ORIGINAL ARTICLE

Multicenter evaluation of stress-first myocardial perfusion image triage by nuclear technologists and automated quantification

Waseem Chaudhry, MD,[a] Nasir Hussain, MD,[a] Alan W. Ahlberg, MA,[a]
Lori B. Croft, MD,[b] Antonio B. Fernandez, MD,[a] Mathew W. Parker, MD,[a]
Heather H. Swales, MD,[a] Piotr J. Slomka, PhD,[c] Milena J. Henzlova, MD,[b] and
W. Lane Duvall, MD[a]

[a] Henry Low Heart Center, Hartford Hospital, Hartford, CT
[b] Mount Sinai Division of Cardiology (Mount Sinai Heart), New York, NY
[c] Cedars-Sinai Medical Center, Los Angeles, CA

Background. A stress-first myocardial perfusion imaging (MPI) protocol saves time, is cost effective, and decreases radiation exposure. A limitation of this protocol is the requirement for physician review of the stress images to determine the need for rest images. This hurdle could be eliminated if an experienced technologist and/or automated computer quantification could make this determination.

Methods. Images from consecutive patients who were undergoing a stress-first MPI with attenuation correction at two tertiary care medical centers were prospectively reviewed independently by a technologist and cardiologist blinded to clinical and stress test data. Their decision on the need for rest imaging along with automated computer quantification of perfusion results was compared with the clinical reference standard of an assessment of perfusion images by a board-certified nuclear cardiologist that included clinical and stress test data.

Results. A total of 250 patients (mean age 61 years and 55% female) who underwent a stress-first MPI were studied. According to the clinical reference standard, 42 (16.8%) and 208 (83.2%) stress-first images were interpreted as ''needing'' and ''not needing'' rest images, respectively. The technologists correctly classified 229 (91.6%) stress-first images as either ''needing'' (n = 28) or ''not needing'' (n = 201) rest images. Their sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were 66.7%, 96.6%, 80.0%, and 93.5%, respectively. An automated stress TPD score ≥1.2 was associated with optimal sensitivity and specificity and correctly classified 179 (71.6%) stress-first images as either ''needing'' (n = 31) or ''not needing'' (n = 148) rest images. Its sensitivity, specificity, PPV, and NPV were 73.8%, 71.2%, 34.1%, and 93.1%, respectively. In a model whereby the computer or technologist could correct for the other's incorrect classification, 242 (96.8%) stress-first images were correctly classified. The composite sensitivity, specificity, PPV, and NPV were 83.3%, 99.5%, 97.2%, and 96.7%, respectively.

Conclusion. Technologists and automated quantification software had a high degree of agreement with the clinical reference standard for determining the need for rest images in a stress-first imaging protocol. Utilizing an experienced technologist and automated systems to screen stress-first images could expand the use of stress-first MPI to sites where the cardiologist is not immediately available for interpretation. (J Nucl Cardiol 2017;24:809–20.)

Key Words: SPECT MPI · stress-first · nuclear technologists · radiation dose reduction · automated perfusion quantification

Abbreviations	
MPI	Myocardial perfusion imaging
CAD	Coronary artery disease
MI	Myocardial infarction
PCI	Percutaneous coronary intervention
CABG	Coronary artery bypass grafting
CZT	Cadmium-zinc-telluride
TPD	Total perfusion deficit
AUC	Area under the curve
PPV	Positive predictive value
NPV	Negative predictive value

See related editorial, pp. 821–825

INTRODUCTION

Myocardial perfusion imaging (MPI) is traditionally composed of two sets of images, rest and stress, with the purpose of the rest images being the determination of whether any stress perfusion defects are reversible (ischemic) or fixed (infarcted or caused by attenuation artifact). When stress MPI is normal, however, the rest image becomes superfluous. There are now data from thousands of patients that the stress-first strategy provides high-quality perfusion data equivalent to a full rest-stress study,[1,2] while saving time in the imaging laboratory and reducing radiation exposure both to the patients and to the laboratory personnel.[3,4] In current clinical practice, the majority of appropriately indicated diagnostic stress MPI studies are found to be normal, especially in patients with no prior history of coronary artery disease (CAD), which has been substantiated by recent articles investigating the temporal trends of abnormal or ischemic MPI study results in large clinical cohorts.[5-7] With the increasing prevalence of normal MPI studies and competition from other non-invasive imaging modalities, it is imperative that the field develops more cost-effective strategies for the initial evaluation of patients, and stress-first protocols represent an attractive option.

Yet stress-first protocols have not been widely adopted perhaps reflecting challenges such as the need for attenuation correction, feasibility of real-time review of stress images, and concerns about reimbursement.[8] For the successful implementation of a stress-first MPI protocol, an experienced reader must be available to timely interpret the stress images and determine the need for rest images, but, in routine clinical practice, the majority of cardiologists who interpret MPI studies have other concurrent clinical responsibilities. One solution would be off-site or remote reading systems to review images on other computers or tablets. Another option to overcome the need for immediate stress image review is to take advantage of the two resources that are always available when perfusion imaging is performed: experienced nuclear technologists and automated computer quantification. There is scant literature assessing the ability of nuclear technologists to determine the need for a rest study with stress-first MPI,[9,10] and only tangential literature on automated quantification.[11,12]

Utilizing the resources at hand, technologists and computer technology, we might be able to overcome one of the hurdles preventing the adoption of stress-first protocols in many laboratories. Hence, the purpose of this multicenter study was to assess technologists' and automated computer quantification's ability to correctly identify stress-first MPI studies requiring rest studies.

