Test Object (PRoTO) and the associated semi-automated analysis program, it was necessary to test it extensively for precision and accuracy, and preliminarily for utility, before its distribution for wider use in MRI system quality control (QC). The test object and analysis program measured the desired quantities reproducibly and they accurately measured predicted changes from intentionally adjusted imaging system parameters, yielding sensitivity of the various test measures to deviation in the system operating parameters. From a single scan of the most recent revision of the test object, multiple quantitative quality control measures were obtained throughout the scanning volume on two MR imaging systems over periods of six and twelve months, respectively. From these and earlier trials, an initial indication was obtained of which performance measures are worth monitoring for QC. This experience suggests that signal-to-noise ratio (SNR) and distortion (including display scale) should be monitored but not necessarily the resolution. The latter was only found to alter at the same time or later than other parameters such as SNR had changed. Slice thickness was found to vary on some units and this measure was also used in normalizing the SNR by voxel volume. SNR, distortion, and resolution measurements using field-echo sequences were less stable than those using spin-echo sequences. Use of this QC program to test a wide variety of image quality measures allowed timely assessment of the long-term variability of the units tested. Long-term variability may become among the most important measures for comparison of system performance and maintenance. Results are still inconclusive on the importance of tracking measures from sequences that are potentially most sensitive to small system misadjustments.

Keywords: **Magnetic resonance imaging (MRI); Quality assurance; Signal-to-noise; Phantom; Image processing.**

INTRODUCTION

Test objects have been developed to measure the quality of images produced by MR systems with much more sensitivity to typical, objectively defined image quality measures than is possible through inspection of clinical images. Typically, these objects are used to verify potential misadjustments of a unit when the clinical images produced by the unit seem degraded to the system users (radiologists and technologists). Although obvious changes would be detected by the users, the effect of subtle changes on ease of diagnosis and treatment planning might be significant, particularly with certain types of performance degradation. One situation where test object quality control would be superior, would be to detect slow degradation in the imaging system that takes place on a time scale that hides it from the system users. If the degradation progressed in this way, a significant reduction could occur in the image quality without the users becoming aware of it.

There is clearly a cost in machine time involved in providing the additional system monitoring required to perform test object-based quality control. The AAPM[1] suggests that resonance frequency, signal-to-noise ratio (SNR), image uniformity, spatial linearity or anisotropy, high-contrast resolution, slice thickness, and slice separation/position be determined for quality control (QC). The time required can be minimized by monitoring only a subset of the image quality parameters determined to be most important[2] or by determining all parameters with a single acquisition. To date most test objects require various inserts and a

RECEIVED 11/16/93; ACCEPTED 4/14/94.
Address correspondence to James H. Ellis, MD, Department of Radiology, B1D520 University of Michigan Hospital, 1500 E. Medical Center Dr., Ann Arbor, MI 48109-0030, USA.

number of scanning sequences before a reasonably complete range of performance quantities can be assessed over the portion of the scan volume appropriate to these measures. A test object has been developed by the authors to measure all of the parameters recommended by the AAPM with a single scan by using computer image analysis. This object and the associated computer code are available currently for testing at MR facilities as part of collaborations with the authors.

The predecessor of the Parallel Rod Test Object (PRoTO) evaluated here was developed and reported previously for use in determining quantitative measures of MR unit performance.[3] Both this and the current device were not of common design since there was only one fixed insert of a checkerboard pattern for each test object, rather than the more usual series of inserts, each designed to test one or a few different performance measures. However, significant questions remained due to the nonstandard geometry. Specifically, the utility of the small image areas used compared to other SNR test objects needed to be tested. The image area used was a compromise between our desire to make regional, geometric-fidelity measures and the need for a large number of pixels over a uniform region for reasonable statistics on measures such as resolution and SNR. Other questions include whether the test-object results were precise and accurate enough for sensitive QC? Could the long-term variability of an MR unit be determined? Specifically, how sensitive was the test object and analysis to small system changes? This study sets out to answer some of these questions.

