Reliability of Echocardiographic Assessment of Left Ventricular Structure and Function
The PRESERVE Study

Vittorio Palmieri, MD, Björn Dahlöf, MD, PhD,* Vincent DeQuattro, MD, FACC,†
Norman Sharpe, MD, FACC;‡ Jonathan N. Bella, MD, Giovanni de Simone, MD, FACC,
Mary Paranicas, BA, Dawn Fishman, BA, Richard B. Devereux, MD, FACC

New York, New York; Goteborg, Sweden; Los Angeles, California; and Auckland, New Zealand

OBJECTIVES The study was done to evaluate reliability of echocardiographic left ventricular (LV) mass.

BACKGROUND Echocardiographic estimation of LV mass is affected by several sources of variability.

METHODS We assessed intrapatient reliability of LV mass measurements in 183 hypertensive patients (68% men, 65 ± 9 years) enrolled in the Prospective Randomized Enalapril Study Evaluating Regression of Ventricular Enlargement (PRESERVE) trial after a screening echocardiogram (ECHO) showed LV hypertrophy. A second ECHO was repeated at randomization (45 ± 25 days later). Two-dimensional (2D)-guided M-mode or 2D linear measurements of LV cavity and wall dimensions were verified by one experienced reader.

RESULTS Mean LV mass was similar at first and second ECHO (243 ± 53 vs. 241 ± 54 g) and showed high reliability as estimated by intraclass correlation coefficient (RHO) = 0.93. Within-patient 5th, 10th, 90th and 95th percentiles of between-study difference in LV mass were −32 g, −28 g, +25 g and +35 g. Mean LV mass fell less from the first to the second ECHO than expected from a formula to predict regression to the mean (2 ± 6 vs. 17 ± 12 g, p < 0.001). Reliability was also high for LV internal diameter (RHO = 0.87), septal (RHO = 0.85) and posterior wall thickness (RHO = 0.83). Substantial or moderate reliability was observed for measures of LV systolic function and diastolic filling (RHO from 0.71 to 0.57).

CONCLUSIONS Left ventricular mass had high reliability and little regression to the mean; between-study LV mass change of ±35 g or ±17 g had ≥95% or ≥80% likelihood of being true change. (J Am Coll Cardiol 1999;34:1625–32) © 1999 by the American College of Cardiology

Cardiovascular mortality and morbidity increase with increasing values of left ventricular (LV) mass independently of other cardiovascular risk factors (1–6). High LV mass also reflects cumulative effects of cardiovascular risk factors (7,8). Therefore, estimation of LV mass represents an important task for cardiovascular risk stratification.

See page 1633

Echocardiography is more sensitive for assessing LV hypertrophy than electrocardiography (9,10). Although the ability of echocardiography to measure cardiac structure and function noninvasively is generally accepted, this modality is affected by several sources of variability (11). Thus, recommendations have been produced to enhance reproducibility of measurements and to facilitate comparison among laboratories (12–15).

Regression of echocardiographic LV hypertrophy has a beneficial impact on cardiovascular mobility and mortality (16). However, the phenomenon of regression to the mean may cause misleading results in studies of regression of LV hypertrophy (17,18). This phenomenon typically occurs when studies select subjects because a variable (i.e., LV mass) exceeds a specified partition. Because of regression to the mean, reevaluation of the selected population may reveal a lower mean value of the variable in question than on first evaluation simply because of random fluctuation of measurements (17,19). This phenomenon may confound interpretation of studies on regression of LV hypertrophy in which patient recruitment is based on LV mass exceeding a partition value. However, the higher the reliability of measurements the less regression to the mean occurs (18).

We assessed the reliability of measurements of LV mass and indices of LV systolic function and diastolic filling in

From the Division of Cardiology, The New York Hospital–Weill Medical College of Cornell University, New York, New York; *University of Goteborg, Goteborg, Sweden; †Los Angeles County/U.S.C. Medical Center, The White Memorial Medical Center, Los Angeles, California; ‡Department of Medicine, Auckland Hospital, Auckland, New Zealand. Supported in part by grant CDSP964 from Merck & Company, Whitehouse Station, New Jersey.

