Refinements in stress echocardiographic techniques improve inter-institutional agreement in interpretation of dobutamine stress echocardiograms

R. Hoffmann¹, T. H. Marwick², D. Poldermans³, H. Lethen⁴, R. Ciani⁵, P. van der Meer³, H.-P. Tries², P. Gianfagna⁵, P. Fioretti⁵, J. J. Bax³, M. A. Katz⁶, R. Erbel⁶ and P. Hanrath¹

¹Medical Clinic I, University RWTH Aachen, Aachen, Germany; ²University of Queensland, Brisbane, Australia; ³Thoraxcenter Rotterdam, Rotterdam, The Netherlands; ⁴DKD Wiesbaden, Wiesbaden, Germany; ⁵Azienda Ospedaliera S Maria della Misericordia, Udine, Italy; ⁶University GHS Essen, Essen, Germany

Aims To determine the degree of inter-institutional agreement in the assessment of dobutamine stress echocardiograms using modern stress echocardiographic technology in combination with standardized data acquisition and assessment criteria.

Method and Results Among six experienced institutions, 150 dobutamine stress echocardiograms (dobutamine up to 40 µg.kg⁻¹.min⁻¹ and atropine up to 1 mg) were performed on patients with suspected coronary artery disease using fundamental and harmonic imaging following a consistent digital acquisition protocol. Each dobutamine stress echocardiogram was assessed at every institution regarding endocardial visibility and left ventricular wall motion without knowledge of any other data using standardized reading criteria. No patients were excluded due to poor image quality or inadequate stress level. Coronary angiography was performed within 4 weeks. Coronary angiography demonstrated significant coronary artery disease (≥50% diameter stenosis) in 87 patients. Using harmonic imaging an average of 5.2 ± 0.9 institutions agreed on dobutamine stress echocardiogram results as being normal or abnormal (mean kappa 0.55; 95% CI 0.50–0.60). Agreement was higher in patients with no coronary artery disease (equal assessment of dobutamine stress echocardiogram results by 5.5 ± 0.8 institutions) or three-vessel coronary artery disease (5.4 ± 0.8 institutions) and lower in one- or two-vessel disease (5.0 ± 0.9 and 5.2 ± 1.0 institutions, respectively; P=0.041). Disagreement on test results was greater in only minor wall motion abnormalities. Agreement on dobutamine stress echocardiogram results was lower using fundamental imaging (mean kappa 0.49; 95% CI 0.44–0.54; P<0.01 vs harmonic imaging).

Conclusion Modern echocardiographic technology in combination with standardized digital image processing and uniform reading criteria results in a higher inter-institutional agreement in the interpretation of dobutamine stress echocardiogram compared to historic reports. (Eur Heart J, 2002; 23: 821–829, doi:10.1053/euhj.2001.2968)

© 2001 The European Society of Cardiology. Published by Elsevier Science Ltd. All rights reserved.

Key Words: Dobutamine echocardiography, coronary artery disease, interpretation, concordance, accuracy.

See page 768, doi:10.1053/euhj.2001.3051 for the Editorial comment on this article

Introduction

Dobutamine stress echocardiography is widely used for the detection of coronary artery disease and the evalu-
In recent years there has been a significant evolution in stress echocardiographic technologies. Digital image storage has allowed standardized image acquisition and data display without loss of image resolution. Second harmonic imaging has afforded unprecedented endocardial border delineation, even in patients previously considered poor candidates for stress echocardiographic examinations\(^4,5\). The combination of these technological advances has resulted in significant improvements in image quality. A further important element in the drive to improve interpretive agreement in stress echocardiography has been the adoption of standardized reading guidelines\(^3,6\). However, while each of the described elements might be anticipated to narrow the interpretive range in the reading of stress echocardiograms, this benefit has not been established, nor has the impact of a combination of the described improvements on inter-institutional variance in interpretation been evaluated.

Thus, the results of an historic analysis\(^1\) are still used to describe inter-observer reader variance in stress echocardiography although this study no longer reflects the current practice of the technique. The present study analysed the agreement within a panel of six expert readers on interpretation of 150 dobutamine stress echocardiograms performed using modern techniques.

