Reproducibility of Echocardiographic Left Ventricular Measurements

DONALD C. WALLERSON AND RICHARD B. DEVEREUX

SUMMARY Serial echocardiograms with acceptable reproducibility of measurements may be produced by careful performance and interpretation of the studies. The following recommendations have been shown to enhance reproducibility. 1) Strict adherence to quality control is necessary to generate echocardiograms of the highest technical quality. Sonographers should be aware of the definition of a technically adequate study — including correct beam or plane angulation and continuous visualization of interfaces — and seek this ideal in every study. Participation by the sonographer in performance of measurements enhances recognition of the requirements for accurate quantitative echocardiography. Regular machine calibration is a prerequisite to accurate quantitative echocardiography. 2) Considerable effort must be made to standardize the position of each acoustic window and angulation from which the patient is imaged — with deviation from these norms being recorded for future reference. If at all possible, measurements should be taken at end expiration. If that is not possible, measurement of several consecutive beats will limit the impact of respiratory variation. 3) A uniform convention of measurement should be adopted. The best candidates for M-mode measurements are the American Society of Echocardiography recommendations for general measurement and the Penn convention for calculation of M-mode left ventricular mass. Further data is needed to determine which approaches to two-dimensional measurements best combine accuracy and reproducibility. 4) Interpretation of echocardiograms may be made most reproducible by measuring pertinent parameters from multiple beats and using the mean as the result and by having at least two readers interpret each echocardiogram, possibly with two separate readings by each reader. (Hypertension 9 [Suppl II]: II-6-II-18, 1987)

KEY WORDS • echocardiography • left ventricle • reproducibility

OVER the past decade echocardiography has become accepted as a valuable noninvasive method for evaluating congenital and acquired cardiac disease. The initial echocardiographic technique, time motion or M-mode echocardiography, produced images with excellent depth and temporal resolution (1 mm and 1 msec, respectively). It was soon recognized, however, that for some patients, the selected narrow portions of the heart visualized by the M-mode beam provided incomplete or even misleading information about the anatomic relationships of cardiac structures or function of the heart as a whole. More recently, two-dimensional echocardiography, a technique with superior spatial orientation but inferior temporal resolution (30 msec) compared to M-mode echocardiography, has extended the usefulness of echocardiography by providing correct representation of cardiac anatomic relationships.

In view of the ability of echocardiography to measure cardiac structure and function noninvasively, investigators soon recognized that standardization of methods was necessary both to enhance reproducibility of measurements and to facilitate their comparison between laboratories. Several sources of variability have been recognized and a number of recommendations to limit this variability have been published over the last 10 years. To date, however, individual publications have addressed only one or a few sources of variability in echocardiographic measurements, and no critical review of variability in quantitative echocardiographic assessment of the left ventricle is available. Therefore, this review examines variability of echocardiographic measurements in a comprehensive manner, identifying the sources of variability in echocardiographic measurement, assessing available approaches to limit such variability, and finally making recommendations to enhance reproducibility of echocardiographic measurements. The following topics are considered in sequence: echocardiographic quality control, equipment calibration, technical factors affecting reproducibility of measurements from serial echocardiograms, physiologic variation of sequential echocardiographic measurements, echocardiographic measurement convention and reproducibility of results, magnitude of intraobserver and interobserver variability, temporal variability, and reproducibility of two-dimensional echocardiographic measurements.
Echocardiographic Quality Control

Every investigation that employs echocardiography is dependent on the technical quality of the tracings. Very rarely are the inclusion or exclusion criteria for echocardiograms specified in published descriptions of methods. Uniformity in the definition of an adequate echogram is necessary if strictly comparable data are to be obtained in different laboratories. This is an especially important factor in cooperative studies or in population surveys. Schieken et al.¹ have precisely defined a technically satisfactory M-mode echocardiogram of the left ventricle as comprising the following:

- Generation of a single dominant line representing each interface being imaged.
- Demonstration of continuous interface lines at least 5 mm in length at the point of measurement.
- Demonstration of interfaces with the motion pattern characteristic of the specific cardiac structure being imaged.

An additional requirement for timing of echocardiographic measurements, particularly of end-diastolic dimensions, is the simultaneous recording of QRS complexes with readily identifiable onset and peak of deflections. These criteria were found by Schieken et al.¹ to result in high reproducibility of measurements of the left ventricle, left atrium, aortic root, and left ventricular (LV) ejection indices. Twenty tracings judged satisfactory by these criteria yielded highly reproducible measurements, whereas measurements from eight echocardiograms not meeting these criteria were described as nonreproducible.¹

For LV mass measurements, analogous criteria were proposed by Devereux and Reichek.² LV echograms had to demonstrate continuous motion of right and left septal surfaces and endocardial and epicardial interfaces of the posterior LV wall throughout the cardiac cycle. Measurements were made on LV views at the level of the LV minor axis, identified at or just below the tips of the mitral leaflets. Measurements were made at the peak of the R wave of the electrocardiogram. LV mass measurements in 24 tracings deemed technically adequate by these criteria exhibited good reproducibility between two experienced observers (r = 0.94, p < 0.001).³

Enforcement of criteria for technical adequacy will exclude tracings from a proportion of patients. The percentage of excluded tracings will decrease to acceptable levels as technicians and physicians performing the studies become familiar with the criteria. It is our impression and that of other investigators that the yield of technically excellent echocardiograms is enhanced by involving sonographers in the performance of measurements on tracings generated by them.