MATERIALS AND METHODS

Study Design and Patients

We prospectively evaluated images from consecutive patients who were undergoing a clinically indicated stress-first Tc-99m gated SPECT MPI at two tertiary care medical centers over a 4-month period from March 2014 through June 2014. Patients were enrolled at Hartford Hospital, Hartford, CT (a 900-bed urban teaching hospital) and Mount Sinai Hospital, New York, NY (a 1200-bed inner city teaching hospital). This study was approved by the Institutional Review Boards at both medical centers. The decision to perform a stress-first protocol was made by the nuclear cardiology attending based on available clinical and ECG data. Patients with known CAD may have been included based on recent (within past 1-2 years) normal MPI or stress echocardiography, normal or non-obstructive cardiac CT angiography or invasive coronary angiography, or abnormal coronary calcium scores alone.

Stress-first images, with raw image datasets and gated information if available, were reviewed immediately after they were acquired and processed. A single technologist (certified nuclear medicine technologist) blinded to the patients' clinical information reviewed the stress images. Ten technologists who all had experience with stress-first MPI in day-to-day clinical practice with an average of 13.4 ± 5.4 years of work experience participated. Subsequently and similarly, a single board-certified (CBNC) nuclear cardiologist blinded to medical history and stress test results reviewed the same images. Six cardiologists participated with an average of 8.6 ± 6.5 years of post-fellowship work experience. Automated quantification analysis was performed on all stress images in a single batch after all patients were enrolled.

Demographic, clinical, and stress testing variables were prospectively collected in all patients at the time of evaluation. Demographic variables recorded included age, gender, height,

weight, and BMI. Clinical variables included history of diabetes, hypertension, hyperlipidemia, smoking (past or present), family history of CAD, congestive heart failure, documented CAD (which included known CAD by diagnostic testing or patient history, history of myocardial infarction [MI], history of percutaneous coronary intervention [PCI], history of coronary artery bypass grafting [CABG]), and stressor (exercise or pharmacologic) used.

Stress Testing and Imaging Protocols

Standard exercise and pharmacologic protocols as endorsed by the American Society of Nuclear Cardiology (ASNC) were employed.[13] A stress-first imaging sequence was employed using Tc-99m agents. The radionuclide doses were weight adjusted based on patient weight with a low dose given for weight <250 lbs and a high dose used for weight ≥250 lbs. A standard imaging protocol as endorsed by ASNC was used for all patients.[14,15] Patients were imaged using one of two SPECT camera systems: a conventional Na-I SPECT camera (Cardio MD, Philips Healthcare, Andover, MA) or a cadmium-zinc-telluride (CZT) high-efficient SPECT camera (Discovery NM 530c, GE Healthcare, Haifa, Israel). The conventional Na-I SPECT camera routinely used a Gd-153 line source (Vantage Pro, Philips Healthcare, Andover, MA) and the high-efficient SPECT camera used prone in addition to supine imaging for attenuation correction.

Endpoints

Technologists indicated their choice for rest images ''needed'' or ''not needed'' for each patient's stress images.
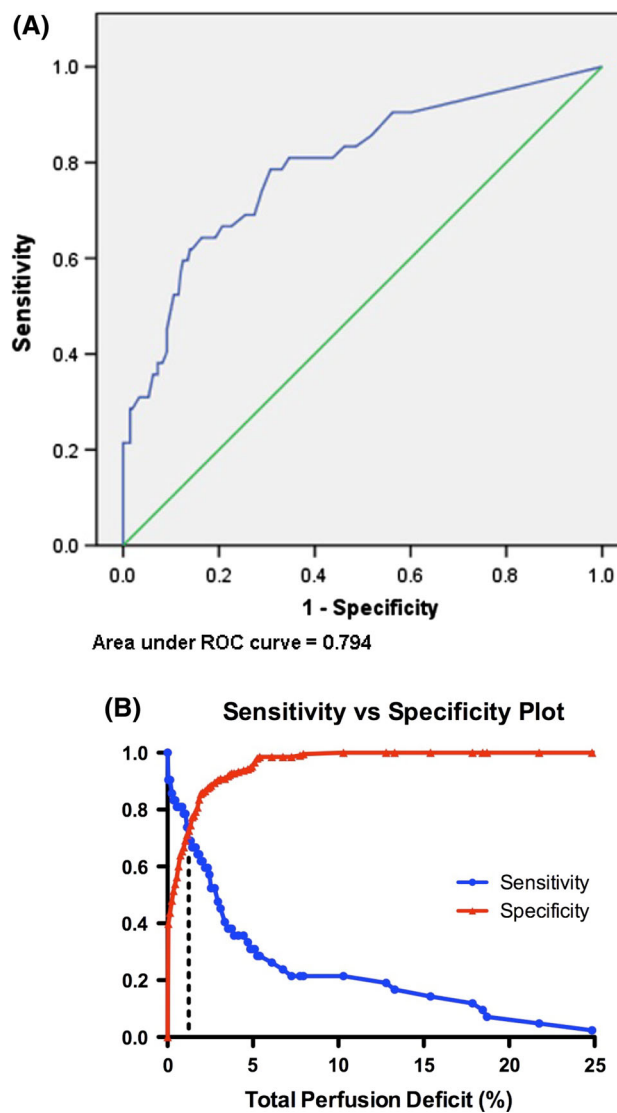
Figure 1. (**A**) ROC curve for the automated stress TPD score compared to the reference standard clinical read's determination of the need for rest images. (**B**) Sensitivity vs specificity plot used to determine the optimal cutoff value for the automated stress TPD.

By definition, rest images were felt to be needed if the combination of stress non-attenuation corrected and attenuation corrected images did not result in normal perfusion by visual inspection. The cardiologists also separately recorded their choice for rest images as ''needed'' or ''not needed'' for each patient. The reference standard for the need for rest images was the decision of the interpreting nuclear cardiologist incorporating patients' clinical presentation, past medical history, and stress test data with visual interpretation of perfusion, raw image data, and gated images. If this clinical interpretation was normal, then rest imaging was not performed, and if it was abnormal, then subsequent rest imaging was performed. If rest images were performed, then they were also incorporated into the reference standard's deliberation as to the need for rest images.

