JACC: CARDIOVASCULAR IMAGING VOL. 13, NO. 3, 2020

PUBLISHED BY ELSEVIER

EDITORIAL COMMENT

Improving Prognostic Accuracy of Myocardial Perfusion Imaging With Quantitative Assessment*

Todd D. Miller, MD,[a] Michael K. O'Connor, PhD[b]

Although quantitative techniques for assessing noninvasive and invasive cardiac imaging have been available for years, the cardiology community has been relatively slow to embrace these techniques. The subjective interpretation of images remains the most commonly applied standard in clinical practice. Two parameters that determine the value of cardiac imaging are reproducibility (shooting an arrow multiple times and hitting the same spot on the target) and accuracy (shooting an arrow and hitting the bullseye). For nuclear myocardial perfusion imaging (MPI), quantitative analysis has been shown to enhance reproducibility (1-6). To assess accuracy, the results of imaging need to be compared with an accepted gold standard such as coronary angiography (7) or patient outcomes.

SEE PAGE 774

In this issue of *iJACC*, Otaki et al. (8) report the results of a comparison between automated MPI and subjective visual analysis for predicting major adverse cardiac events (MACE) in patients with possible or established coronary artery disease. The authors analyzed MPI in 19,495 patients enrolled in the REFINE SPECT (REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT). This registry consists of 5 multinational medical centers that performed technetium-labeled single-photon emission computed tomography (SPECT) MPI using ultrafast imaging (solid-state detectors) camera systems. Subjective visual analysis was performed at each site by experienced board-certified nuclear cardiologists who had access to clinical and functional data and also site-specific quantification software to aid in the generation of the clinical report. Quantitative analysis was performed for all images at the Cedars-Sinai Medical Center (Los Angeles, California) nuclear cardiology core laboratory by experienced laboratory technologists using standard laboratory software. Patients who underwent early revascularization (≤90 days following MPI) were excluded from analysis. Follow-up data were collected over an average of 4.5 years, applying conventional methodology. MACE occurred in 14.2% of the population and included all-cause mortality (n = 1,341), nonfatal myocardial infarction (MI) (n = 187), unstable angina (n = 207), and late revascularization (n = 1,025).

For visual analysis, the authors applied commonly accepted categories of normal, probably normal, equivocal, and abnormal. In a subset of 13,533 patients, in whom the images were scored on-site using the 17-segment model (9), images were also categorized using the stress image summed stress score (SSS) of 0 = normal, 1 = probably normal, 2 to 3 = equivocal, or ≥4 = abnormal. For quantitative analysis, the authors developed approximate similar categories using the stress image total perfusion deficit (TPD) as a percentage of myocardium: where normal = TPD <1%; probably normal = TPD ≥1% to <3%; equivocal = TPD ≥3% to <5%; and abnormal = TPD ≥5%. For some analyses TPD of 0% was considered "super normal," and the abnormal category was further divided into TPD ≥5% to ≤10% and TPD >10% (to approximate SSS categories 4 to 6 and ≥7, respectively).

The results of this study demonstrate that both visual and quantitative interpretation could accurately stratify the risk for this population. For both methods, there were statistically significant, clinically meaningful, progressively higher annual MACE rates across imaging categories: for subjective analysis, 2.0% with normal increasing to 7.4% with

*Editorials published in *JACC: Cardiovascular Imaging* reflect the views of the authors and do not necessarily represent the views of *iJACC* or the American College of Cardiology.

From the [a]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota; and the [b]Department of Radiology, Mayo Clinic, Rochester, Minnesota. Both authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ISSN 1936-878X/$36.00 https://doi.org/10.1016/j.jcmg.2019.04.009

abnormal images ($p < 0.0001$); for quantitative analysis, 1.3% if TPD = 0% increasing to 7.8% if TPD ≥10% ($p < 0.0001$). Similar results were seen for increasing adjusted hazard ratios (HRs) for MACE in Cox multivariable models: for visual assessment of normal, HR 1.00, increasing to abnormal, HR 1.78 (95% confidence interval [CI]: 1.60 to 1.98; $p < 0.001$). For quantitative assessment TPD of 0, HR 1.00 increasing to TPD >10% (HR: 2.46; 95% CI: 1.99 to 3.05; $p < 0.001$). Interestingly, however, neither technique alone performed better for MACE prediction than the other (area under the curve [AUC] for visual was 0.635 vs. 0.641 for TPD; $p = 0.25$), but a combined model incorporating both techniques performed better than either one alone (AUC for visual plus TPD was 0.658; $p < 0.0001$ vs. TPD or visual). Multiple additional analyses using ischemic assessment in place of the stress images alone, incorporating the summed scoring system approach, or restricting the outcome endpoint analysis to hard events only (death or nonfatal MI) produced qualitatively similar results.

