Computerized evaluation of echocardiographic stress tests in patients with poorly visualized endocardium using analysis of color-encoded contrast-enhanced images

Victor Mor-Avi, Lissa Sugenga, Robert J. Weiss, Eran Toledo, Lynn Weinert, Tina Bouchard, Kirk T. Spencer, Roberto M. Lang

Section of Cardiology, Noninvasive Cardiac Imaging Laboratory, University of Chicago MC5084, 5841 S. Maryland Avenue, Chicago, IL 60637, USA
Androscoggin Cardiology Associates, Auburn, Maine, USA

Received 26 January 2005; received in revised form 22 March 2005; accepted 9 April 2005
Available online 23 May 2005

KEYWORDS
Ultrasound imaging; Stress testing; Contrast media; Ventricular function

Abstract
Aims: We hypothesized that real-time color encoding of contrast-enhanced images would allow objective detection of stress-induced wall motion abnormalities (WMA).
Methods: We studied 117 patients with poorly visualized endocardium undergoing dobutamine stress tests. Color-encoded images (Philips, color kinesis) were obtained at rest and peak stress in four standard views during i.v. infusion of Definity. Images were reviewed without color overlays by two expert readers, who graded regional wall motion as normal or abnormal. In 101/117 patients (86%), in whom contrast enhancement allowed endocardial tracking, regional fractional area changes were calculated from the color overlays, and thresholds for calling a stress-induced WMA were optimized in a randomly selected subgroup of 34 patients (ROC analysis) to achieve maximum agreement with expert grades. This computerized detection of stress-induced WMA was then tested prospectively in the remaining 67 patients, using the expert grades as a "gold standard".
Results: 20/67 patients had resting WMA and 13/67 patients developed WMA at peak stress. The automated technique detected stress-induced WMA in at least one vascular territory with a sensitivity, specificity and accuracy of 0.80, 0.65 and 0.69,
while the level of agreement between the two experts was 0.62, 0.91 and 0.85, respectively.

Conclusion: Analysis of color-encoded, contrast-enhanced images allows objective, accurate, automated detection of stress-induced WMA in patients with poor acoustic windows.

© 2005 The European Society of Cardiology. Published by Elsevier Ltd. All rights reserved.

Introduction

Echocardiographic evaluation of regional left ventricular (LV) function is based on visual interpretation of dynamic 2D images, which relies on the reader’s ability to integrate spatial and temporal information and is thus subjective and experience-dependent. One of the approaches used to overcome this problem is real-time color encoding of endocardial motion, which allows quantitative analysis of regional wall motion at rest.1,2 We also showed that LV wall motion can be assessed quantitatively from images obtained during pharmacologically induced stress,3 where the conventional visual interpretation is even more difficult and experience-dependent. However, both visual interpretation and color encoding are fundamentally dependent on image quality, and in many patients, limited by suboptimal endocardial visualization. While contrast enhancement is routinely used to improve endocardial visualization, the conventional visual assessment of LV wall motion with contrast remains as subjective as it is without contrast. We and others have recently described the use of color-encoded contrast-enhanced power modulation images for objective assessment of resting LV wall motion in patients with poor acoustic windows who require contrast enhancement for endocardial visualization.4,5 The goal of the present study was to test the feasibility of using this methodology in this challenging patient population in the setting of stress testing.

Accordingly, this protocol was designed to study patients with poor acoustic windows that severely limit endocardial visualization and thus preclude the assessment of wall motion without contrast enhancement. Our specific goal was to determine in a large group of these patients the accuracy of automated detection of wall motion abnormalities from contrast-enhanced images obtained at rest and at peak stress, as well as stress-induced abnormalities. To achieve this goal, we acquired and analyzed color-encoded, contrast-enhanced images in a large group of patients with poor acoustic windows undergoing dobutamine stress testing, and used consensus visual interpretation of two expert readers of conventional dynamic contrast-enhanced images as a “gold standard” for comparison.