Manuscript received October 21, 1998; revised manuscript received May 18, 1999, accepted July 19, 1999.
Patients with LV hypertrophy [LV mass index during four weeks before the echocardiographic screening. In men and in women seated blood pressure 140 and/or 90 mm Hg if unmedicated during four weeks before the echocardiographic screening. Patients with LV hypertrophy [LV mass index >116 g/m² in men and in women <60 years of age and >104 g/m² in older women (1,21)], on a screening echocardiogram that showed ejection fraction >40% and no severe valvular disease and/or cardiomyopathy received at least one week of placebo before the baseline echocardiographic evaluation at randomization. All patients gave written informed consent to the study.

Full LV and Doppler measurements were obtained on screening studies for approximately the first 60% of patients to evaluate the reliability of echocardiographic parameters between screening and baseline evaluation. Studies were initially blindly read by skilled readers (VP, JNB, GdeS or MP): 183 of 212 subjects with full screening echocardiographic measurements had them verified, and in a majority of patients at least one measurement was corrected by a single, highly experienced reader (RBD).

Echocardiography. Sonographers received extensive training in the PRESERVE protocol including both written material and didactic and hands-on training at the Reading Center in New York. The standardized echocardiographic protocol required recording of ≥10 cycles of two-dimensional (2D) parasternal long- and short-axis LV views and ≥10 cycles of M-mode with optimal cursor beam orientation in each view (20–23). If the 2D-guided M-mode beam could not be optimally oriented, 2D-long axis views were used to obtain linear measurements of LV cavity and walls according to American Society of Echocardiography recommendations (23). The LV mass was calculated by an anatomically validated formula (24). Left ventricular endocardial fractional shortening (FS), midwall shortening (MWS) and circumferential end-systolic stress (ESS) measure of afterload were derived (25). To assess LV systolic performance adjusted for afterload, stress-corrected fractional shortening (c-FS) and stress-corrected midwall shortening (c-MWS) were derived (25). Pulse-wave Doppler flow pattern at mitral anulus was traced electronically to measure peak velocities of early and late diastolic LV filling.

Statistical analysis. Data are reported as mean ± SD. Intraobserver and interobserver components of variability of measurements were not assessed.

Paired-sample t tests were used to evaluate differences between measurements on first and second echocardiogram. The reliability of echocardiographic measurements was investigated using two methods. The intraclass correlation coefficient (RHO), an estimator of variability between replicate measurements (18), was derived by the formula:

$$RHO = \frac{MS_y}{MS_x + MS_w},$$

where MSx was calculated as $MS_x = (MS_y - MS_w)/K_0$, and where MSy is the between-subject mean square of variance, a measure of between-subject variability; MSw is the within-subject mean square of variance; and K0 is the number of repeated measurements (two in this study). Analysis of variance for repeated measures was used to estimate MSb and MSw. If the ratio $MS_w/MS_y$ decreases, RHO increases. The closer RHO is to 1 the higher is the proportion of the variance of measurements free from random errors. Consequently, reliability of measurements increases based on the assumption that MSw is an unbiased estimator of within-subject variability due to random error of measurements of a variable, whereas MSb is an unbiased estimator of the error-free between-subject variability (18). The 95% confidence interval (CI) of RHO was estimated using a two-way random effect model for absolute agreement.