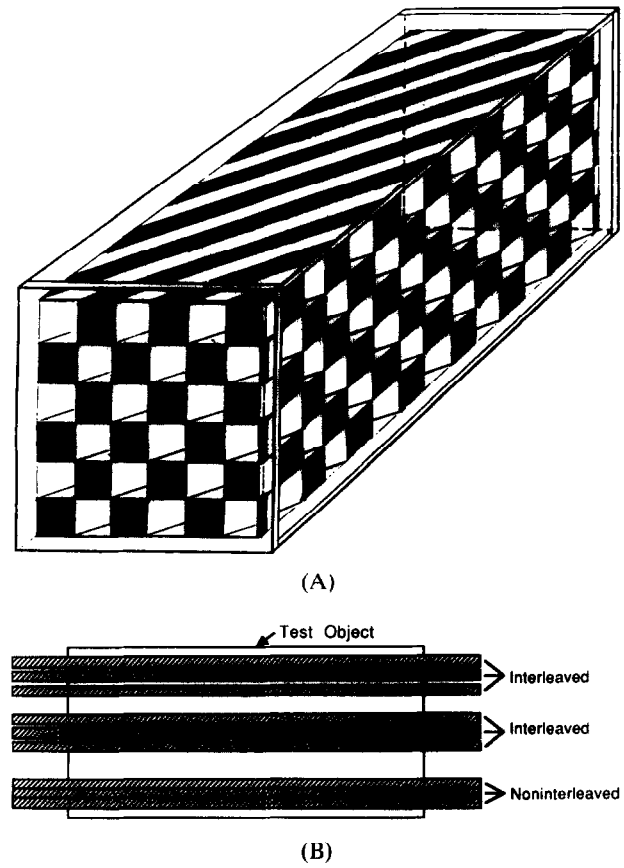


Fig. 1. (A) Test object. The small cross-section approximates the transverse cross-section of a human head and the large cross-section approximates the cross-section of an average human torso. The test object is scanned on a 6° wedge to allow subpixel determination of the edge profiles. (B) Planar view showing slice position of center slice of each group of three.

METHOD

Phantom Specifications

The test object consists of an 18 × 18 × 38 cm rectangular acrylic enclosure (Fig. 1A) containing an array of 60 square acrylic rods. The rods abut the smaller faces and two of the long rectangular faces at a 45° angle. They are parallel to the other two long rectangular faces. The inter-rod spacing is filled with a 400 μM manganese chloride ($MnCl_2$) solution through a fill port on one end of the phantom. The manganese chloride solution produces a high-contrast signal relative to the rods. When properly positioned, the final transverse image of the long axis has a checkerboard appearance (Fig. 2). The test object is imaged on a 6° wedge to increase the sampling rate used for determination of the edge profiles. The rotation is about the axis of the magnet, causing the images of the test object to appear to be rotated 6° counter-clockwise. The exact positioning of the test object and accuracy of the MRI system positioners can be determined optically using centering marks glued to the outside of the phantom or with grooves milled in the inner sides of the test object enclosure. These grooves also allow testing for scan plane tilt. All images acquired in this study are 256 × 256 in size and 10 mm in slice thickness.

Computerized Analysis

The analysis program calculates the mean value of the performance quantities for each individual block as well as for each slice in the image. These performance quantities may be reported as averages over entire slices or the central portion of each slice or individually in the form of functional images of a given quantity.[3] Time variation can also be monitored through variation plots of average values. The performance quantities calculated are spatial anisotropy, distortion and resolution, slice thickness and separation, horizontal and vertical voxel size, and SNR. Slice thickness is calculated from the profile of the block edges where the

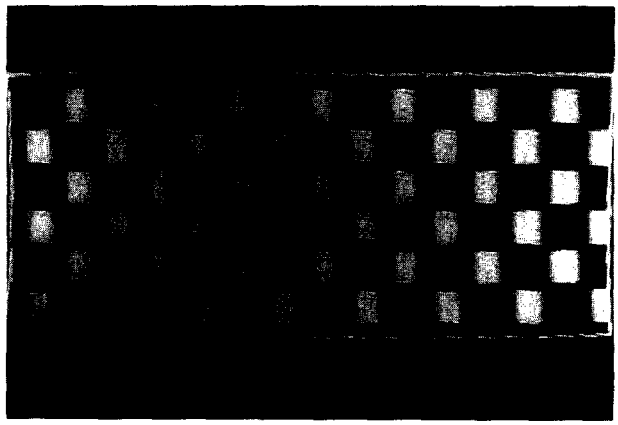


Fig. 2. Checkerboard image of test object showing transverse view of large cross-section.

