EDITORIAL COMMENT

Measuring Myocardial Scar by CMR*

Raymond Y. Kwong, MD, MPH, * Afshin Farzaneh-Far, MD, PhD†‡

Boston, Massachusetts; Chicago, Illinois; and Durham, North Carolina

Late gadolinium enhancement (LGE) using cardiac magnetic resonance (CMR) has emerged as the gold-standard technique for imaging of myocardial scar. The basic principle is inversion-recovery imaging after a 5- to 10-min delay following intravenous administration of gadolinium contrast (1). With appropriate settings, normal myocardium appears nulled or black, whereas nonviable regions appear bright or enhanced. The mechanism under-

lying LGE is not fully understood but is likely based on the inability of gadolinium chelates to cross intact cell membranes (2). In normal myocardium, myocytes are densely packed, and tissue volume is predominately intracellular (~75% to 80%). Therefore, the distribution volume of gadolinium is small, and tissue concentration is low in a typical "voxel" of normal myocardium. With acute necrosis (acute myocardial infarction [MI], myocarditis, etc.), there is membrane rupture, which allows gadolinium to diffuse into myocytes. This results in increased gadolinium concentration (3), shortened T1 relaxation time, and consequent signal enhancement. In the chronic setting, scar has replaced necrotic tissue, and the interstitial space is expanded. This again leads to increased gadolinium concentration (3) and hyperenhancement. In both acute and chronic settings, one can consider viable myocytes as actively excluding gadolinium contrast. Thus, the unifying

mechanism of LGE appears to be the absence of viable myocytes rather than any inherent properties that are specific for acute necrosis, collagenous scar, or other forms of nonviable myocardium.

LGE by CMR was initially validated in animal models of MI, with extensive comparisons demonstrating a nearly exact relationship between the size and shape of infarcted myocardium by LGE to that of histopathology (4,5). These initial studies used a signal intensity threshold of 2 or 3 standard deviations (SD) above normal myocardium to define the infarct. However, these experimental protocols used high-resolution (0.5 × 0.5 × 0.5 mm) ex vivo imaging. Unfortunately, typical in vivo LGE imaging has voxels that are over 100 times larger. In addition, in vivo imaging is complicated by blurring from associated cardiac motion. This results in significantly more partial volume effects compared with ex vivo imaging, creating voxels with intermediate (gray) signal intensity containing an admixture of viable and nonviable myocytes—particularly at the infarct boundary. In this situation, using a low threshold (like 2 SD) may lead to overestimation of infarct size if these gray or intermediate signal intensity voxels are categorized as completely infarcted.

Thus, in vivo quantification of LGE is complicated by unavoidable partial volume effects. In reality, not all myocytes are dead within all infarcted or scarred regions. Thus, significant "patchy" or "gray" regions, presumably due to partial volume effects from a mixture of viable and nonviable myocardium, are not infrequently seen both in the setting of myocardial infarction and in other conditions such as hypertrophic cardiomyopathy. The relative signal intensity of the voxels in these regions will depend on the exact mixture of viable and nonviable myocytes contained within them.

It is on this background that the question of the best method for measurement of scar size using

LGE arises. Higher signal intensity cutoffs have been used in an attempt to avoid overestimation of scar size measurements due to the mechanisms described in the preceding text. However, these may potentially underestimate scar size if there are significant areas of intermediate signal intensity, and are highly dependent on the choice of the remote zone used to calculate the threshold. An alternative method using a threshold of 50% of the maximum intensity within the scar—full-width at half maximum (FWHM)—has been proposed as being more resistant to surface-coil intensity variations, with good correlation to infarct size in an animal model (6). However, this method assumes a bright infarct core and may not be accurate in homogeneously gray scars. In addition, multiple separate islands of scar may cause difficulties with this analysis method (7).

