Comparison of Fully Automated Computer Analysis and Visual Scoring for Detection of Coronary Artery Disease from Myocardial Perfusion SPECT in a Large Population

Reza Arsanjani[1], **Yuan Xu**[1], **Sean W. Hayes**[1], **Mathews Fish**[2], **Mark Lemley Jr**[2], **James Gerlach**[1], **Sharmila Dorbala**[3], **Daniel S. Berman**[1,4], **Guido Germano**[1,4], and **Piotr Slomka**[1,4]

[1]Departments of Imaging and Medicine, and Cedars-Sinai Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA

[2]Oregon Heart and Vascular Institute, Sacred Heart Medical Center, Springfield, OR

[3]Department of Radiology, Division of Cardiology, Noninvasive cardiovascular imaging section, Brigham and Women's Hospital, Boston, MA

[4]David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA

Abstract

We compared the performance of a fully automated quantification of attenuation-corrected (AC) and non-corrected (NC) myocardial perfusion single photon emission computed tomography (MPS) with the corresponding performance of experienced readers for the detection coronary artery disease (CAD).

Methods—995 rest/stress ^{99m}Tc-sestamibi MPS studies, [650 consecutive cases with coronary angiography and 345 with likelihood of CAD < 5% (LLk)] were obtained by MPS with AC. Total perfusion deficit (TPD) for AC and NC data were compared to the visual summed stress and rest scores of 2 experienced readers. Visual reads were performed in 4 consecutive steps with the following information progressively revealed: NC data, AC+NC data, computer results, all clinical information.

Results—The diagnostic accuracy of TPD for detection of CAD was similar to both readers (NC: 82% vs. 84%, AC: 86% vs. 85–87% p = NS) with the exception of second reader when using clinical information (89%, $p < 0.05$). The Receiver-Operator-Characteristics Areas-Under-Curve (ROC-AUC) for TPD were significantly better than visual reads for NC (0.91 vs. 0.87 and 0.89, $p < 0.01$) and AC (0.92 vs. 0.90, $p < 0.01$), and it was comparable to visual reads incorporating all clinical information. Per-vessel accuracy of TPD was superior to one reader for NC (81% vs. 77%, $p < 0.05$) and AC (83% vs. 78%, $p < 0.05$) and equivalent to second reader [NC (79%) and AC (81%)]. Per-vessel ROC-AUC for NC (0.83) and AC (0.84) for TPD were better than (0.78–0.80 $p < 0.01$), and comparable to second reader (0.82–0.84, $p = NS$), for all steps.

Conclusion—For the detection of 70% stenosis based on angiographic criteria, a fully automated computer analysis of NC and AC MPS data is equivalent for per-patient and can be superior for per-vessel analysis, when compared to expert analysis.

Corresponding Author Info: Piotr J. Slomka, PhD, Artificial Intelligence in Medicine Program, 8700 Beverly Blvd, Ste A047N, Los Angeles, CA 90048, USA, Ph: 310-423-4348, piotr.slomka@cshs.org.
First Author Info: Reza Arsanjani, MD (Post-doctoral fellow), Cedars-Sinai Medical Center, 8700 Beverly Blvd, Taper A238, Los Angeles, CA 90048, USA, Ph: 310-423-4332, Reza.Arsanjani@cshs.org

Keywords

Automated Quantification; Coronary Artery Disease; Myocardial Perfusion SPECT; Total Perfusion Deficit

Myocardial perfusion SPECT (MPS) is the most common noninvasive stress imaging modality of choice for diagnosis of Coronary Artery Disease (CAD) (1). Prior studies have demonstrated that quantitative analysis can supplement visual analysis (2–4). Quantitative analysis is more reproducible than visual analysis (5,6). However, despite these advantages, it is currently recommended that quantitative analysis be used only as an adjunct to visual analysis (7)based on the software's inability to explicitly differentiate between perfusion defect and artifact (7). Furthermore, multiple prior studies have shown that attenuation-corrected (AC) MPS assessed either visually or by software analysis resulted in improved diagnostic accuracy as compared to non-corrected (NC) MPS (8–10). However the differences between expert visual and automated analysis of NC and AC data have not been comprehensively evaluated. The aim of this study was to compare the performance of a fully automated analysis of combined NC and AC MPS (10) for the detection of obstructive CAD disease based on angiographic criteria to the visual scoring of experienced readers utilizing NC images, AC images, results of computer-analysis, and clinical information in a progressive and step-wise fashion.

MATERIALS AND METHODS

Patient Population

The subjects who were referred to the Nuclear Medicine Department of Sacred Heart Medical Center, Eugene, Oregon, from March 1, 2003 to December 31, 2006 for rest and stress electrocardiography (ECG)-gated NC and AC MPS were consecutively selected (10). All patients with a prior history of CAD or significant valve disease were excluded. MPS and coronary angiography had to be performed within 60 days without a significant intervening event. The low likelihood (LLk) studies were obtained from patients who performed an adequate treadmill stress test, did not have correlating coronary angiography available, but had < 5% likelihood of CAD using the Diamond and Forrester criteria based on age, sex, symptoms, and ECG response to adequate treadmill stress testing (11). Based on these selection criteria, we identified 650 patients with correlative angiography as described above and 345 patients with a LLk of CAD. The clinical characteristics of the two groups are listed in Table 1. The study protocol was approved by the Institutional Review Board (IRB) and written informed consent was obtained from all subjects.

Image Acquisition and Reconstruction Protocols

Standard ^{99m}Tc-sestamibi rest/stress protocols were employed as previously described with treadmill testing or adenosine infusion with low-level exercise (12). Vertex, dual-detector scintillation cameras with low energy high-resolution collimators and Vantage Pro attenuation correction hardware and software (Philips Medical Systems, Milpitas, CA) were used to acquire MPS.

Tomographic reconstruction was performed by AutoSPECT (13) and Vantage Pro programs (Philips Medical Systems). Emission images were automatically corrected for non-uniformity, radioactive decay, center-of-rotation, and motion during acquisition. Filtered back-projection and Butterworth filters were applied to obtain the NC MPS with an order of 10 and cutoff of 0.50 for rest MPS, and an order of 5 and cutoff of 0.66 for stress MPS. The attenuation maps and the emission data were used to reconstruct the AC images with a

maximum-likelihood expectation maximization algorithm incorporating scatter correction and depth-dependent resolution compensation.

