openheart

A specialty-specific, multimodality educational quality improvement initiative to deimplement rarely appropriate myocardial perfusion imaging

David E Winchester,[1,2] Carsten Schmalfuss,[1,2] Christian D Helfrich,[3] Rebecca J Beyth[4,5]

CrossMark

For numbered affiliations see end of article.

Correspondence to
Dr David E Winchester; david.winchester@va.gov

ABSTRACT

Objective Investigations of Appropriate Use Criteria (AUC) education have shown a mixed effect on changing provider behaviour. At our facility, rarely appropriate myocardial perfusion imaging (MPI) differs by specialty; awareness of AUC is low. Our objective is to investigate if specialty-specific, multimodality education could reduce rarely appropriate MPI.

Methods We designed education focused on the rarely appropriate MPI ordered most often by each specialty. We tracked appropriateness of MPI in three cohorts: pre, post (immediately after) and late-post (4 months after) intervention.

Results A total of 889 MPI were evaluated (n=287 pre, n=313 post, n=289 late-post), 95.3% were men. Chest pain was the most common symptom (n=530, 59.6%), while 14.1% (n=125) had no symptoms. Rarely appropriate testing decreased from 4.9% to 1.3% and remained at 1.4% in the late-post cohort (p<0.0001). In logistic regression, lack of symptoms (OR 31.3, 95% CI 10.3 to 94.8, p≤0.0001) and being in the post or late-post cohorts (OR 0.27, 95% CI 0.11 to 0.68, p=0.006) were associated with rarely appropriate MPI. Preoperative MPI in patients with good exercise capacity was a common rarely appropriate indication. Ischaemia was not observed among patients with rarely appropriate indication for MPI.

Conclusions In certain clinical settings, education may be an effective approach for deimplementing rarely appropriate MPI. The effect of education may be enhanced when focused on improving patient care, delivered by a peer, and needs assessment indicates low awareness of guidelines. Lack of symptoms and preoperative MPI continue to be the predominant rarely appropriate MPI ordered.

INTRODUCTION

The American College of Cardiology Foundation (ACCF) and other cardiology specialty societies first published Appropriate Use Criteria (AUC) for nuclear myocardial perfusion imaging (MPI) in 2005. Since that time, a multitude of studies have evaluated the proportion of testing considered to be 'inappropriate' or 'rarely appropriate'. Meta-analyses of these studies have demonstrated no appreciable change in the proportion of low-value tests over time.[1 2]

A multitude of factors contribute to the delay in adoption of new practice recommendations by physicians and advanced practice providers (APPs).[3] Among the many strategies to encourage change, several have been applied to AUC for cardiac testing. Education is an attractive approach because it requires few resources and physicians and APPs are accustomed to participating in didactics. Education can take many forms and the evidence for effectiveness is mixed. An early attempt to reduce low-value MPI used lectures, meetings and a newsletter but was not found to be effective.[4] Other attempts which combined education with audit and feedback were effective at reducing echocardiograms and MPI.[5 6] A recent meta-analysis of quality improvement (QI) projects for

KEY QUESTIONS

What is already known about this subject?
Medical tests are often ordered in clinical scenarios where the patient is unlikely to benefit. The evidence on whether education can alter such patterns of care delivery is mixed.

What does this study add?
We showed that in certain circumstances, a well-designed educational programme tailored to the needs of the learners may effectively alter patterns of care delivery in the short term.

How might this impact on clinical practice?
Education is a low-cost intervention which, if applied properly, could help curtail the unnecessary use of medical tests and procedures.

AUC found that education was less effective without audit and feedback, whereas another systematic review looking at education for changing behaviour on low-value care overall found the majority of education-based projects to be effective.[7][8]

In previous investigations at our facility, we observed that patterns of rarely appropriate use varied by specialty and that over one-third of providers have never heard of AUC.[9][10] Based on this evidence, we hypothesised that education would effectively reduce rarely appropriate MPI. To test this, we designed and tracked effectiveness of a specialty-specific, multimodality educational QI initiative to deimplement rarely appropriate MPI at our facility.