The second method to assess reliability was the Bland-Altmann (26) within-patient interevaluation interval of agreement of LV mass estimation (as absolute value and as percent change from screening evaluation). The 5th to 95th percentile confidence interval (90% CI) was used to derive thresholds of LV mass change that minimize two-tailed alpha-errors at 5%; 90th, 85th, 80th, 75th, 67th, 33rd, 25th, 20th, 15th, and 10th percentiles of between-study difference in LV mass were also generated, to identify the β- and α-error associated with different thresholds of LV mass change.
Standard deviations (SD) of between-study difference in LV mass and LV mass index were used to estimate sample size needed to detect between-group differences in regression of LV mass with statistical powers at 90% and 80%, and α-errors of 5% and 1%.

Previously reported (19) formulae were used to estimate change in absolute and indexed LV mass. Follow-up LV mass was predicted from screening values by the following formulae:

1. \[ R_g(\%) = \frac{3910}{LVM_g} \]
2. \[ R_{g/m^2}(\%) = \frac{2700}{LVM_{g/m^2}} \]

where \( R_g(\%) \) is the percent reduction expected for absolute LV mass at the second evaluation and \( LVM_g \) is initial LV mass in grams (g).

\( R_{g/m^2}(\%) \) is the percent reduction expected for LV mass/body surface area expected at the second evaluation and \( LVM_{g/m^2} \) is the initial LV mass index.

Paired-sample \( t \) tests were performed to compare predicted LV mass to actual LV mass on the second echocardiogram, in the entire population and in quintiles of the distribution of initial LV mass values.

Relations between interstudy change of LV mass and demographic, clinical and echocardiographic parameters were assessed by Pearson correlations and multiple linear regression analyses. Two-tailed \( p < 0.05 \) was considered significant.

**RESULTS**

**Subjects’ characteristics.** One hundred eighty-three patients (68% men, mean age 65 ± 9 years, from 28 centers) had paired echocardiograms 45 ± 25 days apart at screening and at randomization into the PRESERVE trial. Body mass index (BMI) and body surface area (BSA) were similar at two evaluations (\( r = 0.97, r = 0.99 \), respectively). Systolic blood pressure (BP) rose between screening at baseline visit (163 ± 17 vs. 169 ± 20 mm Hg, \( p = 0.001, r = 0.44 \)), while mean diastolic BP did not rise (96 ± 11 vs. 98 ± 11 mm Hg, \( p > 0.1, r = 0.57 \)). Mean heart rate at both visits was 64 ± 11 beats/min (\( p > 0.1; r = 0.60 \)).

**Reliability and regression to the mean.** The between-study difference in LV mass was −1.7 ± 19.8 g for absolute values, −1.1 ± 11.2 g/m² for LV mass/BSA and −0.4 ± 5.2 g/m².7 for LV mass/Height².7 (Table 1). The relation between LV mass measured at screening and baseline is illustrated in Figure 1. The RHO was 0.90 or higher for absolute LV mass and indexed values. Mean LV internal diameter and wall thickness did not differ between evaluations (\( p > 0.1 \)). Of the primary measures used to derive LV mass, LV diastolic diameter showed the highest RHO.

The within-patient, interevaluation 90% CI of agreement ranged from −14% (5th percentile, −32 g) to +15% (95th percentile, +35 g) for absolute LV mass (Fig. 2), from −14% (−17.8 g/m²) to +14% (+19.3 g/m²) for LV mass/BSA and −14% (−8.7 g/m².7) to +15% (+9.1 g/m².7) for LV mass/Height².7 (Table 2).

Predicted decreases in LV mass from screening to baseline due to regression to the mean were −6.8% for absolute LV mass (−17 g) and −9.7% for LV mass/BSA (−13 g/m²). Predicted LV mass was significantly lower than observed values at baseline evaluation (226 ± 41 g vs. 241 ± 54 g, \( p < 0.0005 \); predicted LV mass/BSA (125 ± 19 g/m²) was