Contour Adjustment and Quality Control

Automatically derived left ventricular (LV) contours were visually verified by an experienced technician from Cedars Sinai Medical Center, and if necessary the valve plane and left ventricular mask were manually adjusted. Additionally, automated software was run without any user intervention (fully unsupervised) and with adjustment performed by a second technologist from Sacred Heart Medical Center with no significant experience with the software, who was provided with simple training instructions. Both technologists were aided by an automated method for quality control (QC) of LV MPS contours (14). This method derives 2 parameters: the shape flag to detect the mask-failure cases, and the valve plane flag to detect the valve plane over- or undershooting. These QC flags were used as a guide by the technologists (more careful assessment of contours with shape QC flag > 3 and valve plane flag > 0.37 or < 0.28) (14); however, the contours were adjusted based on visual judgment when deemed appropriate.

Automated Analysis

Total perfusion deficit (TPD) for NC images was computed with previously developed simplified approach (15). For the AC results, we derived combined NC and AC (+AC) severity, which integrates NC and AC data for improved accuracy (10) similar to the visual AC analysis where readers combine NC and AC data. In short, hypoperfusion severities for NC and AC data were derived and the combined NC and AC (+AC) severity was determined at each polar map location by averaging NC and AC severities computed separately from NC and AC normal limits. Subsequently, +AC-TPD was computed by integrating the +AC severities below polar map normal limits. Threshold of 3% on per-patient basis for TPD was considered abnormal for both NC and +AC analysis (10). The ischemic TPD (ITPD) measurements were calculated as stress TPD minus rest TPD (NC-ITPD and +AC-ITPD), and an ITPD value of 2% on per-patient basis was considered abnormal (10). Partial NC-TPD and +AC-TPD scores for each vascular territory were also obtained with 2% being considered abnormal (10).

Visual Analysis

Visual interpretation of MPS images was based on 17-segment model (16). MPS images were scored independently by two experienced cardiologists, who were both board certified (Reader 1 with 30 years and Reader 2 with 10 years of clinical experience in nuclear cardiology), using a five-point scoring system (0, normal; 1, mildly decreased; 2, moderately decreased; 3, severely decreased; and 4, absence of segmental uptake). Visual reading was performed in 4 consecutive steps with the following information progressively revealed to the readers as follows. During the first step (V1), the readers scored both the NC stress and rest images based on perfusion data, raw projection data and gated function data (17). During the second step (V2), the readers could re-score the stress and rest studies based on additional AC data. During the third step (V3), the readers also had access to the perfusion quantification results obtained by the software. Finally, during the fourth step (V4), the readers were additionally provided clinical information, including age, cardiac risk factors, type of stress, and clinical and ECG responses to stress. During each step, observers could also modify the default assignment of segments to the specific vascular territory as it is possible in QPS/QGS software. Subsequently, summed stress scores (SSS) and summed rest scores (SRS) and summed difference scores (SDS) were calculated from the17-segment scores. All scores were recorded automatically in the batch files, eliminating manual transfer. SSS 4 and SDS 2 were considered abnormal (10). Partial summed scores for

each vascular territory were also obtained with SSS 2 being considered abnormal for any territory (10).

Conventional Coronary Angiography

Conventional coronary angiography was performed according to standard clinical protocols within 60 days of the myocardial perfusion examination. All coronary angiograms were visually interpreted by an experienced cardiologist. A stenosis of 50% for the left main or 70% for other coronary arteries was considered as the gold standard for the detection of CAD. A secondary analysis was performed where any coronary artery stenosis of 50% was considered significant.

Statistical Analysis

Continuous variables were expressed as the mean ± standard deviation, and categorical variables were expressed as percentages (%). Inter-observer agreement between the two readers was compared using Bland-Altman test and Kappa-test. The total estimate of agreement, defined as total cases where the tests agree, was compared between automated and visual reads, as well as between the two visual readers. We also compared the positive percent agreement, defined as total positive cases where the tests agree, and negative percent agreement, defined as total negative cases where the tests agree (18). The overall automated and visual total agreements, positive percent agreements, negative percent agreements, sensitivity, specificity, and accuracy were compared using a Z-test. Automated NC data was compared to visual NC data only (V1), while automated AC results were compared to visual AC data (V2–V4).

For all analyses, p values < 0.05 were considered statistically significant. Receiver Operator Characteristics (ROC) curves were analyzed to evaluate the ability of TPD versus visual scoring for forecasting 70% and 50% stenoses of the coronary arteries. The differences between the ROC Area–Under-Curve (AUC) were compared by the Delong method (19).

RESULTS

Agreement between the Automated and Visual Reads

Table 2 compares the diagnostic agreement (total positive and negative percent agreement) between the two readers as well as each reader and automated quantification. Overall, there was high agreement between the two readers (87% to 91%) and between each reader and the automated results (84% to 89%). The total agreement significantly improved (by at least 3% for both readers and the software) with the addition of +AC data in comparison to NC data. Figure 1 demonstrates the number of cases when the diagnosis was changed during each of the steps. The addition of AC data changed the diagnosis in over 8% of cases for both auto and visual reads. The inter-observer correlations and kappa agreements are shown in Table 3. Inter-observer kappa agreement improved from 0.77 to 0.82 (p = 0.006) with the addition of AC images.

Software versus Reader: Per-Patient Diagnostic Performance

Figure 2 compares diagnostic performance for stress NC-TPD, AC-TPD, and 2 visual readers for detection of 70% stenosis on a per-patient basis. For NC data, the specificities of visual readers were higher, the sensitivity was lower for one reader, and overall accuracy was similar for readers in comparison to the automated analysis. The accuracy and specificity for all the steps with AC data (V2–V4) were similar to the +AC TPD analysis with the exception for the higher accuracy of Reader 2 at V4 incorporating AC, computer and clinical analysis (89% vs. 86%, p < 0.05). The V3 step for Reader 1 incorporating AC and computer analysis increased sensitivity (84% vs. 89%, p < 0.05). Similar results were

noted when comparing NC-TPD, +AC-TPD, and visual reads from both readers for detection of 50% CAD on a per-patient basis. The specificity and accuracy of the automated analysis significantly improved for detection of 70% stenosis (4%) with the addition of +AC-data on per-patient basis. The accuracy for the Reader 1 did not improve at step V4; however the sensitivity and accuracy for the Reader 2 improved significantly when the clinical information (V4) was incorporated, by 5.4% and 2.5%, respectively. There were 25 cases with 70% stenosis, where both expert readers agreed and were correct while the automated analyses were incorrect. On the other hand, there were 8 cases, where the automated analysis was correct, while both experts were incorrect.