Otaki et al. (8) are to be congratulated for assimilating the large REFINE SPECT database and for performing this relevant prognostic study. Their study has generated some important findings. The authors demonstrated that both visual analysis and quantitative assessment individually provide accurate, meaningful risk stratification using the newer ultrafast camera imaging systems. The prognostic accuracy of visual analysis reported in this study was likely due to highly experienced readers working in laboratories with extensive experience in nuclear cardiology. Performance results in smaller, less-experienced laboratories may not be as good. In those settings, quantitative assessment may be more accurate than visual image interpretation. Also, the authors did not attempt to discern any differences in the performance of the 2 types of cadmium-zinc-telluride (CZT) gamma photon detector cameras. Because the D-SPECT (Spectrum Dynamics, Caesarea, Israel) and NM 530c (GE Healthcare, Haifa, Israel) models have dramatically different detector geometries, it is possible that some differences in accuracy may exist between them. From the description provided, neither detector type incorporated attenuation correction, and it is unclear how many patients were imaged in multiple positions (supine or upright for D-SPECT and supine or prone for the NM 530c unit). Imaging in multiple positions has been shown to improve specificity with the D-SPECT system (10). Reduction in imaging artifacts should further improve image accuracy.

Although an extensive body of research demonstrates the prognostic value of MPI (11,12), most studies used older, rotating gamma camera systems. For comparative purposes of the results from the current study performed using CZT detectors to older studies that used NaI detectors, it is useful to focus on the endpoint of hard events. The research consensus is that patients with normal images should have an annual event rate for cardiac death or nonfatal MI no >1%. In this study, the hard event rates for visually normal images were slightly higher and ranged between 1.1% and 1.5% for stress TPD categories <1% to <5% and were 2.6% for TPD ≥5%. These higher event rates in the current study compared to those from older studies can be explained most readily by the use of all-cause mortality in this study versus cardiac death in earlier studies.

Perhaps the most interesting finding from this study is that the prognostic accuracy of neither approach was superior to the other, but both approaches combined were better than either one alone. This finding is best illustrated in the authors' Central Illustration (8), where the annual MACE rate was highest at 8.0% for the patient subset with both subjectively abnormal images and TPD ≥5% and was lowest at 1.6% for the subset with both normal images and TPD <1%. The scoring system that the authors developed trying to match categories of visual analysis to categories of quantitative assessment (e.g., visual normal to TPD <1%, visual probably normal to TPD ≥1% to <3% and so forth) was rational but imperfect. Approximately one-half (50.5%) of the study population did not match to the same category. Although there might not have been much difference for a patient assigned to the category adjacent to the category of designed match (e.g., a patient with visually normal results with a TPD category ≥1% to <3% instead of TPD category <1%), the extremes of mismatched categories are more intriguing. In this study, there were 72 patients (0.4%) who had visually abnormal studies that were determined to be normal (TPD <1%) by quantitative analysis. What were the experienced readers seeing on these images that the quantitative program did not flag as abnormal? There was a much larger number of patients (5.4% [n =1,062]) who had definitely abnormal images (TPD ≥5%) by quantitative analysis that were interpreted as visually normal. Unfortunately, the authors did not provide any details for these mismatched groups at the opposite ends of the visual and quantitative scoring systems. The reader can only speculate as to the reason(s) for the discrepancy in image interpretation in these patients.

Another interesting issue that these data raise is the apparent discrepancy between the results for quantitative analysis of MPI generated on-site versus the results produced by the core laboratory. Recall that the visual analysis performed on-site

JACC: CARDIOVASCULAR IMAGING, VOL. 13, NO. 3, 2020
MARCH 2020:786–9

encompassed not only visual inspection of the images but also awareness of all clinical and hemodynamic data and results from site-specific quantification software. The authors do not state how many patients were studied at each center and what site-specific quantification software programs were used. Presumably some centers, such as Cedars-Sinai, used the same software program that was applied in the core laboratory to provide the on-site nuclear cardiologists with quantification results as part of the visual analysis interpretation. There are several possibilities to explain the approximately 50% mismatch between visual and quantitative analysis category assignments, including the fact that the same software program may generate different results if the images were processed differently (e.g., selection of different contours) by on-site versus core laboratory technologists; the different quantitation software programs used on-site versus the core laboratory program could produce different results; the on-site nuclear cardiologists generating the final clinical report might have been more strongly influenced by other factors (subjective interpretation of the images, clinical and hemodynamic data, and others) that weighed more heavily than the available on-site quantification results. It would be interesting to know how many patients were enrolled at each site and what on-site software quantitation programs were used, but even knowing this information would not provide a complete answer to this issue.

