Improved interpretation of dobutamine stress echocardiography following 4 months of systematic training in patients following acute myocardial infarction

D.V. Anand, I.D. Theodosiadis, R. Senior

Northwick Park Hospital, Watford Road, Harrow HA1 3UJ, UK

Abstract

Objectives  Interpretation of stress echocardiography is subjective, and highly dependent on the experience of the interpreter. We sought to evaluate whether a cardiologist without any previous experience in stress echocardiography could adequately learn the skills of interpreting dobutamine stress echocardiograms (DSE) in post-infarct patients, after a period of systematic training.

Methods  A trainee in cardiology blindly reported 51 consecutive DSEs from a database of post-infarction studies, after 2 and 4 months of systematic training. We compared his interpretation with that of an expert.

Results  Agreement between the trainee and the expert improved significantly from 2 to 4 months of training in the left anterior descending artery territory for the overall scan interpretation (from $k = 0.58$ to $k = 0.73$; $p = 0.03$), wall thickening assessment in individual segments (from $k = 0.40$ to $k = 0.55$; $p < 0.01$) and the diagnosis of viable myocardium (from $k = 0.11$ to $k = 0.43$; $p = 0.01$). Similar improvement was observed in left circumflex, but not in the right coronary artery territory. Agreement in identifying inducible ischaemia also remained poor.

Conclusion  This study suggests that systematic training can significantly reduce interobserver variability in a short time frame (4 months) and may improve the interpretation of DSE by a trainee. But improvements in image quality and use of predefined reading criteria are necessary to improve interobserver agreement further in myocardial regions where conformity in dobutamine stress echocardiographic interpretation is low.

© 2003 The European Society of Cardiology. Published by Elsevier Ltd. All rights reserved.

Introduction

Dobutamine stress echocardiography (DSE) has emerged as an accurate, non-invasive technique to detect coronary artery disease (CAD), and
evaluate the functional relevance of angiographically proven coronary artery lesions.\textsuperscript{1,2} The independent and additive prognostic value of DSE over and above other clinical, resting echocardiography and exercise electrocardiography variables has been well established.\textsuperscript{3,4} After an acute myocardial infarction (MI), DSE can identify patients at high risk of subsequent ischaemic events.\textsuperscript{5} In patients with established CAD and left ventricular dysfunction, DSE can identify dysdyssynergic yet viable myocardium and predict improvement in function after successful revascularisation.\textsuperscript{6,7} However, reading of DSE is dependent on subjective interpretation of visual information, a process which has been shown to be extremely sensitive to the experience of the reader.\textsuperscript{8} Furthermore there may be considerable differences in the interpretation of DSEs, even among experienced readers.\textsuperscript{9}

The aim of this study was to evaluate whether a cardiologist without any previous experience in stress echocardiography could learn the skills of interpretation of DSE in post-infarct patients after 4 months of systematic training.

**Methods**

**Study design**

A trainee in cardiology was given a systematic 2-month course of training in DSE. This included formal lectures on the principles of DSE, practical training in stress echocardiography (12–16 stress echocardiograms are performed in the laboratory weekly), image acquisition and interpretation by a consultant cardiologist with expertise in stress echocardiography. After 2 months of training the trainee blindly reported DSEs from the database, which have been previously reported by a consultant cardiologist with expertise in stress echocardiography. After 2 months of further training, the trainee was asked to report the same echocardiograms, blinded to his initial interpretations.

**Patient selection**

Fifty-one consecutive DSEs from a database of post-infarct studies were analysed. Only haemodynamically stable patients with no post-infarction angina were studied between 3 and 7 days after an acute MI. All patients had their beta-blockers stopped for at least 24 h prior to the test. The hospital research ethics committee approved the study protocol and all patients gave informed consent.

**Dobutamine stress echocardiography**

**Protocol**

After a baseline electrocardiogram and echocardiogram were obtained, DSE was performed as previously described\textsuperscript{10} on an HDI 5000 ultrasound system (Advanced Technological Laboratories, Bothell, WA, USA) with the use of a broadband width (2–4 MHz) transducer in the tissue harmonic mode. Endpoints of the test were the achievement of 85% of target heart rate; development of severe ischaemia (increasing angina, extensive wall motion abnormality or >2 mm ST segment depression); development of hypotension (decrease in systolic BP \(\geq 20 \text{mmHg}\)); significant hypertension (systolic blood pressure \(>240\text{mmHg}\)) and dyspnoea or severe ventricular arrhythmias. Images were acquired in the parasternal (long and short axis) and apical (chambers 2 and 4) views and evaluated from a digitised quad-screen display.

**Interpretation**

A 16 segment LV model, which has been previously well validated in our laboratory was used.\textsuperscript{6,7,10} Systolic wall thickening (SWT) was assessed on a semi-quantitative scale using the following scoring system: 1, normal; 2, mild reduction; 3, severe reduction; and 4, wall thickening absent. Echocardiograms were read at rest and during low and high-dose dobutamine infusion.

