Comparing two methods for determining appropriateness of myocardial perfusion imaging: Criteria from the American College of Cardiology Foundation and the American College of Radiology

Anastasiya Bagrova, BS,[a] Ali Y. Alsamarah, MD, MS,[a,b] and David E. Winchester, MD, MS[a,c]

[a] College of Medicine, University of Florida, Gainesville, FL
[b] Cardiovascular Medicine Department, Boston Medical Center, Boston, MA
[c] Malcom Randall Veterans Affairs Medical Center, Gainesville, FL

Received May 3, 2017; accepted Jun 14, 2017
doi:10.1007/s12350-017-0965-1

Background. **Appropriate use criteria (AUC) developed by the American College of Cardiology Foundation and the appropriateness criteria (AC) developed by the American College of Radiology (ACR) are two existing methods of rating appropriateness of myocardial perfusion imaging (MPI). One study found poor agreement of ratings between the two methods. However, using the most contemporary AUC from 2013, it is unknown if poor agreement still exists.**

Methods. **Retrospective cohort investigation comparing patients undergoing nuclear MPI between June 2011 and September 2014. The appropriateness category was determined based on the 2013 AUC (Appropriate, may be appropriate, rarely appropriate) and the 2010 ACR AC (usually appropriate, maybe appropriate, usually not appropriate). The primary outcome was the degree of the agreement between the two methods.**

Results. **The kappa coefficient between ACR AC and AUC was 0.32, $P < 0.0001$, indicating poor agreement; 40 (8%) patients were classified by the AUC but could not be classified by the ACR AC.**

Conclusion. **The two methods for rating the appropriateness of MPI have poor agreement; a potential for disagreement between providers and payers if only one method is used. (J Nucl Cardiol 2019;26:826–30.)**

Key Words: Single-photon emission computed tomography (SPECT) • Myocardial perfusion imaging (MPI) • American College of Cardiology Foundation (ACCF) • American College of Radiology (ACR)

Abbreviations			
AUC	Appropriate use criteria	CAD	Coronary artery disease
ACCF	American College of Cardiology Foundation	MPI	Myocardial perfusion imaging
ACR	American College of Radiology		

Electronic supplementary material The online version of this article (doi:10.1007/s12350-017-0965-1) contains supplementary material, which is available to authorized users.

The authors of this article have provided a PowerPoint file, available for download at SpringerLink, which summarizes the contents of the paper and is free for re-use at meetings and presentations. Search for the article DOI on SpringerLink.com.

Reprint requests: Ali Alsamarah, MD MS, Cardiovascular Medicine Department, College of Medicine, University of Florida, Gainesville (Bagrova), 88 East Newton Pavilion, Boston, Massachusetts 02218; *Ali.Alsamarah@gmail.com*

1071-3581/$34.00

See related editorial, pp. 831–832

INTRODUCTION

Appropriate use criteria (AUC) for myocardial perfusion imaging (MPI) have been published over a decade now. A growing number of health care providers, as well electronic health software vendors, are incorporating AUC for use at the point of care. AUC developed by the American College of Cardiology Foundation (ACCF) and the appropriateness criteria (AC) developed by the American College of Radiology (ACR) are the main existing methods that assist health care providers in decisions regarding the appropriateness of MPI.[1–6] The two methods use different approaches which may create separate categories determination. Congressional regulations were passed in 2014 mandating health care providers to use appropriate use criteria when ordering advanced imaging in Medicare patients, and linked their payments to the appropriateness of the test.[7] Prior report from a veteran study suggested discordance between the previously published AUC from 2009 and ACR AC.[8,9] However, based on the most contemporary publication of AUC in 2013 and studying the non-veteran population, the degree of agreement is unknown.

METHODS

Study Population

We retrospectively identified a cohort of patients who underwent single-photon emission computed tomography (SPECT) MPI between June 2011 and September 2014 at the University of Florida Health. The Institutional Review Board reviewed the protocol and waived the requirement for informed consent. MPI studies were performed using technetium-99m SPECT in a typical clinical fashion with images obtained at rest and post-stress, induced either by patient's exercising on a treadmill or pharmacologically. Results of the test were obtained from the final report of the MPI; we did not re-evaluate the original study images. Demographic information, medical history, and reasons for ordering MPI information were gathered from notes of emergency department, in-hospital follow-up, and provider's order details. Study data were collected and managed using a Research Electronic Data Capture.[10] CAD was defined as any angiographic evidence of atherosclerotic disease in any of the epicardial coronary arteries or prior myocardial infarction (MI).