The ROC curves comparing NC-TPD, +AC-TPD and visual reads are shown in Figure 3. The NC-TPD and +AC-TPD ROC-AUC was significantly higher ($p < 0.01$) when compared to NC (V1) and AC (V2) reads respectively. The ROC-AUC for both visual readers at the final step (V4) were similar to +AC-TPD analysis (0.91 vs. 0.92, p = NS). Similarly using 50% stenosis cut-off, the ROC-AUC for NC-TPD and +AC-TPD was significantly higher ($p < 0.01$) when compared to the visual NC and AC reads and the visual ROC-AUCs for the final read (V4) for both readers were also similar to automated analysis.

Software versus Readers: Per-Vessel Diagnostic Performance

On a per-vessel basis the diagnostic performance of the automated analysis were comparable to the visual analysis of Reader 2 but the accuracies of NC-TPD and +AC-TPD were superior to the analysis of Reader 1 ($p < 0.05$) (Figure 4). In individual territories, the diagnostic accuracy of the automated analysis for detection of 70% stenosis based on angiographic criteria was higher than Reader 1 and equivalent to Reader 2 in all V1–V4 steps for LAD and LCX territories (see supplemental data). In addition, NC-TPD analysis of RCA territory was more accurate than both readers using NC data only (82% vs. 77–78% %, $p < 0.05$). Per-vessel diagnostic accuracy did not improve with addition of computer and clinical analysis. Neither of the readers had higher diagnostic accuracy than the computer software in any of the territories. The addition of AC information to NC, improved the diagnostic accuracy for automated analysis (83% vs. 81%, $p < 0.05$), which was similar to the pattern seen for per-patient analysis.

The ROC curves comparing automated and visual measurements on a per-vessel basis (70% stenosis) are shown in Figure 5. The ROC-AUC for NC-TPD and +AC-TPD was significantly higher ($p < 0.01$) when compared to the visual NC and AC reads (V1 and V2 respectively) for Reader 2 and was comparable to Reader 1.

Software versus Reader: Ischemic Measurements

We also compared the performance in detection of 70% stenosis on per-patient basis using automated (ITPD) and visual (SDS) ischemic measurements with generally similar results as for the stress measurements. The automated diagnostic accuracy was 82% (NC) and 83% (+AC) ITPD. The ITPD-NC sensitivity was superior to visual scoring (87% vs. 78% for Reader 1 and 63% for Reader 2, $p < 0.001$), the sensitivity of +AC-ITPD (87%) was superior to that of Reader 2 (74–78%) for all steps (V2–V4) ($p < 0.001$). The specificities of both readers were superior for NC (86% and 91% vs. 80%, $p < 0.001$), and the specificity of Reader 2 (V2–V4) was superior for +AC data ($p < 0.001$). The accuracies of NC-ITPD and Reader 1 were similar (82% vs. 83%, $p = 0.56$), but the accuracy of NC-TPD was superior to that of Reader 2 (82% vs. 78%, $p = 0.03$). The accuracy of +AC-ITPD (83%) was similar to both readers for V2–V4 steps (83–86%, $p > 0.06$). The ROC-AUCs for NC (0.90) and +AC ITPD (0.91) were better than for V1 read (0.80–0.84 - both readers), and V2–V4 reads (0.84–0.89 – both readers) ($p < 0.02$), respectively.

Software Analysis: Impact of Manual Contour Adjustment

LV contours were manually adjusted for shape (localize or mask option) or valve plane (constrain option) in 11% of NC cases and 29% of the AC cases (valve plane-only in 58% for NC and 48% for AC of these adjustments) data by an experienced Cedars-Sinai technologist. The shape or valve plane were manually adjusted in 21% of cases for NC and 34% of cases for AC data by a less experienced technologist (valve plane-only in 84% for NC and 71% for AC of these adjustments) ($p < 0.05$ for both versus experienced technologist).

The comparison of the diagnostic performance for unsupervised, less experienced technologist and experienced technologist are shown in Table 4. Overall, sensitivities, specificities, and accuracies for NC and +AC were similar for all three runs. There was a trend toward improved +AC-TPD specificity for the more experienced technologist ($p = 0.059$). Nevertheless, the ROC-AUC for NC and +AC analysis were significantly lower for the inexperienced technologist than those obtained by using contours performed by an experienced technologist. We also compared these results to our expert visual readings (stage 1 and 2). There were no significant differences in overall diagnostic accuracy between the expert readers and these three types of automated analysis. However, the NC and +AC ROC-AUC for both readers were significantly lower than ROC-AUC for all three categories ($p < 0.05$).

DISCUSSION

Visual analysis of MPS, currently the recommended standard clinical practice, is dependent on a subjective interpretation of the data and prone to possible bias related to reader experience, which is also a major shortcoming of other non-invasive stress tests such as stress echocardiography. In this study, the overall diagnostic accuracy of the fully automated computer analysis using NC and AC MPS was at least equivalent to the expert visual reads on a per-patient basis. Furthermore, on a per-vessel basis, the automated analysis for detection of 70% stenosis based on angiographic criteria was at all times at least comparable and in some cases superior to the visual analysis. To our knowledge this is the first study comprehensively comparing automated analysis to the semi-quantitative visual analysis by evaluating the incremental diagnostic value of supplementing NC data by AC images, computer analysis, and clinical information to on per-patient, and per-vessel basis.

The reading experts in our study were attending physicians from premier imaging centers with extensive experience in MPS interpretation. It is therefore likely that a fully automatic analysis could play an integral role as a guide for the less experienced reader who may be less certain about normal variation in uptake (7). Prior studies have demonstrated that less-experienced readers have more variability when compared to experienced readers (20). Our study demonstrated that automated analysis was at least comparable to visual reads on per-vessel territory and at times outperformed it based on diagnostic accuracy and ROC analysis for detection of 70% stenosis based on angiographic criteria. Therefore, although ASNC currently recommends supplementation of quantitative to visual analysis for MPS, it might be feasible that in the future visual analysis might be used only to override the quantitative analysis in a minority of cases.