**Methods**

**Patients**

At each of six institutions, 25 consecutive patients scheduled to undergo coronary angiography for the evaluation of suspected coronary artery disease underwent dobutamine stress echocardiography. A total of 150 patients (111 men, 39 women; mean \(\pm\) SD age 61.4 \(\pm\) 10.8 years, range 26 to 83) were investigated. Patients with previous Q wave myocardial infarction, congestive heart failure, severe congenital or acquired valvular heart disease or documented cardiomyopathy were excluded.

No patient was excluded on the basis of poor echocardiographic image quality. Previous myocardial revascularization had been performed in 46 patients (35 with coronary angioplasty, 15 with coronary artery bypass grafting, five with both revascularization procedures), and 35 patients had a non-Q wave infarction. One hundred and twenty-four patients reported angina within the last 4 weeks before the examination. Antianginal therapy before the examination included nitrates in 85 patients, calcium channel antagonists in 40 and beta-blocker agents in 93 patients. Three patients were on digoxin. Beta-blocker medication was stopped 3 days before stress echocardiography in 33 patients. The study was approved by the local ethical committees. All patients were studied after giving written, informed consent.

**Dobutamine stress protocol**

Dobutamine infusion was started at 5 \(\mu\)g . kg\(^{-1}\) body weight per minute and increased every 3 min to 10, 20, 30 and 40 \(\mu\)g . kg\(^{-1}\) min\(^{-1}\). If 85% of age-predicted maximal heart rate was not attained and the test was negative for ischaemia, additional atropine was administered intravenously in 0-25 mg increments each minute, up to a maximum of 1·5 mg. The pharmacological stress was stopped before reaching the maximal dosage for any of the following reasons: development of new wall motion abnormalities; progressive and severe angina or dyspnoea; severe ventricular arrhythmias; systolic or diastolic blood pressure >240 mmHg or >120 mmHg, respectively, or a decrease in systolic blood pressure \(\geq\) 20 mmHg. A 12-lead electrocardiogram was monitored continuously and blood pressure was measured every 2 min during and up to 10 min after dobutamine stress echocardiography. The rate pressure product was calculated by multiplying systolic blood pressure and heart rate.

**Image acquisition**

Image acquisition was performed with the patient in the left lateral decubitus position. In each patient, imaging was performed in the parasternal long-axis and short-axis as well as in the apical four- and two-chamber views. Imaging was performed throughout the examination. However, only the images at rest and peak stress were acquired and stored for subsequent analysis. All studies were stored in digital format on magneto-optical disks. Both modalities (fundamental imaging and second harmonic imaging) were used to acquire image loops at rest and at peak stress. All centres used an Agilent Sonos 5500 system (Agilent Technologies, Andover, Massachusetts, U.S.A.) with a 4 MHz fundamental transducer (2-1/4-2 MHz for second harmonic imaging).

**Interpretation of stress echocardiograms**

One experienced observer per institution evaluated visibility and wall motion of all dobutamine stress echocardiograms separately, in random order and non-consecutively for the fundamental and second harmonic imaging modus. To base interpretation exclusively on imaging data, the six centres evaluated all 150 dobutamine stress echocardiograms using a standardized report form without knowledge of any patient data apart from the echocardiographic images. No information on the reason for termination of the dobutamine stress test and the maximal pharmacological stress level was given to the interpreting physician. Clinical or angiographic data of the patients were also not made available.

Readers were asked to apply previously defined specific reading criteria for both wall motion analysis and endocardial visibility\(^2\), although no joint reading session was undertaken. Endocardial visibility of each segment was evaluated at rest and at peak dobutamine dose using a scoring system which ranges from 1 = barely
or not visible endocardium, to 2=moderately or partly visible endocardium, to 3=good visible endocardium as described previously[6]. For myocardial segments visible in more than one view the higher visibility score was used for further analysis. For wall motion analysis, the 16 segment model of the American Society of Echocardiography[7] was used. Wall motion was evaluated for each segment as 1=normal, 2=mildly hypokinetic, 3=severely hypokinetic, 4=akinet c, 5=dyskinetic. Rest and peak stress stages were compared side by side.

