In this issue of Circulation, Doukky et al report findings from a cohort of 1511 patients from 11 outpatient community-based practices (20 primary care physicians and 2 cardiologists) in the Chicago metropolitan area. The patients underwent single-photon emission computed tomographic (SPECT) myocardial perfusion imaging and were then followed up for 27±10 months for major adverse cardiac events: death, death or myocardial infarction, and cardiac death or myocardial infarction. The SPECT studies were categorized on the basis of the 2009 appropriate use criteria (AUC) as appropriate, uncertain, or inappropriate. The investigators report that 823 patients (54.5%) underwent SPECT scans that were classified as appropriate or uncertain and 688 patients (45.5%) underwent SPECT scans that were classified as inappropriate. In those patients whose SPECT scans were appropriate or uncertain, abnormal scans were of significant value in predicting major adverse cardiac events with hazard ratios of 3.1 to 3.7 compared with normal scans. However, in those patients undergoing SPECT classified as inappropriate, abnormal SPECT scans did not achieve statistical significance in predicting major adverse cardiac events, although the hazard ratios ranged from 2.3 to 11.8. Regardless of the appropriateness of SPECT, the presence of ischemia on SPECT, reflected in the summed difference score (SDS), predicted subsequent coronary angiography and revascularization. As the investigators indicate, this is the first large study validating the prognostic implications of SPECT AUC, further supporting its clinical utility. In this editorial, we examine both the internal and external validity of this study to place it in context for evidence-based clinicians.

Is the study internally valid? That is, do the data justify the conclusions? This study has a number of strengths, beginning with the size of the cohort and the several-year duration of follow-up. Although follow-up was >99% complete, a number of patients were excluded (n=182) or lost to follow-up (n=14). To their credit, the authors systematically compare these patients with those patients with follow-up data. The differences were modest, suggesting that the exclusion of these patients did not bias the results. The SPECT studies were all acquired on a single dedicated cardiac SPECT camera and interpreted by a single expert nuclear cardiologist, eliminating the variability inherent in different equipment and different interpreters. The SPECT results were categorized into published groupings of the summed stress score and SDS, permitting comparison with extensive previous literature using the summed stress score and SDS.

There were some potential methodological weaknesses in the study. Although the authors define their cohort as prospective, the categorization of appropriateness was based on retrospective chart review and “computer-based logic,” which is not further defined. We are not told how many observers performed this review and whether they were independent of the practices, if the clinicians knew they were being monitored at the time they ordered the study, and exactly how the computer-based logic dealt with missing data. For example, “abnormal ECG” is listed as the primary indication for SPECT in 136 of the patients (9%). How were these patients categorized because “abnormal ECG” does not appear as an indication in the AUC?

There were very few hard cardiac events in the patients with inappropriate studies (5 deaths, 0 cardiac deaths, and 2 myocardial infarctions), reducing the power of the study to detect prognostic value in this group. Appropriate use appears in the prognostic models because inappropriate patients had fewer events. The lack of significant prognostic value for SPECT in the inappropriate patients reflects the decreased statistical power in this group resulting from their favorable prognosis. Many readers may incorrectly conclude that SPECT performs “differently” in inappropriate patients. However, the test for interaction between appropriateness and SPECT was not significant (P>0.43 for all end points), indicating that SPECT did not perform differently in inappropriate patients. For example, for the end point of death or myocardial infarction, abnormal SPECT had a hazard ratio of 3.3 with 95% confidence limits of 1.6 to 6.5 and a value of P=0.001. In contrast, in the inappropriate patients, abnormal SPECT had a hazard ratio of 4.0 but with wider 95% confidence limits of 0.7 to 21.8 and an insignificant value of P=0.11. Are these 2 findings significantly different from one another? The interaction analysis found a value of P=0.70, indicating that they are not.

Although the authors perform a separate analysis showing the impact of ischemia (reflected in SDS) on coronary angiography and revascularization in all of the patients, regardless of the appropriateness of SPECT, the impact of revascularization on the analysis of hard cardiac events is uncertain. The authors appear to have performed their analysis of hard cardiac events.
events without censoring patients at the time of revascularization, which was the customary approach in previous prognostic studies of SPECT. The clinical model constructed by the investigators does not include symptom status, which is particularly important given the much higher prevalence of asymptomatic patients in the inappropriate group (27.9%) compared with the appropriate/uncertain group (8.5%). Symptom status has a major role in AUC ratings. Inclusion of such a variable in the analysis might have replaced appropriate use in the prognostic model. Ejection fraction is included in this model only as a categorical (normal/abnormal) variable, so an ejection fraction of 49% is equivalent to one of 22%. However, the impact of this issue on the results is likely minimal because only 2.4% of the study population had prior myocardial infarction. The clinical model apparently incorporates a single “prior coronary artery disease” variable, which makes previous myocardial infarction equivalent to percutaneous coronary intervention or coronary artery bypass graft surgery. It is difficult to reconcile the numbers provided in the authors’ Table 1 for patients with known coronary artery disease (n=271) with the total of patients with prior myocardial infarction, percutaneous coronary intervention, or coronary artery bypass graft surgery (n=200). Finally, given the modest number of hard cardiac events (a maximum of 45 for the combined variable of death and myocardial infarction), the findings presented for death, cardiac death, and cardiac death and myocardial infarction in Figure 5 violate the statistical rule of thumb that one should not have >1 significant variable for each 10 events.\(^2\) We therefore suspect that these models were all “overfitted” and are unlikely to be reproduced in other studies.