Automated Quantification

Stress supine and supine attenuation corrected (Gd-153 line source) images from the conventional Na-I SPECT camera, and stress supine and prone images from the CZT SPECT camera were quantified separately using their respective, gender-specific normal limits.[16] Automatically generated myocardial contours were evaluated by an experienced technologist, and when necessary, contours were adjusted to correspond to the myocardium. The quantitative perfusion endpoint used was the total perfusion deficit (TPD), which reflects a combination of both the severity and extent of the defect in one parameter.[16] When the stress non-attenuation corrected and stress attenuation corrected (prone or Gd-153 line source) resulted in different TPDs, the lower value was utilized. All quantitative analysis was performed in batch mode of de-identified data by a technologist blinded to any clinical results.

Receiver operating characteristic (ROC) curve analysis was used to determine the sensitivity and specificity of the automated stress TPD score compared to the reference standard clinical read's determination of the need for rest images as well as the total area under the curve (AUC) (Figure 1A). Sensitivity was plotted against specificity to determine the stress TPD which maximizes sensitivity and specificity for determining the need for rest images based on the reference standard clinical read (Figure 1B).

Statistics

Demographic and clinical characteristics were expressed as mean ± standard deviation for continuous variables or as percentages for categorical variables. A two-tailed Student's T test was used to compare means while Chi-squared or Fishers exact test was used to compare proportions. The clinical reference standard interpretation was used as the reference to calculate sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) for the technologists' interpretation, the computer-automated quantification (using the optimal cutoff value based on ROC curve analysis), and the cardiologists' blinded interpretation. In addition, sensitivity, specificity, accuracy, PPV, and NPV were also calculated for TPD score ≥ 2 and ≥ 3. A separate analysis was performed combining the results of the technologist's interpretation and automated computer quantification to simulate the results of a technologist interpreting the stress perfusion images with the aid of automated software. Percent agreement and Cohen's Kappa (κ) statistic were used to analyze agreement between technologists vs clinical reference standard interpretation, automated quantification vs clinical reference standard interpretation, the technologists/computer combined vs clinical reference standard interpretation, and blinded cardiologists vs clinical reference standard interpretation. Kappa values were classified as follows: $<0.20 =$ poor agreement, 0.21 to 0.40 = fair agreement, 0.41 to 0.60 = moderate agreement, 0.61 to 0.80 = good agreement, and 0.81 to 1.00 = excellent agreement. Clinically important agreement was defined a priori as a kappa value >0.50. The criteria for statistical significance were predetermined at $P < .05$. All statistical analyses were performed using SPSS software, version 19 (IBM/SPSS, Armonk, NY, USA).

RESULTS

## Demographics and Clinical Reference Standard Interpretation

A total of 250 stress-first SPECT MPI patients were analyzed (Table 1). The mean age of the patients was 60.9 ± 11.2 years, 113 (45.2%) were male, the average BMI was 29.4 ± 6.8 kg·m^{-2}, and 52.8% underwent exercise stress. The prevalence of hypertension and hyperlipidemia was high in the cohort, at 60.8% and 45.6%, respectively. Consistent with a lower risk stress-first population, the history of diagnosed CAD was only 6% with previous revascularization being infrequent. Of the total cohort, 99 (39.6%) were imaged with a high-efficient SPECT camera, and 151 (60.4%) were imaged with a conventional NA-I SPECT camera.

According to the reference standard nuclear cardiologists' clinical read, 42 (16.8%) out of the 250 stress-first studies were interpreted as ''needing'' rest images, whereas 208 (83.2%) were interpreted as ''not needing'' rest images. The only statistically significant differences between the patients needing rest images and those that did not, were the proportion of patients with a history of diagnosed CAD, and camera type. The greater proportion of patients with history of diagnosed CAD requiring rest images (14.2% vs 4.3%, $P = .02$) consistent with their greater likelihood of having ischemia. The newer high-efficient SPECT camera may benefit from improved count sensitivity and spatial resolution, resulting in the need for rest images less frequently with high-efficient SPECT compared to conventional SPECT (7.1% vs 23.2%, $P = .0009$). BMI was higher in the group needing rest

Table 1. Demographic and clinical characteristics of patients for the entire cohort and based on the need for rest imaging

Characteristic	All patients (N = 250)	Rest-imaging needed (N = 42)	Rest-imaging not needed (N = 208)	P value*
Age	60.9 ± 11.3	62.5 ± 11.7	60.6 ± 11.2	.32
BMI	29.4 ± 6.8	31.2 ± 7.1	29.1 ± 6.7	.07
Gender				.06
Male	113 (45.2%)	25 (59.5%)	88 (42.3%)	
Female	137 (54.8%)	17 (40.5%)	120 (57.7%)	
Cardiac risk factors				
Hypertension	152 (60.8%)	26 (61.9%)	126 (60.6%)	.87
Diabetes	57 (22.8%)	8 (19.0%)	49 (23.5%)	.69
Hyperlipidemia	114 (45.6%)	23 (54.8%)	91 (43.7%)	.26
Family history	59 (23.6%)	9 (21.4%)	50 (24.0%)	.84
H/O CAD	15 (6.0%)	6 (14.2%)	9 (4.3%)	.02
H/O PCI	4 (1.6%)	1 (2.4%)	3 (1.4%)	.52
H/O CABG	1 (0.4%)	0 (0%)	1 (0.5%)	1.0
H/O MI	2 (0.8%)	1 (2.4%)	1 (0.5%)	.07
Stress test factors				
Stressor				.82
Exercise	132 (52.8%)	21 (50.0%)	111 (53.4%)	
Pharmacologic	118 (47.2%)	21 (50.0%)	97 (46.6%)	
Camera type				.0009
Na-I SPECT	151 (60.4%)	35 (83.3%)	116 (55.8%)	
CZT SPECT	99 (39.6%)	7 (16.7%)	92 (44.2%)	