In this issue of *iJACC*, Flett et al. (8) address the issue of intra- and interobserver variability of the various techniques for measurement of myocardial scar in patients with acute and chronic MI, as well as with hypertrophic cardiomyopathy. This is an important question and potentially has significant implications for sample size requirements in clinical trials using scar size as a surrogate outcome. Given the effectiveness of current therapies in acute MI, demonstrating further additive mortality reduction is becoming increasingly difficult, necessitating very large sample sizes. The costs and logistics of conducting such large studies limits the number of new therapies that can be tested. For this reason, there is growing interest in the use of LGE measured infarct size as a surrogate end point in acute MI trials.

Flett et al. (8) demonstrate that scar size varies significantly depending on the method of measurement, with the 2SD method giving significantly larger sizes. There was no statistically significant difference between scar volume by manual, FWHM, 5SD, and 6SD methods. The intra- and interobserver agreement of the different techniques was assessed using the intraclass correlation coefficient (ICC). This showed that the FWHM technique had the lowest ICC, which the authors therefore suggest as being the "most reproducible" technique for measuring LGE size. Interestingly, the ICC was lowest for all techniques in patients with hypertrophic cardiomyopathy compared with MI patients. Finally, the authors estimate sample size requirements for randomized trials with paired and unpaired comparisons using the different techniques. Their calculations suggest significant reduc-

tions in sample size requirements with the use of the FWHM technique.

The authors are to be commended for exploring this understudied but important issue. It should be emphasized at the outset that none of these techniques are truly automated or objective. They all require considerable human input to distinguish bright artifact and/or noise, as well as dark no-reflow regions. More importantly, all require manual tracing of the myocardial borders. This is because there are currently no automated algorithms that reliably distinguish cavity from the endocardial border of infarct. Moreover, since ischemic injury progresses as a "wave front" from subendocardium to epicardium, this border is probably the largest source of variability in infarct size assessment. Given these issues, some investigators have preferred to use manual planimetry or visual estimation on a per-segment basis and tried to account for intermediate-intensity regions by subjectively giving a lower weighting to these regions based on the highest-intensity signal in either the infarct or blood cavity (whichever is higher) (9,10). They have reported low intra- and interobserver variability, for example, SD of differences of 2.6% for interobserver variability and SD of differences of 0.8% for intraobserver variability in the setting of chronic MI. However, it is difficult to compare these numbers with the present study since variability is not reported in the same manner. In the current study, it is unclear how Flett et al. (8) dealt with intermediate-intensity regions in their manual readings. One can speculate that a predefined method for dealing with these regions may have altered the reported variabilities.

These findings echo those of another recent study comparing scar quantitation methods in hypertrophic cardiomyopathy (11). Intraobserver variability was reported as showing lower SD of differences for the FWHM technique compared with manual assessment but with similar ICC for both methods. They also showed that when compared with the manual technique, the 6SD and FWHM methods had the lowest SD of differences. It would have been interesting to see the results of such a comparison in the current study of Flett et al. (8).

Whether these findings are applicable to scar measurements with phase-sensitive inversion-recovery image reconstruction methods would be of interest as these are increasingly used in some centers (12). In addition, development of higher-resolution techniques at increased field strengths

will likely affect variability of LGE size measurements and will require investigation as these technologies evolve (13).

Flett et al. (8) also emphasize the importance of their findings to sample size reductions in clinical trials. However, there is an important difference between trials using paired studies, for example, those with CMR performed before and after a treatment, as compared with unpaired studies where only 1 CMR is performed after randomization to treatment. The variability in paired studies needs to take into account the *interscan* variability (i.e., between 2 separate scans performed on the same individual at different times) as well as the intra/interobserver variability with these techniques. Interscan variability was not assessed in the study of Flett et al. (8), which limits the extension of their findings to paired studies. Clearly, paired studies would be preferable in future trials looking at scar size in hypertrophic cardiomyopathy, since they require significantly lower sample sizes. This is because interpatient variability is no longer a significant cause of variation. In contrast, paired testing is not feasible in most acute MI trials, and infarct size is only assessed after the treatment has been given.