METHODS

We designed an educational initiative to encourage the appropriate use of nuclear MPI. The initiative consisted of two content delivery mechanisms, lectures and printed posters. The lectures were given to groups of providers based on their specialty: primary care, hospital medicine and cardiology. Primary care lectures were given in person at our two largest medical centres and one was broadcast via video link over the intranet to providers throughout our clinical network. Each lecture audience included both physician and APPs. The printed posters were created in Microsoft PowerPoint (Redmond, WA, USA) and printed on 11×17 paper by the medical media office (Supplementary figure). This size was chosen because the primary care clinics have official notice boards which accommodate this poster size. Posters were distributed, two for each clinical site, with the request that they be posted in the providers' work areas for easy reference.

Based on our prior investigation,[9] the content of the lectures was tailored to each specialty and each lasted approximately 1 hour. For example, the lecture for cardiology focused on topics such as (but not limited to) arrhythmias and heart failure assessment, the lecture for primary care focused on initial assessment and screening for heart disease, and the lecture for hospital medicine focused on evaluating acute chest pain and syncope. The poster was geared towards the primary care audience because they order the majority of MPI at our facility.

The tone of the lectures and poster was purposefully one of collaboration. None of the materials focused on shaming or punishing providers for ordering rarely appropriate tests. We informed providers that the goal of the initiative was to work together to achieve the best for our patients. For example, when discussing screening in asymptomatic patients (a common indication which is rarely appropriate), we referred to studies that demonstrated no benefit of this practice and then showed the providers how to apply the ACC Atherosclerotic Cardiovascular Disease Risk Estimator as a more useful clinical approach. A brief, anonymous posteducation email survey was sent the day after the lecture to assess their opinions of the content and its projected impact on their future practice.

The primary outcome of our investigation was the proportion of nuclear MPI performed at our facility that was rated as rarely appropriate. Determinations of appropriateness were made using the 2013 Multimodality AUC for stable ischaemic heart disease or the 2009 AUC for cardiac radionuclide imaging, when needed (ie, for acute chest pain indications).[11][12] Appropriateness rating was performed on all MPI, sequentially, by a nurse at our affiliated research foundation whose effort was supported by grant funding and was not invested in the outcome of the investigation. Our MPI are interpreted by an interdisciplinary group of cardiologists, radiologists and nuclear medicine physicians who generate perfusion sores and study reports by combining the visual and computer-generated findings. In prior investigations at our facility, the proportion of rarely appropriate MPI was 10% to 15%. We estimated that we could detect a 50% relative reduction in rarely appropriate testing with alpha 0.05 and 1-beta of 0.8 using a sample size of 801.[13] Our MPI volume is approximately 150 studies a month. To achieve the desired sample size, we planned to review the appropriateness of 6 months of MPI studies creating three cohorts, each of 2 months' duration: pre, post (immediately after) and late-post (4 months after the intervention). Data were collected between October 2014 and August 2015. This design allowed us to measure the immediate effect of the intervention as well as the durability of the effect.

For each patient we gathered the following data: age, sex, symptoms (chest pain, dyspnoea or fatigue, other, no symptoms) and medical history (prior myocardial infarction (MI) or revascularisation, hypertension, diabetes, current tobacco use). We also gathered information on the provider type (physician or other) and provider specialty (cardiology, primary care, or other). Secondary outcomes included the prevalence of ischaemia and test conclusion (normal or abnormal). For the MPI, we gathered data on the test conclusion, summed stress scores, summed rest scores and summed difference scores (SDSs). Categorical and continuous variables were compared by χ^2 and Student's t-test, respectively. We constructed two logistic regression models to investigate characteristics that were associated with the presence of ischaemia and a test being rarely appropriate. Independent variables for the regression analyses included the patient characteristics and symptoms; the regression model for rarely appropriate testing also included variables on the ordering providers type (physician vs other) and specialty (cardiology vs other) and cohort (post or late-post vs pre). Results are reported as OR with 95% CI. Analysis was performed on SPSS Statistics V.23 (IBM, Armonk, NY, USA).