* Comparison between rest-imaging needed and rest-imaging not needed groups.
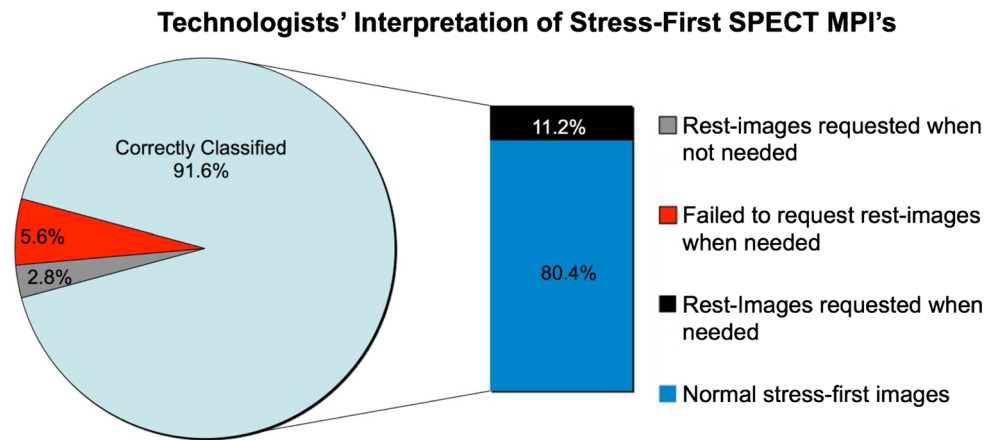
Figure 2. Technologists' interpretation compared to the clinical reference standard.

images (31.2 vs 29.1 kg·m^{-2}), but this did not achieve statistical significance (*P* = .07).

Technologists' Interpretation

Technologists correctly classified 91.6% (229/250) of patient studies as either "needing" or "not needing"

rest images compared to the clinical reference standard (Figure 2). The technologists requested rest images in 35 (14%) of the stress-first studies. They correctly identified 28 (11.2%) studies "needing" rest images and 201 (80.4%) studies "not needing" rest images. They failed to request rest images in 14 (5.6%) studies when needed and wrongly requested rest images in 7 (2.8%)

Table 2. Comparison of correctly and incorrectly classified patients by technologists

Characteristic	Correctly classified (N = 229)	Incorrectly classified (N = 21)	P value
Age	61.0 ± 11.4	59.3 ± 10.5	.51
BMI	29.3 ± 6.9	30.4 ± 5.5	.48
Gender			.12
Male	100 (43.7%)	13 (61.9%)	
Female	129 (56.3%)	8 (38.1%)	
Cardiac risk factors			
Diabetes	55 (24.0%)	2 (9.5%)	
Hypertension	139 (60.7%)	13 (61.9%)	.91
Hyperlipidemia	104 (45.4%)	10 (47.6%)	.85
Family history	56 (24.5%)	3 (14.3%)	.42
H/O CAD	13 (5.7%)	2 (9.5%)	.36
H/O PCI	4 (1.7%)	0 (0%)	1.0
H/O CABG	1 (0.4%)	0 (0%)	1.0
H/O MI	2 (0.9%)	0 (0%)	1.0
Stress test factors			
Stressor			.37
Exercise	123 (53.7%)	9 (42.9%)	
Pharmacologic	106 (46.3%)	12 (57.1%)	
Camera type			.64
Na-I SPECT	137 (59.8%)	14 (66.7%)	
CZT SPECT	92 (40.2%)	7 (33.3%)	

studies when not needed. Their sensitivity, specificity, PPV, and NPV were 66.7%, 96.6%, 80.0%, and 93.5%, respectively. There were no demographic or clinical characteristics significantly associated with correct or incorrect stress-first image interpretation by the technologists (Table 2).

The 21 studies incorrectly classified by the technologists were reviewed. Of the 14 where rest images were not requested, 7 involved small defects of the anterior or inferior wall with a SSS ≤ 3, 5 involved focal defects in the apex, 1 was a patient with normal perfusion but decreased left ventricular systolic function, and 1 involved a medium-sized defect. Of the 7 misclassified cases where rest images were requested by technologists when not needed, 5 had small defects with a SSS of 1 and 2 had defects which normalized with attenuation correction. Of note, half of the 14 patients that the technologists felt did not need the rest images were eventually interpreted as normal with the addition of the rest images.

Blinded Cardiologists' Interpretation

To objectively determine how well the ''expert'' cardiologist adjudicated the need for rest images in a manner similar to the technologists, the clinical cardiologists' interpretations when blinded to clinical and stress test data were compared to the clinical reference standard.

They correctly classified 93.6% (234/250) of patient studies as either ''needing'' or ''not needing'' rest images compared to the clinical reference standard (Figure 3). The cardiologists requested rest images in 40 (16.0%) of the stress-first studies. They correctly identified 33 (13.2%) studies ''needing'' rest images and 201 (80.4%) studies ''not needing'' rest images. The blinded cardiologist failed to request rest images in 9 (3.6%) studies when needed and wrongly requested rest images in 7 (2.8%) studies when not needed. Their sensitivity, specificity, PPV, and NPV were 78.6%, 96.6%, 82.5, and 95.7%, respectively. Agreement between technologists and blinded cardiologists with stress-first image classification was good (91.6% [229/250], κ = 0.671 ± 0.067).

