Quantitative echocardiography presents problems of reproducibility, but less than expected from initial studies performed without 2D guidance. There is evidence that the test-retest “error” of echocardiography is similar to the “normal” error affecting many biological determinations. In particular, in the Reliability of M-mode Echocardiographic Studies (RES) trial, using M-mode tracings from several echocardiography laboratories of different experience, test-retest variability was in the range of 10% to 15% (90% confidence interval). Most error is related to image quality and temporal drift, but, as previously reported, error can be minimized by both appropriate standardization of the examination and in-laboratory self-learning procedures. Reliability can be high, with a coefficient of test-retest concordance of 87% for identification of prognostically adverse left ventricular (LV) hypertrophy (LVH). Thus, together with accuracy (proved by necropsy validation4–9), reproducibility of echocardiographic LV mass8,9 is better than pessimistically stated by Myerson et al.1,7 Moreover, Daniels et al10 showed that the geometric formula used for echocardiographic LV mass determination predicts necropsy LV weight accurately from linear LV wall thickness and chamber dimension measurements.

Myerson et al1 tried to demonstrate the limited “reliability” of echocardiographic LV mass using a magnetic resonance imaging (MRI) reference standard that had several limitations, including use of a machine with low magnetic field (0.5 Tesla), with low spatial resolution and 1 cm slice thickness, whereas use of higher magnetic fields (1.5 Tesla on average, but up to 3 Tesla) are strongly recommended for cardiac cine-MRI. In Myerson’s study,1 measurement was taken “just basal to the papillary muscles,” which, with a 1 cm slice thickness, may not have been standardized exactly at the papillary muscle tips. These pitfalls might be especially aggravated by the intrinsic difficulties of MRI in detecting (and reproducing) the largest LV minor axis, because of great slice thickness and difficulty in aligning acquisitions in near-perfect orthogonal planes with LV structures. These limitations can be offset by echocardiography, using orthogonal images or, especially, long axis with precise anatomic markers.

If these limitations are taken into account, the poor result shown in Figure 3 of the paper1 can be explained by higher tool-related technical variability in obtaining a single tomographic slice at the LV minor axis, as compared with using Simpson’s rule. The test-retest variability reported by Myerson et al1 should in fact be compared with analogous tests performed in Prospective Randomized Enalapril Study Evaluating Regression of Ventricular Enlargement (PRESERVE; mostly 2D) and RES (only M-mode). In the RES, for the same 95% confidence interval, test-retest variability was ±35 g from echocardiograms performed in 16 different centers, compared with the higher values (±45 to 55 g) from applying echocardiographic formulae to single-slice MRI data.

It should be emphasized also that what is clinically important is the ability to detect prognostically adverse values and significant changes,2 both of which are accomplished by echocardiography. Technical variability is <15% at the 90% confidence limit, similar to other biological determinations. Of fundamental importance, we have performed a necropsy comparison study that showed no systematic over- or underestimation of LV mass,4 and numerous investigators have demonstrated the prognostic relevance of echocardiographic LV mass.

Based on results of the PRESERVE study, groups of 41 patients per treatment arm are sufficient to provide statistical power of 90% at an α error level of 1% to detect a between-group difference of at least 10 g/m2 in LV mass. Thus, echocardiographic LV mass remains an excellent bioassay for clinical trials. Although higher magnetic fields and temporal resolution are improving the performance of cardiac MRI, advances in harmonic imaging and in 3D ultrasonography are producing parallel improvements in the ability of echocardiography to image the LV and calculate its mass.

Giovanni de Simone
Richard B. Devereux
Department of Clinical and Experimental Medicine
Federico II University Hospital, Naples, Italy
Division of Cardiology
The New York Presbyterian Hospital
Weill Medical College of Cornell University
New York, NY


Response: What Is Old Is Not Always Best

The points raised by de Simone and Devereux with regard to our paper,1 fall into 2 areas: (1) criticism of the cardiovascular magnetic resonance (CMR) technique and (2) other studies of the reproducibility of echocardiographically determined left ventricular mass.

(1) The spatial resolution of CMR is good, even at 0.5 Tesla (T). Higher field strengths yield better signal-to-noise ratios, but there is no recommendation to use higher field strengths for cine CMR.2,3 T has become routine through availability, and more advanced sequences are programmed for this field strength. However, because improved spatial resolution and signal at higher field strength would only improve the accuracy and reproducibility of the CMR measurement, the argument that the CMR technique was limited is ultimately self-defeating, because, even at lower field strength, CMR exposes the limitations of echocardiography.

The correspondents are incorrect in stating that CMR has difficulty in detecting the true axes of the heart. On the contrary, the strength of CMR is the excellent visualization of any plane within the heart, which makes identification of the true long and short axes more accurate. In particular, the short axis is identified and placed exactly perpendicular to the long axes. Echocardiography has particular problems identifying the true short axes because of the limitations of acoustic access and the inability to determine whether the imaging plane is perpendicular to the long axes.

(2) The reproducibility of echocardiographically determined left ventricular (LV) mass in other studies has been covered elsewhere,
including in our recent review article, and it was not the purpose of our paper to review this again. The correspondents quote 2 papers suggesting 95% confidence intervals of ±35 g for reproducibility but ignore other papers with wider confidence intervals, such as that of Gottdiener et al, which also used 2D-guided M-mode measurements and a single operator and suggested values of ±59 g. Our study did not assess test-retest reliability, because there was a change in LV mass with exercise which varied among individuals. We examined the scatter of values (and confidence limits) of the measured change, and it is not surprising if this was greater than the scatter of 2 measurements of an unchanged LV mass.

Echocardiography has played a fundamental role in assessing LV hypertrophy and determining its prognostic importance. Clearly, it would be possible to continue to examine LV mass with echocardiography in the future. The greater variability of measurements, however, requires more subjects than would be required with CMR, which is why the seminal studies on LV hypertrophy and prognosis needed their large number of subjects to reach their conclusions. Certainly when such large numbers are either not available (gene studies with low allele frequencies or multiple subgroups, for example) or not desirable (cost), then CMR has a significant advantage.

Saul G. Myerson
Dudley J. Pennell
Cardiovascular Magnetic Resonance Unit
Royal Brompton Hospital
London, UK