Several other variables appear to influence the percentage of technically adequate tracings. Specific disease states (especially pulmonary diseases, thoracic deformities, and postsurgical changes in cardiac position) may dramatically reduce the yield of satisfactory studies. The proportion of measurable echocardiograms in population studies has been variably reported as 20 of 28 (71%) by Schieken et al.,¹ 196 of 259 (75%) by Valdez et al.,⁴ and 191 of 236 (81%) by Wong et al.⁵ In a recent survey of normotensive and hypertensive members of an adult employed population, we found LV echograms to be measurable by strict criteria in 621 of 767 (81%).⁶

Equipment Calibration

Echocardiographic measurements depend for their accuracy on correct calibration of the equipment being used. Errors in calibration should be rare when equipment is delivered from the factory but develop commonly thereafter, particularly when machines are heavily used and moved frequently for bedside examinations or travel between different research sites. On some echocardiographs, calibrations can be correct on one output device but not another (i.e., stripchart recorder vs videotape), necessitating regular checks of the accuracy of each method of display.

M-mode echocardiographs, which image in only a single direction, can be calibrated with the aid of simple blocks of materials in which the velocity of ultrasound is known. The distance between opposite sides of the blocks, in different axes, can then be made to correspond to various numbers of centimeters in the body, using the ratio of transmission of ultrasound in body tissues and the plastic of the block. Methods for checking two-dimensional calibration are newer and more complicated since lateral resolution problems have an important effect on the image produced. Calibration equipment must also mimic the acoustic properties of the soft tissues through which the echocardiograms are produced. Phantoms for calibration of two-dimensional echocardiographs, designed with these considerations in mind, are now available and should be used routinely by all echocardiographic laboratories to check horizontal and vertical calibrations as well as resolutions in different axes. An example of the image produced from one of these two-dimensional phantoms is shown in Figure 1. The method allows calibration checks from nylon lines and cylinders set in a tissue-mimicking matrix. It also allows an appreciation of the inferiority of lateral to depth resolution in currently available machines.

Technical Factors in Reproducibility of Echocardiographic Measurements

In addition to the visual quality of echocardiographic tracings, other aspects of echocardiographic technique have been examined with regard to their effect on reproducibility of measurements. This question has been addressed principally by studies in which serial echograms performed on the same individuals were separated by time intervals too short to permit real changes to develop in cardiac structures. This study design delineated aspects of echocardiographic technique that are important to standardize or replicate to maximize reproducibility.

Patient position during echocardiography and transducer location on the body surface greatly influence the echocardiographic image obtained. Clark et al.⁷ documented the need to record accurately both the patient and bed position yielding the best measurable tracings on the initial echocardiographic study. This information is particularly important in serial studies, in which these positions must be reproduced. A number of positions may appear adequate but produce modestly different echocardiographic images of the same structure.⁸ Most laboratories use the partial (30–45 degrees) left lateral decubitus position for recording M-mode echocardiograms as well as long and short axis parasternal and apical four-chamber two-dimensional views. In some patients, optimal recording of apical four- and two-chamber views may require a steeper left lateral position, achieved with a mattress from which a segment has been removed to permit optimal transducer access. A flat supine position is used for two-dimensional imaging from the subcostal acoustic window. The degree of left lateral positioning may be standardized in a laboratory by using a wedge (or one of several wedges of different shapes) placed under the patient’s back. Many patients require varying degrees of repositioning to accomplish adequate imaging. When that occurs, the details of such repositioning should be a part of the echocardiographic record. With regard to bed positioning, a 30-degree upright tilt of the table is most widely used and would be appropriate for uniform adoption to reduce interlaboratory variability. Any deviation from the laboratory’s routine should be a part of the echocardiographic record.
The most important technical issue influencing reproducibility of serial echocardiography is transducer position during imaging. Based on the results of early validation studies, it was recommended that the standard transducer position be "in the third or fourth interspace, left sternal border, to allow simultaneous recording of continuous endocardial echoes from both the left ventricular posterior wall and the interventricular septum." It was soon recognized, however, that LV echograms satisfying the above criteria could be obtained from multiple interspaces with corresponding variations in LV dimensions. Popp et al. showed that the resultant variation in LV dimensions depended on whether the echocardiographic beam paralleled the LV minor axis or was angled obliquely to it, resulting in overestimation of LV internal dimensions and wall thickness. Figure 2A illustrates the excellent reproducibility of LV end-diastolic dimensions and fractional shortening on different strips recorded by Popp et al. from the same chest wall location, whereas considerable scatter is seen when dimensions are obtained from other transducer locations (Figure 2B). Calculations made by Evans et al. revealed that angulation errors...
potentially introduce quantitatively more important errors into ejection phase indices than do side-to-side motion of the heart or rotation about the heart (Table 1).

In some patients the so-called standard interspaces are not usable for adequate echocardiographic imaging due to the effects of disease states or unusual body habitus. In obese patients or patients with ascites and elevated diaphragms, the superior shift of the heart may allow imaging only from the second intercostal space. Conversely, in an individual with a long trunk the heart may lie low in the thorax and imaging may be possible only from the fifth intercostal space.