Automated Computer Quantification

ROC curve analysis demonstrated that an automated stress TPD score ≥1.2 was associated with optimal sensitivity and specificity (Figure 1). Based on a TPD score ≥1.2, the computer correctly classified 71.6% (179/250) of patient studies as either ''needing'' or ''not needing'' rest images compared to the clinical reference standard (Figure 4). The computer requested rest images in 91 (36.4%) of the stress-first studies. The computer correctly identified 31 (12.4%) studies ''needing'' rest images and 148 (59.2%) studies ''not needing'' rest

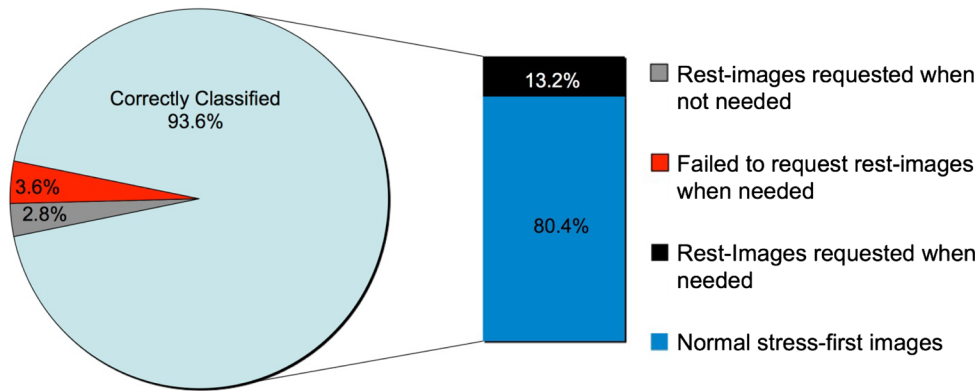**Cardiologists' blinded Interpretation of Stress-First SPECT MPI's**

Figure 3. Cardiologists' blinded interpretation compared to the clinical reference standard.

Computer's Interpretation of Stress-First SPECT MPI's
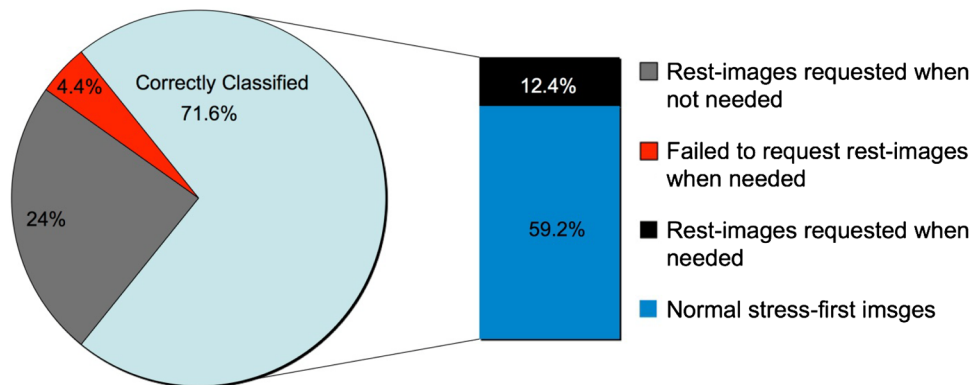

Figure 4. Computer's interpretation compared to the clinical reference standard using an automated TPD score ≥ 1.2.

images. The automated quantification with a TPD cutoff of ≥ 1.2 failed to request rest images in 11 (4.4%) studies when needed and wrongly requested rest images in 60 (24%) studies when not needed. Its sensitivity, specificity, PPV, and NPV were 73.8%, 71.2%, 34.1%, and 93.1%, respectively. Male gender and pharmacologic stress were more frequently classified incorrectly by the automated quantification (Table 3).

We examined the effect of varying the automated TPD cutoff value used for decision making on the overall accuracy. Using an automated stress TPD score ≥ 2, the computer correctly classified 81.6% (204/250) of patient studies. The higher cutoff resulted in lower sensitivity but higher specificity (higher PPV but lower NPV). Its sensitivity, specificity, PPV, and NPV were 61.9%, 85.6%, 46.4%, and 91.8%, respectively (Table 4). Using a cutoff of ≥ 3, the computer correctly classified 83.2% (208/250) of patient studies. The trend of lower sensitivity and higher specificity with higher

cutoffs continued. Its sensitivity, specificity, PPV, and NPV were 47.6%, 90.4%, 50.0%, and 89.5%, respectively.

Composite Analysis of Technologists' and Computer's Classification

To simulate the scenario in which a technologist would have access to the real-time automated computer quantification when making decisions regarding the need for rest images, we created a combined model of assessment. In this model, the computer or technologist was able to correct for the other's incorrect classification such that if either the technologist or the automated quantification chose the correct rest-imaging decision, then it was counted as an agreement with the clinical reference standard. Using these rules, the technologist and computer correctly classified 96.8% (242/250) of patient studies. The composite sensitivity, specificity, PPV, and

Table 3. Comparison of correctly and incorrectly classified patients by automated computer quantification

Characteristic	Correctly classified (N = 179)	Incorrectly classified (N = 71)	*P* value
Age	60.4 ± 11.3	62.3 ± 11.3	.23
BMI	29.5 ± 6.7	29.3 ± 7.3	.84
Gender			.008
Male	71 (39.7%)	42 (59.2%)	
Female	108 (60.3%)	29 (40.8%)	
Cardiac risk factors			
Diabetes	38 (21.2%)	19 (26.8%)	.44
Hypertension	110 (61.5%)	42 (59.2%)	.85
Hyperlipidemia	87 (48.6%)	27 (38%)	.17
Family history	9 (5.0%)	6 (8.5%)	.38
H/O CAD	9 (5%)	6 (8.5%)	.83
H/O PCI	4 (2.2%)	0 (0%)	.58
H/O CABG	0 (0%)	1 (1.4%)	.28
H/O MI	2 (1.1%)	0 (0%)	.56
Stress test factors			
Stressor			.002
Exercise	106 (59.2%)	26 (36.6%)	
Pharmacologic	73 (40.8%)	45 (63.4%)	
Camera type			.49
Na-I SPECT	111 (62.0%)	40 (56.3%)	
CZT SPECT	68 (38.0%)	31 (43.7%)	