RESULTS

Our investigation included a total of 889 patients, 287 in the pre cohort, 313 in the post cohort and 289 in the

Table 1 Characteristics for n=889 patients

Age (mean±SD)	Pre, n=287 64.8±9.1		Post, n=313 66.3±8.6		Late, n=289 67.1±8.9		p Value 0.007
	n	%	n	%	n	%	
Male	270	94.1	299	95.5	278	96.2	0.47
Prior MI	48	16.7	42	13.4	37	12.8	0.35
Prior revascularisation	101	35.2	99	31.6	92	31.8	0.59
Hypertension	246	85.7	266	85	256	88.6	0.4
Diabetes mellitus	124	43.2	135	43.1	122	42.2	0.96
Current tobacco use	76	26.5	72	23	90	32.1	0.24
Symptoms							
Chest pain	168	58.7	187	59.7	175	60.6	0.91
Fatigue and/or dyspnoea	143	49.8	152	48.6	102	35.7	0.001
Other	72	25.2	104	33.4	143	49.7	<0.0001
No symptoms	46	16.1	36	11.5	43	14.9	0.24
Provider characteristics							
Attending	183	63.8	186	59.4	165	57.1	0.2*
Housestaff	44	15.3	63	20.1	48	16.6	
APP	60	20.9	64	20.4	76	26.3	
Inpatient test	78	27.2	106	33.9	95	32.9	0.17
Cardiology	101	35.2	115	36.7	119	41.2	0.02†
Primary care	104	36.2	86	27.5	70	24.2	
Other	82	28.6	112	35.8	100	34.6	

*Single comparison across attending, housestaff and APP.
†Single comparison across cardiology, primary care and other.
APP, advanced practice provider; MI, myocardial infarction.

late-post cohort (table 1). A minority of the patients had prior MI (n=127, 14.3%) or revascularisation (n=292, 32.8%); the substantial majority had hypertension (n=768, 86.4%). Most patients had chest pain (n=530, 59.6%), whereas a minority (n=125, 14.1%) had no symptoms. The majority of studies were classified using the 2013 AUC for stable ischaemic heart disease (622, 70.0%) while the remainder were classified using the indications for acute presentations from the 2009 AUC for MPI. The only patient characteristics that differed across the three cohorts were symptom burden and age.

The proportion of appropriate testing increased from 80.5% to 95.2% after our intervention; the effect persisted and remained at 92.7% late after. Simultaneously, the proportion of rarely appropriate testing decreased from 4.9% to 1.3% and remained at 1.4% (p<0.001, global χ^2) (figure 1). The posteducation survey was completed by 18 providers; 78% of respondents agreed with the statement 'I am likely to change my use of cardiac testing based on this presentation'. In our logistic regression model testing, variables associated with rarely appropriate testing included lack of symptoms (OR 31.3, 95% CI 10.3 to 94.8, p≤0.0001) and being in the post or late-post cohorts (OR 0.27, 95% CI 0.11 to 0.68, p=0.006).

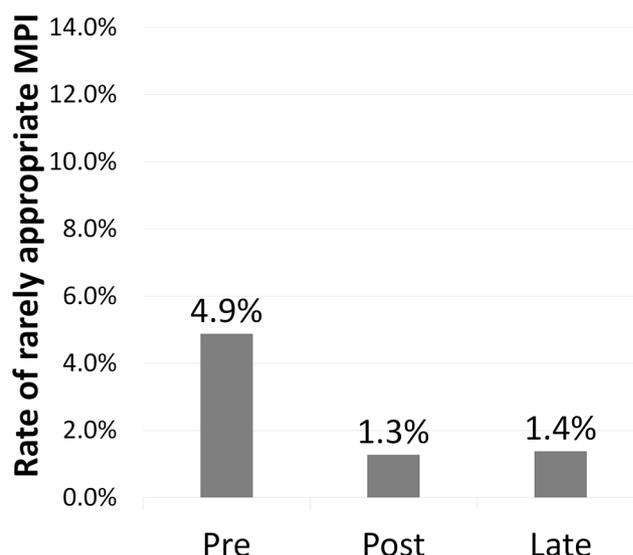Among patients with a rarely appropriate indication for testing, none had myocardial ischaemia (SDS ≥3).