Thus, patient positioning, bed positioning, and transducer location must be stable from study to study. A simple approach to minimizing the variability of these factors is to note patient angulation, bed position, and interspace as part of the echocardiographic record. Clark et al.7 showed that by adhering to these simple recommendations, reproducibility was such that serial measurement of LV end-diastolic dimension producing differences greater than 3 mm could confidently be deemed as a real change (falling outside 95% confidence limits derived as ± 2 SD; Figure 3). Similarly, differences in LV fractional shortening greater than 5.5% could also confidently be deemed a real change (see Figure 3).

More elaborate techniques have been proposed to minimize variability due to body position and transducer location. One approach adopted by Wong et al.3 utilized an inclinometer to assure reproducibility of transducer angulation. Stefadouros and Canedoe12 developed an elaborate triangulation device to achieve the same goal. Both of these approaches resulted in good reproducibility but prolonged time for performance of the study and added some difficulty in the use of cumbersome equipment. In Table 2 the reproducibility of LV measurements obtained using these devices is compared to that obtained by Clark et al.7 and other investigators in otherwise similar short-term studies in which reproducibility of patient position and intercostal space were accomplished manually. The modest increment in reproducibility attained by more elaborate methods does not justify their use in large-scale echocardiographic studies.

Physiologic Variability and Echocardiographic Measurement

The heart is a dynamic organ, constantly changing its size and function in response to variations in heart rate, preload (influenced by body fluid balance and venous tone), and afterload (influenced by the relationship between blood pressure and LV geometry). Variations in these parameters occur both in the course of normal fluctuations in body physiology over time and in response to exogenous factors. The latter may include changes in diet, medications, and emotional stress with its effects on a- and b-adrenergic tone.

DeMaria et al.13 examined the effect of heart rate on echocardiographic measurements of LV internal dimension (LVID). Right atrial pacing was used to increase heart rate in increments of 10 beats/min to a maximum rate of 150 beats/min. They were able to show that a 2.7% decrease in LVID occurred, in a virtually linear fashion, for each 10 beats/min increment in heart rate (Figure 4). On the other hand, Felner et al.14 and Bellenge15 did not detect any significant change in the LVID in pa-
### Table 2. Reproducibility of Echocardiographic Left Ventricular Measurements

<table>
<thead>
<tr>
<th>Reference</th>
<th>Subject characteristics</th>
<th>Coefficient of variation (%)</th>
<th>Index of reproducibility</th>
<th>Re-performance variability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LVIDd</td>
<td>LVIDs</td>
<td>PWT</td>
</tr>
<tr>
<td>With special equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wong et al.5</td>
<td>Normal size LV</td>
<td>1.8</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(49 ± 5.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dilated LV</td>
<td>4.6</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(73 ± 8.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stefanadouros and Canedoe12</td>
<td></td>
<td>14.5</td>
<td>46.6</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(50 ± 6.4)</td>
<td>(35 ± 8.0)</td>
<td></td>
</tr>
<tr>
<td>Without special equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clark et al.7</td>
<td>Valvular and coronary</td>
<td>2.9</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>disease</td>
<td>(61 ± 0.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felner et al.14</td>
<td>Males</td>
<td>3.1</td>
<td>4.8</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(48 ± 1.5)</td>
<td>(30.9 ± 1.7)</td>
<td>(8.0 ± 0.4)</td>
</tr>
<tr>
<td>Pollick et al.43</td>
<td>1 beat</td>
<td>3.1</td>
<td>4.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(45.7 ± 1.4)</td>
<td>(29.2 ± 1.3)</td>
<td>(7.3 ± 0.5)</td>
</tr>
<tr>
<td>Lapido et al.44</td>
<td>15 Normal subjects</td>
<td>3.9</td>
<td>4.39</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>and patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pietro et al.45</td>
<td>10 Independent M-mode</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2D guided</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gordon et al.61</td>
<td>Outpatients</td>
<td>5.9</td>
<td>5.2</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(68 ± 4)</td>
<td>(58 ± 3)</td>
<td></td>
</tr>
<tr>
<td>MacMahon et al.54</td>
<td>Normal subjects</td>
<td>2.2</td>
<td>3.1</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Congestive myopathy</td>
<td>3.5</td>
<td>4.3</td>
<td>—</td>
</tr>
</tbody>
</table>

Mean values ± SD are in parentheses.

LVIDd = end-diastolic left ventricular internal dimension; LVIDs = end-systolic LVID; PWT = posterior wall thickness; IVST = interventricular septal thickness; LV = left ventricle; SD = standard deviation; 2D = two-dimensional.

### Figure 4. Relationship of left ventricular (LV) end-diastolic dimension (left panel) and end-systolic dimension (right panel) to heart rate. The LV dimension is expressed for each subject as a percentage of the measurement recorded at a heart rate of 100 beats/min, which serves as the baseline of 100%. The percentage of LV dimensions in each individual subject is then correlated with heart rate at 10-beat increments through a range of 50 to 150 beats/min. (Reprinted from DeMaria et al.12 with permission.)
The degree to which variation in heart rate needs to be considered an important determinant of the reproducibility of echocardiographic measurements appears to depend on the expected fluctuations in heart rate among a group of patients. Based on our own experience in annual reassessments of apparently normal individuals, the mean intrapatient difference in heart rate between examinations is less than 10 beats/min. This suggests that no more than a 2 to 3% variation in LVID would be induced by fluctuation in heart rate. Even so, this estimate of introduced variability undoubtedly overstates the magnitude of the expected effect since reflex factors that increase heart rate would also tend to enhance venous return to the heart.