Methods

The protocol was approved by the Institutional Review Board. One hundred and seventeen patients were referred for clinically indicated dobutamine stress testing, whose endocardium could not be visualized, in at least two contiguous segments in each apical view (2-, 3- and 4-chamber) without contrast enhancement despite the use of harmonic imaging, were enrolled in this protocol. After obtaining written informed consent, an i.v. line was placed in the brachial vein for infusion of contrast and dobutamine, and the subject was imaged in a left lateral decubitus position.

Stress protocol and image acquisition

After blood pressure reading and 12-lead electrocardiogram were obtained, ultrasound imaging was performed (Philips SONOS 5500) in the parasternal short-axis (SAX) as well as apical 2-, 3- and 4-chamber (A2C, A3C and A4C) views using a 2.5 or 3.5-MHz transducer in the power modulation mode. Contrast enhancement was achieved by continuous infusion of Definity (Bristol-Myers Squibb; 1.3 ml in 50 ml saline; i.v. drip). The rate of infusion was optimized to produce visible LV cavity opacification without significant acoustic shadowing. Contrast enhancement was achieved by continuous infusion of Definity (Bristol-Myers Squibb; 1.3 ml in 50 ml saline; i.v. drip). The rate of infusion was optimized to produce visible LV cavity opacification without significant acoustic shadowing. After optimizing imaging settings for endocardial tracking, color kinesis was activated to color-encode systolic endocardial motion1,2,6 within a region of interest containing the left ventricle. Image settings and regions of interest were digitally saved for each view at baseline, and retrieved to allow repeated imaging of the same view during subsequent phases of the protocol with the same settings. After baseline images were acquired, intravenous dobutamine infusion was initiated at 5 μg/kg/min and then gradually increased to 10, 20, 30 and 40 μg/kg/min. At each dose, blood pressure and 12-lead electrocardiogram...
were monitored. If target heart rate was not reached with the highest dose of dobutamine (40 μg/kg), atropine (0.2–1.0 mg i.v.) was administered. Imaging was repeated at peak dose of dobutamine. In each view and each phase of the protocol, images were acquired during end-expiration and stored digitally for off-line review and computer analysis of the color overlays.

**Conventional interpretation of wall motion**

Contrast-enhanced images were independently reviewed without the color overlays by two experienced echocardiographers (mean of 10 years of experience, 5000 dobutamine stress tests). Wall motion was graded segment-by-segment as normal, abnormal (including hypokinetic, akinetic or dyskinetic) or uninterpretable, at each phase of the protocol using the official guidelines for the segmentation of standard views. Dobutamine-induced wall motion abnormality was defined as a change in segment classification from normal at rest to abnormal at peak stress. Discordant interpretations were then settled in a separate session, in which the two experts reviewed the images jointly. This session resulted in consensus interpretation of LV regional wall motion, which was subsequently used as a reference for comparisons.

**Analysis of color overlays**

Images obtained at rest and at peak stress were reviewed off-line on a personal computer, and end-systolic color overlays were analyzed using previously described custom software. Briefly, images in each view were segmented based on coronary perfusion territories using segmentation schemes (Fig. 1, bottom) that were previously described in detail. The number and size of segments used for the automated analysis of color overlays in each view followed the same conventions used for the visual interpretation of wall motion. From both resting and peak dobutamine color overlays, pixels of each color and pixels marked as blood were counted in each segment. Pixel counts were used to quantify the magnitude of regional endocardial motion in terms of incremental fractional area change (RFAC, in percent of segmental end-diastolic area), and displayed as stacked histograms. The computation time required to analyze one cardiac cycle is <1 s, once the end-systolic frame is selected and the anatomic landmarks necessary for segmentation are identified.