Appropriateness Determination

Appropriateness category was determined by using the 2013 ACCF AUC criteria:[1] appropriate, maybe appropriate, rarely appropriate and the ACR AC:[2–6] usually appropriate, maybe appropriate, usually not appropriate. Of the 500 patients in the cohort, nine were excluded due to incomplete or missing information regarding the rationale of stress testing. Subjects who could not be classified under the ACR criteria were assigned to the ''cannot be determined'' category and were excluded only from the agreement statistical analysis, not the entire study. ACCF appropriateness classification was performed using a hierarchical approach as advised by the ACC AUC and was facilitated by using a smartphone application [© 2014 Astellas Pharma US, Inc. Northbrook, IL], whereas the ACR appropriateness was performed using materials found on ACR website (https://acsearch.acr.org/list), which included published material from 2011 to 2014. For each patient, to minimize bias, the chart would be reviewed and then a determination of appropriateness would be made separately by two investigators. Ratings for AUC were determined by author AA, whereas ratings for ACR were determined by SB.

Definitions of Outcomes and Statistical Analysis

The primary outcome of this study was to establish the agreement between the ACCF and ACR appropriateness ratings, using the kappa statistic. The secondary outcome was the test result: normal vs abnormal. We compared the outcomes between the each ACCF and ACR appropriateness ratings by Chi-square analysis. P value of <0.01 was considered significant. Statistical analyses were performed using SAS software [SAS version 9.3, SAS Institute, Cary NC].

RESULTS

Of the total 500 MPIs reviewed, 302 (60.4%) were females, and the average age of our patient population was 58 years. There was a high prevalence of obesity (mean BMI = 32), hypertension (77.6%), hyperlipidemia (62%), and history of active or prior tobacco use (74.6%). (Table 1) The most common presenting symptom was chest pain or discomfort (71.4%), followed by dyspnea (7%).

Primary Outcomes

Out of 491 patients, 40 (8%) could be classified by the ACCF AUC but had no correlated rating in the ACR AC system. Such scenarios included a variety of common presentations such as syncope, newly diagnosed heart failure, atrial fibrillation, and pre-operative cardiac assessment. Those 40 patients were excluded from the kappa calculation.

The kappa correlation coefficient between ACR and ACCF appropriateness ratings was 0.32, $P < 0.0001$, indicating poor agreement. (Table 2) When using ACCF AUC, we found that 73.4% of MPI studies fell under ''appropriate'' category, 8% under ''maybe appropriate,'' and 18.6% were ''rarely appropriate.'' Applying the ACR criteria, 73% of studies were classified as ''usually appropriate,'' 20.2% as ''may be appropriate,'' and 6.8% as ''usually not appropriate.'' Of the

Table 1. Baseline characteristics

Variable	N = 500 (%)
Age (mean ± SD)	58.2 ± 12.7
BMI (mean ± SD)	32 ± 13.6
Gender	
Male	198 (39.6)
Female	302 (60.4)
Presenting symptoms	
Chest pain/discomfort	357 (71.4)
Dyspnea	35 (7.0)
Fatigue	8 (1.6)
Syncope	9 (1.8)
Other	91 (18.2)
Medical history	
Hypertension	388 (77.6)
Hyperlipidemia	310 (62.0)
Diabetes mellitus	178 (35.6)
CAD	115 (23.0)
Family history of early CAD	56 (11.2)
CKD	83 (16.6)
Stroke	61 (12.2)
CMP or CHF	82 (16.4)
Active tobacco use	197 (21.4)
Former tobacco use	176 (35.2)

SD, standard deviation; *BMI*, body mass index; *CAD*, coronary artery disease; *CKD*, chronic kidney disease; *CMP*, cardiomyopathy; *CHF*, congestive heart failure

discordant ratings, 31 patients (6.8%) were two appropriateness categories apart: 7 patients were appropriate by ACCF AUC but usually not appropriate by ACR AC, and 24 patients were usually appropriate by ACR AC but rarely appropriate by ACCF AUC. Common reasons for such difference were found in symptomatic patients with low pretest probability of CAD who were unable to exercise, or symptomatic patients with intermediate-or-high-risk profile who had a normal stress test study within 2 years of presentation.