rod intersects the imaging plane at a 45° angle, while the slice separation is calculated from the relative positions of blocks in adjacent slices. Distortion and anisotropy are calculated from the principal deformation of the image (corresponding to the major and minor axes of the elliptical image of the unit circle). The deviation from the value of 1, of the most affected axis, is the distortion. The ratio of the larger to smaller principal distortions is the anisotropy. High anisotropy shows that one direction on the image is more deformed than the other, while distortion can measure changes in scale. Both distortion and anisotropy are calculated for each signal block from the locations of the outer corners of the four adjacent signal blocks. The resolution is calculated from the sharp transition edge of a signal block. Measures of the radiofrequency (RF) scale factor and signal attenuation are also noted with each scan. Horizontal and vertical pixel sizes are the measured pixel sizes in the horizontal and vertical directions. The SNR is calculated as the ratio of pixel values in signal areas to nonsignal areas, normalized against measured voxel volume that equals the product of the pixel sizes and the slice thickness. Definitions and calculation procedures of these quantities were outlined elsewhere.[3] The analysis program is semi-automatic and, once running, involves no operator intervention. This has removed subjective decisions and made the analysis operator independent. For routine QC tracking, the performance of the unit is determined from time plots of the mean value of each performance quantity over the central area of each slice.

Tests Performed

Precision of this QC method was assessed with reproducibility studies. In these studies, the test object was scanned three times consecutively in a Picker Vista 0.5 T scanner (Picker International, Cleveland, Ohio), with and without repositioning the test object in the magnet between each scan. A spin-echo imaging sequence was used with a TR of 1000 ms and a TE of 30 ms (SE1000/30). Two signal averages were used. A total of nine slices were collected, with a slice gap of 10 mm for six of the slices and zero gap for the other three slices. The field-of-view (FOV) was 40 cm. The image matrix size is 256 × 256. Performance quantities of each scan were extracted from the generated images using the QC analysis system. These performance quantities were then analyzed to study their variation across repeated scans. The variations were reported as percent standard deviation of the mean of each performance quantity. This represents uncertainty in determining these performance quantities under normal conditions due to random, short-term variation in imaging conditions and machine performance. A change in these performance quantities beyond this level could be reasonably attributed to a consistent change of machine performance. Depending on the rate of variation, the machine could be serviced to maintain the optimum level of performance when the image quality was measured to be significantly degraded.

To estimate the accuracy of the QC method, an indirect approach was taken. Instead of comparing the performance quantities obtained using this QC method against hand measured values of these performance quantities, which might suffer from lack of sufficient statistical averaging and, in some cases, be subjective, we compared the performance quantity changes induced by intentional alteration of the nominal MRI scanner settings with the amount of change of the nominal MRI scanner settings. Image quality data, used for this comparison, was obtained from the three center slices of the test object scanned using a spin-echo sequence (SE600/30) on the 0.5 T system. The field of view was 45 cm and slices were separated from each other by 5 mm. There were two signal averages. Images were acquired at the normal settings and with the tip angle of the excitation RF pulse of the spin-echo sequence and the $x/y/z$ gradient amplitudes settings being varied one at a time. The results were reported as the average across all slices of the unbiased r.m.s. deviation of the performance quantities from the theoretically predicted values based on the hardware adjustments. These r.m.s. deviations are considered to be measures of inaccuracy of the machine and the analysis system in tracking the intentional parameter adjustments. It was expected that the SNR will track the changes of the tip angle variation and spatially related performance quantities will follow the gradient amplitude adjustments. Specifically, the x and y gradients alter the voxel size in the direction of the gradient be-

ing varied and, therefore, alter the distortion and anisotropy. The distortion changes proportionally to the gradient since it equals the largest linear elongation or contraction of the image. The anisotropy, however, has an irregular functional form since it is calculated as the ratio of the larger principal distortion to the smaller, which corresponds to the major and minor axes of the elliptical image of a circle. Because the relationship is dependent on the relative magnitudes of the principal distortions, the dependence of the anisotropy on the varied gradient changes from inverse to linear as the gradient surpasses the nominal value. The measured resolution changes also correspond to changes in the y gradient, while the thickness and separation change in proportion to the z gradient. The average measured SNR differences from the expected SNR values corresponding to different tip angle settings of the excitation RF pulse was translated back to degrees. It should be interpreted as the average uncertainty in determining tip angle changes if SNR value is solely relied upon to predict this change.

In order to assess the ability of the test object system to detect local, inhomogeneous distortion, the test object was scanned in the head coil of the 0.5 T system with a small piece of steel wire attached to the outer shell of the test object. The metal piece measured 0.9 mm in diameter and was 1 cm long. It was used to produce a local distortion in the image. The performance quantities were calculated from the generated images to evaluate their response to this type of distortion in comparison with normal measurement variation.