How should the results of this study be applied in clinical practice? Since the combination of visual and quantitative analysis provided more accurate prognostication than either approach alone, the logical answer is that both techniques should be used. By visual analysis the majority of patients (82% [n = 16,037]) had normal or nearly normal images (visual categories of normal, probably normal, and equivocal). Stress MPI has been shown to function as an effective gatekeeper for coronary angiography and revascularization (13). Very few patients with normal or nearly normal images are referred for coronary angiography. In this study, most of the patients in these visual categories did not have events. It is likely that many of the mild abnormalities seen on studies graded as probably normal or equivocal represented imaging artifacts. However, some of these patients had coronary artery disease that was not visually apparent or only minimally or mildly detectable. The addition of TPD resulted in improved risk stratification, with annual MACE rates progressively increasing across TPD categories and being 2 to 3 times higher for TPD ≥5% versus TPD <1%. For these 3 visual categories, annual hard event rates were clinically meaningful between 2.5% and 3.2% for the TPD category ≥5%, indicating that visual analysis alone failed to detect some high-risk patients among these generally low-risk patient subsets. Although unproven, correct identification of these patients offers the possibility of applying different treatment strategies that could potentially lower their event rates. This group of authors has previously reported that quantitative assessment of MPI can improve reproducibility (4,6) and diagnostic accuracy in patients referred for coronary angiography (7). In this study they demonstrate the additional value of quantitative analysis for accurate prognostication.

ADDRESS FOR CORRESPONDENCE: Dr. Todd D. Miller, Mayo Clinic, 200 First Street Southwest, Rochester, Minnesota 55905. E-mail: miller.todd@mayo.edu.

REFERENCES

1. Mahmarian JJ, Moyé LA, Verani MS, Bloom MF, Pratt CM. High reproducibility of myocardial perfusion defects in patients undergoing serial exercise thallium-201 tomography. Am J Cardiol 1995;75:1116–9.

2. Cerqueira MD, Nguyen P, Staehr P, Underwood SR, Iskandrian AE. Effects of age, gender, obesity, and diabetes on the efficacy and safety of the selective A2A agonist regadenoson versus adenosine in myocardial perfusion imaging: integrated ADVANCE-MPI trial results. J Am Coll Cardiol Img 2008;1:307–16.

3. Mahmarian JJ, Cerqueira MD, Iskandrian AE, et al. Regadenoson induces comparable left ventricular perfusion defects as adenosine: a quantitative analysis from the ADVANCE MPI 2 trial. J Am Coll Cardiol Img 2009;2:959–68.

4. 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 Nuclear Cardiol 2009;16:45–53.

5. Iskandrian A, Garcia E, Faber T, Mahmarian J. Automated assessment of serial SPECT myocardial perfusion images 2009;16:6–9.

6. Xu Y, Hayes S, Ali I, et al. Automatic and visual reproducibility of perfusion and function measures for myocardial perfusion SPECT. J Nucl Cardiol 2010;17:1050–7.

7. Arsanjani R, Xu Y, Hayes SW, et al. Comparison of fully automated computer analysis and visual scoring for detection of coronary artery disease from myocardial perfusion SPECT in a large population. J Nucl Med 2013;54:221–8.

8. Otaki Y, Betancur J, Sharir T, et al. 5-Year prognostic value of quantitative versus visual MPI in subtle perfusion defects: results from REFINE SPECT. J am Coll Cardiol Img 2020;13:774–85.

9. Cerqueira MD, Weissman NJ, Dilsizian V, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the Cardiac Imaging Committee of the Council

on Clinical Cardiology of the American Heart Association. J Am Coll Cardiol 2002;105:539–42.

10. Nakazato R, Tamarappoo BK, Kang X, et al. Quantitative upright-supine high-speed SPECT myocardial perfusion imaging for detection of coronary artery disease: correlation with invasive coronary angiography. J Nucl Med 2010;51:1724–31.

11. Shaw LJ, Iskandrian AE. Prognostic value of gated myocardial perfusion SPECT. J Nucl Cardiol 2004;11:171–85.

12. Schuijf JD, Poldermans D, Shaw LJ, et al. Diagnostic and prognostic value of non-invasive imaging in known or suspected coronary artery disease. Eur J Nucl Med 2006;33:93–104.

13. Miller TD, Hodge DO, Milavetz J, Gibbons RJ. A normal stress SPECT scan is an effective gatekeeper for coronary angiography. J Nucl Cardiol 2007;14:187–93.

KEY WORDS major adverse cardiac event, myocardial perfusion imaging, prognosis, quantification