The effect of respiratory variation on echocardiographic measurements of LVID was evaluated by Brenner and Waugh after it was initially described by Feigenbaum. Brenner and Waugh showed a 6% decrease in end-diastolic LVID at end-inspiration compared to end-expiration and recommended that recording be made at end-expiration. An alternative approach, when patients have difficulty in cooperating with the recording of an end-expiration strip maneuver, is to obtain long records of the best LV views. The pertinent measurements are then obtained as the mean of up to six consecutive beats, usually including all phases of a respiratory cycle.

Other physiologic variables beyond the control of the echocardiographer, including preload, afterload, and systemic volume status, are especially likely to influence measurements of LV function. While neither preload nor circulatory volume status can be measured noninvasively, it is possible to measure accurately an index of LV afterload (LV end-systolic meridional wall stress) by use of echocardiographic measurement with simultaneous determination of cuff blood pressure. We have introduced variability undoubtedly overstates the magnitude of the expected effect since reflex factors that increase heart rate. Even so, this estimate of introduced variability undoubtedly overstates the magnitude of the expected effect since reflex factors that increase heart rate would also tend to enhance venous return to the heart.

The effect of respiratory variation on echocardiographic measurements of LVID was evaluated by Brenner and Waugh after it was initially described by Feigenbaum. Brenner and Waugh showed a 6% decrease in end-diastolic LVID at end-inspiration compared to end-expiration and recommended that recording be made at end-expiration. An alternative approach, when patients have difficulty in cooperating with the recording of an end-expiration strip maneuver, is to obtain long records of the best LV views. The pertinent measurements are then obtained as the mean of up to six consecutive beats, usually including all phases of a respiratory cycle.

Other physiologic variables beyond the control of the echocardiographer, including preload, afterload, and systemic volume status, are especially likely to influence measurements of LV function. While neither preload nor circulatory volume status can be measured noninvasively, it is possible to measure accurately an index of LV afterload (LV end-systolic meridional wall stress) by use of echocardiographic measurement with simultaneous determination of cuff blood pressure. We have shown that resting LV systolic function is inversely related to end-systolic stress in both normotensive subjects and hypercontractility.

Echocardiographic Measurement Convention and Reproducibility of Results

To obtain echocardiographic measurements that are reproducible between either observers or studies, the method of making the measurement must be defined. The several sets of measurement conventions that have been recommended are based either on logic or data. Three of these are particularly important because of their use in quantitative studies of LV mass and function (Table 3). The first of these to be introduced was the National Institutes of Health (NIH) convention. Early studies utilized measurements by this convention to delineate anatomic features of hypertrophic cardiomyopathy, athletic LV hypertrophy, and systemic hypertension, as well as to assess the relationships between demographic variables and LV anatomy. The fact that LV chamber and wall dimensions were measured at different times in the cardiac cycle, underlying the logical basis of LV mass calculations, as well as the lack of necropsy correlation data have caused this method to be superseded for measurement of LV mass and most other variables by two other conventions, whose introduction was based on carefully collected data of different types.

The Penn convention, devised in 1977, was based on a study in which LV mass was calculated from echocardiograms examining two alternative assumptions about each of three variables: LV shape, wall segments to be measured to determine mean myocardial thickness, and identification of endocardial surfaces. These calculations yielded eight alternative echocardiographic estimates of LV muscle mass, which were systematically compared to necropsy LV mass in 34 patients. The set of assumptions that yielded the most accurate echocardiographic measurements of LV mass was introduced as the Penn convention (Figure 5; Table 4). Penn convention end-diastolic measurements are taken at the peak of the R wave, and the thickness of endocardial interfaces is excluded from wall thickness measurements. Although development of the Penn convention was not based on assessment of interobserver and intertest reproducibility, subsequent study has revealed that reproducibility is acceptably high (Table 5).

The reverse sequence was followed by Sahn et al. in developing the M-mode measurement recommendations of the American Society of Echocardiography (ASE). In formulating these recommendations, major emphasis was placed on the reproducibility of measurements performed by 76 readers on a set of echocardiograms from the Multi-Ethnic Study of Atherosclerosis (MESA) study.
of sample echocardiographic tracings. The most important features of the ASE measurement convention are that dimensions are taken from the leading edge of one interface to the leading interface of the next; identification of end diastole was at the onset of the QRS complex for all measurements; and identification of end systole for LV measurement was at the time of the most posterior displacement of the interventricular septum in systole (Figure 6).

Subsequent studies have assessed both the validity and reproducibility of measurement by the ASE convention in comparison with angiographic or necropsy reference standards. Crawford et al. compared LV measurements by the ASE recommendations with the echocardiographic measurement convention previously established in their laboratory using angiographic LV volumes as the reference standard. Volumes derived using the cube function formula and echocardiographic LV measurements by the ASE recommendations consistently overestimated volume relative to angiographic determinations, but the overestimations were systematic and highly reproducible. The measurement convention previously used in Crawford's laboratory afforded a better correlation with angiographically determined volumes. Although both echocardiographic methods overestimated angiographic LV ejection fraction, the ASE recommendations produced a better correlation with angiographically determined values.