**Objective detection of regional wall motion abnormalities**

Wall motion was evaluated in each segment by comparing the calculated RFAC with a predetermined threshold value, which was initially set arbitrarily for each view at rest and stress. Assuming that RFAC value below the threshold in a segment could be used to classify wall motion in that particular segment as abnormal, these automated classifications were compared with the consensus expert grades. Receiver Operating Characteristic (ROC) analysis was then used, as described below, to optimize the RFAC thresholds for automated detection of wall motion abnormalities. The data were also analyzed on a vascular territory basis, rather than segment-by-segment basis. To this effect, an abnormality in a vascular territory of one of the three major coronary arteries (LAD, LCX and RCA) was called if the number of segments with abnormal wall motion in that territory was larger than a preset number. This threshold number of segments was also optimized for each coronary artery using ROC analysis. To allow automated detection of dobutamine-induced wall motion abnormalities, data obtained in each patient at rest were used as the reference. Stress-induced wall motion abnormalities were automatically detected in individual segments when reduction in RFAC of more than a preset percentage of the baseline value was noted. This percent change threshold was also optimized in each view by means of ROC analysis.

**Receiver operating characteristic analysis**

All patients enrolled in the study were randomly divided into two groups. First, all threshold values were progressively changed one-by-one in a Study Group consisting of one third of the patients, and the agreement levels with the consensus expert interpretation were calculated for each set of thresholds. Then the set of thresholds that resulted in the highest levels of agreement was applied prospectively to a larger Test Group, consisting of the remaining two thirds of the patients. Agreement levels with the reference interpretation were calculated in the Test Group.

**Statistical analysis**

In each patient, the level of agreement between the quantitative technique and the consensus
Figure 1  Top: resting power modulation images obtained in a patient with poorly visualized endocardium in the four standard views. Bottom: segmentation schemes used for both visual interpretation of the gray-scale power modulation images and quantitative segmental analysis of color overlays containing information on systolic endocardial motion in each view (SAX — parasternal short-axis; AP2 — apical 2-chamber; AP3 — apical 3-chamber; AP4 — apical 4-chamber).
expert interpretation was assessed by counting concordant grades (true positive and true negative) as well as discordant grades (false positive and false negative) assigned by the automated technique at rest and during peak stress, as well as for detecting stress-induced abnormalities. Segments classified by the experts as uninterpretable were not counted as either concordant or discordant. Initially, the counts of concordant and discordant grades obtained in the Study Group were used to calculate sensitivity, specificity, positive and negative predictive values (PPV and NPV) and overall accuracy, which were used for optimization of thresholds by the ROC analysis. Subsequently, we also calculated the sensitivity, specificity, PPV, NPV and accuracy obtained in the Test Group with these optimized thresholds. To put these values in perspective, we also calculated the sensitivity, specificity, PPV, NPV and accuracy in the Test Group from the concordant and discordant grades between the two readers for the abnormalities noted at rest and during peak stress, as well as for the detection of stress-induced abnormalities.

Results

Fig. 1 (top) shows an example of four standard views of the left ventricle obtained in a patient whose poor endocardial definition precluded the interpretation of wall motion without contrast enhancement. Fig. 1 (bottom) shows the corresponding segmentation schemes used for both the visual interpretation of regional wall motion and the automated analysis of color overlays. Of the 117 patients, 15 patients (13%) were excluded from analysis because the improvement in endocardial visualization with contrast enhancement was not sufficient to allow adequate endocardial tracking and color encoding. One patient was excluded because of incomplete image acquisition. The remaining 101 patients were randomly divided into the Study Group ($N = 34$) and the Test Group ($N = 67$).

Study Group

Expert interpretation noted abnormal wall motion in 79/816 (9.7%) segments at rest and in 90/816 (11%) at peak stress. While according to the inclusion criteria, at least 25% segments (6 out of 24) had to be uninterpretable without contrast, with contrast enhancement, the number of uninterpretable segments was only 39/816 (4.8%) at rest and 43/816 (5.3%) at peak stress. ROC analysis resulted in maximal agreement levels with the consensus expert interpretation with RFAC thresholds of 59% in the SAX view and 52% in the apical views at rest, and 69% in the SAX view and 66% in the apical views at peak stress.

