

EDITORIAL COMMENT

Quality Improvement Using the AUC: Is it This Easy?*

Raymond J. Gibbons, MD, J. Wells Askew, MD

Rochester, Minnesota

The rapid increase in cardiac imaging services throughout the 1990s was the source of increasing concern on the part of third-party payers. In response, the American College of Cardiology Foundation, working with other partners, developed a series of appropriate use criteria (AUC) for cardiac imaging modalities as well as coronary revascularization. These AUC were intended to provide guidance to practicing clinicians about specific patient situations and to allow them to potentially compare their own practices with published data

[See page 545](#)

and national benchmarks in an effort to reduce inappropriate imaging. Published data regarding the use of AUC in quality improvement projects are scant and, until recently, focused on AUC for single-photon emission computed tomography imaging (1) and coronary computed tomography angiography (2). In this issue of *JACC*, Bhatia et al. (3) report on a successful quality improvement project using the transthoracic echocardiography AUC to reduce the rate of inappropriate transthoracic echocardiograms (TTEs) ordered by medical house officers on the teaching general medical service at Massachusetts General Hospital. Using an intervention that consisted of a lecture regarding the AUC for transthoracic echocardiography, a pocket card applying these AUC to common scenarios, and biweekly feedback on ordering behavior

via e-mail, these investigators reduced the number of inappropriate TTEs from 13% to 5% ($p < 0.001$). The authors are to be congratulated for their implementation of a successful quality improvement project.

On the basis of this project, it is tempting to conclude that simple educational tools will dramatically reduce the rate of inappropriate imaging studies and thereby increase the value of cardiac imaging. However, as evidence-based clinicians, we would like to respectfully offer several caveats regarding both the internal and external validity of these findings. Do the data presented justify the conclusions (i.e., are they internally valid)? The classification of the TTEs was performed independently by 2 study investigators who were therefore unblinded during the intervention period and knew that the study was designed to show improvement. The paper reports that these investigators never disagreed in the classification of 345 TTEs; they therefore had a kappa value of 1.0, representing perfect agreement. In contrast, Ballo et al. (4) reported a kappa value of 0.83 in 931 patients, and Aggarwal et al. (5) reported an even lower kappa value of 0.55 in 529 patients. The authors report that they were able to perform “data collection, classifications of TTEs, and feedback to house staff in 25 hours.” They do not indicate whether this was a total time for all investigators or for each investigator. Assuming the latter and that *each* investigator spent 80% of his or her time on classification of TTEs, each had 1,200 min to classify 345 patients or ~ 3.5 min per patient, which seems remarkably short compared with our own experience. The authors’ description of their classification system lists very few assumptions with respect to overlapping indications or the order of application of tables, which have clearly concerned other investigators applying the AUC for echocardiography (4).

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

From the Division of Cardiovascular Diseases and Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, Minnesota. Dr. Gibbons is a consultant for Lantheus Medical Imaging. Dr. Askew has reported that he has no relationships relevant to the contents of this paper to disclose.

We are concerned that this perfect agreement, as well as the speed of classification, may reflect the use of standardized language by ordering house staff to ensure that an echocardiogram was categorized as appropriate rather than a true change in the rate of inappropriate echocardiograms.

The authors compared their intervention period of February to June with January to November of the preceding year. House staff in the second half of any academic year may potentially perform better than in the first half of the year, which may have contributed to the authors' positive finding. Obviously, the physicians who were ordering studies during the pre-intervention period were different physicians from those in the intervention period, another uncontrolled factor. External forces, such as increasing public discussion of health care reform, may have improved physician performance independent of the authors' educational intervention. We showed this effect previously (1) and argued for the importance of "contemporary controls" to make certain that measured effects are actually due to the quality improvement project rather than external forces. The rate of admissions to the medical services decreased significantly between the pre-intervention period (16.8 per day) and the intervention period (13.9 per day), a 17% decrease that may have decreased the "time pressure" on the medical house staff and also improved their ordering patterns. The 26% reduction in the number of TTEs ordered per day reported by the authors must be adjusted for the reduction in the number of admissions; the result is a less impressive 10% reduction in TTEs ordered per day per admission.

Based on these concerns, we are not certain that the measured improvement in appropriateness can be solely attributed to the educational intervention project initiated by the authors. Changes in the external environment, the rate of admission, and the use of standardized terminology by the house staff may have contributed substantially to the measured improvement.