More recently, Woythaler et al. and Devereux et al. have assessed the validity of LV muscle mass determination using echocardiographic measurements by the ASE convention. Both groups found that LV mass calculated by the ASE convention correlated well with necropsy LV mass but systematically overestimated it. This overestimation can be corrected by a simple regression equation (see Table 4, equation 3). Other formulas used to estimate LV mass from M-mode and two-dimensional echocardiographic measurements are also shown in Table 4 and additional two-dimensional measurements used to calculate LV muscle mass are shown in Figure 7.

In view of these findings, it appears that both the ASE and Penn conventions provide adequate reproducibility of echocardiographic LV measurements and can be used to estimate LV muscle mass accurately using necropsy-validated regression equations. For most purposes, the ASE convention appears to be preferable because its wide dissemination makes it suitable for standardizing measurements among clinical and research laboratories. For research studies in which a principal goal is evaluation of LV muscle mass, the Penn convention appears to

### Table 4. Formulas for Echocardiographic Measurement of Left Ventricular Mass

<table>
<thead>
<tr>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-mode echocardiography</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Muscle cross-sectional area (CSA) = $\pi \left( \frac{LVIDd}{2} + PWTd + IVSTd \right)^2 - \pi \left( \frac{LVIDd}{2} \right)$ | Gaash et al. 
| Penn convention: LV mass (g) = $1.04 \left( LVIDd + PWTd + IVSTd \right)^3 - \left( LVIDd \right)^3 - 13.6$ g | Devereux and Reichek 
| ASE convention: LV mass (g) = $0.80 \left[ 1.04 \times \left( IVSTd + LVIDd + PWTd \right)^3 - \left( LVIDd \right)^3 \right] + 0.6$ g | Devereux et al. |
| Two-dimensional echocardiography |
| Simpson's rule method: LV volume = $\frac{N-1}{2} A_T N^3 + \frac{M}{6} \times \left( \frac{A_p}{L} \right) \times L$ | Helak and Reichek 
| Area-length method A: LV volume = $\frac{2}{3} \left( a + b + d \right)^3$ | Helak and Reichek, Reichek et al. |
| Area-length method B: LV volume = $1.05 \times \left( a + b + d \right)^3$ | Schiller et al. |

All dimensions are in centimeters. A = short axis area; IVSTd = end-diastolic intraventricular septal thickness; L = longest epicardial or endocardial length of the left ventricle; N = number of sections; p = papillary muscle level; PWTd = end-diastolic posterior wall thickness; T = thickness of each section; a, b, d, i in last equation corresponds to the labels in Figure 7. Other abbreviations as in Table 2.
Table 5. Interobserver and Interstudy Reproducibility of Penn Convention Measurements

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Interval between studies</th>
<th>n</th>
<th>Correlation coefficient</th>
<th>Standard deviation</th>
<th>Mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVIDd</td>
<td>0</td>
<td>60</td>
<td>0.98</td>
<td>1.4 mm</td>
<td>1.4 mm</td>
</tr>
<tr>
<td>PWT</td>
<td>0</td>
<td>60</td>
<td>0.91</td>
<td>0.9 mm</td>
<td>0.8 mm</td>
</tr>
<tr>
<td>LV mass</td>
<td>0</td>
<td>24</td>
<td>0.94</td>
<td>29 g</td>
<td>26 mm</td>
</tr>
<tr>
<td>LV mass</td>
<td>0</td>
<td>24</td>
<td>0.84</td>
<td>42 g</td>
<td>38 g</td>
</tr>
<tr>
<td>LV mass</td>
<td>3 Mo</td>
<td>8</td>
<td>0.98</td>
<td>28 g</td>
<td>26 g</td>
</tr>
<tr>
<td>LV mass</td>
<td>15 Mo</td>
<td>89</td>
<td>0.78</td>
<td>29 g</td>
<td>26 g</td>
</tr>
</tbody>
</table>

Abbreviations as in Table 2.

be slightly preferable because it yielded a closer correlation than the ASE measurements between echocardiographic and necropsy LV muscle mass.3-36

Intraobserver and Interobserver Variability

Echocardiographic measurements are also subject to intraobserver or measurement variability (i.e., fluctuation of measurements on the same echocardiogram when it is measured multiple times). Additionally, in attempts to enhance objectivity and accuracy of echocardiographic measurements, studies often employ multiple readers to interpret the same echograms, the mean of the values produced by all readers being accepted as the final one. This introduces an additional source of variation, called interobserver variability. This is a compound variable since there is also concurrent measurement variability. Vignola et al.41 described 5.2, 10.0, and 16.4% interobserver errors (ex-

Figure 6. American Society of Echocardiography recommendations for M-mode measurements. Diastolic measurements are made at the onset of the QRS complex of the electrocardiogram (EKG); cavities and walls are measured at the level of the chordae below the mitral valve. The illustration and the elliptical inserts a, b, c, d and e show the leading edge method as well as measurements using the thinnest continuous echo lines: ARV = right ventricular anterior wall; RV = right ventricle; LV = left ventricle; PLV = posterior left ventricular wall; S = septum; PPM = papillary muscle; AMV, PMV = anterior and posterior mitral valve leaflets; A, B, C, D, E and F = points of mitral valve motion; EN = endocardium; EP = epicardium; Ao = aortic root; AV = aortic valve; LA = left atrium. The extra line in insert B, which is excluded from the septal measurement, represents a portion of tricuspid valve apparatus. (Reprinted from Sahn et al.28 with permission of the American Heart Association.)