Figure 1 Rate of rarely appropriate myocardial perfusion imaging (MPI) for three patient cohorts. In this bar graph, the rate of rarely appropriate MPI is displayed, significantly decreasing from 4.9% (n=14, pre cohort) to 1.3% (n=4, post cohort) and persisting at 1.4% 4 months after (n=4, late-post cohort) (p<0.0001 for trend).

Table 2 Myocardial perfusion imaging results

	Appropriate, n=780		Maybe appropriate, n=65		Rarely appropriate, n=22		p Value
	n	%	n	%	n	%	
SDS ≥3	57	7.3	1	1.5	0	0.0	0.09
SDS ≥7	19	2.4	1	1.5	0	0.0	0.69
Abnormal	203	25.6	15	22.7	1	4.5	0.07

Four tests which were not successfully rated were excluded.
SDS, summed difference score.

(table 2) The prevalence of ischaemia and abnormal testing decreased from appropriate to maybe appropriate, to rarely appropriate test indications, although not significantly (p=0.09 for SDS ≥3; p=0.07 for abnormal test). In our logistic regression model of ischaemia on MPI, the only predictive variable was prior MI (OR 2.47, 95% CI 1.34 to 4.55, p=0.004).

Ten indications accounted for 74% of all the MPI performed (table 3). The most common indication overall was 2013 indication #64, ischaemic equivalent in a patient with prior revascularisation (n=163). Asymptomatic patients who can exercise (2013 indication #8) and preoperative assessment in patients with good exercise capacity (2013 indication #71) were among the most common rarely appropriate indications for testing. We evaluated the most common rarely appropriate indication

for each of the three time cohorts (pre, post and late-post); however, the sample sizes were too small to make any conclusive observations about change in individual indications for testing.

DISCUSSION

We designed a specialty-specific, multimodality educational intervention to test the hypothesis that education could effectively deimplement rarely appropriate MPI testing. Our results demonstrate that in immediate and short-term follow-up at a facility such as ours, education may be effective in this regard.

Education is perceived as being a weak intervention with a mixed track record for changing the behaviour of physicians' and APPs' ordering patterns. One of the first

Table 3 Most common overall and rarely appropriate MPI indications

Rank	n	AUC version	Indication*	Rating	Description
Most common test indications overall					
1	163	2013	64	A	Postrevascularisation with ischaemic equivalent
2	121	2013	3	A	Symptomatic with intermediate pretest CAD risk and able to exercise
3	128	2009	9	A	Acute chest pain, possible ACS, high TIMI risk, negative troponin
4	76	2009	8	A	Acute chest pain, possible ACS, low TIMI risk, negative troponin
5	59	2013	4	A	Symptomatic with intermediate pretest CAD risk and unable to exercise
6	45	2013	58	A	Non-obstructive CAD on coronary angiography
7	19	2013	24	A	Abnormal ECG with intermediate/high CAD risk
8	19	2013	25	A	Abnormal exercise ECG test
9	15	2013	74	M	Poor/unknown functional capacity, intermediate risk surgery ≥1 risk factor
10	14	2013	76	A	Poor/unknown functional capacity, kidney transplant evaluation
Most common rarely appropriate indications†					
1	8	2013	8	R	Asymptomatic with intermediate pretest CAD risk and able to exercise
2	3	2013	71	R	Moderate or good functional capacity prior to any surgery
3	2	2013	67	R	Asymptomatic and <5 years after coronary bypass surgery

*Indication is the number assigned of the individual AUC clinical scenario described in the table.
†The remainder of rarely appropriate tests were each ordered only once.
ACS, acute coronary syndrome; AUC, Appropriate Use Criteria; CAD, coronary artery disease; TIMI, thrombolysis in myocardial infarction.