Conclusions

All methods for assessment of scar size require considerable user input and are not as objective as they may initially sound. Flett et al. (8) have confirmed previous studies showing significant differences in scar size when measured by different techniques. They also suggest that the FWHM technique has the lowest intra- and interobserver variability. Although direct comparisons are difficult, others have reported very low intra- and interobserver variability using a refined visual technique method that tries to subjectively account for partial volume effects. Future studies should directly compare these methods. In addition, interscan variability needs to be assessed and compared by these differing methods since it has significant impact on sample sizes for paired studies. Future developments include refinement and testing of a number of sophisticated computerized algorithms that have been proposed based on assigning a weighting to each voxel, depending on image intensity or regional feature analysis (14–16). However, at the present time, these techniques are not widely available, and most studies will continue to rely on 1 or more of the current methods.

Reprint requests and correspondence: Dr. Raymond Y. Kwong, Brigham and Women's Hospital, Cardiovascular Division, Department of Medicine, 75 Francis Street, Boston, Massachusetts 02115. *E-mail: rykwong@partners.org.*

REFERENCES

1. Simonetti OP, Kim RJ, Fieno DS, et al. An improved MR imaging technique for the visualization of myocardial infarction. Radiology 2001;218:215–23.
2. Weinmann HJ, Brasch RC, Press WR, Wesbey GE. Characteristics of gadolinium-DTPA complex: a potential NMR contrast agent. AJR Am J Roentgenol 1984;142:619–24.
3. Rehwald WG, Fieno DS, Chen EL, Kim RJ, Judd RM. Myocardial magnetic resonance imaging contrast agent concentrations after reversible and irreversible ischemic injury. Circulation 2002;105:224–9.
4. 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;36:1985–91.
5. Kim RJ, Fieno DS, Parrish TB, et al. Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. Circulation 1999;100:1992–2002.
6. Amado LC, Gerber BL, Gupta SN, et al. Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. J Am Coll Cardiol 2004;44:2383–9.
7. Kim HW, Farzaneh-Far A, Kim RJ. Cardiovascular magnetic resonance in patients with myocardial infarction: current and emerging applications. J Am Coll Cardiol 2009;55:1–16.
8. Flett AS, Hasleton J, Cook C, et al. Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance. J Am Coll Cardiol Img 2011;4:150–6.
9. Kim HW, Klem I, Shah DJ, et al. Unrecognized non-Q-wave myocardial infarction: prevalence and prognostic significance in patients with suspected coronary disease. PLoS Med 2009;6:e1000057.
10. Kim RJ, Farzaneh-Far A. The diagnostic utility of cardiovascular magnetic resonance in patients with chest pain, elevated cardiac enzymes and non-obstructed coronary arteries. Rev Esp Cardiol 2009;62:966–71.
11. Spiewak M, Malek LA, Misko J, et al. Comparison of different quantification methods of late gadolinium enhancement in patients with hypertrophic cardiomyopathy. Eur J Radiol 2010;74:e149–53.
12. Kellman P, Arai AE, McVeigh ER, Aletras AH. Phase-sensitive inversion recovery for detecting myocardial infarction using gadolinium-delayed hyperenhancement. Magn Reson Med 2002;47:372–83.
13. Schelbert EB, Hsu LY, Anderson SA, et al. Late gadolinium-enhancement cardiac magnetic resonance identifies postinfarction myocardial fibrosis and the border zone at the near cellular level in ex vivo rat heart. Circ Cardiovasc Imaging 2010;3:743–52.

14. Heiberg E, Ugander M, Engblom H, et al. Automated quantification of myocardial infarction from MR images by accounting for partial volume effects: animal, phantom, and human study. Radiology 2008;246: 581–8.

15. Hsu LY, Ingkanisorn WP, Kellman P, Aletras AH, Arai AE. Quantitative myocardial infarction on delayed enhancement MRI. Part II: clinical application of an automated feature analysis and combined thresholding infarct sizing algorithm. J Magn Reson Imaging 2006;23:309–14.

16. Hsu LY, Natanzon A, Kellman P, Hirsch GA, Aletras AH, Arai AE. Quantitative myocardial infarction on delayed enhancement MRI. Part I: animal validation of an automated feature analysis and combined thresholding infarct sizing algorithm. J Magn Reson Imaging 2006;23:298–308.