Secondary Outcomes

Using ACCF AUC, studies classified as maybe appropriate were least likely to be abnormal (4.9%), as compared to appropriate (22.5%) and rarely appropriate (10%), ($P < 0.001$). Using ACR AC, usually non-appropriate studies were least likely to be abnormal (6.5%), as compared to maybe appropriate (8.9%) and usually appropriate (22.7%), ($P = 0.002$).

DISCUSSION

We found poor agreement between the AUC and ACR AC, confirming prior reports.[8] Our study used the most contemporary 2013 AUC publication with key difference to the prior 2009 AUC. First, the 2009 guideline used the terms ''appropriate,'' ''uncertain,'' and ''inappropriate,'' while the 2013 guidelines used the terms ''Appropriate,'' ''Maybe Appropriate,'' and ''Rarely Appropriate.'' The 2009 AUC uses 80 clinical scenarios vs 64 clinical scenarios in the 2013 AUC, therefore leaving less number of studies which would otherwise be of undetermined appropriateness category. In addition, some studies considered appropriate by 2009 AUC may not be appropriate by 2013 AUC and vice versa. Prior study by Winchester et al. was conducted at a veteran's hospital. Our cohort had more females, and younger population, making the results more generalizable.

ACR AC have notably fewer clinical scenarios where the appropriateness of MPI is rated, leaving many specific indications for imaging without an appropriateness rating. Many clinical scenarios detailed in the ACCF AUC but absent from ACR AC are appropriate such as: pre-operative evaluation, new heart failure or cardiomyopathy, syncope, and atrial fibrillation. All of these scenarios were effectively excluded from the ACR AC, which might have impacted the overall results and the correlation between the two systems. In fact, in our cohort, only 8% of the AUC indications could not be paired by ACR AC, compared to 18.8% unmatched

Table 2. Concordance between the ACCF AUC and ACR MPI appropriateness rankings

		ACR AC			
		Usually app	**Maybe app**	**Usually not app**	**Total**
AUC	Appropriate	293 (64.9%)	31 (9.5%)	7 (1.5%)	331 (73.4%)
	Maybe app	12 (2.6%)	12 (2.6%)	12 (2.6%)	36 (8%)
	Rarely app	24 (5.3%)	48 (10.6%)	12 (2.6%)	84 (18.6%)
	Total	329 (73%)	91 (20.2%)	31 (6.8%)	451

The kappa correlation coefficient between ACR and ACCF appropriateness ratings was 0.32, $P < 0.0001$

indications in a prior cohort.[8] One possible explanation is the difference in baseline characteristics of the cohorts; our study had younger population with more female predominance. Other possibilities include that the investigators who reviewed the appropriateness ratings, have ''overfit'' those indications into one of the ACR ratings. Furthermore, 31 patient categories (6.8%) were discordant and two appropriateness categories apart: appropriate by ACCF AUC but usually not appropriate by ACR AC, or rarely appropriate by ACCF AUC but usually appropriate by ACR AC. Additionally, the poor agreement between the two methods was a result of one appropriateness category apart; in such one test can be considered rarely appropriate (or usually not appropriate) by one system but maybe appropriate by the other.

Our study emphasizes not only the differences in applying appropriateness criteria by providers, but it highlights potential concerns as we are approaching the implementation of Protecting Access to Medicare Act (PAMA) in 2018. The new Act directs the Centers for Medicare and Medicaid (CMS) to promote the use of appropriate use criteria for advanced imaging services.[7] The discordance in the appropriateness category could result in disagreement between payors such as the Centers for Medicare and Medicaid Services and healthcare providers.