---

Table 1. Primary and Derived Echocardiographic Left Ventricular Measurements

<table>
<thead>
<tr>
<th></th>
<th>First Study (Screening)</th>
<th>Second Study (Baseline)</th>
<th>Intraclass Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>LV mass (g)</td>
<td>243 ± 53</td>
<td>241 ± 54</td>
<td>0.93 (0.91–0.95)</td>
</tr>
<tr>
<td>LV mass/heigh².7 (g/m².7)</td>
<td>63 ± 15</td>
<td>63 ± 15</td>
<td>0.94 (0.92–0.95)</td>
</tr>
<tr>
<td>LV mass/BSA (g/m²)</td>
<td>135 ± 25</td>
<td>134 ± 25</td>
<td>0.90 (0.87–0.93)</td>
</tr>
<tr>
<td>IVS (cm)</td>
<td>1.2 ± 0.2</td>
<td>1.2 ± 0.2</td>
<td>0.85 (0.80–0.88)</td>
</tr>
<tr>
<td>LVID (cm)</td>
<td>5.5 ± 0.5</td>
<td>5.5 ± 0.5</td>
<td>0.87 (0.82–0.90)</td>
</tr>
<tr>
<td>PWT (cm)</td>
<td>1.1 ± 0.1</td>
<td>1.1 ± 0.1</td>
<td>0.83 (0.77–0.86)</td>
</tr>
</tbody>
</table>

BSA = body surface area; IVS = interventricular septal thickness; LVID = LV internal dimension; PWT = posterior wall thickness. Results are mean ± SD. Paired-differences are nonsignificant.

---

Figure 1. Relation between screening values of LV mass (horizontal axis) and LV mass measured at subsequent baseline echocardiogram (vertical axis).
also lower than observed baseline LV mass index (134 ± 26 g/m², \( p < 0.005 \)).

Measured LV mass at screening and baseline was compared in each quintile of screening LV mass (Fig. 3). No differences were observed between screening and baseline mean LV mass in the 1st to 4th quintile (1st: 186 vs. 187 g; 2nd: 215 vs. 214 g; 3rd: 234 vs. 234 g; 4th: 256 vs. 256 g, all \( p > 0.1 \)), whereas in the 5th quintile LV mass at baseline was lower than at screening (318 g vs. 326 g, \( p < 0.05 \)). Similar results were obtained using LV mass/BSA (data not shown). We compared LV mass measured at baseline to LV mass predicted by formulae to assess regression to the mean in each quintile of screening LV mass. At the lowest quintile, the difference between observed and predicted LV mass did not reach statistical significance (187 vs. 182 g; \( p = 0.06 \)), but in all higher quintiles mean observed LV mass at baseline significantly exceeded predicted LV mass (2nd: 214 vs. 204 g; 3rd: 234 vs. 219 g; 4th: 256 vs. 236 g; 5th 318 vs. 291 g, all \( p < 0.001 \)) (Fig. 3). Similar results were obtained for LV mass/BSA (data not shown).

Table 2. Likelihood of True Changes in LV Mass for Different Thresholds of Intraindividual Variability