Table 4. Sensitivity, specificity, accuracy, positive predictive value, and negative predictive value associated with various automated quantification TPD cutoff values (≥1.2, ≥2, and ≥3)

| | Automated quantification cutoff value | | |
	TPD ≥ 1.2	TPD ≥ 2	TPD ≥ 3
Sensitivity (%)	73.8	61.9	47.6
Specificity (%)	71.2	85.6	90.4
Accuracy (%)	71.6	81.6	83.2
Positive predictive value (%)	34.1	46.4	50.0
Negative predictive value (%)	93.1	91.8	89.5

NPV were 83.3%, 99.5%, 97.2%, and 96.7%, respectively. The technologists misclassified 21 stress-first studies, and the automated quantification corrected 13 (61.9%) of them. They classified 6 studies as ''needing'' rest images and 7 studies as ''not needing'' rest images, the automated quantification corrected them in those 13 cases. On the other hand, the computer-misclassified 71 stress-first studies and technologists corrected 60 (84.5%) of them. All 60 of those studies were classified as ''needing'' rest images according to the automated quantification and technologists corrected those studies.

Comparison Between Conventional and High-Efficient SPECT Cameras

As two SPECT camera types with two different methods of attenuation correction were used in this study, it was possible that correct categorization of stress images could vary based on the camera type and assessment technique. Both groups of patients were similar, except for slightly higher age, lower BMI, and greater proportion of pharmacologic stress in the high-efficient camera patient group (Table 5). Compared to the clinical

Table 5. Demographic and clinical characteristics of patients based on SPECT camera type

Characteristic	Conventional Na-I camera (N = 151)	High-efficiency CZT camera (N = 99)	P value
Age	59.2 ± 11.4	63.5 ± 10.7	.003
BMI	30.7 ± 7.7	27.5 ± 4.6	.0002
Gender			.95
Male	69 (45.7%)	44 (44.4%)	
Female	82 (54.3%)	55 (55.6%)	
Cardiac risk factors			
Diabetes	31 (20.5%)	26 (26.3%)	.37
Hypertension	91 (60.3%)	61 (61.6%)	.94
Hyperlipidemia	65 (43.0%)	49 (49.5%)	.38
Family history	43 (28.5%)	16 (16.16%)	.04
H/O CAD	6 (4.0%)	9 (9.1%)	.11
H/O PCI	2 (1.3%)	2 (2.0%)	.65
H/O CABG	1 (0.6%)	0 (0%)	1.0
H/O MI	2 (1.3%)	1 (1.0%)	1.0
Stressor			.005
Exercise	91 (60.3%)	41 (41.4%)	
Pharmacologic	60 (39.7%)	58 (58.6%)	

Table 6. Classification of the stress images compared to the clinical reference standard based on SPECT camera type

Classification	High-efficiency camera (N = 99)	Conventional camera (N = 151)	P value
Technologists			.64
Correct	92 (92.9%)	137 (90.7%)	
Incorrect	7 (7.1%)	14 (9.3%)	
Automated quantification			.47
Correct	68 (68.7%)	111 (73.5%)	
Incorrect	31 (31.3%)	40 (26.5%)	
Blinded cardiologist			.38
Correct	91 (91.9%)	143 (94.7%)	
Incorrect	8 (8.1%)	8 (5.3%)	

reference standard, there was no difference in the ability of technologist ($P = .64$), automated quantification ($P = .47$), and blinded cardiologist ($P = .38$) to correctly classify the stress images (Table 6).

DISCUSSION

Our results demonstrate that nuclear technologists had a high degree of agreement with the clinical reference standard read performed by the board-certified nuclear cardiologists. Technologists were able to correctly classify 91.6% of stress-first SPECT studies, resulting in a good degree of agreement with a kappa value of 0.68. This performance by the technologists compares favorably with the 93.6% of studies (a kappa value of 0.77) correctly identified by the cardiologists blinded to patients' stress test data and clinical presentation. While the automated computer quantification alone was less successful with an accuracy of 83.2%, when it was used in conjunction with the technologists' visual assessment, an accuracy of 96.8% was achieved. This high degree of accuracy by technologists and automated quantification highlights the possibility of overcoming the hurdle of the need for immediate image review by a physician which may prevent the adoption of stress-first protocols in many laboratories.

To our knowledge, only two previous studies have investigated a technologist's role in determining the need for rest images in a stress-first protocol, and none have employed automated computer quantification. A retrospective study of 121 patients found an accuracy of 88.8% for the technologists agreement with a consensus read reference standard for determining the need for rest images in a stress-first MPI protocol.[10] However, 14 of the patients were excluded from analysis as the technologists were able to choose an ''uncertain'' categorization if they wanted a physician's input. In a very interesting epilog, this particular laboratory in Malmö, Sweden converted to the routine of having the technologists determines the need for rest images in general practice and found that in a specified period of time, 56.4% of studies required rest images compared to 56.2% during an earlier period of time when physicians made the determination. The technologists in this study received formalized training for reading the stress-first images similar to that of physicians which represents an excellent approach to further standardize quality and improve accuracy in the laboratory. A second large, retrospective study of 532 patients from Göteborg, Sweden compared technologists and physicians to a reference standard interpretation.[9] The technologists received a brief tutorial and were given guidelines for interpretation of the stress images. There was excellent agreement to the reference standard in patients who required a rest study for the technologists (99%) and the physicians (98%). However, in patients who did not need a rest study based on the reference standard evaluation, technologists only correctly classified 21% of patients and physicians 32%. This study had several limitations including the lack of attenuation correction and variable access to clinical data during interpretations. In addition, the reference standard was based on the past clinical interpretation of the rest and stress images based on the presence of ischemia or infarction, which may misclassify fixed defects felt to be attenuation as not needing rest images, when if the stress images alone were evaluated, they would need rest images.