To evaluate MRI system variability using the set of performance quantities obtained from the QC method and to determine which performance quantities are more appropriate for routine QC, the test object was scanned regularly with the 0.5 T scanner and a 1.5 T Signa (General Electric Medical Systems, Milwaukee, WI) over a period of 1 yr and 6 mo, respectively. The time between each scan varied from a week to a few weeks. In the tests performed on the 0.5 T unit and the 1.5 T unit, spin-echo (SE1000/300) and gradient-echo (FE500/18) sequences were used. Nine transverse slices were collected, three at each end and three in the middle (Fig. 1B). One of the end groups was not interleaved and did not have any gap between slices. The other image slices were interleaved and had a slice gap equal to half of the slice thickness of 10 mm. The FOV was 40 cm for the 0.5 T system and 45 cm for the 1.5 T system. Two signal averages were used.

RESULTS

Reproducibility

Table 1 gives the variation of the image parameters (used as performance quantities) of the 0.5 T unit during the reproducibility study. Two groups of data were recorded: one group with the test object repositioned between scans and one without repositioning. These data show the upper limits on the precision with which the performance quantities could be determined. By comparing these two groups of data, the variability associated with positioning the test object could also be estimated. In most performance quantities, there were no significant differences between these two groups of data. However, slice separation and spatial resolution obtained with the spin-echo image sequence seemed to be quite sensitive to repositioning of the test object. Aside from SNR, slice separation and spatial resolu-

Table 1. Reproducibility study of performance quantity measures

	No repositioning		Repositioning between tests	
Performance quantity	Spin-echo	Field-echo	Spin-echo	Field-echo
Horizontal voxel size	0.1	0.1	0.1	0.2
Vertical voxel size	0.2	0.3	0.1	0.3
Separation	1.8	2.3	3.6	2.6
Thickness	0.3	0.3	0.4	1.2
SNR	3.5	3.2	3	4.5
Anisotropy	0.2	0.2	<0.1	0.4
Distortion	0.1	0.2	0.3	0.1
Resolution	1.8	1.9	4.7	1.7

Upper limits on the precision of the performance quantities are calculated as their variability, i.e., percent standard deviation (SD) of multiple measurements of a performance quantity, both with and without repositioning the test object in the magnet between measurements. The standard deviation of the distortion value is given because distortion is defined as a percentage difference.

tion, all quantities could be reproduced to within 1%. Field echo measurements were similar to the spin-echo data.

Accuracy

The accuracy of the QC method was measured in terms of the sensitivity of the performance quantities to changes in nominal MRI system settings. These results are reported in Table 2, where an operational definition of sensitivity is given. The quantified misadjustments of nominal MRI system settings were used to artificially induce a change of performance in the MR unit. How well performance measures follow the expected relationship to the misadjusted machine parameter gave us an estimation of the accuracy of the QC system. Because the nominal tip angle of the excitation RF pulse is 90°, the SNR measurement which tracks this tip angle change does not seem very sensitive, at least in the sense of changes of the tip angle in degrees. This is because the image intensity is proportional to sine of this tip angle, which varies least at angles near 90°. Figure 3, showing the SNR as a function of tip angle, is representative of the agreement between theory and experiment achieved. Using distortion measurement, changes of x or y gradient below 1% were detected. If anisotropy measurement was used, the minimum change to the gradient that could be detected was 2–2.4%. This indicates distortion was a more sensitive indicator than anisotropy to spatial changes in images. Changes in the z gradient of 3% and 5.5% can be tracked with slice thickness and separation respectively. In general, comparing Table 1 and Table 2 shows that the measurement precision for a performance quantity

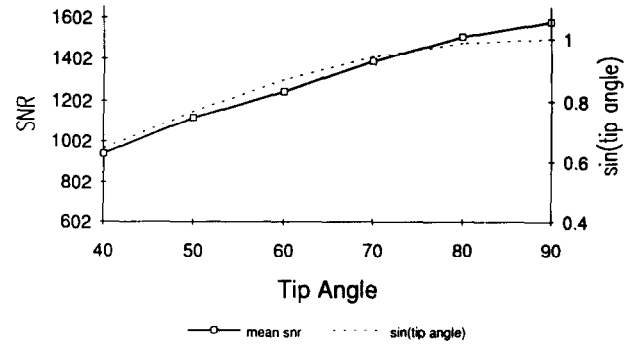


Fig. 3. Measured variation in SNR as function of tip angle. The dashed line is the theoretical sine curve.

determines the minimum change to the MR system setting that can be discriminated from normal variation of the MR unit and measurement device.