We also assessed the nuclear technologist role in contour verification during automated analysis. Surprisingly, there were no significant differences between sensitivity, specificity, and accuracy for fully-unsupervised analysis, less-experienced, and more experienced contour adjustments for NC and AC data. However, the ROC-AUC for NC-TPD and +AC-TPD for unsupervised analysis and analysis by a less experienced technologist were slightly lower as compared to those generated based on contours checked by an experienced

technologist. Therefore, there could be some potential advantage of contour checking by an experienced technologist, but it is likely small. Nevertheless, the ROC-AUC in all three cases was higher than the blinded visual reads by both expert readers. Furthermore, the less-experienced technologist was adjusting the contours more frequently, especially the valve plane, which may be related to the fact that the QC flag (especially valve flag) typically has low specificity and high sensitivity for detection of contour failures and can over-indicate the need for adjustment.

Importantly, the visual analysis of stress NC and AC data included gated function information, as well as stress and rest perfusion images, while the stress automated analysis was solely based on stress TPD and was therefore a truly stress-only perfusion analysis. Additionally, the comparison of the visual versus automated analysis using ischemic parameters did not reveal any significant differences in this patient population with suspected CAD but no previous documented history of coronary disease. Therefore, the automatic performance results can be extrapolated to any stress-only studies, which has the promise to reduce patient imaging time and radiation dose (21).

We also evaluated the benefits of adding AC information to the NC information for both automated and visual analysis. The addition of AC information to the NC data in our study resulted in improved diagnosis of 70% stenosis based on accuracy and ROC evaluation for automated analysis, on a per-patient and per-vessel basis which is consistent with prior studies(10,22–24). Additionally, the overall agreement between the two readers, as well between readers and TPD, improved with the addition of AC information. Therefore, our findings suggest that if AC is available it should routinely be used when interpreting MPS. It is possible that similar benefits could be obtained by prone-supine analysis instead of the use of AC data (25).

Prior studies have demonstrated that the knowledge of clinical information may results in significant change in interpretation of MPS studies (26). However, in our analysis the addition of clinical information did not routinely improve diagnosis of CAD from MPS and the degree of agreement between the two readers did not significantly improve when they were provided with clinical information. However, the diagnostic accuracy for Reader 2 did improve by 3 % ($p < 0.05$) when clinical information was considered. It should also be noted that although these differences are small, the potential advantages of experienced visual observer integrating clinical information cannot be easily dismissed.

The overall agreement between the two expert readers in our study was good (kappa = 0.77–0.83). Although the overall accuracy was similar amongst the two readers, one expert had consistently higher sensitivity while the other had higher specificity. This highlights a common difficulty in providing a definite diagnosis from visual analysis of MPS. The data illustrates that different readers operate at different sensitivity/specificity thresholds and this in part is the cause of the inter-observer variability.

This study has several limitations. Visual coronary angiography interpretation was used as the gold standard for this study with its known limitations. LLk data were included in the analysis and were considered to have normal angiograms. However, prior studies have indicated that patients with normal angiographic data can often have an abnormal MPS scans due to referral bias, affecting the overall diagnostic accuracy of MPS scans (27). Our comparisons are based solely on diagnostic performance for obstructive coronary artery disease; however, the visual analysis and software also provide information regarding the extent of the myocardium involved, which may have prognostic implications. Further prognostic studies are needed to clearly demonstrate the superiority of the automated method. Finally, the results were obtained on only one particular camera and attenuation

correction system; therefore, further multicenter evaluation will be required in the future to confirm these results.

CONCLUSION

Automated computer analysis utilizing NC and AC MPS data with contours checked by an experienced technologist is at least equivalent to visual analysis in terms of detection of coronary angiographic findings of 70% stenoses even when the reader is provided clinical and LV function information, and can outperform the experienced reader on a per-territory basis. Furthermore, attenuation-correction improves the diagnostic accuracy of automated analysis as well as improving the inter-observer agreement for visual analysis between readers.

Acknowledgments

REFERENCES

1. Sharir T, Ben-Haim S, Merzon K, et al. High-speed myocardial perfusion imaging initial clinical comparison with conventional dual detector anger camera imaging. JACC Cardiovasc Imaging. 2008; 1:156–163. [PubMed: 19356422]

2. Berman DS, Kang X, Van Train KF, et al. Comparative prognostic value of automatic quantitative analysis versus semiquantitative visual analysis of exercise myocardial perfusion single-photon emission computed tomography. J Am Coll Cardiol. 1998; 32:1987–1995. [PubMed: 9857883]

3. Leslie WD, Tully SA, Yogendran MS, et al. Prognostic value of automated quantification of 99mTc-sestamibi myocardial perfusion imaging. J Nucl Med. 2005; 46:204–211. [PubMed: 15695777]

4. Tamaki N, Yonekura Y, Mukai T, et al. Stress thallium-201 transaxial emission computed tomography: quantitative versus qualitative analysis for evaluation of coronary artery disease. J Am Coll Cardiol. 1984; 4:1213–1221. [PubMed: 6334109]

5. Berman DS, Kang X, Gransar H, et al. Quantitative assessment of myocardial perfusion abnormality on SPECT myocardial perfusion imaging is more reproducible than expert visual analysis. J Nucl Cardiol. 2009; 16:45–53. [PubMed: 19152128]

6. Iskandrian AE, Garcia EV, Faber T, et al. Automated assessment of serial SPECT myocardial perfusion images. J Nucl Cardiol. 2009; 16:6–9. [PubMed: 19152123]

7. Holly TA, Abbott BG, Al-Mallah M, et al. Single photon-emission computed tomography. J Nucl Cardiol. 2010; 17:941–973. [PubMed: 20552312]

8. Grossman GB, Garcia EV, Bateman TM, et al. Quantitative Tc-99m sestamibi attenuation-corrected SPECT: development and multicenter trial validation of myocardial perfusion stress gender-independent normal database in an obese population. J Nucl Cardiol. 2004; 11:263–272. [PubMed: 15173773]

9. Thompson RC, Heller GV, Johnson LL, et al. Value of attenuation correction on ECG-gated SPECT myocardial perfusion imaging related to body mass index. J Nucl Cardiol. 2005; 12:195–202. [PubMed: 15812374]