Figure 7. Left ventricle as a truncated ellipsoid. The internal dimensions used in this study are shown in the short axis (left) and the long axis (right). Four semiminor axes or radii (b) are shown in the short axis and two are shown in the long axis. Note that placement of the minor diameter (equivalent to two semiminor axes) determines the division of semimajor axes. The full semimajor axis (a) and the truncated semimajor axis (d) appear in the long axis representation of the left ventricle. (Reprinted from Schiller et al.38 with permission of the American Heart Association.)
pressed as the mean difference between observers divided by the average measurement) for measurement of the end-diastolic LVID, interventricular septal thickness, and LV posterior wall thickness, respectively (Table 6). Monoson et al. obtained a correlation coefficient of .91 between measurements of both the interventricular septal and posterior LV wall thickness of the same echo tracing by two observers. Sahn et al. evaluating measurements for end-diastolic LV posterior wall thickness, interventricular septal thickness, and LVID on the same echocardiograms by 76 observers, showed respective percentage uncertainties of 23.4%, 19.5% and 8.2% when the ASE convention was used for measurement. Percentage uncertainty was obtained as follows: "the mean and standard deviation for each measurement on each recording were determined, first combining all measurement criteria. The 95th percentile confidence ranges were considered to be 1.97 standard deviations. The percentage uncertainty was the 95th percentile confidence limit divided by the mean for the measurement times 100. Percentage uncertainty is normalized by the mean for the measurement, allowing comparison of the ranges of errors between the echograms which differed in the absolute measurements." Table 7 shows the percentage uncertainty generated by different measurement conventions for end-diastolic interventricular septal thickness, posterior wall thickness, and septal LVID as well as end-systolic LVID. The conventions that produced the smallest percentage uncertainty were selected as the ASE recommendations.

Notably, measurement of end-diastolic LVID exhibited the least interobserver variability in each of these studies. The measurements with the greatest interobserver variability included those of the mitral valve E-F slope, LV posterior wall and septal thickness, and amplitude of posterior wall excursion. This observation is not surprising since the boundaries of these structures are often represented by multiple lines rather than clear, single, continuous lines. Reader experience appears to diminish, but not eliminate, interobserver variability, as illustrated by the superior performance of readers with at least 2 years of experience in a laboratory with greater than 800 echo-tracings. Sahn et al. evaluating serial echo-tracings done on the same patient in a time period considered too short to allow for any real change in the measured parameters. A number of investigations have attempted to define and limit potential sources of temporal variability. These studies were designed to develop confidence limits to detect real changes only those differences that fall outside the established confidence limits. Clark et al. defined a coefficient of variation as "the standard deviation of all the presented data in a given case divided by the mean of those measurements." Two standard deviations on either side of the mean were calculated from the estimates of reproducibility, and the 95% confidence interval impact of interobserver variation. The above recommendations were essentially satisfied by the study’s protocol, including use of the ASE's measurement convention. Their intraobserver and interobserver agreement was excellent (Table 8).

**Temporal Variability**

Attempts to define the temporal variability of echocardiographic measurements have utilized serial echograms done on the same patient in a time period considered too short to allow for any real change in the measured parameters. A number of investigations have attempted to define and limit potential sources of temporal variability. These studies were designed to develop confidence limits to detect real changes only those differences that fall outside the established confidence limits. Clark et al. defined a coefficient of variation as "the standard deviation of all the presented data in a given case divided by the mean of those measurements." Two standard deviations on either side of the mean were calculated from the estimates of reproducibility, and the 95% confidence interval

**TABLE 2. Quantification of Interobserver Variance for Echocardiographic Measurements**

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Standard error</th>
<th>Error as percentage of mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVID</td>
<td>±2.35</td>
<td>5.2</td>
</tr>
<tr>
<td>LVd</td>
<td>±2.32</td>
<td>7.5</td>
</tr>
<tr>
<td>IVST</td>
<td>±0.80</td>
<td>10.0</td>
</tr>
<tr>
<td>PWT</td>
<td>±1.35</td>
<td>16.4</td>
</tr>
</tbody>
</table>

Abbreviations as in Table 2. Reprinted from Vignola et al. with permission.

**TABLE 7. Relationship of Percentage Uncertainty for Left Ventricular Dimensions to Measurement Convention**

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Onset QRS</th>
<th>Peak R wave</th>
<th>Peak posterior wall motion</th>
<th>Nadir septal motion</th>
<th>Smallest LV dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVID</td>
<td>8.2</td>
<td>11.8</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>LVd</td>
<td>—</td>
<td>15</td>
<td>14</td>
<td>16</td>
<td>—</td>
</tr>
<tr>
<td>IVST</td>
<td>19.5</td>
<td>23.8</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>PWT</td>
<td>23.4</td>
<td>23</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Adopted for recommendation by the American Society of Echocardiography. Abbreviations as in Table 2. Adapted from Sahn et al. with permission.