<table>
<thead>
<tr>
<th>LV Mass Change</th>
<th>Likelihood of</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g ) (%)</td>
<td>( (g/m^2) )</td>
</tr>
<tr>
<td>≥+35 (14.9)</td>
<td>≥19.3</td>
</tr>
<tr>
<td>≥+25 (10.8)</td>
<td>≥15.0</td>
</tr>
<tr>
<td>≥+19 (9.3)</td>
<td>≥11.2</td>
</tr>
<tr>
<td>≥+16 (6.6)</td>
<td>≥8.0</td>
</tr>
<tr>
<td>≥+13 (5.6)</td>
<td>≥6.6</td>
</tr>
<tr>
<td>≥+7 (3.3)</td>
<td>≥4.1</td>
</tr>
<tr>
<td>≥+1 (1.6)</td>
<td>≥2.1</td>
</tr>
<tr>
<td>≤-10 (-4.2)</td>
<td>≤-6.4</td>
</tr>
<tr>
<td>≤-14 (-5.7)</td>
<td>≤-7.9</td>
</tr>
<tr>
<td>≤-17 (-6.2)</td>
<td>≤-9.3</td>
</tr>
<tr>
<td>≤-21 (-7.7)</td>
<td>≤-12.3</td>
</tr>
<tr>
<td>≤-26 (-8.7)</td>
<td>≤-14.8</td>
</tr>
<tr>
<td>≤-28 (-10.8)</td>
<td>≤-16.0</td>
</tr>
<tr>
<td>≤-32 (-14.2)</td>
<td>≤-17.8</td>
</tr>
</tbody>
</table>
Confidence intervals of LV mass changes: Likelihood of true changes. As a consequence of the observed between-evaluation variability, in a single patient there is a 5% chance that an LV mass variability greater than \(13.5\) g does not represent a true increase (Table 2). This likelihood rises to 10% for an increase of 25 g, 15% for an increase of 16 g, 20% for an increase of 13 g, 25% for an increase of 7 g and 33% for an increase of 5 g. In the other direction, for a between-evaluation LV mass variability of \(-10\) g there is a 33% chance that LV mass is not truly decreased, 25% for a decrease of 14 g, 20% for a decrease of 17 g, 15% for a decrease of 21 g, 10% for a decrease of 28 g and 5% for a decrease of 32 g.

Impact of variability of LV mass estimation on detection of LV hypertrophy regression. Table 3 shows the sample size needed per group to detect between-group difference in LV mass change with statistical power at 80% and 90%, and alpha-errors of 5% and 1%, using SDs of between-study differences in LV mass of 19.8 g, 11.2 g/m² and 5.3 g/m².\(^{7}\) A between-group difference of 10 g/m² would be detected with 28 or 40 patients per group with 90% statistical power at 5% and 1% \(\alpha\)-errors. Sample sizes needed per group for other levels of between-group difference in LV mass and LV mass indices are reported in Table 3. Using a between-echocardiogram difference SD of 20 g/m² over a long term (\(\geq 1\) year) (27) instead of 11.2 g/m² in the present analysis, 86 or 121 patients per arm are required to detect a between-arm difference of 10 g/m² with 90% power and \(\alpha\)-errors of 5% and 1%.

Correlates of the variability of LV mass. Differences in LV mass between evaluations (Table 4) were positively related to differences in LV internal diameter (\(r = 0.38\)), interventricular septal thickness (\(r = 0.43\)) and posterior wall thickness (\(r = 0.59\)) (all \(p < 0.0005\)). Gender, age and between-study change in weight did not significantly affect LV mass. Differences in systolic or diastolic blood pressure did not correlate significantly with differences in LV mass, possibly owing to the narrow range of their changes. Additionally, no significant associations were found between variability in LV mass and screening levels of body weight, blood pressure, LV internal diameter or Doppler-stroke volume.

Prevalence of LV hypertrophy at the second study. Predefined criteria for LV hypertrophy were met by 159 patients (87%) on the second evaluation. Comparing patients in whom LV hypertrophy was or was not confirmed, the latter group had lower LV mass at screening (120 vs. 137 g/m², \(p < 0.001\)) and lower systolic (157 vs. 165 mm Hg, \(p < 0.05\)) but not diastolic blood pressure (97 vs. 96 mm Hg, \(p > 0.1\)). The two subgroups were comparable in proportion of women (24% vs. 33%) and age (66 ± 8 vs. 65 ± 8 years). Body weight did not change between evaluations in either subgroup (\(p > 0.1\)).