An abnormal stress echocardiogram was defined as an increase in score from rest to stress in at least one segment. In addition to these general reading criteria, previously[2] defined standardized reading guidelines were applied which required that (a) basal inferior and basal septal hypokinesia are not identified as being abnormal unless an adjacent segment is also affected by a new dyssynergy or there is a clear deterioration of function to akinesia or dyskinesia, (b) induced delayed contraction is an index of ischaemia in the absence of conduction disturbance, (c) identification of ischaemia is based on anticipated coronary territories, (e.g. a new wall motion abnormality affecting only the midseptum does not make sense) (d) significant resting wall motion abnormalities (hypokinesia in at least three segments or akinesia in at least one segment) also indicates an abnormal test and the presence of coronary artery disease.

**Coronary angiography and quantitative coronary angiography**

Coronary angiograms were stored digitally and the severity of coronary stenosis was determined quantitatively by a blinded core lab using commercially available software (QuantCor, CASS II, Siemens, Erlangen, Germany). Significant coronary artery disease was considered present when ≥50% reduction of vessel diameter was observed in at least one major coronary artery. The severity of coronary artery disease was classified as one-, two- or three-vessel disease. Bypass graft stenosis was counted as equivalent to stenosis of the grafted native vessel.

**Statistical analysis**

All continuous variables are expressed as mean ± (SD). Comparison between different modalities was performed using paired and unpaired Student’s t-tests, analysis of variance and Pearson’s linear correlation for continuous variables, and chi-square test or Spearman’s rank test for categorical variables as appropriate. In addition to analysing the overall agreement on interpretation of dobutamine stress echocardiograms, clinical, procedural and echocardiographic variables which impacted the inter-institutional agreement were studied. Echocardiographic variables consisted of the size of the wall motion abnormality during peak stress as defined by the number of abnormal segments, the severity of the induced dyssynergy (mild or severe hypokinesia vs akinesia or worse) and image quality of stress echocardiograms.

The Cohen-κ coefficient for inter-observer agreement was calculated to test the hypothesis that agreement was greater than chance alone[8]. Average coefficients of agreement (kappa) were computed for the six readers of the different institutions. The coefficient of agreement (kappa) was graded as follows: 0 to 0·2=poor to slight; 0·21 to 0·4=fair; 0·41 to 0·6=moderate; 0·61 to 0·8=substantial; 0·81 to 1·0=nearly perfect. Analysis was performed separately for fundamental and second harmonic imaging, although special analyses relate to dobutamine stress echocardiography results obtained with harmonic imaging if not otherwise stated. Significant differences were defined as a P value <0·05.

**Results**

**Coronary angiography**

Coronary angiography identified 87 patients as having coronary artery disease; 47 had one-vessel disease (54% of those with significant coronary artery disease), 33 (38%) had two-vessel disease and seven (8%) had three-vessel disease. In 53 patients (35%) the left anterior descending coronary artery was affected, in 42 (28%) the circumflex artery and in 36 (24%) the right coronary artery.

**Dobutamine stress test**

The dobutamine stress was stopped prematurely in 36 patients (severe ventricular extrasystolic beats: nine patients, hypotension: four patients, hypertension: three patients, new atrial fibrillation or supraventricular tachycardia: two patients, significant angina pectoris: 18 patients). The average maximal dobutamine dosage was 38 ± 5 μg . kg⁻¹ . min⁻¹. Additional atropine was given in 93 (62%) patients with a mean dosage of 0·8 ± 0·5 mg. The frequency of additional atropine administration differed between the institutions, ranging from 35% to 91% (P<0·001). Atropine dosage ranged between the centres from 0·5 ± 0·4 to 1·2 ± 0·6 mg (P<0·001). Centres more frequently using atropine also applied higher doses.