Long-Term Variability

Results of the long-term variability study of the various performance quantities calculated using the test object in the 0.5 T Vista (Picker) unit and in a 1.5 T Signa (GE) imager are listed in Table 3. The variability was calculated as the standard deviation of the quantity as a percentage of the mean with the exception of the distortion, where the standard deviation is quoted, since distortion is already a percent difference. Performance quantities extracted from spin-echo studies were more stable temporally than the values obtained from field echo studies. All performance quantities extracted from spin-echo studies have less than 10% average vari-

Table 2. Accuracy estimation of the QC method under study, measured as the sensitivity of performance quantities to changes of MRI scanner settings

Machine parameter	Performance measure	Functional relationship	Sensitivity
Tip angle	SNR	sin(tip angle)	16°
x Gradient	Horizontal voxel size	1/(x-gradient)	1.1%
	Distortion	1/(x-gradient)	0.7%
	Anisotropy	x-gradient, 1/(x-gradient)	2.4%
y Gradient	Vertical voxel size	1/(y-gradient)	0.9%
	Distortion	1/(y-gradient)	0.7%
	Anisotropy	y-gradient, 1/(y-gradient)	2.0%
	Resolution	1/(y-gradient)	1.6%
z Gradient	Slice thickness	1/(z-gradient)	3.0%
	Slice separation	1/(z-gradient)	5.5%

Sensitivity values are defined operationally as the average across three slices of the unbiased r.m.s. deviation of the measured performance quantities from their theoretically predicted value evaluated using their functional relationship with the machine parameters. Variations larger than the stated sensitivity are detectable with a probability exceeding 68%.

Table 3. Long-term variability is the percent standard deviation of a performance measure obtained in studies over a period of several months and separated by at least several days

	0.5 T		1.5 T
Performance quantity	Spin-echo	Field-echo	Spin-echo
Horizontal voxel size	0.6	0.7	1
Vertical voxel size	0.4	0.4	0.5
Separation	4	4	3.9
Thickness	3	12	1.1
SNR	6	20	13
Anisotropy	1	2	0.4
Distortion	0.8	2.8	0.5
Resolution	4	8.3	2.8

The standard deviation of the distortion value is given because distortion is a defined as a percentage difference.
All slices: slice thickness = 10 mm, slice gap = 5 mm. Spin echo sequence: TR = 1000 ms/TE = 30 ms, field echo sequence: TR = 500 ms/TE = 18 ms.

ation over the year in which the data were collected on the 0.5 T unit. Data were compiled from the 1.5 T scanner for 6 mo. There were 25 and 8 data points collected for the 0.5 T and the 1.5 T scanners, respectively. Although studying only two scanners makes this data somewhat anecdotal, the scanners were in clinical operation and had routine manufacturer service (under contract) for the entire test period, making their performance representative of the performance of other clinical machines.

Figure 4 shows the long-term variability in the distortion for the 0.5 T unit. The other performance measures could also be displayed in this manner. Control limits (running mean ±2 standard deviations) were calculated to determine if the unit varies from the norm. The horizontal voxel size and anisotropy results follow trends similar to that of the distortion data. The shift in distortion in December corresponded to an increase in voxel size that was also noted by the service personnel. The voxel size was decreased in April. The field-echo results had a similar time history as the spin-echo results, but tended to show more week-to-week variation. Although there was occasionally a spurious measurement, the test object and analysis generally tracked the performance of the MRI unit. Figure 5 shows the variation in the SNR for the 0.5 T unit as measured with the PRoTO, compared with the Picker SNR test object results measured by the Picker service engineers. The different scale is due to the PRoTO data being normalized against voxel volume.

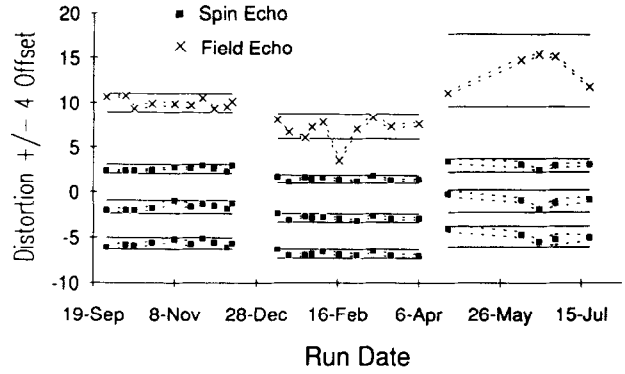


Fig. 4. Distortion data collected over a year. The plotted values are the mean distortion value for the slice, SD of mean (dashed line), and running mean ±2 SD (solid line) over periods in which the performance was stable. The distortion value is plotted for the center slice of each group of three for spin-echo and for one center slice for field echo. The slices have been offset (±4%) for clarity.