10. Xu Y, Fish M, Gerlach J, et al. Combined quantitative analysis of attenuation corrected and non-corrected myocardial perfusion SPECT: Method development and clinical validation. J Nucl Cardiol. 2010; 17:591–599. [PubMed: 20387137]

11. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. N. Engl J Med. 1979; 300:1350–1358. [PubMed: 440357]

12. Slomka PJ, Fish MB, Lorenzo S, et al. Simplified normal limits and automated quantitative assessment for attenuation-corrected myocardial perfusion SPECT. J Nucl Cardiol. 2006; 13:642–651. [PubMed: 16945744]

13. Germano G, Berman DS. On the accuracy and reproducibility of quantitative gated myocardial perfusion SPECT. J Nucl Med. 1999; 40:810–813. [PubMed: 10319755]

14. Xu Y, Kavanagh P, Fish M, et al. Automated quality control for segmentation of myocardial perfusion SPECT. J Nucl Med. 2009; 50:1418–1426. [PubMed: 19690019]

15. Slomka PJ, Nishina H, Berman DS, et al. Automated quantification of myocardial perfusion SPECT using simplified normal limits. J Nucl Cardiol. 2005; 12:66–77. [PubMed: 15682367]

16. Germano G, Kavanagh PB, Slomka PJ, et al. Quantitation in gated perfusion SPECT imaging: the Cedars-Sinai approach. J Nucl Cardiol. 2007; 14:433–454. [PubMed: 17679052]

17. Wolak A, Slomka PJ, Fish MB, et al. Quantitative myocardial-perfusion SPECT: comparison of three state-of-the-art software packages. J Nucl Cardiol. 2008; 15:27–34. [PubMed: 18242477]

18. Shoukri, MM. Measures of Interobserver Agreement and Reliability. Second Edition. Boca Raton, FL: CRC Press; 2010. p. 11-15.

19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 44:837–845. [PubMed: 3203132]

20. Golub RJ, Ahlberg AW, McClellan JR, et al. Interpretive reproducibility of stress Tc-99m sestamibi tomographic myocardial perfusion imaging. J Nucl Cardiol. 1999; 6:257–269. [PubMed: 10385181]

21. Heller GV, Bateman TM, Johnson LL, et al. Clinical value of attenuation correction in stress-only Tc-99m sestamibi SPECT imaging. J Nucl Cardiol. 2004; 11:273–281. [PubMed: 15173774]

22. Ficaro EP, Fessler JA, Shreve PD, et al. Simultaneous transmission/emission myocardial perfusion tomography. Diagnostic accuracy of attenuation-corrected 99mTc-sestamibi single-photon emission computed tomography. Circulation. 1996; 93:463–473. [PubMed: 8565163]

23. Kluge R, Sattler B, Seese A, et al. Attenuation correction by simultaneous emission-transmission myocardial single-photon emission tomography using a technetium-99m-labelled radiotracer: impact on diagnostic accuracy. Eur J Nucl Med. 1997; 24:1107–1114. [PubMed: 9283102]

24. Shotwell M, Singh BM, Fortman C, et al. Improved coronary disease detection with quantitative attenuation-corrected Tl-201 images. J Nucl Cardiol. 2002; 9:52–62. [PubMed: 11845130]

25. Nishina H, Slomka PJ, Abidov A, et al. Combined supine and prone quantitative myocardial perfusion SPECT: method development and clinical validation in patients with no known coronary artery disease. J Nucl Med. 2006; 47:51–58. [PubMed: 16391187]

26. Simons M, Parker JA, Donohoe KJ, et al. The impact of clinical data on interpretation of thallium scintigrams. J Nucl Cardiol. 1994; 1:365–371. [PubMed: 9420719]

27. Miller TD, Hodge DO, Christian TF, et al. Effects of adjustment for referral bias on the sensitivity and specificity of single photon emission computed tomography for the diagnosis of coronary artery disease. Am J Med. 2002; 112:290–297. [PubMed: 11893368]