Reliability of measurements of LV systolic function. Mean FS and MWS did not differ between evaluations (\(p > 0.1\) (Table 5). The RHO was 0.68 for MWS, and 0.65 for FS. Circumferential ESS rose minimally between echocardiograms (183 to 188 kdyne \(\times\) cm⁻², \(p = 0.06\), RHO = 0.60). Changes in ESS were negatively related to changes in FS (\(r = -0.56\)) and MWS (\(r = -0.28\)). The RHO for stress-corrected MWS was higher than for stress-corrected FS (0.71 vs. 0.56).
Reliability of left ventricular diastolic filling. Left ventricular filling was measured at mitral annulus on both examinations in 111 patients. Peak velocity of early LV diastolic filling “E” wave (54 ± 17 vs. 53 ± 21 cm/s, p > 0.1, between-study difference SD = 17.5 cm/s, RHO = 0.58) and peak velocity of late LV diastolic filling “A” wave (73 ± 20 vs. 71 ± 20 cm/s, p > 0.1, between-study difference SD = 18.6 cm/s, RHO = 0.57) showed moderate reliability. Interstudy change “E” peak velocity was related to change of LV internal diameter (r = 0.25, p < 0.05), but not of body weight, heart rate, systolic or diastolic blood pressure (all p > 0.1). Change of peak “A” velocity was not related to any considered variable.

DISCUSSION

Reliability of estimation of LV mass. In the present study, reliability of echocardiographic measurements of LV mass, internal dimension and wall thickness was relatively high (28). To our knowledge, this is the first study of this question in a large cohort of hypertensive patients with LV hypertrophy. In accord with the high reliability of LV mass, mean LV mass measured at baseline was significantly higher than LV mass predicted from screening values by a formula to estimate regression to the mean (19). Slight regression to the mean was detected only in the highest quintile of screening LV mass values (Fig. 3). However, the regression to the mean was less than expected, because baseline LV mass significantly exceeded values predicted by the equations of Herpin and Demange (19).

Findings of this study demonstrate that homogeneity of echocardiographic readings can be achieved in a multicenter study on regression of LV mass in patients with LV hypertrophy. Centralized reading with a single and highly experienced final arbiter of readings, use of standardized echocardiographic protocol and a “hands-on” training program for sonographers (20) may have contributed to reduced variability and regression to the mean of LV mass estimation. Substitution of linear measurements of LV wall thickness and internal dimension from the 2D-long axis view whenever the M-mode beam was not ideally oriented may also have contributed to the relatively low variability and regression to the mean observed in this study. Because of the greater mean age and BMI of participants in PRESERVE, the majority of measurements were made from 2D views. Potential limitations of the M-mode to obtain reproducible LV mass measurements have been addressed (29–31). Although 2D echocardiography can also be affected by difficulties in obtaining correct long-axis views or in delineating endocardial and epicardial borders, our findings suggest that a protocol emphasizing obtaining long-axis views that maximize LV chamber size achieves high reliability of LV mass whichever echocardiographic display (M-mode or long-axis 2D) best demonstrates correct interface definition and measurement orientation (32). Additionally, the fact that screening and baseline studies from multiple centers were read interspersed with each other as they arrived at the Reading Center attenuated potential bias due to recall of the first evaluation at the reading of the second.

Our study shows that an intraindividual increase or decrease of calculated LV mass greater than 18 g/m² (or 34 g) is needed to be at least 95% certain that there has been a true change in that direction. Because of the study design, this variability includes intrapatient biologic variability, as well as intrareader and intrasonographer variabilities. It is worthy of consideration that for other clinically relevant variables exhibiting measurement variability, like arterial blood pressure, changes that are 80% or even 75% likely to represent true decrease, corresponding to LV mass changes of 17 g or 14 g, are commonly used to guide clinical

Table 4. Relation Between Change of LV Mass and Dimensions and Hemodynamic Parameters

