Coronary Artery Disease

Variability of Myocardial Perfusion Defects Assessed by Thallium-201 Scintigraphy in Patients With Coronary Artery Disease Not Amenable to Angioplasty or Bypass Surgery

Daniel Burkhoff, MD, PhD,* James W. Jones, MD, PhD,† Lewis C. Becker, MD, FACC‡
New York City, New York; Columbia, Missouri; and Baltimore, Maryland

OBJECTIVES
We sought to assess the variability of results obtained with thallium scintigraphy as a method for tracking the extent of myocardial ischemia in medically refractory patients with angina who are not suitable for coronary artery bypass graft surgery or percutaneous transluminal coronary angioplasty.

BACKGROUND
New therapies are being evaluated for patients with “no option” angina in whom medical therapy has failed. Nuclear techniques, like thallium scintigraphy, are used in multicenter trials to evaluate whether such therapies improve myocardial perfusion. However, the variability of test results is unknown in this patient group in a multicenter study.

METHODS
The Angina Treatments: Lasers And Normal Therapies In Comparison (ATLANTIC) study was a randomized trial of transmyocardial laser revascularization (n = 182). Patients underwent dipyridamole thallium stress tests at baseline and 3, 6 and 12 months after enrollment. The control group (n = 90) was treated with constant medical therapy during the study and is a relevant group to investigate test variability. Test variability over time was quantified by the mean absolute change in the percentage of reversible perfusion defects between baseline and follow-up.

RESULTS
Baseline percent myocardium with ischemia averaged 17.0 ± 13.7% and did not change during follow-up. However, variations in the percent myocardium with reversible perfusion defects over time amounted to an average of 6 to 8 percentage points, or 43% to 55% of the baseline value. Only ~13% of this variability was attributable to variability in image reconstruction and analysis.

CONCLUSIONS
As demonstrated in the ATLANTIC study, percent myocardial ischemia in control subjects receiving constant medical therapy varied in individual patients by an average of ~50%. This may limit the utility of thallium scintigraphy to detect improved myocardial perfusion over time in response to therapy. (J Am Coll Cardiol 2001;38:1033–9) © 2001 by the American College of Cardiology

Angiogenic growth factor therapies (1) and transmyocardial revascularization (TMR) (2–5) are being investigated for treating patients with otherwise untreatable, medically refractory angina (“no option” patients). The goal of these therapies is to relieve symptoms by improving blood flow to ischemic regions through induction of new blood vessel growth. Although symptom relief and improved quality of life and exercise tolerance are important therapeutic goals, investigators have sought objective evidence of improved myocardial perfusion, as seen in animal models of chronic ischemia (6), using nuclear techniques such as thallium scintigraphy (7–9). However, the variability of myocardial perfusion defects assessed by such techniques in the target patient population has not been determined within the context of a multicenter study (10).

The Angina Treatments: Lasers And Normal Therapy In Comparison (ATLANTIC) study (3) was a multicenter, prospective, randomized study comparing TMR plus continued medical therapy with continued medical therapy alone in otherwise untreatable patients with refractory angina. Dipyridamole thallium stress tests performed at baseline and 3, 6 and 12 months after study entry showed no significant change in either group from baseline at any time point in the percent myocardium exhibiting reversible (ischemic) defects. However, the standard deviation of the mean percentage point change of ischemic myocardium in both groups was relatively large, almost equaling the mean baseline value of this variable. When focusing on the control group, whose results were expected to be reasonably consistent from one test to the next, this observation suggests that test results may vary significantly in individual patients over the study period. This raises questions as to the consistency of the amount of ischemia present in these patients and the variability inherent in the dipyridamole thallium stress

From the *Department of Medicine, Columbia University, New York City, New York; †Department of Surgery, University of Missouri, Columbia, Missouri; and the ‡Department of Medicine, The Johns Hopkins University, Baltimore, Maryland. This study was supported by Eclipse Surgical Technologies, Inc., Sunnyvale, California. None of the authors have any financial relationship with Eclipse Surgical Technologies.

Manuscript received January 17, 2001; revised manuscript received May 31, 2001, accepted June 19, 2001.
testing over time, which has never been assessed in a quantitative fashion in the context of a multicenter study over a relatively long time frame. Previous efforts to examine the variability of thallium scintigraphic results are limited to evaluations performed at single centers with small numbers of patients over short follow-up periods (11–14). Furthermore, these studies were performed in patients with routine coronary artery disease (i.e., patients encountered more commonly in daily practice with fewer and more discrete coronary lesions), so that the results may not apply to the “no option” patients being targeted for angiogenic therapies.