Results of DSE were considered normal if all the segments judged normal at baseline showed a hyperdynamic response with increased SWT. The development of new or worsening regional wall motion abnormality during dobutamine stress was considered indicative of ischaemia. Initial improvement in contractility by one grade in an asynergic segment during low-dose dobutamine infusion (5–20 \(\mu\)g) with subsequent worsening at peak dose was considered a biphasic response.\textsuperscript{10}

Each of the segments was assigned to one or three coronary artery distributions; the left anterior descending (LAD) artery distribution included the anterior, anteroseptal and the mid, apical inferoseptal walls. The right coronary artery (RCA) distribution included the inferior, posterior and basal inferoseptal walls. The left circumflex artery (LCX) distribution included the lateral wall. Because of the varying vascular supply, the apex was allocated to any other involved territory. If the apex alone was involved, the LAD was implicated.

**Statistical analysis**

The results of the interpretations by the trainee following 2 and 4 months of training were
compared with the assessment made by the expert on the following criteria:

1. Agreement between the trainee and the expert in the assessment of SWT (normal vs abnormal) was analysed in each coronary territory (LAD, LCX and RCA) on three different aspects.
   (a) Overall scan interpretation.
   (b) Interpretation before and during dobutamine stress.
   (c) Assessment of SWT in each individual myocardial segment.

2. Agreement between the trainee and the expert was also evaluated based on disease severity (SWT grades 1 and 2 vs SWT grades 3 and 4).

3. Finally agreement in the evaluation of viable and ischaemic myocardial segments was assessed.

Interobserver agreement was assessed by calculating the kappa (κ) value which was considered poor when the κ value was <0.4, fair when the κ value was between 0.4 and 0.59, good when the κ value was between 0.6 and 0.79 and excellent when the κ value was >0.8. The difference between the first and second evaluation by the trainee was assessed by McNemar’s test. A p value <0.05 was considered significant. The Bonferroni correction was applied to account for multiple testing.

Results

Agreement in overall scan interpretation (Table 1)

Agreement between the trainee and the expert for the overall scan interpretation (normal vs abnormal) improved significantly in the LAD (from κ = 0.58 at 2 months to κ = 0.73 at 4 months; p = 0.03) territory. Similar improvement was observed in the LCX territory (from κ = 0.51 to κ = 0.61; p = 0.02) but not in the RCA territory.

<table>
<thead>
<tr>
<th>Coronary artery territory</th>
<th>Agreement (κ value)a</th>
<th>First reading</th>
<th>Second reading</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAD</td>
<td>80% (0.58)</td>
<td>87% (0.73)</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Circumflex</td>
<td>84% (0.51)</td>
<td>86% (0.61)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>RCA</td>
<td>70% (0.28)</td>
<td>63% (0.13)</td>
<td>NS</td>
<td></td>
</tr>
</tbody>
</table>

a Agreement is expressed in %, with κ value in brackets.

Agreement in interpretation before and during dobutamine stress (Table 2)

Interobserver agreement in both the LAD was fair at all stages (rest, low dose, high dose) of the test at 2 months but improved significantly to good agreement at 4 months (p = 0.03). Agreement in the left circumflex territory also showed similar improvement from fair at 2 months to good at 4 months. However, agreement in the RCA territory remained poor.

Agreement in interpretation in each individual myocardial segment

Agreement in the interpretation of wall thickening in individual myocardial segments improved from 72% at the first reading (κ = 0.40) to 78% (κ = 0.55) at the second reading (p<0.01).

Agreement in results according to severity of wall thickening abnormality

The concordance in interpretation of normal and mildly abnormal segments (grades 1 and 2), vs severely abnormal segments (grades 3 and 4) improved from 82% (κ = 0.5) at the first reading to 86% (κ = 0.63) at the second reading (p<0.01).

Agreement in the evaluation of viable myocardium

We also analysed myocardial viability in all 51 patients. There were 21 patients with viable myocardium according to the expert reader. The trainee, at first reading identified five patients with viable myocardium (κ = 0.11), while at second reading identified 13 of 21 patients (κ = 0.43) with viable myocardium (p = 0.01).

Agreement in identifying ischaemia

Analysis of data, concerning the existence of ischaemic and non-ischaemic myocardium, revealed nine patients who developed an ischaemic response, majority of which (five patients) was in the RCA territory and 42 patients without an ischaemic response. Of these, the trainee at first reading identified one of nine patients with an ischaemic response and 41 of 42 patients without it. At second reading the trainee did not identify any patients with ischaemia but identified 40 of 42 patients without ischaemia. Although interobserver agreement in the interpretation of a non-ischaemic
response was good (98% at the first reading and 95% at the second reading), there was poor concordance in the identification of an ischaemic response, which was mainly seen in the RCA territory.