<table>
<thead>
<tr>
<th></th>
<th>ΔSBP</th>
<th>ΔWeight</th>
<th>ΔLVID</th>
<th>ΔIVS</th>
<th>ΔPWT</th>
<th>ΔSV</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔLVM</td>
<td>r = 0.03</td>
<td>r = −0.05</td>
<td>r = 0.38</td>
<td>r = 0.43</td>
<td>r = 0.59</td>
<td>r = 0.09</td>
<td>r = 0.05</td>
<td>r = −0.09</td>
</tr>
<tr>
<td>(p &gt; 0.1)</td>
<td>(p &gt; 0.1)</td>
<td>(p &lt; 0.0005)</td>
<td>(p &lt; 0.0005)</td>
<td>(p &lt; 0.0005)</td>
<td>(p &gt; 0.1)</td>
<td>(p &gt; 0.1)</td>
<td>(p &gt; 0.1)</td>
<td></td>
</tr>
</tbody>
</table>

Δ = difference between first and second study; SBP = systolic blood pressure; SV = Doppler-derived stroke volume. Other abbreviations as in Table 1.

Table 5. Echocardiographic Measurements of LV Systolic Function

<table>
<thead>
<tr>
<th></th>
<th>First Study (Screening)</th>
<th>Second Study (Screening)</th>
<th>Intraclass Correlation Coefficient</th>
<th>RHO (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS (%)</td>
<td>34 ± 5</td>
<td>34 ± 5</td>
<td>0.65 (0.55 ± 0.75)</td>
<td></td>
</tr>
<tr>
<td>MWS (%)</td>
<td>16 ± 2</td>
<td>16 ± 02</td>
<td>0.68 (0.60–0.75)</td>
<td></td>
</tr>
<tr>
<td>ESS (kdyne·cm⁻²)</td>
<td>183 ± 42</td>
<td>188 ± 44</td>
<td>0.60 (0.49–0.70)</td>
<td></td>
</tr>
<tr>
<td>c-FS (%)</td>
<td>137 ± 17</td>
<td>138 ± 17</td>
<td>0.56 (0.44–0.66)</td>
<td></td>
</tr>
<tr>
<td>c-MWS (%)</td>
<td>103 ± 14</td>
<td>102 ± 14</td>
<td>0.71 (0.62–0.78)</td>
<td></td>
</tr>
</tbody>
</table>

Results are mean ± SD. All paired-differences are nonsignificant.
decisions. The relevance for risk stratification of change in LV mass in hypertensive patients has been recently reported (16) and implications of using LV mass change as a surrogate end point have been discussed (33). Based on the present findings, groups of 41 patients per treatment arm need to complete the period of randomized treatment to provide statistical power of 90% at an alpha-error level of 1% to detect a between-group difference of at least 10 g/m² in LV mass. Assuming 20 g/m² as SD of between-study difference in LV mass over a larger between-study interval, 86 patients per treatment arm should provide 90% power at an alpha-error of 5% to detect a between-group difference of 10 g/m² in LV mass. Thus, the PRESERVE trial (in which over 200 patients completed the year of treatment) has better-than-anticipated power to detect the hypothesized difference between treatment arms in LV hypertrophy regression (20).

**Reliability of assessment of LV systolic function and diastolic filling.** Mean MWS and FS did not differ between studies, and the reliability of measurements of systolic function fell in the “substantial” range (28). The LV systolic performance is a dynamic process dependent on a beat-to-beat balance among preload, afterload and myocardial contractility. When afterload was taken into account, c-MWS had a higher RHO than did unadjusted values (0.71 vs. 0.68), whereas RHO for c-FS was lower than for FS itself (0.56 vs. 0.65).

Doppler peak velocity of “E” and “A” waves, parameters of LV diastolic filling, showed “moderate” reliability. Heart rate, systolic BP and circumferential ESS, variables related to LV filling, showed moderate between-study correlations ($r = 0.44$ to $r = 0.60$).

**Study limitations.** Evaluation of interobserver variability of LV dimension and LV mass measurements was not a goal of our study. Our readings were verified by a single experienced reader, which may have reduced the variability of estimation of LV mass by eliminating potential interobserver variability among readers in a single center and, even more, between readers in different centers.