Pharmacological stress resulted in an increase in heart rate from 67 ± 13 beats . min⁻¹ (range 45 to 107) at baseline to 136 ± 17 beats . min⁻¹ (range 69 to 176) at peak stress. The rate-pressure product increased from 9180 ± 2395 mmHg min⁻¹ to 21 629 ± 5367 mmHg min⁻¹ with maximal pharmacological stress. Forty-two patients developed angina during the stress test, and electrocardiographic changes developed in 38 patients.
Echocardiographic results

Average endocardial visibility was best at rest using harmonic imaging. Endocardial visibility was significantly worse during stress conditions and when fundamental imaging was used (Table 1). The six institutions submitted dobutamine stress echocardiograms with significant differences in endocardial visibility, as demonstrated by the mean scoring from all six readers (Table 1). This difference in endocardial visibility was evident for fundamental and second harmonic imaging as well as for rest and peak stress conditions. There was also a significant difference between the six institutions in the assessment of endocardial visibility (Table 2). Institutions that submitted echocardiograms with the lowest endocardial visibility tended to assess the visibility of dobutamine stress echocardiograms highest (Fig. 1).

Echocardiographic test results

The number of dobutamine stress echocardiograms identified as positive or negative differed between the six institutions. With harmonic imaging, an average of 74 dobutamine stress echocardiograms, with a range from 62 to 87, was evaluated as positive (Fig. 2(a)). The variation between centres in the interpretation of stress echocardiograms as positive was significant ($P=0.006$). Significant resting wall motion abnormalities as defined by the evaluation of at least five readers were present in six patients (4%). Assessment of images acquired with fundamental imaging resulted in an average of 69 positive dobutamine stress echocardiograms with a range from 55 to 83 between the six institutions (Fig. 2(b)). There was also a significant variation between centres in the interpretation of stress echocardiograms as positive; this was significant ($P=0.002$).

Agreement between institutions

In terms of the results of harmonic imaging, complete agreement on dobutamine stress echocardiograms as being positive or negative was obtained in 67 (45%) cases. Agreement of five, four or three out of the six institutions was achieved in 47 (31%), 32 (21%) and four (3%) cases, respectively. For each of the 150

---

Table 1  Endocardial visibility of echocardiograms as assessed by all readers depending on the institution submitting the stress echocardiograms. Endocardial visibility assessment is calculated as mean of all 16 segments

<table>
<thead>
<tr>
<th>Echocardiograms submitted by</th>
<th>Endocardial visibility</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Harmonic rest</td>
<td>Harmonic stress</td>
<td>Fundamental rest</td>
<td>Fundamental stress</td>
</tr>
<tr>
<td>Centre A</td>
<td>2.81 ± 0.28</td>
<td>2.68 ± 0.36</td>
<td>2.45 ± 0.46</td>
<td>2.36 ± 0.48</td>
</tr>
<tr>
<td>Centre B</td>
<td>2.72 ± 0.38</td>
<td>2.64 ± 0.40</td>
<td>2.53 ± 0.44</td>
<td>2.39 ± 0.54</td>
</tr>
<tr>
<td>Centre C</td>
<td>2.46 ± 0.44</td>
<td>2.40 ± 0.47</td>
<td>2.16 ± 0.50</td>
<td>2.14 ± 0.54</td>
</tr>
<tr>
<td>Centre D</td>
<td>2.46 ± 0.44</td>
<td>2.35 ± 0.49</td>
<td>2.10 ± 0.47</td>
<td>2.09 ± 0.49</td>
</tr>
<tr>
<td>Centre E</td>
<td>2.51 ± 0.43</td>
<td>2.32 ± 0.53</td>
<td>2.27 ± 0.49</td>
<td>2.10 ± 0.54</td>
</tr>
<tr>
<td>Centre F</td>
<td>2.58 ± 0.42</td>
<td>2.30 ± 0.52</td>
<td>2.30 ± 0.46</td>
<td>2.08 ± 0.53</td>
</tr>
<tr>
<td>P-ANOVA</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean</td>
<td>2.59 ± 0.50*†</td>
<td>2.45 ± 0.49‡</td>
<td>2.31 ± 0.41§</td>
<td>2.20 ± 0.54</td>
</tr>
</tbody>
</table>

* $P<0.001$ harmonic rest vs harmonic stress, † $P<0.001$ harmonic rest vs fundamental rest, ‡ $P<0.001$ harmonic stress vs fundamental stress, § $P<0.001$ fundamental rest vs fundamental stress.