studies of education to reduce rarely appropriate MPI used the same basic methods as in our study: lectures and printed materials.[4] In this study, education was not effective, which set a negative tone for future application of education-only strategies. Looking at education on a broader sale, a recent systematic review on interventions to change provider behaviour suggests that the majority of studies on clinician education are actually effective.[8] We believe there are a few reasons why our educational intervention may have changed provider behaviour in this particular setting. First, prior data demonstrated that a substantial portion of providers at our facility were unaware of AUC. This served as a needs assessment which indicated that an underlying cause of overuse might be lack of knowledge about appropriateness. Because we observed a durable reduction in rarely appropriate imaging, it may be reasonable to conclude that lack of knowledge was a key driver more so than fixed practice habits or false beliefs about the benefit of testing. Second, prior data showed that the patterns of rarely appropriate testing differed by specialty. Using education tailored to the audience may have enhanced the effect. Third, we adopted a collaborative tone focused on improving patient care and used a peer to deliver a multimodality approach. A collaborative tone, peer delivery, focus on patient care and not on criticism are characteristics of effective audit and feedback programmes.[14] Similar observations have been made for educational interventions geared towards training physicians in delivering high value care.[15] This conclusion is indirectly supported by the survey data from providers that indicated their intent to change MPI ordering habits based on the education received.

Our data suggest that education may still be a worthwhile endeavour, particularly in settings where awareness of a best practice is low. In contrast, a stronger intervention, such as audit and feedback, may not be well suited for deimplementing rarely appropriate MPI at our facility. Over a 6 month time frame, only 22 rarely appropriate MPI were observed. Given that we have dozens of providers who can order MPI, a periodic report for individual providers would not likely give timely enough feedback to have an appreciable impact on ordering behaviour. An alternative approach, immediate feedback with a point-of-care decision support tool, is being considered at our facility. We must also acknowledge that given the low rate of rarely appropriate testing at baseline, further efforts to reduce low-value tests may have detrimental effects or be discouraging providers from ordering tests that may benefit individual patients. While the AUC do not include a target proportion for rarely appropriate testing, the majority of providers at our facilities feel that 0% to 5% is an acceptable proportion.[10]

An alternative explanation to our observations is that providers did not change their ordering habits, but rather just changed their documentation when ordering MPI to meet the AUC. It is possible that MPI use was largely appropriate all along and providers are now accurately documenting the appropriate reasons for ordering MPIs. We also cannot distinguish between a true education effect of the intervention versus what might be termed a Hawthorne effect that simply raised awareness about MPI appropriateness. However, we did not make any public declaration that the appropriateness of MPI was being tracked at our facility, and furthermore, the persistently low rate of rarely appropriate testing months after the education concluded suggests that the effect is a true one and due to a change in providers' knowledge about appropriate use of MPIs.

While not a primary focus of this investigation, these data add to the body of literature on the low prevalence of ischaemia among rarely appropriate tests. This observation provides further confidence to the application of AUC for cancelling or eliminating rarely appropriate tests a priori. We have also duplicated the finding that lack of symptoms continues to be a strong predictor of a MPI test being rarely appropriate. While acknowledging that women, diabetics, and post-transplant patients may have a more subtle symptom profile, physicians and APPs thinking of ordering a MPI for a patient without symptoms should consult the AUC given the robust association with rarely appropriate indications. In should also be acknowledged that MPI can be used for risk stratification, but also for diagnosing CAD. In the latter, some additional leeway for rarely appropriate testing may be in order. Lastly, we observed the common use of MPI for preoperative risk assessment in patients with good functional capacity. This continues to be an area for improvement in the application of AUC for MPI.