**FIGURE 8. Analysis of the percentage of responses falling within the 95% confidence interval for a given echocardiographic measurement made by one, two, or three readers making readings on one or two occasions. Each response was the average of three to five measurements and each reading was made on separate occasions. (Reprinted from Clark et al. with permission of the American Heart Association.)**
was determined for each echocardiographic measurement (see Figure 3). Lapido et al. used a similar method to assess temporal variability. Pietro et al. expressed temporal variation (termed re-performance variability) as a "percentage obtained by the absolute difference between measurements in study 1 and study 2 divided by the measurement in study 1.").

Clarke et al. determined that by their approach a change in end-diastolic LVID of 0.3 cm or greater represented a biologically significant (see Figure 3). The two studies noted above obtained acceptable levels of temporal variability (overall reproducibility). In contrast to those studies, Monoson et al. reported very poor correlations between two sets of readings by the same observer on echocardiograms performed on the same patient on different days. It is not clear why this group obtained such poor reproducibility, but this study stands out as an exception compared to other studies that have addressed this issue and to our own experience.

One application of echocardiographic methods is the evaluation of LV mass in hypertension. As is indicated in Table 2, moderately good reproducibility of LV muscle mass determinations have been obtained in serial echocardiograms on the same subject in our laboratory. An important factor in these measurements is that the echocardiograms on each subject were measured totally independently of each other, maximizing the potential for variability. From data recently reported by other investigators, we anticipate that paired reading of serial tracings on the same subject would reduce this variability by approximately 50%. This approach introduces the possibility of systematic observer bias, however, when any clue is available as to the circumstances under which recordings are taken. Particularly striking examples occur in studies to determine the effects of ß-blockade or heart valve replacement on LV mass or function, where the presence of a reduced heart rate or a prosthetic heart valve would indicate which study followed therapeutic intervention.

With the introduction of two-dimensional echocardiographic imaging, many echocardiographers have concluded, in advance of any data, that two-dimensional guidance of M-mode echocardiograms would result in greater interstudy reproducibility of measurements. This was examined by Pietro et al., who were unable to detect any benefit from two-dimensional guidance. For example, in their patient population the re-performance variability for all structures measured was 8.7 ± 0.9% and 9.4 ± 0.7% for independent M-mode and two-dimensional guided M-mode studies, respectively (Figure 9). A similar lack of any clear advantage obtained by two-dimensional guidance was reported by Panidis et al.

### Table 8: Blind Duplicate Measurements of 10 Tracings by Three Observers

<table>
<thead>
<tr>
<th>Tracing no</th>
<th>LVIDd</th>
<th>LVIDs</th>
<th>PWT</th>
<th>PWT-A</th>
<th>IVST</th>
<th>IVST-A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>1</td>
<td>4.8</td>
<td>4.9</td>
<td>4.9</td>
<td>3.3</td>
<td>3.4</td>
<td>3.3</td>
</tr>
<tr>
<td>2</td>
<td>5.2</td>
<td>5.1</td>
<td>5.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>3</td>
<td>4.6</td>
<td>4.5</td>
<td>4.5</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>4</td>
<td>5.0</td>
<td>4.9</td>
<td>4.9</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>5</td>
<td>4.7</td>
<td>4.5</td>
<td>4.6</td>
<td>2.8</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>6</td>
<td>5.2</td>
<td>5.1</td>
<td>5.1</td>
<td>2.8</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>7</td>
<td>5.2</td>
<td>5.1</td>
<td>5.1</td>
<td>3.7</td>
<td>3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>8</td>
<td>4.9</td>
<td>5.0</td>
<td>4.9</td>
<td>3.6</td>
<td>3.6</td>
<td>3.6</td>
</tr>
<tr>
<td>9</td>
<td>5.2</td>
<td>5.1</td>
<td>5.1</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>10</td>
<td>4.8</td>
<td>4.7</td>
<td>4.8</td>
<td>3.1</td>
<td>2.5</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Reproducibility of Two-Dimensional Echocardiographic Measurements

Two-dimensional echocardiographic imaging has been introduced more recently than M-mode echocardiography, and hence its quantitative applications have been less extensively studied. It provides the advantage of a second dimension in which to view cardiac structures, thus eliminating some errors in ultrasound beam angulation that otherwise might go unrecognized. But two-dimensional echocardiography also has some inherent, unique limitations that affect reproducibility. Difficulties common to two-dimensional and M-mode echocardiography include measurement or intrarater variability, interobserver variability, and temporal variability. Similarly, the necessity for strict quality control and the importance of standardization of technique are pertinent considerations. A com-
mittee of the ASE has published clear recommendations concerning standard transducer positions and imaging planes in an attempt to standardize the views from which measurements are made. It must be recognized, however, that the two-dimensional tomographic planes may be off axis or misaligned in the third direction, as shown most clearly by Erbel et al. in a study in which echocardiographic views from the so-called apical window were shown to miss the true LV apex (as defined by angiography) by a wide margin in many patients. Delineation of technical precautions to prevent this and comparable errors is urgently needed.