Local Distortions

The images produced by placing a 1-cm length wire on the outside of the test object are clearly distorted as shown in Fig. 6. The worst, average distortion measured by the program over any nine-block region in Fig. 6A was −1.0% and in Fig. 6B was −4.1%. The visible and measured distortion varied in each slice and from slice to slice. Only four of the slices were distorted and the sides of these images opposite the wire were normal.

DISCUSSION AND CONCLUSIONS

Estimates of the sensitivity of the Parallel Rod Test Object (PRoTO) and the associated semi-automated

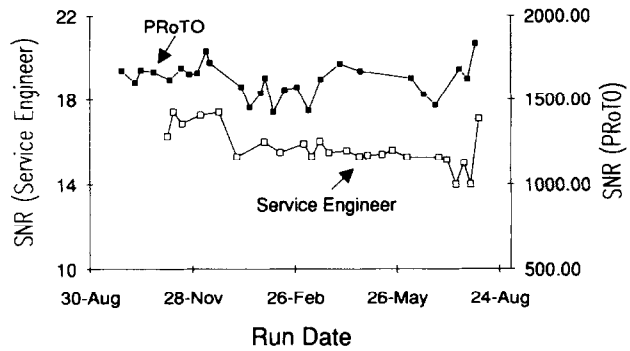
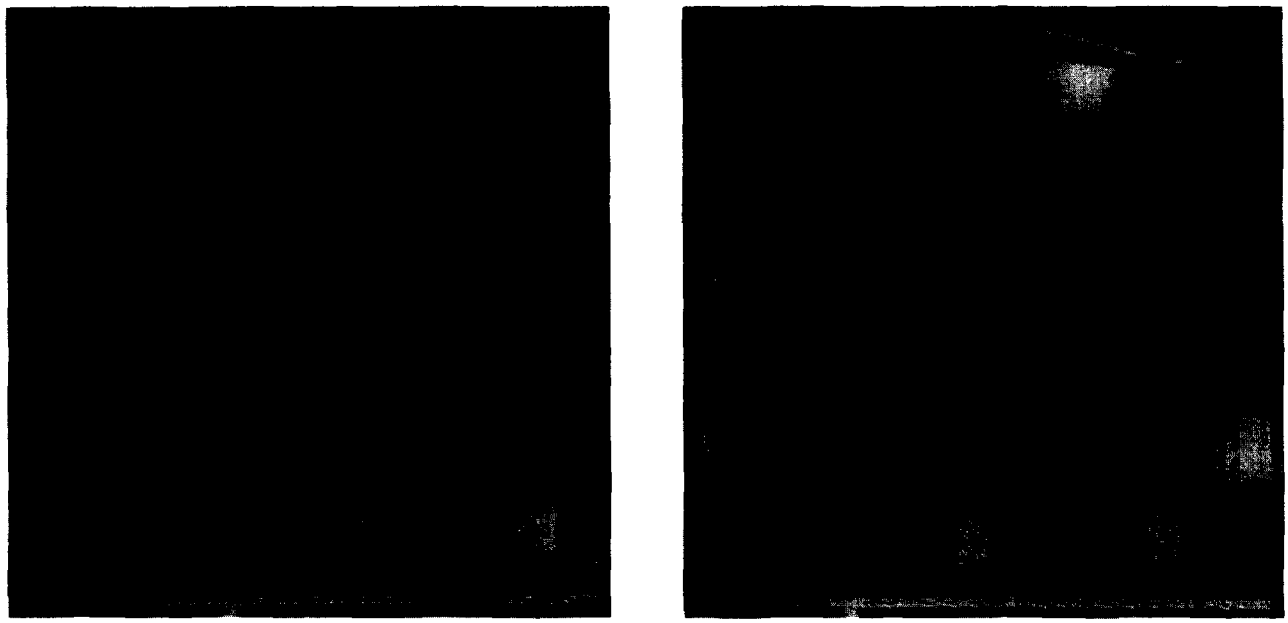


Fig. 5. Percentage difference from the mean of the signal to noise ratio obtained for spin-echo sequences with the parallel rod test object (PRoTO) and the Picker SNR test object. SNR determined with the PRoTO is volume normalized to cm^3.



(A) (B)

Fig. 6. Images of test object in 0.5 T head coil in a plane parallel to square face, when a 1-cm length of steel wire was attached to the head coil frame. The plane of (A) is 1.4 cm from location of the wire, which is in the plane of (B).

analysis program were determined from the smallest change of a performance measure which could be attributed to a consistent change in the unit rather than testing variability. It was bounded by the precision and accuracy with which the performance measures could be determined.