☐ **Total changed to positive**   ■ **Total changed to negative**
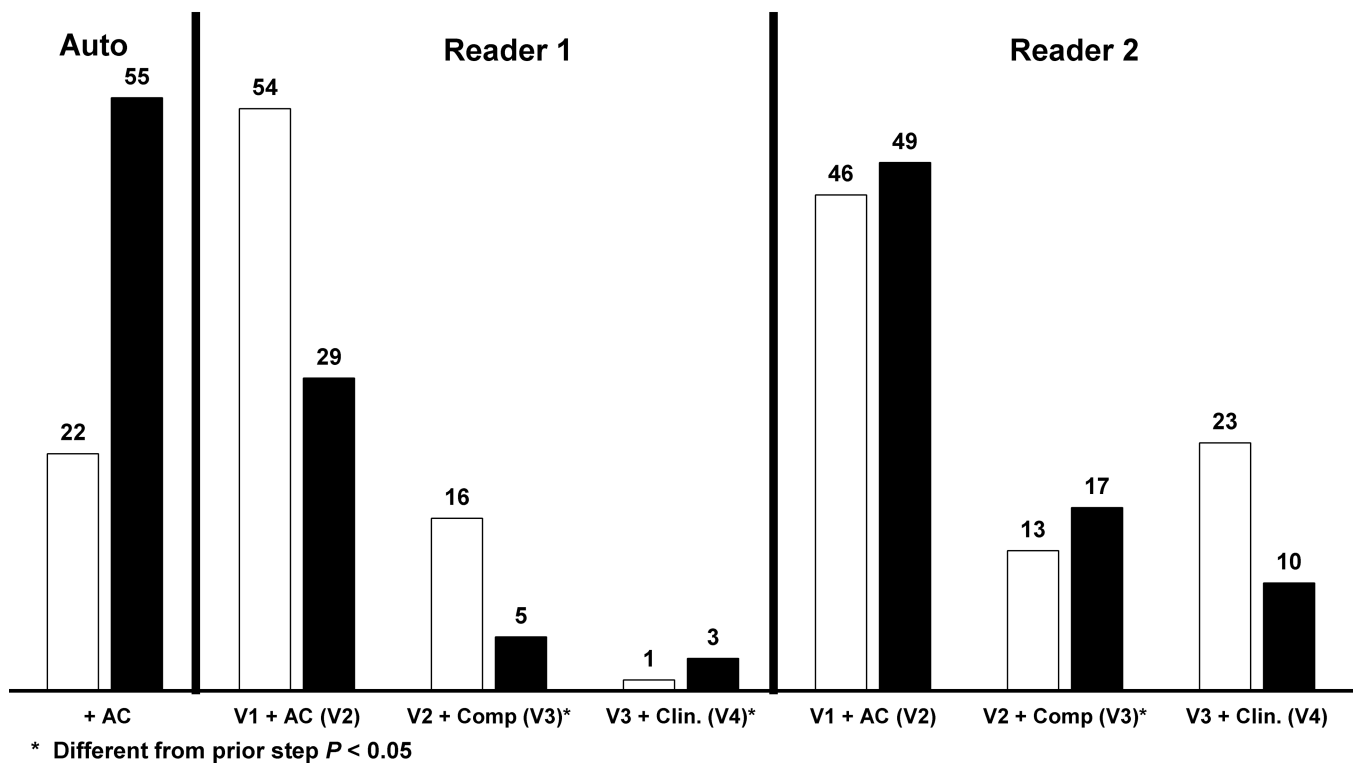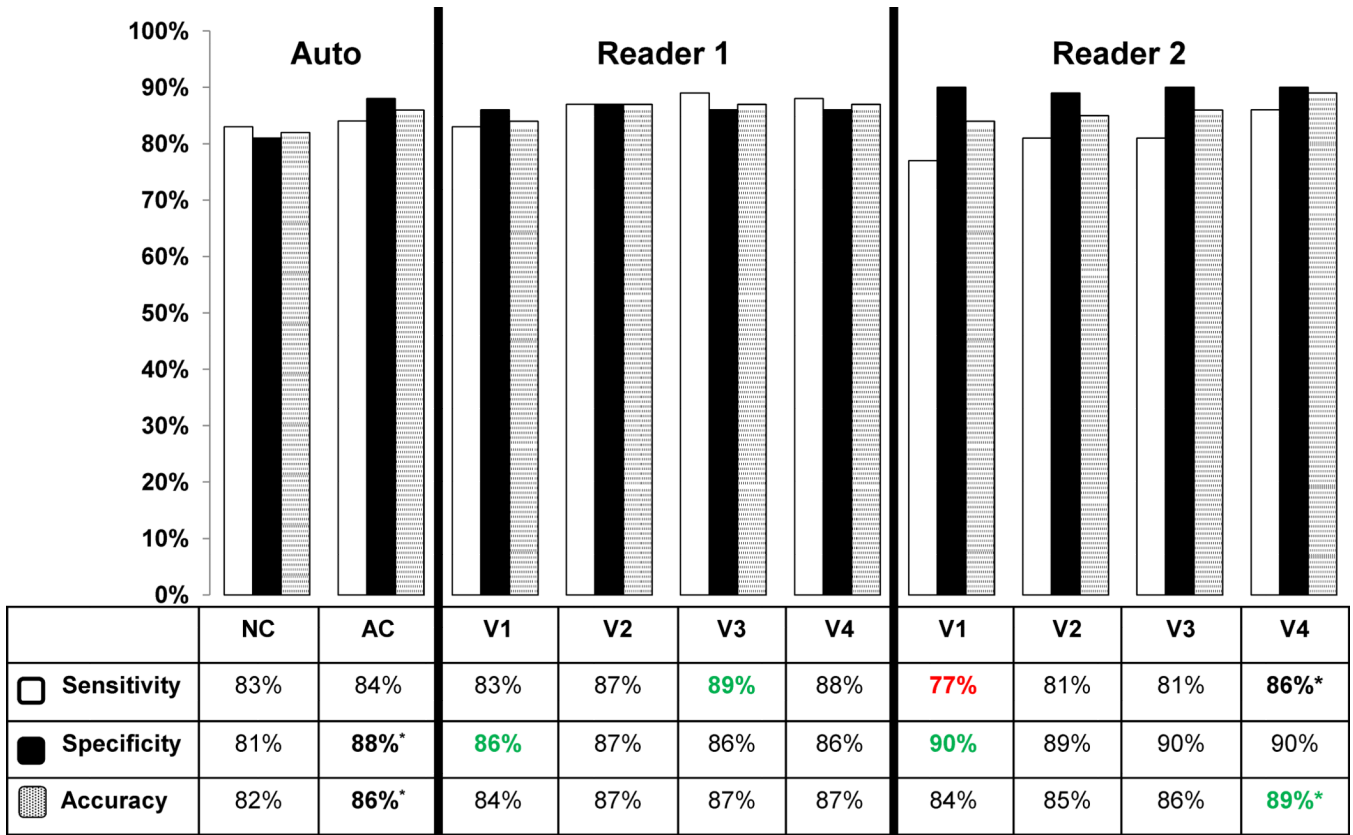

Figure 1.
Number of cases with changed diagnosis in each subsequent step for both automated and visual analysis. * Indicates significant difference compared to a prior step ($p < 0.05$).

		Auto		Reader 1				Reader 2			
		NC	AC	V1	V2	V3	V4	V1	V2	V3	V4
☐	Sensitivity	83%	84%	83%	87%	**89%**	88%	**77%**	81%	81%	**86%***
■	Specificity	81%	**88%***	**86%**	87%	86%	86%	**90%**	89%	90%	90%
▒	Accuracy	82%	**86%***	84%	87%	87%	87%	84%	85%	86%	**89%***

*Different from prior step *P* < 0.05
Green = visual better than auto *P* < 0.05   Red = auto better than visual *P* < 0.05
V1 = NC only   V2 = NC + AC
V3 = NC + AC + computer   V4 = NC + AC + computer + clinical