Experience with two-dimensional echocardiographic imaging has identified some internal checks, available to the sonographer, which will limit the error introduced by lack of knowledge regarding the third dimension in the planar images produced. For example, in determining the correct plane for a short axis LV view, rigorously producing the smallest circular cross section at a standardized level (just below the tips of the mitral valve or at the level of the papillary muscles) will minimize errors of oblique angulation of the beam. A second useful internal check pertains to the apical four- and two-chamber views. The true long axis of the left ventricle may be checked by producing what appears in a four-chamber view to be a maximal long axis and transverse dimension and then turning the transducer 90 degrees to record the two-chamber view. If the measured long axis is longer in this view, then one can be certain that the initial four-chamber plane did not maximize the LV long axis. In all long-axis or apical views, it is also important to maximize the transverse dimension of the left ventricle (or other chamber). Otherwise, the oblique planes may underestimate chamber size and overstate wall thickness and motion.

Important quantitative applications of two-dimensional echocardiography include measurement of valve areas and LV muscle mass, both of which are supported by careful validation studies with comparison to appropriate reference standards. LV wall motion and volumes in asymmetric chambers can also be assessed by two-dimensional echocardiography although results of numerous proposed approaches are not entirely consistent.

LV volume and ejection fraction estimation are probably the most frequent measurements for which two-dimensional echocardiography is used since, theoretically, this mode of imaging should be superior to M-mode in that asymmetric chambers may be more accurately represented. Reproducibility and accuracy share a common limitation with M-mode echocardiography; however: endocardial surface identification. Manipulation of gain settings may help, but even in the newer generation of equipment, endocardial surface identification is a difficult task. Compounding this problem is the observation by Erbel et al. in a study of Helak and Reichek that considerable overestimation of interface thickness was produced by a variety of two-dimensional echocardiographs. This overestimation appeared to be less prominent in a study by Wyatt et al. Efforts to enhance interface recognition have generated considerable literature. Computer-assisted endocardial surface identification and approximating the LV outline using a calibrated ellipsoid to obviate the painstaking tracing of the endocardial surface outline have both been recommended. For serial studies it has also been proposed to record and replicate gain settings to alleviate variability in interface thickness and location. Although reasonable correlations between angiographic and two-dimensional echocardiographic determinations of LV volumes have been obtained, two-dimensional echocardiographic methods have systematically underestimated angiographic LV volumes.

Despite these problems, acceptable reproducibility of LV volume measurements have been obtained by newer two-dimensional echocardiographic methods. For example, Gordon et al. showed that for group data (30 subjects), two-dimensional echocardiographic determination showing changes of 2% for LV end-diastolic volume and ejection fraction and 5% for LV end-systolic volume would be significant. For serial measurements in an individual patient, however, one would need changes of at least 15% for LV end-diastolic volume, 25% for LV end-systolic volume, and 10% for ejection fraction to be confident that the observed changes exceeded 95% confidence limits.

These data suggest that two-dimensional echocardiographic LV measurements are sufficiently reproducible to be utilized in hypertension and epidemiologic research. At the same time, however, the limited availability of data concerning variability between serial two-dimensional echocardiograms on the same subject (as opposed to the more extensive information substantiating reasonable interobserver and intraobserver variability of measurements of the same echocardiogram needs to be corrected. In addition, the reproducibility of two-dimensional measurements needs to be assessed in patients with specific disease states, as has been done for M-mode measurements in patients with aortic regurgitation or congestive cardiomyopathy.

Acknowledgment
We thank Miss Virginia Burns for her assistance in preparation of this manuscript.
REPRODUCIBILITY OF ECHOCARDIOGRAPHIC MEASUREMENTS/Wallerjon and Devereux

References

5. Wong M, Shah PM, Taylor RD. Reproducibility of left ventricular internal dimensions measurement. Am J Cardiol 1980;65:505-510
15. Belsakie1 Bead to beat variability of echocardiographic measurements of left ventricular end-diastolic diameter and performance JCU 1979;7:26-28.1
28. Devereux RB, Savage DD, Drayer JM, Laragh JH. Left ventricular hypertrophy and function in high, normal, and low-renin forms of essential hypertension. Hypertension 1982;65:99-108
34. Gardin JM, Henry WL, Savage DD, Ware JH, Burn C, Borer JS. Echocardiographic measurement in normal subjects: evaluation of an adult population without clinically apparent heart disease. JCU 1979;7:447-447
42. Vignola PA, Bloch A, Kaplan AD, Walker HJ, Chetiellis PN, Myers GS. Interobserver variability in echocardiography. JCU 1975;7:238-247
43. Monson PA, O'Rourke RA, Crawford MH, White DH. Measurements of left ventricular wall thickness and systolic thickening by M-mode echocardiography: interobserver and intrarow variability. JCU 1978,6:252-258
46. Devereux RB, Savage DD, Drayer JM, Laragh JH. Left ventricular hypertrophy and function in high, normal, and low-renin forms of essential hypertension. Hypertension 1982;4:524-531
48. Umaili MSA, Prakash RR. Magnitude of daily variations in serial echocardiographic measurements of septal thickness, left ventricular wall thickness and estimated left ventricular mass [Abstract]. Clin Res 1984;32:212A
Reproducibility of echocardiographic left ventricular measurements.
D C Wallerson and R B Devereux

Hypertension. 1987;9:II6
doi: 10.1161/01.HYP.9.2_Pt_2.II6

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://hyper.ahajournals.org/content/9/2_Pt_2/II6

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Hypertension can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Hypertension is online at:
http://hyper.ahajournals.org//subscriptions/