Table 2  Endocardial visibility of all 150 submitted echocardiograms depending on the institution analysing the visibility. Endocardial visibility assessment is calculated as mean of all 16 segments

<table>
<thead>
<tr>
<th>Echocardiograms analysed by</th>
<th>Endocardial visibility</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Harmonic rest</td>
<td>Harmonic stress</td>
<td>Fundamental rest</td>
<td>Fundamental stress</td>
</tr>
<tr>
<td>Centre A</td>
<td>2.34 ± 0.43</td>
<td>2.13 ± 0.46</td>
<td>1.98 ± 0.45</td>
<td>1.81 ± 0.44</td>
</tr>
<tr>
<td>Centre B</td>
<td>2.49 ± 0.45</td>
<td>2.29 ± 0.51</td>
<td>2.20 ± 0.44</td>
<td>2.09 ± 0.50</td>
</tr>
<tr>
<td>Centre C</td>
<td>2.51 ± 0.36</td>
<td>2.44 ± 0.35</td>
<td>2.19 ± 0.36</td>
<td>2.16 ± 0.41</td>
</tr>
<tr>
<td>Centre D</td>
<td>2.77 ± 0.27</td>
<td>2.59 ± 0.37</td>
<td>2.50 ± 0.34</td>
<td>2.26 ± 0.44</td>
</tr>
<tr>
<td>Centre E</td>
<td>2.67 ± 0.41</td>
<td>2.57 ± 0.51</td>
<td>2.42 ± 0.59</td>
<td>2.39 ± 0.62</td>
</tr>
<tr>
<td>Centre F</td>
<td>2.76 ± 0.40</td>
<td>2.70 ± 0.45</td>
<td>2.55 ± 0.48</td>
<td>2.49 ± 0.52</td>
</tr>
<tr>
<td>$P$</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
dobutamine stress echocardiograms, agreement on test results as normal or abnormal was obtained by 5·2 ± 0·9 institutions. The average coefficient of agreement (kappa) among readers from different institutions for normality or abnormality of the dobutamine stress echocardiograms was 0·55 (95% CI 0·50–0·60; range 0·40–0·75 for any two of the six institutions). The average kappa coefficient for any of the six institutions, related to the other five institutions ranged from 0·49 to 0·63. Average agreement was 5·7 ± 0·4 on the six patients with significant resting wall motion abnormalities as defined by the majority assessment of at least five readers. For the 144 patients without significant resting wall motion abnormalities agreement was 5·2 ± 0·9. Agreement for each of the 16 segments ranged from 5·3 ± 1·2 to 5·8 ± 0·5 institutions with similar reading of test results as normal or abnormal. There was a tendency for lower agreement in the basal inferior

Figure 1  Endocardial visibility of echocardiograms using harmonic imaging at peak stress conditions. Solid bars relate assessment of endocardial visibility to the centre by which echocardiograms were submitted. Assessment is calculated as average of all six institutions for all 16 segments. Open bars relate assessment of endocardial visibility of all 150 dobutamine echocardiograms to each analysing institution.

Figure 2  Evaluation of positivity (solid bars) versus negativity (open bars) for the 150 dobutamine stress echocardiograms by institution. Numbers above bars=number of patients. The upper panel (a) relates to assessment using harmonic imaging, the lower panel (b) relates to fundamental imaging.
and basal posterior septal segments (Fig. 3). However, this did not reach statistical significance.

Agreement between institutions on the results of fundamental imaging was lower. For each of the 150 dobutamine stress echocardiograms, agreement on test results as normal or abnormal was obtained by 4·8 ± 0·9 institutions (P = 0·02 vs harmonic imaging). The average coefficient of agreement (kappa) among all readers from the six institutions was 0·49 (95% CI 0·44–0·49; P < 0·01 vs harmonic imaging) with a range from 0·37 to 0·68.