Regarding the clinical outcomes of the study, the likelihood of a study being abnormal was highest when rated as appropriate by either the ACCF AUC (22.5%) or ACR AC (22.7%). These findings are in line with those reported in a multitude of prior studies and a recent meta-analysis on appropriateness of MPI.[11]

Limitations

The study limited due to its retrospective feature, as such, the assessment of appropriateness was limited to the data available in the documents in the electronic health record. Given the single-centered nature, the generalizability of the study is restricted, and may not be applied to different populations. We limited the number of adjudicators who reviewed the chart to minimize probable intra- or inter-observer variability, however, that still can be a potential confounder, which has been reported to in prior studies of AUC.[12] The study was underpowered to detect hard clinical outcomes.

NEW KNOWLEDGE GAINED

In this study, we found poor agreement between the most contemporary AC, AUC, and ACR, consistent with prior reports comparing the 2 methods, but has uses different methodology based on the most contemporary published literature of appropriate criteria. Potential conflicts between payors and provider may arise if regulators use one method for reimbursement purposes.

CONCLUSIONS

The two systems for rating the appropriateness of MPI, developed by the ACCF and ACR, have poor agreement between them, which can create disagreement between payors and providers if only one AC is used.

Acknowledgment

We want to thank the REDCap staff, the Clinical Translational Science Institute (CTSI), specifically Dr Jonathan Shuster at the University of Florida for assistance in gathering and analyzing the data.

Disclosures

None.

References

1. Wolk MJ, Bailey SR, Doherty JU, Douglas PS, Hendel RC, Kramer CM, et al. ACCF/AHA/ASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 multimodality appropriate use criteria for the detection and risk assessment of stable ischemic heart disease: A report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, and Society of Thoracic Surgeons. J Am Coll Cardiol 2014;63:380-406.
2. Hoffmann U, Venkatesh V, White RD, Woodard PK, Carr JJ, Dorbala S, et al. ACR Appropriateness Criteria((R)) acute non-specific chest pain-low probability of coronary artery disease. J Am Coll Radiol 2012;9:745-50.
3. Earls JP, White RD, Woodard PK, Abbara S, Atalay MK, Carr JJ, et al. ACR appropriateness criteria(R) chronic chest pain–high probability of coronary artery disease. J Am Coll Radiol 2011;8:679-86.
4. Earls JP, Woodard PK, Abbara S, Akers SR, Araoz PA, Cummings K, et al. ACR appropriateness criteria asymptomatic patient at risk for coronary artery disease. J Am Coll Radiol 2014;11:12-19.
5. Mammen L, White RD, Woodard PK, Carr JJ, Earls JP, Hendel RC, et al. ACR appropriateness criteria(R) on chest pain, suggestive of acute coronary syndrome. J Am Coll Radiol 2011;8:12-18.
6. Woodard PK, White RD, Abbara S, Araoz PA, Cury RC, Dorbala S, et al. ACR appropriateness criteria chronic chest pain-low to intermediate probability of coronary artery disease. J Am Coll Radiol 2013;10:329-34.
7. HR4302 RPJ. Protecting Access to Medicare Act of 2014. 2016(113th Congress (2013-2014)).

8. Winchester DE, Wolinsky D, Beyth RJ, Shaw LJ. Discordance between appropriate use criteria for nuclear myocardial perfusion imaging from different specialty societies: A potential concern for health policy. JAMA Cardiol 2016;1:207-10.
9. Hendel RC, Cerqueira M, Douglas PS, Caruth KC, Allen JM, Jensen NC, et al. A multicenter assessment of the use of single-photon emission computed tomography myocardial perfusion imaging with appropriateness criteria. J Am Coll Cardiol 2010;55:156-62.
10. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; 42:377-81.
11. Elgendy IY, Mahmoud A, Shuster JJ, Doukky R, Winchester DE. Outcomes after inappropriate nuclear myocardial perfusion imaging: A meta-analysis. J Nucl Cardiol 2015.
12. Lin FY, Rosenbaum LR, Gebow D, Kim RJ, Wolk MJ, Patel MR, et al. Cardiologist concordance with the American College of Cardiology appropriate use criteria for cardiac testing in patients with coronary artery disease. Am J Cardiol 2012;110: 337-44.