The successful utilization of stress-first protocols relies on accurate identification of the studies requiring rest imaging, and therefore, it is crucial not to miss patients who would require rest imaging as they would need to be recalled to the laboratory on a later date. Both the technologists and automated quantification had high NPV, 93.5% and 93.1%, respectively. In addition, the classification of stress-first studies by the technologists compared favorably to that of the blinded cardiologists. Of the misclassified studies by technologists, most were cases of not requesting rest images when needed and perfusion findings in these studies frequently were small/mild defects. Of the few misclassified cases where rest images were requested by technologists when not needed, perfusion in these studies frequently was made ''normal'' enough for the cardiologist with attenuation correction but not for the technologist. The discrepancy between technologists and the clinical cardiologists' interpretation may be due to several factors such as availability of stress test data and clinical risk factors as well as a higher skill level with interpretation, particularly with poor-quality images and the addition of attenuation correction. With clinical information available, readers may be more likely to overlook small perfusion defects as artifact in low-likelihood patients. The converse might be true in high-risk patients. The pre-test likelihood clearly affects a reader's comfort in passing on, or requesting, rest images. While very small perfusion defects are unlikely to appreciably affect a patient's prognosis, they may affect the initiation of downstream medical therapy.

Notably, in the composite analysis, when technologists corrected the computer and vice versa, the NPV increased to 96.7%. Only 8 out of 250 stress-first studies were incorrectly classified by the composite model highlighting the ability of nuclear technologists using appropriate computer technology to appropriately screen stress-first images. Of those 8 incorrectly classified studies, 7 studies needed rest images and were missed by both the computer and technologists. Individually the technologists failed to request rest images 5.6% of the time and the computer 4.4% of the time. There were no clinical or demographic variables significantly associated with an incorrect classification by technologists. The computer incorrectly classified more men than women and more patients undergoing pharmacologic stress than exercise stress. This brings up the need for appropriate stress-first, gender-specific, and perhaps stressor-specific normal limits for the automated quantification. There was less of a need for rest images in patients imaged with the high-efficient SPECT camera, which suggested improved image quality and less attenuation artifact from improved count sensitivity and spatial resolution. However, camera type and method of attenuation correction did not affect the technologists', the computer's, or the blinded cardiologist's ability to correctly classify stress-first images as there was no difference in the proportion of correct and incorrect classification between cameras, $P = .64$, $P = .47$, and $P = .38$, respectively. Implementing a formal education program for the technologists as utilized in previous studies is likely to further improve their ability to accurately screen stress-first studies.

There are several important advantages to stress-only protocols. A stress-only study can be completed, processed, and read in <90 minutes as opposed to the usual 3-4 hours required for a rest-stress study, saving

time for both the patient and for the laboratory. There is marked reduction in patient radiation exposure by eliminating the need for a rest Tc-99m dose. Compared to a traditional 10 mCi/30 mCi rest-stress protocol, there is a 27%-76% reduction in effective dose with stress-first protocols.[8] These improvements come with the same prognostic information which we have come to expect from traditional rest-stress studies.[1,2] There is also a cost savings to the system by eliminating the second Tc-99m dose and time recovered from eliminating subsequent rest imaging. Technologist and computer review of stress-first images may help overcome one of the major hurdles to broader adaptation of stress-first protocols, the need for real-time review of stress images. Two other substantial hurdles remain in the need for attenuation correction and concerns about low reimbursement in the non-hospital setting.

Advances in the accuracy of the automated quantification software are also possible making it even more clinically useful. In particular, this study does not use a combined supine and prone or combined non-attenuation corrected and attenuation corrected total perfusion defect quantitative analysis, which has been shown to increase the diagnostic accuracy of the technique.[17,18] Even more advanced machine learning techniques could be employed in this setting to augment the computer's accuracy by incorporating other quantitative or clinical information.[19,20] The TPD cutoff value used with the automated quantification can be chosen to maximize either negative or PPV. The lower the TPD value, the higher the NPV, meaning that more rest images will be requested than are likely needed. Using a TPD cutoff of $\geq 1.2\%$ in this study resulted in 60 unnecessary rest studies but very few ''missed'' rest images. The computer, as well as the technologist's, threshold in determining the need for rest images can be adjusted, so that the sensitivity and specificity are adjusted to laboratory preferences. This preference is likely to depend on the laboratory's patient population, such as its mix of inpatients and outpatients, average travel time to the laboratory for outpatients, the pre-test likelihood of disease based on cardiac risk factors, and the expected normalcy rate in the laboratory, and may even change on a patient to patient basis. One may be willing to accept a low sensitivity but high specificity for an inpatient who is not expected to leave the hospital soon, while a high sensitivity and lower specificity may be desirable for an outpatient with a long commute to the imaging laboratory. A prospective evaluation of this next-generation software alone and in conjunction with technologists' visual assessment would be the next step in the development of this processes.

LIMITATIONS

The results may not be immediately applicable to all locations as the nuclear technologists who participated in this study individually had many years of experience with stress-first imaging and inexperienced or new technologists may not have the same facility with stress-first image interpretation. However, in practice, a standardized educational curriculum for technologists on the subject matter may be able to compensate for lack of clinical experience. The study design was such that the clinical decision to perform rest imaging did not hinge on the technologists stress-first image classification and may have influenced their decision regarding the need for rest imaging. The workflow employed in this study may not be the same as that used in general nuclear medicine laboratories and may vary from that used in private outpatient cardiology practices. No external standard such as coronary angiography or prognostic data was used as a reference standard to assess the accuracy of the clinical MPI interpretations. The model used for the composite technologist and automated reads was a post hoc analysis and represents a best-case scenario for use and not necessarily a real-world implementation. In a real-world assessment of the composite approach, one would need to determine on a case-by-case basis, which technique to use in the case of discrepancies. However, the combination of unbiased technology and extensive practical experience appears to be the most effective method for successfully implementing a stress-first protocol and should be the subject of future studies. Information on patients' current or past smoking status was incomplete and not included in the analysis. While significant differences between both technologists and blinded cardiologists vs automated computer quantification in determining the need for rest imaging were observed, our study was underpowered to detect significant differences between technologists and blinded cardiologists and to do so would require approximately 2675 patient MPI studies.