**Agreement corresponding to disease severity**

Agreement was dependent on disease severity. The number of institutions with agreement that the harmonic imaging result was positive or negative was 5·5 ± 0·8 for patients with no coronary artery disease and 5·4 ± 0·8 institutions for patients with three-vessel disease. For patients with two- and one-vessel disease, agreement was less, being recorded in 5·0 ± 0·9 and 5·2 ± 1·0 institutions, respectively (P = 0·041) (Fig. 4).

**Clinical and procedural variables influencing interpretive agreement**

Age, patient gender and medication on the day of dobutamine stress echocardiography had no significant impact on interpretive agreement. The maximum dobutamine dose and the maximum achieved rate pressure product were not found to be significantly different between the observed agreement levels. Atropine administration was most frequent in dobutamine stress echocardiograms with lowest agreement (agreement by only 3/6 centres) and least frequent in tests with highest agreement (6/6 centres). Tests on which the highest agreement was obtained had induced wall motion abnormalities of greater extent and severity (Table 3). This also resulted in a greater wall motion score index for dobutamine stress echocardiograms with the greatest agreement.

**Sensitivity, specificity, diagnostic accuracy**

Sensitivity, specificity and diagnostic accuracy of each institution in the detection of angiographically proven coronary artery disease in terms of stress echocardiograms acquired with harmonic imaging are given in Table 4. Sensitivity for detection of significant coronary artery disease ranged from 62% to 85%. Comparison of sensitivity and accuracy for detection of significant
coronary artery disease demonstrated significant differences between the six institutions.

**Discussion**

This study demonstrates that (1) a higher inter-institutional agreement in the interpretation of dobutamine stress echocardiograms can be achieved using modern echocardiographic technology in combination with standardized reading criteria than historic reports; (2) the use of second harmonic imaging contributes importantly to this improvement; (3) greater severity and extent of induced wall motion abnormality results in better agreement on interpretation of dobutamine stress echocardiograms; (4) in spite of improvements in reader agreement compared to historic reports, there is still a significant variance which resulted in different diagnostic accuracy; (5) institutions most critical of the assessment of endocardial visibility performed dobutamine stress echocardiograms of highest image quality.

**Previous studies**

An analysis of five expert institutions performed in 1994 demonstrated low agreement in the interpretation of dobutamine stress echocardiograms. An average coefficient of agreement among investigators from different institutions for normality or abnormality of dobutamine stress echocardiograms of 0.37 was reported. This analysis has been disappointing for the echocardiographic community and frequently mentioned by opponents of the technique as an example of the low reliability of stress echocardiographic results. However, the result reflected the stress echocardiographic practice of that time. Several factors causing the high variability in stress echocardiographic reading could be identified. Most importantly, some readers were more conservative in reading stress echocardiograms as positive and others who tended to over-interpret studies, indicating that reading criteria were different. Obviously, the highly subjective process of stress echo interpretation was insufficiently constrained by strict interpretation guidelines. Furthermore, low image quality (characterized by insufficient endocardial border delineation) and induced wall motion abnormalities of minor severity resulted in low reader agreement. In a subsequent study, standardized reading criteria were defined to narrow the interpretive variations. A reduction in inter-institutional variance in interpretation was demonstrated if dobutamine stress echocardiograms were read against the standardized reading criteria.

**Current study**

Stress echocardiography has experienced a rapid evolution in recent years owing to improved technology and methodology. Second harmonic imaging, which allows more reliable wall motion analysis especially in cases with poor fundamental image quality, has rapidly been extended to stress echocardiography. In single centre studies, image quality, diagnostic accuracy as well as inter-observer agreement on stress echocardiographic interpretation have all been improved by this modality.

This study sought to combine all the improvements which have been incorporated into the practice of stress echocardiography to allow an assessment of inter-institutional agreement with current technology. Thus,
compared to an analysis of inter-institutional agreement performed by almost the same group of investigators several years ago, second harmonic imaging, digital image processing, a slightly higher stress level and standardized reading criteria were applied. The higher level of pharmacological stress resulted in a slightly higher rate-pressure product and may have induced more pronounced wall motion abnormalities, fostering easier differentiation of normal and pathological responses in some borderline situations[12,13].