Part of the variability of serial estimations of LV dimensions and mass detected in our study is potentially due to intraindividual biological variability over the mean interval of six weeks between echocardiograms. A short between-study interval (e.g., on the same day or within a few days) would better relate observed measurement variability to intra- or interreader variability rather than biological variability. The inability to show significant associations of between-study echocardiographic measurement variability with changes in body weight, blood pressure or heart rate may reflect imprecision in clinical as well as echocardiographic measurements. Of note, the small increase in systolic blood pressure between screening and baseline echocardiograms may have attenuated the reduction of LV mass that might otherwise have occurred due to the regression to the mean. Finally, this study does not address the important issue of long-term measurement “drift” even within a single reading center. However, we have previously reported that change in LV mass in normotensive subjects over three to six years was as low as 6 g, after accounting for change in BMI, systolic blood pressure and 24-h urinary sodium excretion (34).

**Conclusions.** The reliability of 2D echocardiographic measurements of primary LV structures and LV mass between screening and baseline phases of the PRESERVE trial was fairly high, showing little regression to the mean. Short-term between-study variability in LV mass of $\pm 34$ g or $\pm 18$ g in single patients have, respectively, $\approx 90\%$ or $\approx 80\%$ likelihood to be true changes in the identified direction. Thus, echocardiographic evaluation of LV mass using a standardized protocol and a Central Reading Center is able to detect even small changes of LV mass in modest-sized populations. The clinical relevance of LV mass change at different thresholds of confidence for likelihood of true change in single patients needs further study.

**APPENDIX**

List of investigators (by country): Argentina, Martin Lombardero (Istituto Cardiovascular de Buenos Aires), Rafael Diaz (Istituto Cardiovascular de Rosario); China, Wan-Chun Chen, Xiao-Pei Xue, (Shanghai 6th People’s Hospital); Germany, Kilian Glänzer (Medizinische Universität Poliklinik, Bonn), Georg Meurers (Krankenhaus Vencenti-uum Intesiv-Station, Ruhpolding), Peter Schuster (St. Marien Krankenhaus Kardiologische Abteilung, Siegen), Gerd Fröhlig (Universitätskliniken des Saarlandes, Homburg/Saar), Susanne Mohr-Kahaly (Universitätsklinik Mainz), Bernhard Maisch (Klinikum der Philipps Universität Marburg); Hong Kong, Chu-Pak Lau, Hung-Fat Tse (Queen Mary Hospital), Yuk-Kong Lau (Ruttonjee Hospital); Israel, Esther Paran (Hypertension Unit, Soroka Medical Center), E. Podjarny (Nephrology Department, Meir Medical Center); Italy, Eugenio Nannini (Ospedale Castelnuovo Garfagnana, Lucca), Oreste de Divitiis (Università “Federico II,” Napoli), Enrico Agabiti-Rosei, Maria L. Muienes (Università degli Studi di Brescia), Carlo Porcellati, Paolo Verdecchia (Ospedale “R. Silvestrini,” Perugia), Guido Gigli (Ospedale di Rapallo); New Zealand, Norman Sharpe, Fiona Stewart (Auckland Hospital); Norway, Halfden Ihlen (Med. Avd. B. Rikshospitalet, Oslo); Philippines, Rody G. Sy (Philippine General Hospital); Portugal, Mario Lopes (Lisboa); Sweden, Björn Dahlöf, Filip Jacobson (University Hospital, Göteborg); USA, Michael H. Alderman (Albert Einstein College of Medicine, New York), P.A.N. Chandraratna, Vincent DeQuattro, Willa Hsueh (LAC/USC Medical Center, Los Angeles), Donald Schmidt (Mt. Sinai Samaritan Medical Center, Milwaukee, Wisconsin).
REFERENCES

2. Koren MJ, Devereux RB, Casale PN, et al. Relation of left ventricular mass and geometry to morbidity and mortality in uncomplicated essential hypertension. Ann Intern Med 1991;114:345–52.