Figure 2.
Diagnostic performance of automatic analysis versus visual analysis for detection of   70% coronary artery lesions on per-patient basis (Number of patients with   70% stenotic lesion on cardiac catheterization = 463), * indicates significant difference compared to a prior step (*p* < 0.05). The automated analysis was also compared to visual analysis (NC vs. V1 and AC vs. V2–V4). Green color signifies that visual analysis was better than automated, while red color signifies that automated analysis was better than visual analysis (*p* < 0.05).
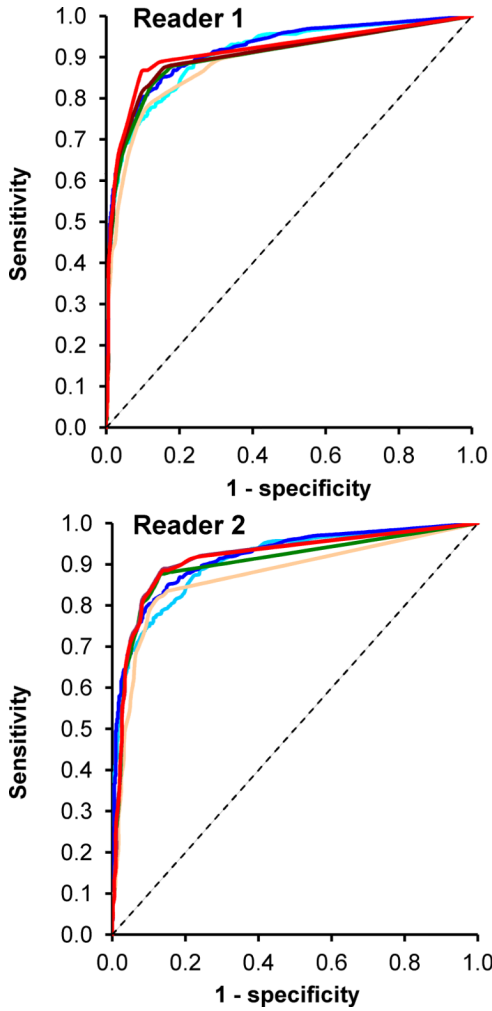
		ROC-AUC	ROC-AUC
🟦	NC-TPD	0.91	-
🟦	AC-TPD	0.92 *	-
		Reader 1	Reader 2
🟧	V1	0.87 #	0.89 #
🟩	V2	0.90 *#	0.90 #
🟫	V3	0.92 *	0.90 #
🟥	V4	0.91	0.91 *

*Comparison to prior step *P* < 0.01
#Auto vs. visual *P* < 0.01
  Red = auto better than visual
  V1 = NC only
  V2 = NC + AC
  V3 = NC + AC + computer
  V4 = NC + AC + computer + clinical

Figure 3.
The receiver operating characteristic (ROC) curves comparing the automated versus visual reads on per-patient basis 2 readers for detection of   70% stenosis. * Indicates significant difference compared to a prior step ($p < 0.05$). The automated analysis was also compared to visual analysis (NC vs. V1 and AC vs. V2–V4), # indicates significant difference ($p < 0.01$). Red color signifies that automated analysis was better than visual analysis ($p < 0.01$).
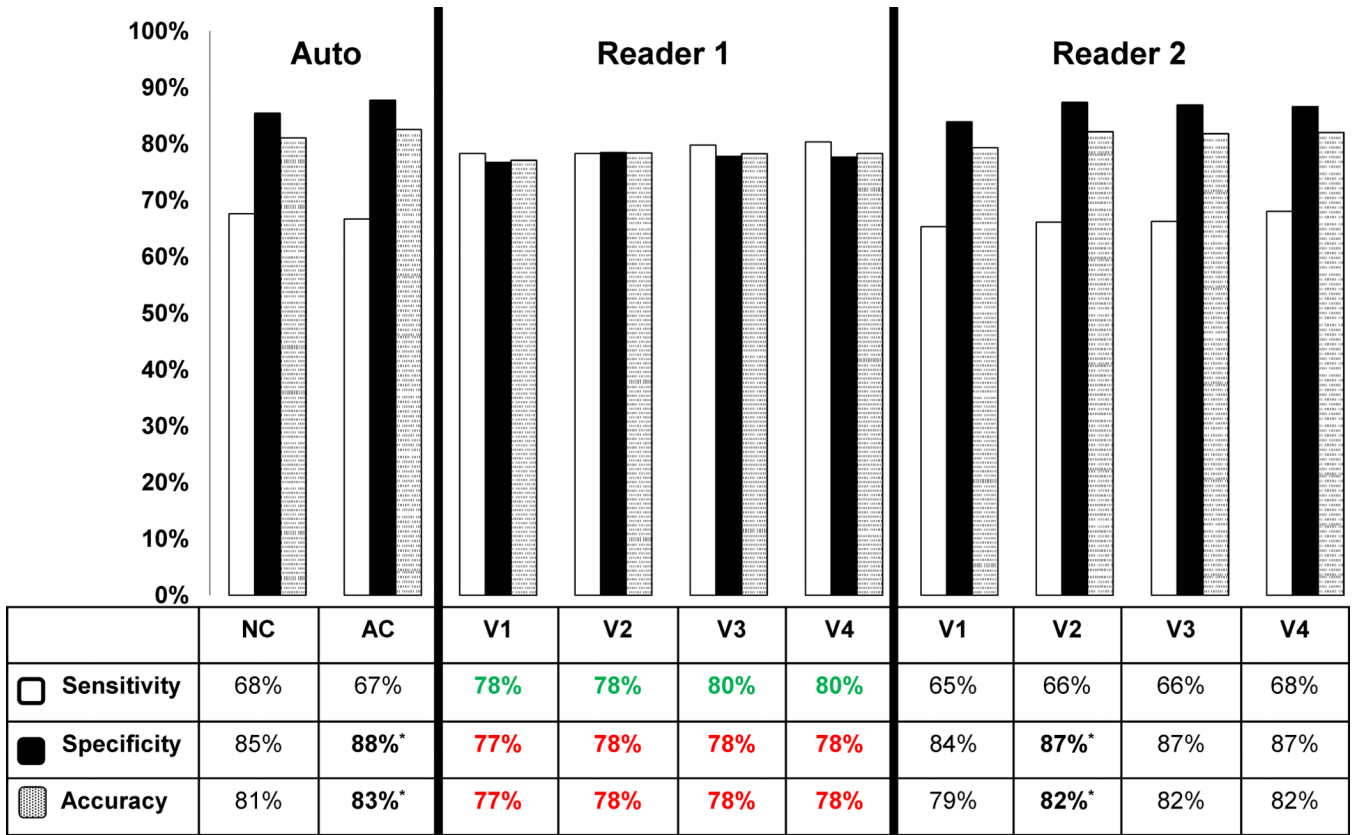
		NC	AC	V1	V2	V3	V4	V1	V2	V3	V4
□	Sensitivity	68%	67%	78%	78%	80%	80%	65%	66%	66%	68%
■	Specificity	85%	88%*	77%	78%	78%	78%	84%	87%*	87%	87%
▨	Accuracy	81%	83%*	77%	78%	78%	78%	79%	82%*	82%	82%