Are the results consistent with previous literature? Consistent with prior studies, summed stress score was more strongly associated with major adverse cardiac events, whereas SDS was the major predictor of revascularization. The overall rate of an abnormal SPECT study was low at 11%. The prevalence of severe ischemia (SDS >7, which converts to ischemia >10% of the left ventricle) was even lower at 2.8%. This was a low-risk population. Only 7% proceeded to coronary angiography and 4% to coronary revascularization. Only 2.3% died in a mean of 2.25 years of follow-up. These findings are all consistent with the Study of Myocardial Perfusion and Coronary Anatomy Imaging Roles in Coronary Artery Disease (SPARC) study,\(^3\) as well as the long-term trends in SPECT published by the Cedars-Sinai group.\(^4\) As previously mentioned, the inappropriate patients were an even lower-risk subset of this low-risk group.

SPECT has well-recognized prognostic value. The multiple studies demonstrating this were well known to the expert panels that crafted the AUC. Presumably, the absence of evidence for the prognostic value of SPECT was considered by these panels in categorizing certain indications as inappropriate. Thus, the overall findings on prognostic value that are demonstrated in this study are predictable from the previous literature.

Although the authors did not design this study to examine AUC rates, the rate of inappropriate studies is surprising
and not consistent with the previously published literature on AUC.\textsuperscript{5-9} As shown in Table 1, the inappropriate rate demonstrated in this study (45.5\%) is more than twice as high as any previously published inappropriate rate. This finding has major implications if it is representative of other community practices in the country. The prevalence of specific inappropriate indications is also quite different from prior studies (Table 1). In this study, 28\% of the inappropriate patients were asymptomatic and therefore presumably either low risk or intermediate risk by the Framingham score. This percentage is considerably lower than that reported in 3 of 4 previous study cohorts. Preoperative testing was much less frequent (3\%) than in 4 of 5 previous study cohorts. The authors suggest that previously published data based on tertiary-care center experience likely included higher-risk populations. We agree. As shown in Table 2, the patients included in the present study were fundamentally different from previously published groups. They were younger with a lower rate of previous myocardial infarction, previous percutaneous coronary intervention, and previous coronary artery bypass graft surgery. All of these factors define a lower-risk population and likely account for the lower cardiac event rates found in this study.

Are these findings generalizable? That is, should they be viewed as representative? We think not. First, the rate of appropriate/inappropriate studies, rate of subsequent cardiac event rates, and the prognostic value of SPECT are heavily dependent on the patient population studied. Because this study included many low-risk patients, these results likely apply to only similar low-risk patients. More important, the present study reports substantial differences in AUC categories between primary care physicians and cardiologists and even greater variability in the overall spectrum of inappropriate studies between physicians. At one end of the spectrum, 1 physician had an inappropriate rate of only 10\%; at the other end, 1 physician had an inappropriate rate of 77\%. It is unlikely that the results of the present study apply to either of these 2 individuals. To the degree that there is similar, if not greater, variability in physicians across the country, the “group” results presented here may not apply to many individual physicians or many practices with higher or lower rates of inappropriate studies.

Despite our concerns, we believe that this study makes a very important contribution to the literature. It is the second largest study of the AUC and the first to examine the prognostic value of the AUC. This study demonstrates that inappropriate patients have fewer cardiac events. The results have limited generalizability because of the low-risk patient population studied and the observed variability in inappropriate rates across physicians. However, this issue of limited generalizability is not unique to AUC or this study. There is tremendous variability in many cardiac procedures, including the rate of percutaneous coronary intervention and coronary artery bypass graft surgery in different Medicare referral regions,\textsuperscript{10} the rate of inappropriate percutaneous coronary intervention across hospitals,\textsuperscript{11} and the rate of evidence-based implantable cardioverter-defibrillator use in hospitals, among others. We need more studies such as this one to better define the variability in inappropriate rates, the most common inappropriate indications, and the prognostic value of AUC in other populations. This study is one step along the road to quality improvement, the ultimate goal of the AUC.

Disclosures

Dr Gibbons is a consultant for Lantheus Medical Imaging and has received speaking honoraria from AstraZeneca. Dr Miller is a consultant for Astellas Pharma.

References


Key Words: Editorials ◼ cardiac imaging techniques
Single-Photon Emission Computed Tomography Appropriateness: Does It Matter for Patient Outcomes?
Raymond J. Gibbons and Todd D. Miller

_Circulation_. 2013;128:1595-1597; originally published online September 10, 2013; doi: 10.1161/CIRCULATIONAHA.113.005595
_Circulation_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/128/15/1595

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation_ is online at:
http://circ.ahajournals.org//subscriptions/