The reproducibility studies on the 0.5 T unit (Table 1) show upper limits on the precision with which the performance quantities can be determined. Except for SNR, slice separation and resolution, all the measured quantities could be reproduced to within 0.4% for spin-echo sequences and 1.2% for field-echo sequences, with or without the phantom repositioned between scans. For most measurements, variability over multiple studies, with repositioning the PRoTO between scans, were similar to results obtained when the test object was not repositioned between scans. This indicated that the test object can be scanned reliably over time in the MR unit by demonstrating that experimental error in position is dominated by statistical error of the measurement system. Repositioning error may have been small due to accurate placement of the PRoTO in the MR unit bore using laser alignment or because the analysis program was tolerant of imprecise positioning. It also indicated that motion of the fluid in the relatively confined cells of the test object was not too great after rapidly repositioning the test object.

Slice separation varied more than thickness, distortion, and anisotropy (with or without repositioning of the test object). The difference could be explained partially by the way in which slice separation is determined. Slice separation was calculated as the distance between the middle of a slice thickness edge of a signal block of an image slice, to the middle of the slice thickness edge of the corresponding signal block of the next image slice. Each extra step taken to determine slice separation in addition to the determination of slice thickness will introduce some error. Calculation of slice separation as the difference of two quantities will also enhance the error by square root of two. If, in the future more precision is required, then the variability in slice separation measurements could be reduced by significant reprogramming.

By scanning the PRoTO with an MR unit that was purposefully misadjusted, we were able to confirm the expected functional dependence of performance measures on machine parameters (Table 2) and produce an accuracy estimate of the QC system. The values quoted are the average across all slices of the unbiased r.m.s. deviation from the theoretically predicted value of the measured quantity. These values are calculated using theoretical relationships of the performance quantities to the machine parameters. Because this estimate is based on greater fluctuations in the machine operat-

ing parameters than are encountered during normal QC and on the accuracy with which these parameters could be controlled, this measure probably underestimates the accuracy of the test system.

From this group of data, it can be concluded that a higher degree of accuracy can be achieved with our system by using distortion instead of using anisotropy or spatial resolution to determine spatial changes in the image plane. Similarly, it is better to use slice thickness instead of slice separation to detect z gradient-related changes, at least when the QC system under study is used. Given that the accuracy of the performance measuring system is reasonably determined, and that its reproducibility is well characterized, the test object can be employed to evaluate the accuracy or at least relative accuracy of the imaging system under various imaging conditions that are employed clinically and that might be most desirable for quality control measures.

Local distortion produced by attaching a steel wire to the head coil was easily visible on this very regular test object pattern. The distortion measured by the test object analysis system was significantly different at the location of the wire, but was relatively small because of the spatial averaging employed for stability in routine quality control. Special automated processing for detection of small local distortion may not be necessary, but at least visual inspection should be performed of images made throughout a large, coil-filling test object of regular shape and preferably regular pattern to detect small local distortions that can be caused by accumulation in the magnet of very fine magnetic materials.

The long-term variation of image parameters was greater than the variability of the test object and program (Table 3), which allowed for detection of machine changes over time (Figs. 4 and 5). As expected, the spin-echo parameters were more stable than the field-echo parameters, but large differences between the pulse sequences were noted only for SNR, distortion and resolution. The larger variation in the SNR with the field-echo sequence was due, at least in part, to structure noise originating from ringing at the signal/non-signal edge transition region resulting from finite sampling of the sharp block edges.[4] The SNR variability results from the 0.5 T unit have been compared to the service engineer's weekly figures. The latter were acquired using a Picker 19-cm uniformity phantom from which the signal is calculated from the central 75% of the area of the test object image and the noise from a similar area of an image acquired with no RF. The variability of the service engineer's spin-echo SNR results, measured with the rigorous but more time consuming standard method, was 2% as compared to the 6% we measured. For field-echo sequences, the Picker service engineer measured the long-term variability of the SNR to be 8%, again significantly lower than the 20% we measured. Figure 4 shows the normalized variation of the SNR measured with the PRoTO and with the Picker test object using similar spin-echo sequences. Clearly there was greater variation in the measurements made with the PRoTO, possibly resulting from structure noise. However, a long-term reduction from November to February was apparent in both data sets. If, in the future, it is concluded that structure noise is affecting the SNR too greatly, then a uniform section could be added to the end of the test object. After positioning this new insert near the center of the magnet, repeat scanning with no RF would give the noise value. To date that has not been necessary.