NEW KNOWLEDGE GAINED

With the increasing prevalence of normal MPI studies and competition from other non-invasive imaging modalities, it is imperative that the field of nuclear cardiology develops more efficient, cost effective, and radiation sparing strategies for the initial evaluation of patients, and stress-first protocols represent an attractive option. A combination of technologist's review and automated quantification software could potentially greatly expand the use of stress-first imaging to sites where the cardiologist is not immediately available to interpret stress images, thereby overcoming one of the hurdles preventing the adoption of stress-first protocols.

CONCLUSION

Technologists and automated quantification software had a high degree of agreement with the clinical reference standard for determining the need for rest images in a stress-first imaging protocol with NPVs of 93.5% and 93.1%, respectively. Combining the two resources, further improved the accuracy. Utilizing an experienced technologist and automated systems to screen stress-first images could potentially greatly expand the use of stress-first imaging to sites where the cardiologist is not immediately available to interpret stress images while maintaining accurate diagnostic capabilities.

References

1. Chang SM, Nabi F, Xu J, Raza U, Mahmarian JJ. Normal stress-only versus standard stress/rest myocardial perfusion imaging: Similar patient mortality with reduced radiation exposure. J Am Coll Cardiol 2010;55:221-30.
2. Duvall WL, Wijetunga MN, Klein TM, Razzouk L, Godbold J, Croft LB, et al. The prognosis of a normal stress-only Tc-99m myocardial perfusion imaging study. J Nucl Cardiol 2010;17:370-7.
3. Gowd BM, Heller GV, Parker MW. Stress-only SPECT myocardial perfusion imaging: A review. J Nucl Cardiol 2014;21:1200-12.
4. Duvall WL, Guma KA, Kamen J, Croft LB, Parides M, George T, et al. Reduction in occupational and patient radiation exposure from myocardial perfusion imaging: Impact of stress-only imaging and high-efficiency SPECT camera technology. J Nucl Med 2013;54:1251-7.
5. Rozanski A, Gransar H, Hayes SW, Min J, Friedman JD, Thomson LE, et al. Temporal trends in the frequency of inducible myocardial ischemia during cardiac stress testing: 1991 to 2009. J Am Coll Cardiol 2013;61:1054-65.
6. Duvall WL, Rai M, Ahlberg AW, O'Sullivan DM, Henzlova MJ. A multi-center assessment of the temporal trends in myocardial perfusion imaging. J Nucl Cardiol 2015. doi:10.1007/s12350-014-0051-x.
7. Thompson RC, Allam AH. More risk factors, less ischemia, and the relevance of MPI testing. J Nucl Cardiol 2015;22:552-4.
8. Hussain N, Parker MW, Henzlova MJ, Duvall WL. Stress-first myocardial perfusion imaging. Cardiol Clin 2015. doi: 10.1016/j.ccl.2015.06.006.
9. Johansson L, Lomsky M, Gjertsson P, Sallerup-Reid M, Johansson J, Ahlin NG, et al. Can nuclear medicine technologists assess whether a myocardial perfusion rest study is required? J Nucl Med Technol. 2008;36:181-5.
10. Tragardh E, Johansson L, Olofsson C, Valind S, Edenbrandt L. Nuclear medicine technologists are able to accurately determine when a myocardial perfusion rest study is necessary. BMC Med Inform Decis Mak 2012;12:97.
11. Sharir T, Pinsky M, Pardes A, Rochman A, Prochorov V, Kovalski G, et al. Comparison of the diagnostic accuracy of very low stress-dose to standard-dose myocardial perfusion imaging: Automated quantification of one-day, stress-first SPECT using a CZT camera. J Nucl Cardiol 2015. doi:10.1007/s12350-015-0130-7.
12. Nakazato R, Berman DS, Hayes SW, Fish M, Padgett R, Xu Y, et al. Myocardial perfusion imaging with a solid-state camera: Simulation of a very low dose imaging protocol. J Nucl Med 2013;54:373-9.
13. Henzlova MJ, Cerqueira MD, Hansen CL, Taillefer R, Yao SS. ASNC Imaging Guidelines for Nuclear Cardiology Procedures: Stress protocols and tracers. J Nucl Cardiol 2009;16:331.
14. Hansen CL, Goldstein RA, Berman DS, Churchwell KB, Cooke CD, Corbett JR, et al. Myocardial perfusion and function single photon emission computed tomography. J Nucl Cardiol 2006;13:e97-120.
15. Holly TA, Abbott BG, Al-Mallah M, Calnon DA, Cohen MC, DiFilippo FP, et al. Single photon-emission computed tomography. J Nucl Cardiol 2010;17:941-73.
16. Slomka PJ, Nishina H, Berman DS, Akincioglu C, Abidov A, Friedman JD, et al. Automated quantification of myocardial perfusion SPECT using simplified normal limits. J Nucl Cardiol 2005;12:66-77.
17. Nishina H, Slomka PJ, Abidov A, Yoda S, Akincioglu C, Kang X, 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;47:51-8.
18. Xu Y, Fish M, Gerlach J, Lemley M, Berman DS, Germano G, et al. Combined quantitative analysis of attenuation corrected and non-corrected myocardial perfusion SPECT: Method development and clinical validation. J Nucl Cardiol 2010;17:591-9.
19. Arsanjani R, Xu Y, Dey D, Fish M, Dorbala S, Hayes S, et al. Improved accuracy of myocardial perfusion SPECT for the detection of coronary artery disease using a support vector machine algorithm. J Nucl Med 2013;54:549-55.
20. Arsanjani R, Xu Y, Dey D, Vahistha V, Shalev A, Nakanishi R, et al. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. J Nucl Cardiol 2013;20:553-62.