To address these questions, we assessed the variability of test results over time in the control subjects of the ATLANTIC study. This is an appropriate population in which to address the question of test result consistency, because these patients are fully representative of those being targeted for angiogenic therapies, and because the patients were followed for one year without any other interventions or significant changes in medical therapy.

METHODS

The ATLANTIC study. The ATLANTIC study inclusion and exclusion criteria and results have been presented previously (3). Briefly, this was a 16-site, prospective, randomized study that enrolled 182 subjects who provided written, informed consent, with medically refractory class III or IV angina deemed untreatable by coronary artery bypass graft surgery (CABG) or percutaneous transluminal coronary angioplasty (PTCA) and with an ejection fraction ≥30%. Ninety-two subjects underwent TMR plus continued maximal antianginal therapy (TMR + MEDs), and 90 subjects received just continued medical therapy (MEDs only). Subjects were followed for one year (no cross over), and they continued treatment with the maximal antianginal medical regimen. Subjects receiving TMR + MEDs demonstrated significant improvements in angina, exercise tolerance and quality of life, whereas subjects treated with MEDs only showed little change in any of these variables.

Thallium scintigraphy. An experienced core laboratory was used to standardize the procedures. Core laboratory representatives visited the sites to review the protocol and standardize acquisition variables. The collimator was set at low energy and high resolution. The matrix size was set at 64 × 64 pixels; the zoom was set at 1:1.5; the number of angles was 32 over a 180° range; the dwell time was set at 60 s; and the center of rotation and uniformity corrections were applied during acquisition without attenuation correction. Drinks containing caffeine, sublingual and oral nitroglycerin and aminophylline were all withheld on the morning of the study. Repeat studies were performed at the same time of day and on the same scanner as baseline studies. A standard dipyridamole dose (0.14 mg/kg body weight per min for 4 min, 0.056 mg/kg) and a standard thallium dose (2.5 to 3.0 mCi initially, with an additional 1 mCi for re-injection) were used. Raw data were sent to the core laboratory for image reconstruction and analysis. Image reconstruction was accomplished using SMV commercial software (GE/SMV America, Twinsburg, Ohio), which included correction for patient motion, Wiener filtering of the projection data and filtered backprojection using a ramp filter. The images were re-oriented for short-axis alignment. The Stanford University bull’s-eye protocol was used to create polar maps and quantitative analysis of activity distribution (15,16).

The maps were compared with gender-specific normal maps derived from stress/redistribution thallium studies in 28 men and 40 women. Two types of polar map quantitative analyses were performed. In the first analysis, pixel intensities >2 SD below the normal mean value were considered as abnormal. The image was analyzed pixel by pixel to determine the percentage of pixels with a reversible perfusion defect (normal flow at rest and abnormal flow during stress, ischemic), a fixed defect (abnormal flow at rest and stress, infarct) or normal flow.

In the second analysis, an ischemic score was calculated by weighting the abnormal pixel count by the severity of the perfusion defect on the stress image. Each pixel was assigned a grade as follows: 0 when the pixel intensity was within 2 SD of normal; 0.5 when the pixel intensity was ≥2 and <3 SD below normal; 0.75 when the pixel intensity was ≥3 and <4 SDs below normal; and 1.0 when the pixel intensity was ≥4 SD below normal. The total ischemic score was the sum of the individual pixel scores; ischemic and infarct-related pixels (based on the previous definition) were summed separately.

The reproducibility of core laboratory procedures was assessed by complete reprocessing of 15 randomly chosen studies, one from all but one of the participating centers. These scans were sent from the sites on four occasions distributed over a one-year period to the core laboratory under different identification numbers and were analyzed along with all other scans without previous knowledge of the core laboratory.

Thallium scan results: variability analysis. The main analysis performed to assess test variability is illustrated in Figure 1. In this example, the changes in the percent myocardium with a reversible perfusion defect (which we will define as “ischemia” in the present study) from baseline

### Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>ATLANTIC</td>
<td>Angina Treatments: Lasers And Normal Therapies In Comparison study</td>
</tr>
<tr>
<td>CABG</td>
<td>coronary artery bypass graft surgery</td>
</tr>
<tr>
<td>MEDs</td>
<td>continued maximal antianginal therapy</td>
</tr>
<tr>
<td>PTCA</td>
<td>percutaneous transluminal coronary angioplasty</td>
</tr>
<tr>
<td>SAQ</td>
<td>Seattle Angina Questionnaire</td>
</tr>
<tr>
<td>TMR</td>
<td>transmyocardial revascularization</td>
</tr>
<tr>
<td>VI</td>
<td>variability index</td>
</tr>
</tbody>
</table>
were −10, +10 and +5 percentage points at the three follow-up studies, so that the average change was only −1.7 percentage points, <10% of the baseline value of 25%. However, the variation from test to test is large, so taking a simple mean value obscures random fluctuations in the results. Accordingly, it is more appropriate to quantify variability based on the mean absolute change from baseline, as suggested in a previous study (13).

\[
VI = \left( \frac{1}{3} \sum_{i=1}^{3} |I_i - I| \right) / 3
\]

where VI is the variability index; the subscript B designates the baseline study; \(i\) indicates from 1 to 3 (corresponding to the three follow-up studies); and I is the variable being used to quantify the extent of the myocardial perfusion defect. Variables that will be discussed include percent myocardium with a reversible perfusion defect, percent myocardium with a fixed perfusion defect and the total ischemic score. For example, in Figure 1, the mean VI equals −8 percentage points \((10 + 10 + 5)/3\), which amounts to 32% (i.e., 8/25) of the baseline value. Values of VI were determined for all subjects and were expressed in absolute percentage points.

All data are expressed as the mean value ± SD. Comparisons between results obtained at baseline and follow-up studies were compared with repeated measures analysis of variance (ANOVA). Specific paired observations were compared with the paired \(t\) test. A p value <0.05 was considered statistically significant.

RESULTS

Patient group. The baseline characteristics of the control group (MEDs only) of the ATLANTIC study are summarized in Table 1. The subjects were predominately men in their mid-60s, with a high prevalence of cardiovascular risk factors; most had a history of myocardial infarction and had undergone at least one previous CABG or PTCA, or both. Ejection fraction was well preserved, but exercise tolerance was poor. The results from the Seattle Angina Questionnaire (SAQ) indicated a poor quality of life (each SAQ index ranges from 0 to 100, with 100 indicating the best quality of life) (17).

During the one-year follow-up period, there were nine deaths in the MEDs-only group (10% mortality rate): seven deaths due to cardiac causes and two of unknown causes. Seven subjects withdrew from the study, and there were several missed visits, which lowered the number of available studies. If all available data are included, baseline and three-month scans were available from 60 subjects, baseline and six-month scans from 64 subjects and baseline and 12-month scans from 56 subjects. Scans were available from 47 subjects at all four time points.

It is important to point out that detailed analyses showed that in subjects receiving MEDs only who completed the study, there were no new Q wave myocardial infarctions, and cardiovascular medication use did not change significantly during the follow-up period (3).

Thallium stress test variability. A scattergram showing the percent myocardial ischemia at baseline versus the percent ischemia at three months (n = 60) is shown in Figure 2A. Although there is a statistically significant linear relationship (solid line, \(y = 0.7x + 5.3; p < 0.001\)), the data are scattered widely around both the regression line and the line of identity (dashed line); both the correlation coefficient \((r^2 = 0.43)\) and concordance correlation coefficient \((\rho_c = 0.70)\) (14) were relatively low, indicating a high degree of variation from one test to the other. These same data,
plotted to show the change in percent ischemia (three-month minus baseline) as a function of baseline percent myocardial ischemia for the same subjects. The points scattered approximately equally above and below the horizontal line of 0 change with large variability. The dashed line shows the maximal possible reduction in ischemia for a given baseline value. See text for further details.

For the 60 patients who had scans at baseline and at three months of follow-up, the mean percent myocardial ischemia at baseline was 17 ± 16% (Fig. 3). The mean (±SD) percent ischemia on the follow-up scans was nearly the same, indicating that the mean change was ~0 percentage points. However, if the mean absolute change is analyzed instead of the mean change in percent ischemia, there is a variation of ~9 percentage points in the test results (Fig. 3, cross-hatched bars). This degree of variability, which amounts to ~50% of the baseline amount of ischemia, was similar for the subset of 64 patients having scans at baseline and six months, as well as for the 56 patients having scans at baseline and 12 months (Fig. 3). Repeated measures ANOVA showed that there was no change in the mean percent myocardial ischemia over time (p = NS). Furthermore, the interclass correlation coefficient determined from this ANOVA was 0.75, which confirms a relatively weak correlation between the serial test results. Similarly, the value of VI between follow-up visits did not vary significantly (p = NS).