Fig. 2 shows an example of color-encoded images obtained at rest and peak stress in a subject with normal wall motion and normal response to dobutamine. Fig. 3 shows the corresponding stacked histograms, wherein RFAC in all segments is above threshold values both at rest and stress. In contrast, Figs. 4 and 5 show in the same format data obtained in a patient with wall motion abnormalities both at rest and peak stress, which are reflected by sub-threshold RFAC values in multiple segments.

ROC analysis also showed that the number of segments with abnormal wall motion had to be at least 2 in each, LCX and RCA, territory for that territory to be called abnormal. In the LAD territory, this number had to be at least 4, to reach maximum agreement with the expert grades. With these threshold values, at rest, the quantitative analysis resulted in agreement with the reference method in 614 of the 777 interpretable segments, and yielded accuracy of 0.79 (Table 1, line 1). At peak stress, the rate of agreement was 669 of the 773 interpretable segments, resulting in accuracy of 0.87 (Table 1, line 2).

In this group, 11/34 (32%) patients had resting abnormalities in at least one coronary perfusion territory according to the expert consensus interpretation, and 18/34 showed resting abnormalities by quantitative analysis with accuracy of 0.62 (Table 1, line 3). At peak stress, abnormal wall motion was noted in at least one territory in 10/34 patients according to expert interpretation and in 11/34 patients by quantitative analysis with accuracy of 0.91 (Table 1, line 4). Expert interpretation determined that 8/34 (24%) patients showed stress-induced abnormalities in at least one vascular territory. ROC analysis resulted in minimum segmental percent change of 34% in the SAX view and 30% in the apical views for optimal automated detection of induced abnormalities. With these values, the quantitative analysis showed induced abnormalities in 14/34 (41%) patients with accuracy of 0.76 (Table 1, line 5).

Test Group

The expert consensus graded wall motion as abnormal in 84/1608 (5.2%) segments at rest
and in 85/1608 (5.3%) at stress. With contrast enhancement, the number of uninterpretable segments was 83/1608 (5.2%) at rest and 85/1608 (5.3%) at peak stress. At rest, the quantitative analysis based on the use of ROC-optimized threshold values, resulted in agreement with the reference values in 1275 of the 1525 interpretable segments, and yielded accuracy of 0.84 (Table 1, line 6). In the same patients, the two experts agreed between them in 1533/1608 (95%) segments.

Figure 2 Color-encoded, contrast-enhanced power modulation images obtained at rest (left) and at peak dobutamine stress (right) in a Test Group patient who had normal resting wall motion and showed normal response to dobutamine. Note the increase in the thickness of the color band and the increased relative contribution of the "early" colors (orange and yellow), reflecting more vigorous early systolic contraction at peak stress.
with accuracy of 0.95 (sensitivity 0.69, specificity 0.96, PPV 0.54, NPV 0.98). At peak stress, the rate of agreement between the quantitative analysis and the expert consensus was 1301 of the 1523 interpretable segments, resulting in accuracy of 0.85 (Table 1, line 7). In comparison, the level of agreement between the two experts was: 1534/1608 (95%) with accuracy of 0.95 (sensitivity 0.59, specificity 0.96, PPV 0.42, NPV 0.98).

In this group, 20/67 (30%) patients had wall motion abnormalities in at least one coronary perfusion territory according to the expert consensus interpretation both at rest and peak stress. The quantitative analysis showed resting abnormalities in at least one perfusion territory at rest in 29/67 (43%) patients with accuracy of 0.78 (Table 1, line 8), and at peak stress in 26/67 (39%) patients with accuracy of 0.76 (Table 1, line 9). Both experts

**Figure 3** Stacked color histograms of regional fractional area change (RFAC) in % of regional end-diastolic area (REDA) obtained from images shown in Fig. 2. The threshold RFAC values for automated detection of wall motion abnormalities in each view are shown as black horizontal lines across each histogram. Note that in all views, in all segments, the RFAC values reached at end-systole are above the corresponding threshold, reflecting normal wall motion.
detected stress-induced abnormalities in at least one vascular territory in an equal number of patients (13/67 or 19%), but their interpretations were discordant in 10/67 additional patients, resulting in accuracy 0.85 (sensitivity 0.62, specificity 0.91, PPV 0.62, NPV 0.91). Joint review of these cases pointed at suboptimal image quality as the main cause of discrepancy. By consensus, 15/67 (22%) patients had stress-induced abnormalities. The quantitative analysis showed induced abnormalities in 30/67 (45%) patients, with accuracy of 0.69 (Table 1, line 10).

Figure 4 Color-encoded, contrast-enhanced power modulation images obtained at rest (left) and at peak dobutamine stress (right) in a Test Group patient who had wall motion abnormalities in multiple segments, both at rest and peak stress, reflected by relatively thin color bands in hypokinetic segments.
The subjective nature of the conventional visual interpretation of echocardiographic stress tests is a known subject of concern. In particular, in patients with poor acoustic windows, the diagnostic accuracy of the interpretation not only relies on the reader’s experience and ability to effectively integrate dynamic information on regional endocardial motion and myocardial systolic thickening, but is also affected by the limited views of the underlying anatomic structures and their motion. The ultimate challenge in this regard is the interpretation of images obtained in these patients during stress testing, which can be difficult even in patients with optimal images because of the elevated heart rate and other stress-induced changes in ventricular function. Accordingly, the need for new methods for objective evaluation of LV wall motion suitable for this difficult patient

**Figure 5** Stacked color histograms of regional fractional area change (RFAC) in % of regional end-diastolic area (REDA) obtained from images shown in Fig. 4 with threshold RFAC values shown in the same format as in Fig. 3. Note that in multiple segments, the RFAC values that reached end-systole are below the corresponding threshold, reflecting abnormal wall motion.

**Discussion**

The subjective nature of the conventional visual interpretation of echocardiographic stress tests is a known subject of concern. In particular, in patients with poor acoustic windows, the diagnostic accuracy of the interpretation not only relies on the reader’s experience and ability to effectively integrate dynamic information on regional endocardial motion and myocardial systolic
population is obvious. Over the past decade, improvements in the diagnostic accuracy of dobutamine stress echocardiography in patients with poor image quality with the use of contrast enhancement were reported by several investigators.\textsuperscript{11–19}

We hypothesized that: (i) the use of power modulation imaging with real-time color encoding of endocardial motion during steady-state contrast enhancement could be successfully extended to stress testing in these difficult patients, and (ii) quantitative segmental analysis of these images might allow accurate, automated, objective and experience-independent assessment of LV wall motion. To test these hypotheses, we enrolled only patients who fell into this challenging category and tested our methodology under these unfavorable conditions.

The automated detection of wall motion abnormalities relies on pre-existing threshold values, which needed to be determined first. This goal could be achieved by establishing normal values in a group of subjects with no wall motion abnormalities either at rest or at peak stress, or by optimizing the existing degrees of freedom using the ROC curves analysis in a subgroup of randomly selected patients and then applying the optimized algorithm prospectively to the remaining patients. Although both approaches are legitimate, they have their advantages and disadvantages. While the former approach would theoretically allow us to establish a normal value for individual segments, such "normal values" could be misleading in the absence of corroborative coronary angiography data. Also, this approach would require isolating a relatively large number of patients with normal wall motion, resulting in a "referral bias" in the remaining patients and thus confounding our results. Accordingly, we chose the alternative ROC approach, limited to a smaller number of degrees of freedom to optimize, i.e. one threshold value per view, but had the advantage of no referral bias. This study design allowed us to test the feasibility, optimize the existing degrees of freedom, and subsequently establish the accuracy of objective, automated detection of wall motion abnormalities from contrast-enhanced, color-encoded power modulation images obtained at rest and at peak dobutamine stress, as well as stress-induced abnormalities, in patients with poor acoustic windows.

First, our results confirmed in a large group of patients with poor acoustic windows the feasibility of automated detection of resting wall motion abnormalities with accuracy levels similar to that previously described in our initial study targeting this patient population.\textsuperscript{4} However, in the present study, the analysis was performed both on a segment-by-segment basis as well as on a vascular territory basis, assuming that the latter would be more clinically relevant and also might result in higher accuracy levels, as physiologically unlikely positive findings in single segments would be automatically discarded. This was reflected by a two- to three-fold increase in PPV for the territory-by-territory analysis at the expense of a small drop in NPV in both groups of patients, despite the slight decrease in sensitivity, specificity and accuracy values. Importantly, these values were high enough to endorse the use of this methodology in patients whose images were otherwise uninterpretable. Of note, these values were clinically acceptable when compared to those calculated between the two expert readers, despite their considerable expertise.

In addition to resting images, our results demonstrated the feasibility of automated detection of wall motion abnormalities from images obtained at
peak dobutamine stress, despite the confounding effects of increased heart rate and cardiac translation on image quality and color-encoding. Surprisingly, in the Study Group, the accuracy of the automated detection at peak stress was higher than at rest, when assessed either by segment or by vascular territory. This can probably be explained by the relatively small size of this group. In the larger Test Group, the accuracy levels achieved at rest and peak stress were virtually identical, on both segment-by-segment and territory-by-territory basis. This finding supports the conclusion that color-encoded, contrast-enhanced images obtained at peak stress in patients with poor acoustic windows can be used for automated detection of wall motion abnormalities, with levels of accuracy not lesser than resting abnormalities.

We expected the detection of stress-induced abnormalities, which is the main goal of stress testing, to be less accurate, since it relies on accurate detection of both resting and peak stress abnormalities, and is thus subject to compounding errors inherent to these two steps. Indeed, the accuracy of the automated detection of stress-induced abnormalities on a segment-by-segment basis was low. Nevertheless, despite the compounding errors, the results of our analyses of individual vascular territories showed that stress-induced abnormalities can be automatically detected with reasonable level of accuracy, not unlike some of the previous reports on the accuracy of subjective assessment of wall motion during stress testing. Importantly, the sensitivity of this detection was 0.80 and above in both groups of patients.

One might argue that detection of wall motion abnormalities from images of poor quality is a poor standard for the diagnosis of coronary artery disease, and that other techniques such as coronary angiography would be far superior. However, the goal of our study was not to test the feasibility of automated diagnosis of coronary artery disease, but to test the hypothesis that wall motion abnormalities identified by expert readers in consecutive patients with poor acoustic windows undergoing dobutamine stress testing could be detected objectively using automated computer analysis. Coronary angiography data could not serve as a reference to test this hypothesis. Theoretically, we could have studied only patients who had angiography data available, but such a study design would not have allowed us to determine the success rates of this new methodology in consecutive patients with poor image quality.

Undoubtedly, the main limitation of this methodology is its dependence on image quality, which directly determines the quality of endocardial tracking and color encoding. Nevertheless, it is important to remember that our patients must have had at least 25% uninterpretable segments without contrast enhancement to be included in this study. With contrast enhancement, the number of uninterpretable segments dropped to only 5% in both groups, allowing color encoding of endocardial motion suitable for analysis and accurate detection of wall motion abnormalities.

In summary, analysis of color-encoded, contrast-enhanced power modulation images obtained in patients with poor acoustic windows allows automated detection of resting, peak stress and stress-induced wall motion abnormalities. In addition to being objective and experience-independent, this technique was found to have similar accuracy to that of expert interpretation of wall motion from conventional contrast-enhanced images. Therefore, this simple technique may aid in the interpretation of pharmacological stress tests in this challenging patient population.

References