Discussion

Most diagnostic techniques in cardiology have been evaluated for interobserver agreement including clinical examination, electrocardiography, resting echocardiography, exercise electrocardiography, radionuclide myocardial perfusion imaging and coronary angiography. There is consensus that appreciation of wall motion abnormalities is one of the most difficult skills to master and teach.

Current recommendations for training of physicians in stress echocardiography

Specialized training in stress echocardiography requires participation in at least 100 supervised stress echo examinations and interpretations (direct ‘hands-on’ supervision of 50 exercise stress and 50 pharmacologic stress studies). This training should be done in a laboratory that performs a minimum of 40 stress echocardiograms per month under the close supervision of a fully qualified expert (>200 stress echocardiography studies) who regularly performs and interprets these procedures.

Previous studies evaluating training of physicians in stress echocardiography

Varga et al. demonstrated that a 2-day joint reading session of 50 stress echocardiograms could result in a significant, albeit clearly suboptimal improvement in diagnostic accuracy and interobserver variability (kappa value improving from 0.14 to 0.39). The usefulness of short-term intensive training and supervised reading of stress echocardiography examinations was further emphasised by Imran et al. who showed that the accuracy of interpretation of stress echocardiograms can be substantially improved, through reinforcement of an overtly conservative reading policy.

Present study

This study evaluates the time required to train a physician, with experience in resting echocardiography, in stress echocardiography. We examined patients who had undergone DSE following MI. These patients are more likely to have a combination of normal and dysynergic segments with variety of responses during dobutamine.

Our results show that a significant improvement in the diagnostic accuracy of beginners in DSE can be achieved in a much shorter time frame (4 months) by systematic training. This strategy improved the accuracy of the trainee in the interpretation of overall stress echocardiography results, assessment of SWT in each myocardial segment and identification of viable myocardium. In our study we found that the diagnostic accuracy of the trainee in identifying abnormalities in the RCA territory was poor. Although his ability to interpret resting wall thickening abnormalities in the RCA territory improved at 4 months of training, interobserver agreement in identifying stress induced abnormalities in this territory remained poor. This explains the discordance between the trainee and the expert in accurately identifying inducible ischaemia, since a majority of the patients in our study (five of nine) had ischaemia in this territory. Previous studies have reported significantly low interobserver agreement in the inferoposterior regions, which correspond, to the RCA territory. This has been attributed to decreased endocardial border definition and less optimal image quality in the above segments.

Sources of interobserver variability in stress echocardiography

In addition to the expertise of the physician, there are several factors influencing interobserver
variability in stress echocardiography which can either be technology-related, stress-related or patient-related. Several groups have previously reported their interobserver variability in the stress echocardiographic interpretation. Hoffmann et al. found only a fair (mean kappa value 0.37) agreement in interpretation of stress echocardiograms among five experienced centres when no unified image display and reading criteria were applied. The factors impacting most on interinstitutional observer agreement in the interpretation of DSE were image quality, severity of induced wall motion abnormalities, and the obtained rate-pressure product. Other studies have reported a high intra-institutional observer agreement in DSE, indicating that implicit agreement on reading criteria could result in improved homogeneity in interpreting test results. Since there are substantial variations in segmental left ventricular wall motion to dobutamine stress even in healthy individuals, the use of standardised reading criteria is highly necessary to reduce interobserver variability.

We have previously shown that tissue harmonic imaging, which improves endocardial border delineation substantially, improved interobserver variability for the assessment of SWT abnormality. With the advent of contrast echocardiography, which has been shown to improve endocardial border delineation over and above tissue harmonic imaging, the interobserver variability is likely to improve further. Indeed, Lindner et al. showed that the confidence of trainees in reporting wall motion abnormalities improved with contrast echocardiography.

Limitations of the study

We did not prospectively assess various time points during training which may have allowed us to identify more definitely the time when the physician reached a plateau in training. Furthermore, only one trainee was assessed. A group of trainees may have defined the requirements much better. Furthermore, interpretation of unselected series of patients with and without resting wall thickening abnormalities would have been more appropriate for a trainee.

Conclusion

Previously it has been shown that a learning curve of greater than 100 stress echo studies read with expert supervision is necessary to optimise the diagnostic yield of the stress echo technique. A high degree of interobserver agreement on the interpretation of stress echo studies is vital to clinical decision-making. The present study reinforces that systematic training in the principles and practice of stress echocardiography over a period of 4 months can significantly improve stress echocardiographic interpretation by a trainee. However, improvements in image quality and use of predefined reading criteria are necessary to improve interobserver agreement further in myocardial regions where conformity in DSE interpretation is low.

Acknowledgements

We thankfully acknowledge the help of Mr Ian Grace, Department of Statistics, Imperial College School of Medicine in the analysis of our study results.

References

10. Senior R, Lahiri A. Enhanced detection of myocardial ischaemia by stress dobutamine echocardiography utilizing
the "biphasic" response of wall thickening during low and high dose dobutamine infusion. J Am Coll Cardiol 1995; 26:26–32.