For illustrative purposes, the results from individual subjects shown in Figure 4 graphically illustrate the variations associated with different VI values; percent ischemia and infarction are both shown. A subject having a VI similar to the mean value for the group (9.5) is shown in Figure 4A; a subject with a smaller than average VI (5.0) is shown in Figure 4B; and a subject with a worse than average VI (20.0) is shown in Figure 4C. Dramatic changes—both improvements and worsening—are illustrated in these examples.

It is also evident in these examples that the percent myocardium with a fixed perfusion deficit also varies significantly over time. For the 47 subjects with data at all time points, the percent myocardium with a fixed perfusion defect ("infarction") at baseline was 14.1 ± 10.8%, with an overall variability index of 6.9 ± 4.8 percentage points, or ~50% of the baseline value. Furthermore, if fixed and reversible defects are considered (i.e., added) together in this same group of subjects, the mean percent myocardium with abnormal perfusion is 29.7 ± 16.9% at baseline. This variable also did not vary significantly at any follow-up time point (p = NS by repeated measures ANOVA), and the
The overall variability index was 9.4 ± 4.7 percentage points, which amounted to ~30% of the baseline value.

The total ischemia score (see Methods) was also subjected to a similar analysis. The baseline value of this unitless variable was 390 ± 266, and the respective overall variability index derived from the 47 subjects having data at all time points was 154 ± 92, ~39% of the baseline value.

Thus, the large variability in the results was not limited to the quantification of percent ischemia, per se, but was also apparent in several other indexes of myocardial ischemia.

Reproducibility of image reconstruction and analysis. Fifteen scans selected at random were sent to the core laboratory under different identification numbers and were analyzed four times without the core laboratory’s knowledge that they were duplicates. The results of this analysis showed that, with few exceptions, the results obtained on subsequent repeat analysis were very similar to those of the original analysis. The regression line between the first analysis and the three subsequent analyses was indistinguishable from the line of identity (y = 1.0x – 0.9, r² = 0.94, p < 0.0001). The variability index (refer to the equation in the Methods) amounted to only 2.9 ± 2.0 percentage points. Thus, the variability of image reconstruction contributed to the test’s variability, but this was a minor component.

Stratification of subjects by the amount of ischemia at baseline. To test whether variability is less in subjects with more ischemia, data from the 56 subjects in the MEDs-only group with test results at baseline and 12 months were stratified according to the percent ischemia at baseline and classified into three groups of an approximately equal number of subjects, as shown in Figure 5A. There was a slight increase in the overall VI from ~5 percentage points in the group with <7% baseline ischemia to a maximum of ~11 points in the group with >25% baseline ischemia. However, because VI (absolute percentage points) increased far less than the mean value of baseline ischemia in the different groups, the percent variability decreased substantially as the amount of baseline ischemia increased. This suggests that the test could be more reliable in detecting changes in perfusion in subjects with a greater amount of baseline ischemia.

To test this hypothesis, data from subjects treated with TMR in ATLANTIC (TMR + MEDs) were studied. The baseline characteristics of the 92 original TMR + MEDs subjects were similar to those of the MEDs-only group (3). After accounting for mortality, subject withdrawal and missed visits, 58 of these subjects had scans at both baseline and 12 months. If data from these subjects were stratified according to the same criteria as described earlier, the subjects with the greatest amount of ischemia at baseline had a substantial and highly statistically significant decrease in percent ischemia 12 months after TMR (Fig. 5B). Patients in the middle group showed no change, and patients with little baseline ischemia showed a small increase in the perfusion defect. These findings initially suggested that in the subgroup with the greatest amount of baseline ischemia, TMR might have resulted in significantly improved perfusion. However, data from the 56 subjects in the MEDs-only group with scans at both baseline and 12 months analyzed in this same way (Fig. 5C) revealed the same trends (even with regard to statistical significance). This behavior is an example of the statistical phenomenon of “regression to the mean” when random noise is superimposed on baseline and follow-up “true” test values.
DISCUSSION

Our analysis indicates that, as implemented in the ATLANTIC study, the results of dipyridamole thallium stress testing yielded quantitative assessments of myocardial ischemia that were variable from one test to another. The magnitude of the variability amounted to approximately half of the mean baseline value of ischemia. Such variability was observed with several different variables used to quantify the ischemic burden. Stratification of subjects according to the degree of baseline ischemia did not improve the situation. Rather, regression to the mean value on follow-up studies was observed—that is, scans with low amounts of ischemia at baseline increased during follow-up, whereas scans with high amounts of ischemia at baseline decreased.