If the authors are correct, and their educational effort was primarily responsible for the measured improvement, will similar educational efforts succeed in other settings? As acknowledged by the authors, the generalizability of their results, which focused on medical house staff, to older physicians who have completed training is questionable. We agree with the authors; such older physicians may have more ingrained practice patterns and resent "being told what to do." During our own unsuccessful practice improvement project at the Mayo

Clinic (1), we heard this direct quote on several occasions. A more recent study also failed to demonstrate a significant change in appropriateness ratings of stress echocardiograms ordered by cardiologists after an educational initiative consisting of a cardiology grand rounds lecture, printed stress echocardiogram AUC materials, and a follow-up e-mail (6). Multiple components and strategies are necessary for successful intervention, only some of which were used here (7). It is possible that the residents had an incentive (such as a favorable evaluation) to reduce their ordering of inappropriate studies. The successful quality improvement project on coronary computed tomography angiography in Michigan (2) included a major financial incentive. It may be difficult for other quality improvement projects to identify and use similar incentives. The investigators used timely feedback in the form of biweekly e-mails. This is very time-consuming and feasible only if the ordering frequency on the part of individual physicians is quite high during the period in question. It may be far more difficult to use in an outpatient noncardiology practice where TTEs are ordered less frequently. Finally, the echocardiography AUC do not include patient scenarios reflecting the class III indication for echocardiography from the 2011 American College of Cardiology Foundation/American Heart Association Stable Ischemic Heart Disease Guideline: "Echocardiography, radionuclide imaging, CMR, and cardiac CT are not recommended for routine assessment of LV function in patients with a normal ECG, no history of MI, no symptoms or signs suggestive of heart failure, and no complex ventricular arrhythmias." We cannot tell how many of the echocardiograms performed in this study and labeled as appropriate fell within the scope of this class III recommendation. In a practice setting where many such unnecessary echocardiograms are performed, a quality improvement project focused on the echocardiographic AUC may miss a large number of unnecessary, low-value echocardiograms.

What can an evidence-based clinician conclude from this study? Quality improvement efforts focused on the AUC may possibly succeed using modest educational interventions, particularly if they are focused on physicians in training, who may be more flexible in their behavior. However, the success of such interventions with older physicians with more established patterns of behavior is not guaranteed. Whenever possible, quality improvement projects should be designed with a control group to make certain that the external environ-

ment and the workload in the internal environment are not independently responsible for any favorable measured change. Prompt feedback is important, although this may be logistically difficult in outpatient settings with a low frequency of tests. Incentives, such as favorable evaluations for residents or financial incentives surrounding reimbursement, are potentially important to success. Despite these demanding requirements, it is imperative that the

clinical practice community make every effort to improve the value of imaging using quality improvement projects such as this one.

Reprint requests and correspondence: Dr. Raymond J. Gibbons, Division of Cardiovascular Diseases and Internal Medicine, Department of Medicine, Mayo Clinic, Gonda 6, 200 First Street SW, Rochester, Minnesota 55905. *E-mail:* gibbons.raymond@mayo.edu.

REFERENCES

1. Gibbons RJ, Askew JW, Hodge DO, Kaping B, Carryer DJ, Miller TD. Appropriate use criteria for stress single-photon emission computed tomography sestamibi studies: a quality improvement project. *Circulation* 2011;123:499–503.
2. Chinnaiyan KM, Peyser PA, Goraya T, et al. Impact of a continuous quality improvement initiative on appropriate use of coronary computed tomography angiography. Results from a multicenter, statewide registry, the Advanced Cardiovascular Imaging Consortium. *J Am Coll Cardiol* 2012; 60:1185–91.
3. Bhatia RS, Milford CE, Picard MH, Weiner RB. An educational intervention reduces the rate of inappropriate echocardiograms on an inpatient medical service. *J Am Coll Cardiol Img* 2013;6:545–55.
4. Ballo P, Bandini F, Capecchi I, et al. Application of 2011 American College of Cardiology Foundation/American Society of Echocardiography appropriateness use criteria in hospitalized patients referred for transthoracic echocardiography in a community setting. *J Am Soc Echocardiogr* 2012;25:589–98.
5. Aggarwal NR, Wuthiaropas P, Karon BL, Miller FA, Pellikka PA. Application of the appropriateness criteria for echocardiography in an academic medical center. *J Am Soc Echocardiogr* 2010;23:2677–74.
6. Willens HJ, Nelson K, Hendel RC. Appropriate use criteria for stress echocardiography: impact of updated criteria on appropriateness ratings, correlation with pre-authorization guidelines, and effect of temporal trends and an educational initiative on utilization. *J Am Coll Cardiol Img* 2013;6:297–309.
7. Arrighi JA. Educational initiatives for quality improvement projects: can you teach an old dog new tricks? *Circulation* 2011;123:471–3.

Key Words: appropriateness ■ echocardiography ■ quality improvement.