Although this paper does not attempt to prove the utility of this object system at detecting relevant system errors as they occur in practice, the reduction in SNR apparent from November to February (see Fig. 4) is an example of a case when a real problem was detected with this test object. During this period, the clinical users of the system complained of a reduction in image quality. System service personnel were called and an RF amplifier was replaced, bringing the performance of the unit back to pre-event levels. The ability of the test object to detect clinically relevant image degradation was shown by this event, although the reduction in SNR was sufficient, in this case, to be detected by the clinical users of the unit.

In another example, on the 1.5 T unit the alternation of the SNR between significantly high and low levels on odd and even slices was documented by performance measurements[4] and this documentation was instrumental in getting the system adjusted. As is often the case, some radiologists recalled after the fact that they had noted or suspected that situation under some operating conditions, but had not taken strong action. Availability of a documented record for the system and even comparisons with other systems can be helpful in deciding whether an observed condition or change is abnormal or should be tolerated.

From Table 3, it can be seen that there is some difference in the long-term variability of the different parameters in the machines tested. We suspect that this may prove to be a valid form of comparison between various units from the same manufacturer or between comparable units of different manufacturers. Parameter variability may reflect differences in field strength, engineering choices and design, and differing architecture.

From our experience obtained from scanning the test object and from the results obtained using the semi-automated analysis system, it seems necessary to mon-

itor a selection of, but not all, the suggested QC measures regularly. SNR is sensitive to many abnormalities of an MR system and is usually deemed a very important parameter in any QC program. Our result is consistent with this judgment. Slice thickness was more sensitive than slice separation to changes in the z-direction with our analysis algorithm. It appears, therefore, unnecessary to monitor both for the same purpose.

It was also found necessary to monitor the distortion. With the semi-automated analysis system, anisotropy and resolution tended to follow trends seen in the distortion data but were less sensitive measures. This conclusion contrasts with some common procedures, where only the anisotropy (often called the distortion) is recorded. In that practice, overall display scale is ignored. This is done primarily because the proportions of imaged objects are deemed more important than the overall size (within limits), although both quantities may be important in some medical applications (e.g., radiation therapy planning).

While some of these conclusions are specific to the methods we employed to measure these performance quantities, nevertheless, the test object and analysis program have been shown to be capable of determining quantitative measures of an MRI unit. The QC protocol is quick and easy to perform and takes less than 15 min of machine time. The processing takes more time but is semi-automatic and therefore does not need operator intervention after the initial setup. Image processing software is run on a PC and can run on other computers that support ANSI C programming language with slight modification. To analyze obtained images, it is necessary to remove the image header first. A limited knowledge of the vendor image format is therefore required.

The test object and analysis have been shown to be sensitive to small system changes. As is shown in Table 2, the system was able to detect gradient changes above 0.7% with 68% probability. The test object and analysis program allow the precision and long-term variability of a unit to be determined. The measurement uncertainty was shown to be relatively independent of the test object positioning and much smaller than the variability of the unit tested. The variability measurements, in particular, would be a reasonable way to compare machines. Our results indicate that a subset of the recommended image quality parameters could be monitored without a reduction in the rate of detected machine deviations. By periodically performing a single scan of the test object described here, the performance of the MRI unit could be monitored and the most common system malfunctions detected.

Acknowledgments—This material is based upon work supported in part by the Office of Research and Development, Department of Veterans Affairs Merit Review Grant "Automated Analysis of Magnetic Resonance Imaging System Performance." The authors thank Hong Yeung, PhD, and Tom Chenevert, PhD, for many helpful consultations. We thank Brian Murphy for the preliminary work with this test object, Picker engineer Gregory Kushba for his assistance, and Elaine Grech and Olga Matlega for assistance in manuscript preparation.

REFERENCES

1. Quality assurance methods and phantoms for MR imaging: Report of AAPM NMR Task Group No. 1. *Med. Phys.* 17:287–295; 1990.
2. Rasmusson J.J.; Gray, J.E. Automated MR imaging quality control with clinical pulse sequences (abstract). *Radiology* 185(P):392; 1992.
3. Covell, M.M.; Hearshen, D.A.; Carson, P.L.; Chenevert, T.P.; Shreve, P.; Aisen, A.M.; Bookstein, F.L.; Murphy, B.W.; Martel, W. Automated analysis of multiple performance characteristics in magnetic resonance imaging systems. *Med. Phys.* 13:815–823; 1986.
4. Murphy, B.W.; Carson, P.L.; Ellis, J.H.; Zhang, Y.T.; Hyde, R.J.; Chenevert, T.L. Signal-to-noise measures for magnetic resonance imagers. *Magn. Reson. Imaging* 11: 425–428; 1993.