Our study is limited by the lack of a control group and longer term follow-up. Assessment of appropriateness is a time-consuming process when performed by a third party after the test has been ordered and continual tracking is not readily achievable. Our study is strengthened by the fact that appropriateness ratings were performed by someone with no vested interest in seeing the project succeed. The sample size of our investigation was too small to make conclusive statements about changes in the individual rarely appropriate test indications. There likely are also important characteristics of the VA study setting that may limit generalisability, such as the lack of a financial incentive either for the provider or for the institution to order MPIs. Future investigations should focus on replicating our findings in other settings using more a rigorous design with a control condition; assessing whether the effects of the education intervention appear to persist long term at the pilot site and ascertainment of which educational elements were effective and which could be omitted without diluting the effect.

CONCLUSION

While education has a mixed track record of effectiveness, it appears to be useful in selected circumstances. Future studies of education for changing behaviour should pay

close attention to the environmental conditions that may favour or oppose education as a change strategy.

Author affiliations
[1]Cardiology Section, Medical Service, Malcom Randall VA Medical Center, Gainesville, Florida, USA
[2]Division of Cardiovascular Medicine, Department of Medicine, College of Medicine, University of Florida, Gainesville, Florida, USA
[3]VA Puget Sound Health Care System, Seattle, Washington, USA
[4]Geriatric Research Education and Clinical Centers (GRECC), Malcom Randall VA Medical Center, Gainesville, Florida, USA
[5]Division of General Internal Medicine, Department of Medicine, College of Medicine, University of Florida, Gainesville, Florida, USA

REFERENCES

1. Elgendy IY, Mahmoud A, Shuster JJ, et al. Outcomes after inappropriate nuclear myocardial perfusion imaging: a meta-analysis. *J Nucl Cardiol* 2016;23:680–9.
2. Fonseca R, Negishi K, Otahal P, et al. Temporal changes in appropriateness of cardiac imaging. *J Am Coll Cardiol* 2015;65:763–73.
3. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;282:1458–65.
4. Gibbons RJ, Askew JW, Hodge D, et al. Appropriate use criteria for stress single-photon emission computed tomography sestamibi studies: a quality improvement project. *Circulation* 2011;123:499–503.
5. Bhatia RS, Milford CE, Picard MH, et al. An educational intervention reduces the rate of inappropriate echocardiograms on an inpatient medical service. *JACC Cardiovasc Imaging* 2013;6:545–55.
6. Bohossian HB, Park AW, Holcroft C. The impact of individual variation analysis on myocardial perfusion imaging utilization within a hospitalist group. *J Hosp Med* 2015;10:190–3.
7. Chaudhuri D, Montgomery A, Gulenchyn K, et al. Effectiveness of quality improvement interventions at reducing inappropriate cardiac imaging: a systematic review and meta-analysis. *Circ Cardiovasc Qual Outcomes* 2016;9:7–13.
8. Colla CH, Mainor AJ, Hargreaves C, et al. Interventions aimed at reducing use of low-value health services: a systematic review. *Med Care Res Rev* 2016. epub ahead of print.
9. Winchester DE, Hymas J, Meral R, et al. Clinician-dependent variations in inappropriate use of myocardial perfusion imaging: training, specialty, and location. *J Nucl Cardiol* 2014;21:598–604.
10. Kline KP, Plumb J, Nguyen L, et al. Patient and provider attitudes on appropriate use criteria for myocardial perfusion imaging. *JACC Cardiovasc Imaging* 2016. epub ahead of print.
11. Hendel RC, Berman DS, Di Carli MF, et al. ACCF/ASNC/ACR/AHA/ASE/SCCT/SCMR/SNM 2009 appropriate use criteria for cardiac radionuclide imaging. *Circulation* 2009;119:e561–87.
12. Wolk MJ, Bailey SR, Doherty JU, 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. *J Am Coll Cardiol* 2014;63:380–406.
13. Faul F, Erdfelder E, Lang AG, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;39:175–91.
14. Ivers NM, Grimshaw JM, Jamtvedt G, et al. Growing literature, stagnant science? systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. *J Gen Intern Med* 2014;29:1534–41.
15. Stammen LA, Stalmeijer RE, Paternotte E, et al. