Original Research

Quantitative Myocardial Infarction on Delayed Enhancement MRI. Part II: Clinical Application of an Automated Feature Analysis and Combined Thresholding Infarct Sizing Algorithm

Li-Yueh Hsu, DSc, W. Patricia Ingkanisorn, MD, Peter Kellman, PhD, Anthony H. Aletras, PhD, and Andrew E. Arai, MD*

Purpose: To compare global and regional myocardial infarction (MI) measurements on clinical gadolinium-enhanced magnetic resonance (MR) images using human manual contouring and a computer algorithm previously validated by histopathology, and to study the degree to which visual assessment and human contouring of infarct extent agreed with the computer algorithm.

Materials and Methods: Infarct size in 20 patients was measured by human manual contouring and with an automated feature analysis and combined thresholding (FACT) computer algorithm. Short-axis slices were divided into myocardial sectors for regional analysis. Extent of infarction was also graded visually by consensus of expert readers and compared to human and computer contouring.

Results: Despite good correlations (R = 0.93–0.95) between human contouring and the FACT algorithm, human contouring overestimated infarct size by 3.8% of the left ventricle (23.8% of the MI) area ($P < 0.001$). Human contouring also overestimated the circumferential extent, transmural extent, and extent of infarction within a sector by 7.1%, 18.2%, and 27.9%, respectively (all $P < 0.001$). Both consensus reading and human contouring overestimated infarct grades compared with the FACT algorithm ($P = 0.002$ and $P < 0.001$).

Conclusion: Clinically relevant overestimation of MI can occur in visual interpretation and in human manual contouring, particularly with respect to extent of infarction on a regional basis.

Key Words: myocardial infarction; myocardial viability; magnetic resonance imaging; computer aided diagnosis; contrast agent; gadolinium
J. Magn. Reson. Imaging 2006;23:309–314.
Published 2006 Wiley-Liss, Inc.†

MAGNETIC RESONANCE (MR) delayed contrast-enhanced methods for imaging myocardial infarction (MI) are currently the highest resolution techniques to assess viability clinically. Validation studies showed that the contrast distribution accurately reflects the underlying pathology of the myocardium (1,2). However, most studies have relied on visual interpretation or used simple intensity thresholding to quantify contrast enhancement area. Subjective factors may lead to user-related measurement errors. There is a need for objective methodology, such as the knowledge-based computer algorithm that was presented in Part I of our study, to improve the accuracy and reproducibility of quantifying MI on contrast-enhanced MR images.

Wellnhofer et al (3) in a recent study suggested that dobutamine MR studies more accurately predict recovery of function after coronary artery revascularization than delayed enhancement imaging. Perhaps some of the difficulty in predicting recovery of function related to lack of standardization in quantifying the extent of myocardial infarction. Kim and Manning (4) concluded that many viability disagreements occurred in segments where the transmural extent of infarction was intermediate in severity—a range where most of the clinical uncertainty exists. These intermediate gray levels of signal intensity are the regions in which computerized quantification can yield the most significant improvement.

Human visual perception has fundamental difficulty determining the middle gray across a range of intermediate brightness, a potential problem when interpreting gray scale MR images at a lower resolution. The lightness perceived in the retina follows a logarithmic response and is more receptive to the bright side (5). Other factors such as variations in ambient light and

Laboratory of Cardiac Energetics, National Heart Lung and Blood Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, USA.

Contract grant sponsor: National Heart, Lung, and Blood Institute, National Institutes of Health.

*Address reprint requests to: A.E.A., Laboratory of Cardiac Energetics, National Heart, Lung, and Blood Institute, National Institutes of Health, 10 Center Dr., MSC 1061, Building 10, Room B1D-416, Bethesda, MD 20892-1061. E-mail: araia@nih.gov

Received February 19, 2005; Accepted October 28, 2005.
DOI 10.1002/jmri.20495
Published online 31 January 2006 in Wiley InterScience (www.interscience.wiley.com).

the computer display settings further impair human interpretation since visual perception depends strongly on the illumination level.

In addition to visually discriminating signal intensity changes, the expert human uses empirical knowledge to decide where to select infarct territories. One example of this would be ignoring random bright pixels or moderate-sized bright patches interpreted to represent blood as opposed to non-transmural infarction adjacent to the right ventricular cavity. Thus, computer algorithms could offer the following benefits: 1) objective analysis of infarct regions, 2) elimination of user empirically defined thresholds, and 3) image processing techniques based on expert reader criteria to reduce false positive bright regions.

MATERIALS AND METHODS

Study Design

In this study, we compared human manual contouring of MI to a quantitative computer algorithm previously validated by animal triphenyltetrazolium chloride (TTC) in Part I of our study. We first applied this automated feature analysis and combined thresholding (FACT) computer algorithm on a series of clinical cases to study human contouring of infarct size in a global (per volume and per slice) basis. Since clinical viability assessment is interpreted on a regional (per segment) basis, we then compared the extent of MI within a sector as measured by human and computer contouring. Finally, we studied the degree to which visual assessment and human contouring of infarct grades agreed with the FACT algorithm.

Patients

Twenty patients (17 men, mean age 60 ± 13 years, range 41–79 years, 11 acute and nine chronic MIs) referred from a local community hospital with known MI (eight left anterior descending, six right coronary artery, and six left circumflex artery) were studied to compare infarct size measured by visual assessment, human manual contouring, and the FACT algorithm.

Image Acquisition

Infarct imaging was performed on a 1.5-T scanner (GE Medical Systems, Waukesha, WI, USA) approximately 20 minutes following intravenous injection of gadolinium diethyltriaminepentaacetic acid (Gd-DTPA) (0.2 mmol/kg) using an inversion recovery fast gradient-echo sequence triggered every other heartbeat. Typical imaging parameters of the study included TE 3.4 msec, TR 7.8 msec, approximately optimized inversion time (TI typically 300 msec), and bandwidth ± 31.25 kHz. Images were acquired at end-expiration and lasted about 12 heartbeats. The average in-plane image resolution was $1.4 \times 2.8 \times 8.0$ mm with a slice separation of 3 mm. All images were corrected for surface coil intensity variation and used the phase sensitive inversion recovery (PSIR) reconstruction method (6).

Image Analysis and Statistics

For each patient, all short-axis images (excluding the most basal and apical slices) were used for volumetric coverage of the left ventricle (LV), which included a total of 161 slices (mean 8.1 ± 0.9 slices, range 7–10 slices). All short-axis slices of each patient were divided into 18 sectors (six basal, six mid, and six apical sectors) resulting in a total of 360 sectors for analysis on a per sector basis. Infarct size was measured by manual planimetry and with the FACT computer algorithm. One set of epicardial and endocardial contours were manually traced on all MR images to define myocardial regions for both methods. These myocardial contours served as the initial input for the FACT algorithm. All images in this study were processed using the same computer parameter settings as in Part I of our validation.

Comparisons of the human manual contouring vs. the FACT algorithm were performed using linear correlation (7) and Bland Altman analysis (8). A paired t-test with Bonferroni correction for multiple comparisons was performed to determine if there was a significant difference ($P < 0.05$) between different methods. The result was reported as percent of the LV volume (defined as the mass of LV myocardium) infarcted, percent of the slice infarcted, and percent of the sector infarcted.

In addition, the transmural extent of infarction was graded visually by consensus of three expert readers (level III trained cardiologists under SCMR guidelines) and reported using a clinical scale for transmural extent of infarction (grade 0 = 0%, grade 1 = 1–25%, grade 2 = 26–50%, grade 3 = 51–75%, and grade 4 = 76–100%) (9). The results of human and computer infarct contouring on each slice were divided into sectors and then translated to these grades according to actual sector wall thickness and compared to the consensus reading.

The volume and slice based global measurements as studied previously in Part I of our animal validation provide a method to reduce registration errors in the intermodality comparison between MR and histopathology. These volume and slice based measurements have been commonly reported as infarct size indices in prior patient studies (10–14). The sector based regional assessment, however, allowed a direct comparison of clinically relevant viability measures within a sector including circumferential extent, transmural extent, and percent of sector infarcted.

RESULTS

Examples of clinical infarctions demonstrate that the FACT algorithm can delineate infarct areas in three coronary territories and at different short-axis locations (Fig. 1). Ischemic hypo-enhanced myocardium such as microvascular obstruction areas in Fig. 1 are considered as part of the infarct region by both the FACT algorithm and human manual contouring.

While good correlations of human manual contouring and the FACT algorithm were achieved on volume-by-volume (R = 0.95), slice-by-slice (R = 0.93), and sector-

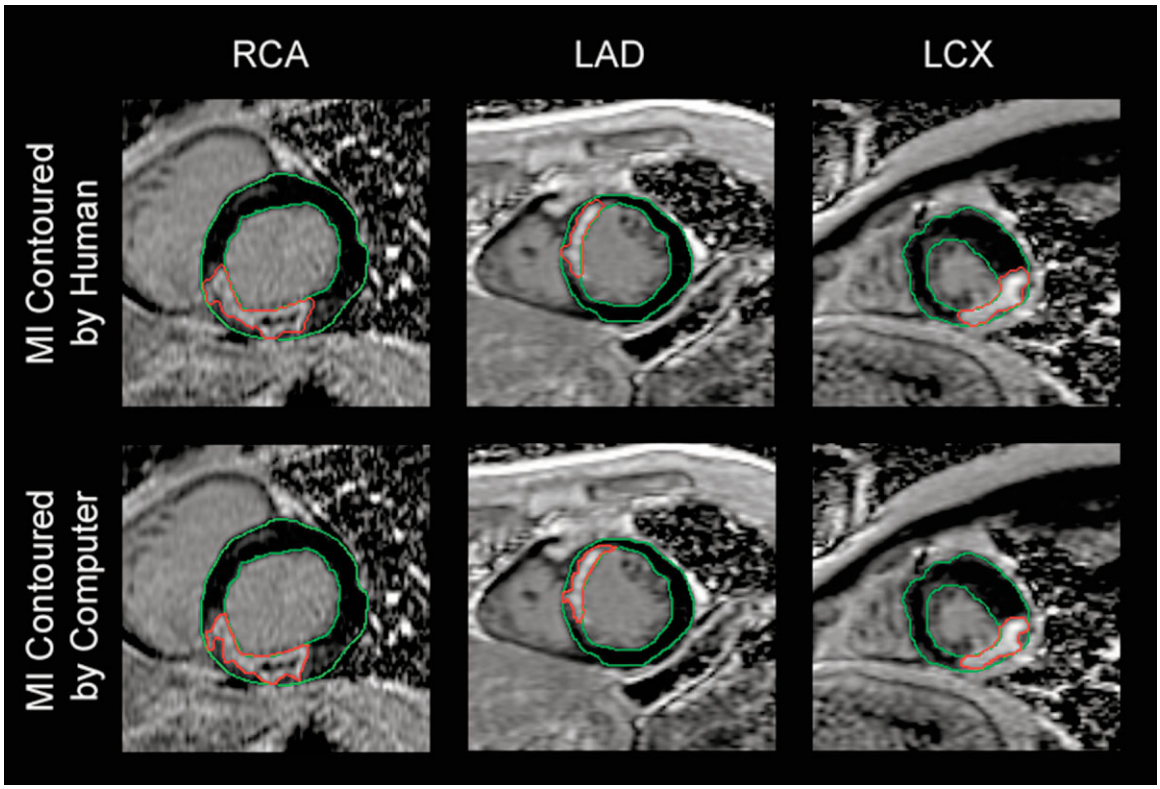

Figure 1. Examples of infarct contours in three different patients demonstrated the FACT computer algorithm works in all three coronary territories and at different short axis locations. (LAD: left anterior descending, RCA: right coronary artery, LCX: left circumflex artery.) The results are comparable to human manual contouring and show the algorithm considered areas of microvascular obstruction as part of the infarction (in the RCA example).

by-sector (R = 0.93) based analysis (Fig. 2), Bland Altman analysis showed that human contouring overestimated infarct size relative to the computer algorithm. In addition, the scatter of errors increased from volume-by-volume based analysis to sector-by-sector based analysis. While human contouring overestimated infarct size by 3.8% of the LV area, this error was equivalent to 23.8% of the MI area ($P < 0.001$) (Fig. 3). The overestimation error measured in percent of LV was considerably smaller than using the percent of MI as the denominator. This is because 84.1% of the patient myocardium was not infarcted and using the percent of LV measurement suppressed the degree of overestimation.

Since viability assessment is typically performed on a segmental basis, we analyzed how the relatively small global (per volume and per slice) overestimation in measuring infarct size propagated into regional (per sector) errors in determining circumferential, transmural, and percent of sector infarction. Using percent of sector infarcted as the denominator, the errors in the transmural extent of infarction were 2.6 times larger than the errors in the circumferential direction (18.2% vs. 7.1%). Circumferential and transmural errors multiplied to result in an even larger overestimation of percent of a sector that appeared infarcted (27.9%, $P < 0.001$). The Bland Altman plot showed a wider scatter and larger error on per sector analysis compared to per volume and per slice analysis (Fig. 2). This was most evident in segments where the transmural extent of infarction ranged from 25 to 75%.

The contingency table of regional viability assessment tabulates the agreement in the transmural extent of infarction using the clinical grading scale previously defined (Fig. 4). For differentiating normal from infarcted segments, there was excellent agreement between the computer and human contouring on 93.9% of sectors (338/360). For the 161 sectors where both the computer and human identified an MI, the transmural extent of infarction agreed completely in 96 sectors (59.6%). However, the transmural extent of infarction was one or two grades higher by human contouring in 57 of the 161 infarcted sectors (35.4%).

When a clinical infarct image is displayed, it can be difficult to perceive borders of the middle gray value due to limited contrast levels and different computer monitor settings (Fig. 5). Patches of intermediate signal intensity due to partial volume effect led to a larger apparent infarct size on some of the human contoured results compared to the computer algorithm. Both consensus reading and human contouring overestimated infarct grades compared with the FACT algorithm ($P = 0.002$ and $P < 0.001$, respectively; Fig. 6).

DISCUSSION

To objectively measure the size of MI on clinical contrast-enhanced MR images, a FACT computer algo-

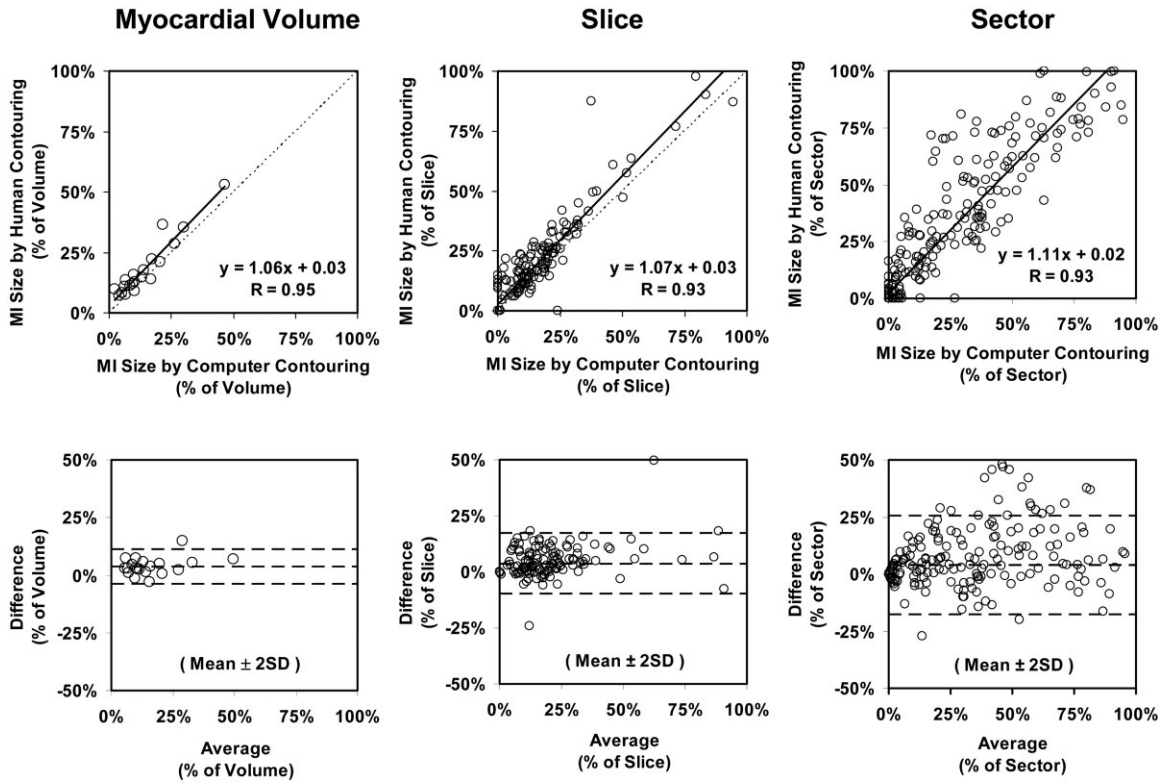

Figure 2. Results of computer and human infarct size measurements are shown in per volume (left), per slice (middle), and per sector (right) based comparisons. "Volume" refers to combining all slices for a given heart into a single volumetric measurement. "Slice" indicates analysis on a slice-by-slice basis. "Sector" refers to analysis performed at the level of dividing a slice into six 60° sectors. Comparable to the previous animal validation, human contouring also overestimates infarct size on the clinical cases. This overestimation and scatter of errors are more evident when the total myocardial area is smaller as in the per sector based comparison.

rithm previously validated by canine histopathology was used as the reference standard to assess the accuracy of human planimetry. Clinically relevant errors in determining regional viability (Fig. 2, right panel) were present despite apparently good correlations in measuring global infarct size (Fig. 2, left and middle panels). Thus, analysis methods as used in prior patient studies may have introduced errors in estimating the degree of delayed enhancement (10–14).

There are several potential clinical implications of the current study. While global infarct size is important, viability assessment attempting to predict recovery of function is usually performed on a segmental basis. Using five clinical MI grading scales, human contouring of the transmural extent of infarction were one to two grades more extensive than the computer algorithm in 35.4% of infarcted segments. These one to two grade discrepancies in transmural extent of infarction (Fig. 4) translate to 25–50% differences in the likelihood of functional recovery after revascularization (9). One should note that most regional viability disagreements occurred in segments where the transmural extent of infarction was intermediate in severity—a range where most of the clinical uncertainty exists (4). Thus, our findings may also explain some of the paradoxical results in the literature where ex vivo high resolution images and correlations look remarkably good (1,2) but

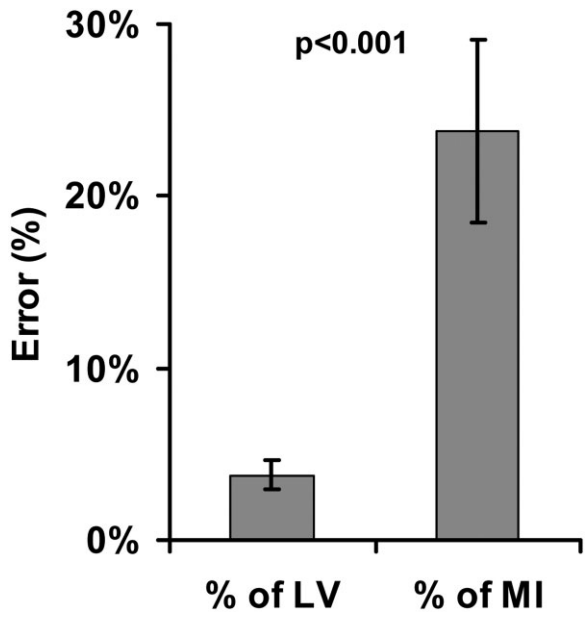

Figure 3. The average error of infarct size overestimation by human manual contouring with respect to the FACT algorithm shows a similar pattern to the prior animal validation study. Large errors in the infarct size ($P < 0.001$) are hidden when divided by the much larger fraction of normal myocardium in calculating percent of LV.

N=360	Human 0%	Human 1-25%	Human 26-50%	Human 51-75%	Human >75%
Computer 0%	**177**	10	0	0	0
Computer 1-25%	11	**45**	18	5	0
Computer 26-50%	1	4	**27**	21	3
Computer 51-75%	0	0	2	**12**	10
Computer >75%	0	0	0	2	**12**

Figure 4. Contingency table for regional viability assessment shows that human manual contouring overestimated 57 of 161 infarcted sectors (35.4%) by one to two grades compared to the FACT computer algorithm.

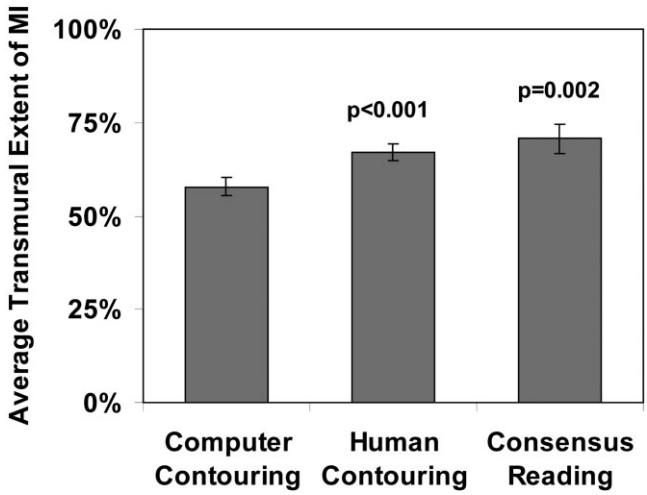

Figure 6. MI as assessed by the FACT computer algorithm, human manual contouring, and consensus reading in 20 patients. Average transmural extent of infarction is graded on all infarct sectors across three short axis locations (apex, mid, and base). (Grade 0 = 0%, grade 1 = 1–25%, grade 2 = 26–50%, grade 3 = 51–75%, and grade 4 = 76–100%.)

clinical studies find it difficult to predict regional viability (3).

For monitoring progression of disease or regression following treatment, improving the precision of the infarct size measurements should only increase the diagnostic and prognostic accuracy of MR infarct imaging. The improved precision of computer infarct sizing may also allow significant decreases in clinical trial sample sizes. For example, these methods may be a useful and reproducible technique across multiple centers to study novel therapies for infarct size reduction.

The current study is the third recent study that indicates human manual contouring overestimates MI size (see Ref. 15 and Part I of our study). We demonstrated in Part I of our animal study that human contouring overestimated infarct size while the FACT computer algorithm was quite accurate on the same images. The degree to which human manual contouring overestimated animal infarcts was similar to the amount of overestimation of clinical infarcts in this study (5.4% in animal in vivo vs. 3.8% in clinical). In addition, another study (15) showed human contouring also overestimated infarct size by a similar degree (8.6%). We suggest this is due to the partial volume effects that appear most prominent on the infarct border. The geometry of a MI sector results in a situation where a one- to two-pixel overestimation of the infarct border in the shorter transmural direction becomes a large percentage error than in the longer circumferential direction. This explains why the error in the transmural direction across the LV wall was twice as large as in the circumferential direction.

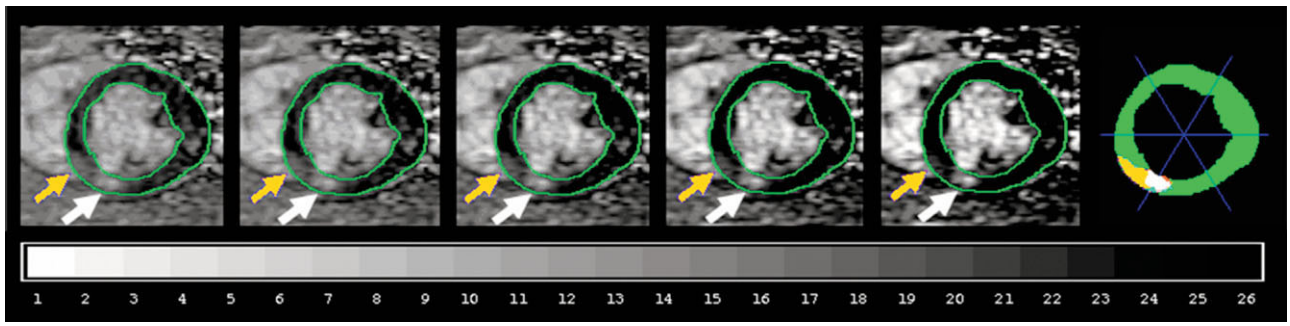

Figure 5. A clinical infarct displayed at five different brightness and contrast levels to illustrate difficulties with the visual interpretation of intermediate gray patches on MR images. The yellow arrow points to a gray patch that is brighter than normal myocardium but not bright enough to exceed the full width half maximum threshold determined by the FACT computer algorithm. The color diagram displays pixel agreement of human and computer measurements: green was normal by both; white was infarcted by both; yellow was infarcted by the human but not the computer; and orange was infarcted by the computer but not the human. The discrepancy between human and computer contouring represents a large fraction of a typical myocardial sector (blue lines). To properly visualize this figure, the computer monitor should be adjusted to exhibit 26 shades on the gray scale bar.

Limitations

Delayed contrast-enhanced MR imaging in principle only provides an anatomical extent of non-viable myocardium. Even with the proposed objective analysis of delayed enhancement images, additional functional information such as dobutamine exam (3) may still be necessary for accurate prediction of recovery of function. In addition, with improved contrast and spatial resolution of delay enhancement imaging in the future, visual interpretation and human manual contouring may improve.

CONCLUSION

In conclusion, despite accurate representation of MI by gadolinium delayed enhancement, clinically relevant overestimation of the infarct size can occur in visual interpretation and in human manual contouring, particularly with respect to the transmural extent of infarction. Using the previously validated FACT infarct sizing algorithm as a reference, human contouring was shown to overestimate regional measurements of MI despite good correlations in global and regional analysis. The magnitude of these errors and the frequency of misclassification of the transmural extent of infarction are large enough to impact clinical viability assessment. When performing infarct imaging, accurate analysis and high quality image acquisition are equally important.

ACKNOWLEDGMENT

The authors thank Christine H. Lorenz, PhD, Siemens Corporate Research, Princeton, NJ, for many helpful suggestions regarding the method and study design. The authors are also grateful for the technical assistance of Glenn Meredith (Siemens Corporate Research, Princeton, NJ) for integrating the computer algorithm into the Argus cardiovascular MR software package.

REFERENCES

1. Kim RJ, Fieno DS, Parrish RB, et al. Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. Circulation 1999:100:1992–2002.
2. Fieno DS, Kim RJ, Chen EL, Lomasney JW, Klocke FJ, Judd RM. Contrast-enhanced magnetic resonance imaging of myocardium at risk: distinction between reversible and irreversible injury throughout infarct healing. J Am Coll Cardiol 2000:36:1985–1991.
3. Wellnhofer E, Olariu A, Klein C, et al. Magnetic resonance low-dose dobutamine test is superior to scar quantification for the prediction of functional recovery. Circulation 2004:109:2172–2174.
4. Kim RJ, Manning WJ. Viability assessment by delayed enhancement cardiovascular magnetic resonance: will low-dose dobutamine dull the shine? Circulation 2004:109:2476–2479.
5. Gonzalez RC, Woods RE. Digital image processing. Reading, MA: Addison Wesley; 1992.
6. Kellman P, Arai AE, McVeigh ER, Aletras AH. Phase sensitive inversion recovery for detecting myocardial infarction using Gadolinium delayed hyperenhancement. Magn Reson Med 2002:47:362–383.
7. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979:86:420–428.
8. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986:1:307–310.
9. Kim RJ, Wu E, Rafael A, et al. The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction. N Engl J Med 2000:343:1145–1453.
10. Petersen SE, Mohrs OK, Horstick G, et al. Influence of contrast agent dose and image acquisition timing on the quantitative determination of nonviable myocardial tissue using delayed contrast-enhanced magnetic resonance imaging. J Cardiovasc Magn Reson 2004:6:541–548.
11. Lund GK, Stork A, Saeed M, et al. Acute myocardial infarction: evaluation with first-pass enhancement and delayed enhancement with MR imaging compared with 201ST1 SPECT imaging. Radiology 2004:232:49–57.
12. Setser RM, Bexell DG, O'Donnell TP, et al. Quantitative assessment of myocardial scar in delayed enhancement magnetic resonance imaging. J Magn Reson Imaging 2003:18:434–441.
13. Mahrholdt H, Wagner A, Holly TA, et al. Reproducibility of chronic infarct size measurement by contrast-enhanced magnetic resonance imaging. Circulation 2002:106:2322–2327.
14. Ingkanisorn WP, Rhoads KL, Aletras AH, Kellman P, Arai AE. Gadolinium delayed enhancement cardiovascular magnetic resonance correlates with clinical measures of myocardial infarction. J Am Coll Cardiol 2004:43:2253–2259.
15. LC Amado, BL Gerber, SN Gupta, et al. Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. J Am Coll Cardiol 2004:44:2383–2389.