Comparison with previous studies. Variability of nuclear stress test results has been examined previously in a few single-center studies in small numbers of patients (11–13). In one of these studies (13), 16 stable subjects studied prospectively (one month between studies) and data from 23 subjects examined retrospectively (up to 13 months between studies) demonstrated a low mean absolute change of 4.5% (identical to the VI used in the present study) and a high concordance correlation coefficient of 0.94 (a statistical index of the tests’ similarity) (14) between percent ischemia determined from polar map analysis on baseline and follow-up studies; values of both indexes are substantially better than achieved in the present study. The investigators concluded that in stable patients, the percent hypoperfused myocardium is reproducible over the course of a year. Conclusions provided by other studies have been similar (11,12).

The present study is different from previous studies in that it is the first to test variability in a prospective, multicenter trial, analyzed in a completely blinded manner, from a significantly large number of subjects, over a one-year period, with four tests being performed over the follow-up period. Unlike earlier studies, the subjects in this study are fully representative of those being targeted for angiogenic therapy (Table 1), with medically refractory angina, angiographically documented, severe coronary artery disease that cannot be revascularized and a high incidence of previous myocardial infarction.

Potential sources of variability. Many potential sources of variability could contribute to the present findings. The conditions of the patients could have changed significantly over time. Subtle changes in medications could have occurred, which were not detected during follow-up. To standardize the degree of intersubject and intrasubject stress, the protocol indicated that the tests be performed at the same time of day, on the same scanner; that certain medications be withheld before the study; and that standard doses of dipyridamole and thallium be administered. However, deviations from these protocols could have occurred. Records documenting the physiologic response to dipyridamole (i.e., changes in heart rate and systemic blood pressure) would have been helpful, but were not reported to the core laboratory. Intrinsic variability in test results could

![Figure 5](image-url)

**Figure 5.** (A) Results (mean ± SD) stratified according to percent baseline ischemia. Test variability, indexed by variability index (VI) (refer to the equation in Methods) increases as baseline ischemia increases, so that VI as a proportion of baseline ischemia decreases. (B) Baseline and 12-month follow-up thallium results from subjects who had transmyocardial laser revascularization (TMR) in the ATLANTIC study, stratified according to percent baseline ischemia. Although the data in B suggest TMR may improve myocardial perfusion in subjects with a greater amount of baseline ischemia, data from the ATLANTIC subjects in the continued maximal anti-anginal therapy (MEDs)-only group (C) exhibit the same trends. This suggests a “regression to the mean” phenomenon, as discussed in the text. The p values were derived from paired t tests.
also be a factor. The data collected within the context of the ATLANTIC study were not sufficient to define the major source(s) of test result variation. The contribution of variations in the techniques used to reconstruct scans at the core laboratory were investigated and shown to contribute only a small amount to overall test variability, however. It is possible that more careful attention to technical details, which is difficult in practice within the context of large-scale, multicenter studies, could reduce test result variability, rendering the test more useful in detecting blood flow changes. Future studies could also include pilot investigations to determine the specific degree of variability at the centers involved in order to make more appropriate calculations of sample size.

Conclusions. Irrespective of whether the observed variability reflects changes in the patients’ status over the follow-up period or technical limitations related to how dipyridamole thallium was implemented in the ATLANTIC study, the present findings provide important and practical information for future studies. Several therapies are under development for “no option” patients with medically refractory angina. The present results suggest that if nuclear scintigraphy is employed in this patient population to quantify changes in myocardial perfusion in multicenter studies, it is important to ensure adequate test reproducibility by demonstrating sufficiently low variability and by including enough subjects in the statistical analysis. These results should be considered to apply to the “no option” group of patients with severe coronary artery disease evaluated in the ATLANTIC study, and not necessarily to the patients encountered more commonly in daily practice with fewer and more discrete coronary lesions. The results (Figs. 4A and C) further indicate that anecdotal reports of dramatic improvements in myocardial perfusion after TMR or angiogenic therapy should be avoided. Unless adequate test variability is documented and is sufficiently low, similar dramatic improvements can be observed in untreated patients over the course of a one-year follow-up period.

Acknowledgments
Please see the original reference (3) for a list of the ATLANTIC investigators who contributed to this study.

References

Reprint requests and correspondence: Dr. Daniel Burkhoff, Columbia University, Black Building 812, 650 W. 168th Street, New York, New York 10032. E-mail: db59@columbia.edu.