EDITORIAL COMMENT

Appropriate Use Criteria
Lessons From Japan*

Olivia Y. Hung, MD, PhD,† Habib Samady, MD,† H. Vernon Anderson, MD‡

Professional societies, including the American College of Cardiology, have for many years developed and periodically updated formal guidelines that attempt to provide a comprehensive review of available evidence for management of certain clinical conditions. These guidelines are distinctly structured documents and include recommendations both for and against various practices. For some treatments, including percutaneous coronary intervention (PCI), it was observed that there are wide variations in practice across many geographic regions. It was further recognized that there are gaps in a guidelines approach to understanding how PCI and other therapies could vary so widely. In order to address these gaps and investigate whether there might be overuse or underuse of various procedures that carry both potential benefit as well as risk, the American College of Cardiology established an appropriateness criteria working group. This group's first task was to develop a methodology for evaluating appropriateness of cardiovascular imaging procedures [1]. Other efforts quickly followed, and in 2009, the initial appropriate use criteria (AUC) for PCI were published [2]. A revised update of these came out in 2012 [3].

The AUC were developed as a schematic way to provide guidance to clinicians on appropriately selecting patients for PCI. The workgroup methodically compiled a set of clinical scenarios and then assigned scores (1 to 9) of relative "appropriateness" for PCI to each scenario. This was an initial attempt to encapsulate a complex clinical decision-making process that assimilates the patient's symptom status, medical therapy, noninvasive stress test results, and angiographic data. By pre-specified design, the graded scenarios were grouped into 3 general categories, with scores 7 to 9 called "Appropriate," scores 4 to 6 called "Uncertain," and scores 1 to 3 called "Inappropriate." However, the choice and strength of the clinical scoring factors have raised many questions. How refractory to medical therapy should angina symptoms be before one can offer a patient symptom relief with PCI? As thorny as that question might be, the issue of noninvasive imaging as the ultimate determinant of significant ischemia and, by extension, general clinical risk has sparked even greater disagreement among thoughtful clinicians. Although there is little debate that at a population level identification of ischemia on stress tests confers an adverse prognosis, the problem for individual patients is that interpretation of these tests for each person is highly variable and has substantial false positive and false negative rates [4-7].

The AUC are beset by numerous additional difficulties. To list just a few: 1) the evidence base is not strong, with most of the criteria based upon "expert opinion" derived from surveys of the clinical literature rather than from hard comparative science using the specific written scenarios themselves; 2) the scenarios are descriptive, but nonspecific, with only about one-half of cases examined able to be classified using them; 3) many or even most individual cases are classified as Uncertain, which further underscores the lack of clear guidance the AUC might otherwise provide; and 4) the AUC are not validated (neither the scored scenarios nor the group categories), and there currently are no metrics nor method by which to validate them.

When these AUC were applied retrospectively to clinical practice in the United States using >600,000 cases from the National Cardiovascular Data Registry (NCDR), several important discoveries were made (8). Acute PCI procedures were the majority of cases analyzed (71%), with only a miniscule 1,770 cases of >350,000 in the sample that could not be classified. Essentially all acute PCI procedures (99%) were classified as Appropriate. For nonacute PCI cases the situation was different. Fully 41% (>100,000) of cases in the sample could not be classified at all, and therefore, nonacute PCI formed a minority (29%) of cases analyzed. For nonacute PCI cases, 50% were classified as Appropriate, 38% were Uncertain, and 12% were Inappropriate. The variation between hospitals was quite broad, with the Inappropriate category ranging from 0% of cases to 55%. Another analysis of hospital-level data from NCDR using >426,000 nonacute PCI procedures from 1,199 hospitals confirmed these findings (9). Here, 51% of the procedures could not be classified, and the overall rates for those that could were Appropriate in 50%, Uncertain in 36%, and Inappropriate in 12%. The variation between hospitals for the Inappropriate classification was again extremely broad (0% to 59%). Other analyses using smaller datasets of nonacute PCI in Washington State and New York State are consistent with the findings from NCDR (10,11). Also noted in all these datasets is that lack of stress test results and not matching any written clinical scenarios were the main reasons for inability to classify nonacute PCI cases.

Now comes a new study, in this issue of *JACC: Cardiovascular Interventions,* that examines application of the AUC to a large registry of 10,050 PCI patients treated in Japan over a 5-year period between 2008 and 2013 (12). Two interesting features of this Japanese work are: 1) classification of cases by both 2009 and 2012 AUC criteria; and 2) examination of the trend in classifications over the 5-year interval. These investigators found that 96% of PCI procedures performed in acute settings were rated as Appropriate using the 2009 AUC criteria, but this fell to 78% using the 2012 criteria. Changes in clinical scenario ratings accounted for this decrease. Nevertheless, the high rates of Appropriate for the acute PCI procedures in Japan are consistent with U.S. reports. For nonacute PCI in Japan, the Inappropriate classification was given to 15% according to AUC 2009, and this doubled to 31% according to AUC 2012. The analysis by the 8 time subintervals from 2008 to 2013 shows steady increases in the percentage of Inappropriate

PCI. So the major finding in this report is that a substantial and increasing number of nonacute PCI cases in Japan are classified as Inappropriate using these U.S. criteria. The reason appears to be that common methods to assess the significance of coronary disease in contemporary Japanese practice are coronary computed tomographic angiography (CCTA) and/or measurement of fractional flow reserve (FFR) during invasive diagnostic angiography, both of which are being used increasingly often. The AUC are predicated on performing stress tests for determining functionally significant ischemia and not on these alternative technologies. This situation requires comment.

The first issue is the evolving role of FFR. FFR is based upon pressure wire measurements in a coronary artery during a diagnostic catheterization procedure. Originally, these measurements were validated against myocardial perfusion imaging as a correlative assessment for ischemia in the distribution of the tested coronary artery. However, FFR has now gone beyond correlations with perfusion imaging-assessed ischemia and arguably is the gold standard for identifying lesion-specific ischemia (13). It has now been validated in its own right as predictive of future clinical events, and elective PCI performed on lesions found significant by FFR are recognized as fully warranted procedures (13-15). Yet, the current collection of guidelines for evaluating patients for coronary disease specify that FFR measurements are only justified in patients who are undergoing Appropriate diagnostic catheterization (i.e., only after pre-catheterization stress tests have been found abnormal), and then only on intermediate lesions in arteries that correlate with the stress test-derived ischemic territory (3,16,17). Japanese cardiologists, like many of their U.S. counterparts, would appear to disagree with this approach. Patients with suspicion of CAD or with known stable ischemic heart disease may get a diagnostic angiogram without preprocedure stress testing, with FFR performed on significant or borderline lesions, and then subsequent PCI for FFR-positive lesions. Interestingly, when these investigators (12) reclassified their CCTA-based PCI cases from Inappropriate to Appropriate, they found that the proportion of Inappropriate PCI still increased over time. Because FFR use also increased over the interval, it is likely that FFR-based cases account for this finding. All of this suggests that Japanese practice may be shifting as both technology and the evidence base evolve (18).

The second issue of importance is the introduction and evolution of CCTA. This is now recognized as a powerful anatomic tool with excellent