This study is important in that it shows that inter-institutional reader agreement can be improved above a level reported using old techniques[1]. However, there remained a significant variability in reading of stress echocardiograms. This resulted in a variability of the diagnostic accuracy for detection of coronary artery disease. Centres which had the highest rates of positive stress echocardiograms tended to have the highest sensitivity and the lowest specificity for detection of coronary artery disease, while centres with the lowest rates of positive stress echocardiograms had the lowest sensitivity and the highest specificity.

There was an interesting reverse relationship, independent of the applied imaging modality, between the image quality (assessed as endocardial border visibility) of echocardiograms provided by an institution and the assessment of image quality by the same institution. Institutions which were more critical in the assessment of endocardial visibility provided dobutamine stress echocardiogram of the highest image quality as assessed by all six readers. Thus, in order to obtain high image quality it is important to retain a high level of criticism towards the level of achieved image quality.

The burden of a considerable inter-observer variability in the interpretation of test results has been described for other imaging techniques and could be reduced by different approaches. The agreement in interpretation of thallium-201 imaging, as demonstrating myocardial ischemia, could be increased from a kappa value of 0·36 without standardization on interpretation to 0·71 with standardization of display and quantification[14]. Thus, the current level of variation on dobutamine stress echocardiogram reading is still greater than those described for scintigraphic techniques. The known variations in wall motion response to dobutamine stress within each patient induce difficulties in the differentiation of normal and pathological segments[15,16]. The effect of this difficulty is especially pronounced in the basal inferior and septal segment[17]. To counter this difficulty, an additional reading criteria which stipulates that basal inferior and septal segments should be considered in combination with the adjacent wall segments has been defined.

Reading of stress echocardiograms in clinical practice is still performed without a modality allowing objective quantitative analysis of wall motion and wall motion response to stress. Recently several techniques have been described for quantitative analysis of stress echocardiograms. These include automatic endocardial border detection, tissue Doppler imaging and strain rate imaging[18–22]. There are still significant limitations to these methods, in particular the time requirements of their application, which have prohibited their integration into standard clinical practice. However, further refinements of these techniques are likely to facilitate acceptance by clinicians.

Limitations of the study
In this study all institutions used the same echocardiographic equipment to allow reading of digital image material at all study sites. Thus, the situation created in this study does not reflect the ongoing difficulties encountered during digital data transfer and reading using currently available echocardiographic equipment of different providers.

Harmonic imaging was always performed after fundamental imaging. We did not change the order of the imaging sequence in order to prevent comparison of the wrong images during the reading process. Although imaging at peak stress did not require more than 1 min for each imaging modality the acquisition of second harmonic imaging after fundamental imaging might have resulted in a slightly higher pharmacological stress level and consequently a slightly greater ischaemic response during second harmonic imaging.

Four percent of patients had significant resting wall motion abnormalities, as defined by the majority of readers, which eased reading of the dobutamine stress echocardiograms as pathological. We used a Q wave myocardial infarction as exclusion criteria instead of resting wall motion abnormalities because an exclusion criterion based on resting wall motion abnormalities would have been more affected by the subjective interpretation process of the reader including the patient. Furthermore, we wanted to keep patient inclusion criteria similar to those used in a multicentre reading study performed several years ago to allow better comparison of test results. The number of patients with three-vessel disease was low in this study limiting the conclusion on the dependency of reader agreement on severity of coronary artery disease.

Conclusions and clinical implications
An improved inter-institutional observer agreement in the interpretation of dobutamine stress echocardiograms can be achieved using modern stress echocardiographic technology in combination with standardized reading criteria. Second harmonic imaging has a significant impact on endocardial visibility and contributes to improved observer agreement. The application of the factors described in this study correlate with the maturation of dobutamine stress echocardiography into a more applicable clinical technique with narrower reader variability.

We acknowledge the support of Agilent Technologies for providing the echocardiographic equipment, the data entry performed by Ellen Severin, and the help of Dr. Bernardi, Udine,
Dr. Pasquale Gianfagna, Udine, and Prof. Heinz Lambertz, Wiesbaden, in performance of the study.

References