Figure 4.
Diagnostic performance of automatic analysis versus visual analysis for detection of ≥70% coronary artery lesions on per-vessel basis in all vessels. * Indicates statistically significant difference compared to a prior step ($p < 0.05$). The automated analysis was also compared to visual (NC vs. V1 and AC vs. V2–V4). Green color signifies that visual analysis was better than automated analysis; red color signifies that automated analysis was better than visual analysis ($p < 0.05$).
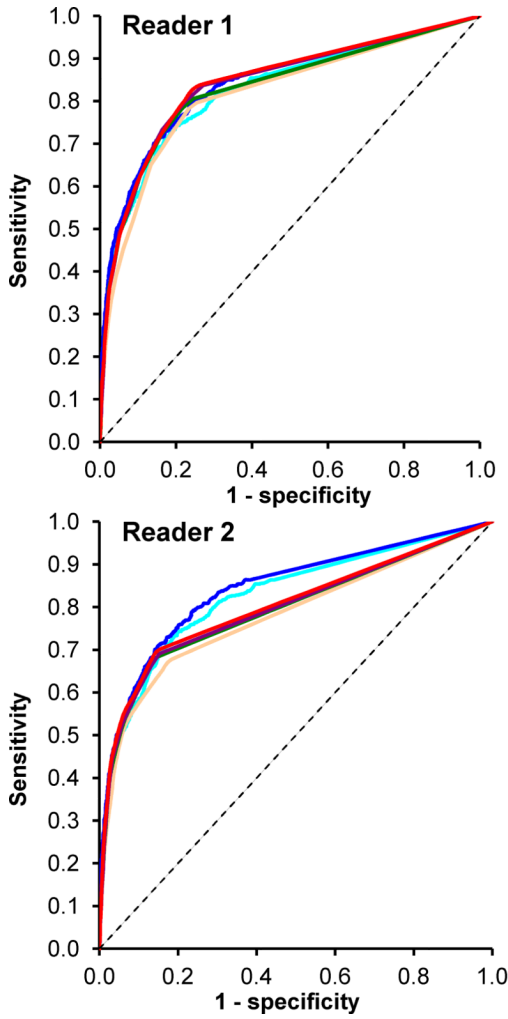
		ROC-AUC	ROC-AUC
	NC-TPD	0.83	-
	AC-TPD	0.84 *	-
		Reader 1	Reader 2
	V1	0.82	0.78 #
	V2	0.83 *	0.79 *#
	V3	0.84 *	0.80 #
	V4	0.84	0.80 *#

*Comparison to prior step *P* < 0.01
#Auto vs. visual *P* < 0.01
Red = auto better than visual
V1 = NC only
V2 = NC + AC
V3 = NC + AC + computer
V4 = NC + AC + computer + clinical

Figure 5.
The receiver operating characteristic (ROC) curves comparing the automated versus visual reads on per-vessel basis for both readers for detection of 70% stenosis. * Indicates statistically significant difference compared to a prior step ($p < 0.01$). The automated analysis was also compared to visual (NC vs. V1 and AC vs. V2–V4) and # signifies significant difference ($p < 0.01$). Red color signifies that automated analysis was better than visual analysis ($p < 0.01$).

Table 1

Baseline characteristics of the patients.

	Angiography Group	LLk Group	P Value
Number	650	345	N/A
Age (Years)	64 ± 12	52 ± 11	P < 0.01
Male %	55 %	39 %	P < 0.01
Female %	45 %	61 %	P < 0.01
Diabetes Mellitus %	27 %	0 %	P < 0.01
Hypertension %	64 %	39 %	P < 0.01
Hyperlipidemia %	51 %	39 %	P < 0.01
Smoking %	19 %	23 %	P = 0.08
Exercise SPECT	34 %	100 %	P < 0.01
Adenosine SPECT	66 %	0 %	P < 0.01
Ejection Fraction (%)	61.7 ± 12.2	64.0 ± 10.5	P = 0.47
Cath: 0-Vessel Disease	181 (28%)	N/A	N/A
Cath: 1-Vessel Disease	206 (32%)	N/A	N/A
Cath: 2-Vessel Disease	148 (23%)	N/A	N/A
Cath: 3-Vessel Disease	115 (17%)	N/A	N/A

N/A = Not applicable.

Table 2

Diagnostic agreement between automated analysis and each individual reader, as well as inter-observer agreement.

	Auto Vs. reader 1			Auto vs. reader 2			Reader 1 vs. reader 2		
Positive % agreement	85%	84%	84%	86%	87%	88%	87%	89%	93%*	95%	94%
Negative % agreement	82%	91%*	92%	92%	83%	89%*	90%	91%	86%	89%*	88%	89%
Total agreement	84%	88%*	88%	88%	85%	88%*	89%	89%	87%	91%*	91%	91%

*
Indicates significant difference compared to a prior step (P < 0.05).

Table 3

Inter-observer agreement comparison between 2 readers at each visual step (V1–V4).

Reader 1 vs. Reader 2	Correlation	Kappa statistics
Step V1	0.88	0.77
Step V2	0.87	0.82
Step V3	0.83	0.83
Step V4	0.85	0.82

V1 = NC only, V2 = NC+AC, V3 = NC + AC + Computer, V4 = NC + AC + Computer + Clinical. All correlations and kappa agreements had p-value < 0.001.

Table 4

Comparison between automated diagnostic performance for fully automated analysis, analysis with contours analyzed and if needed changed by a less experienced technologist, and contours performed by an experienced technologist.

	Fully unsupervised		Less experienced		More experienced	
	NC	+AC	NC	+AC	NC	+AC
Sensitivity	83%	83%	83%	81%	83%	84%
Specificity	78%	84%	81%	88%	81%	88%
Accuracy	80%	84%	82%	85%	82%	86%
ROC-AUC	0.90*	0.91*	0.90*	0.91*	0.91	0.92

*
Indicates significant difference compared to a more experienced ($p < 0.05$).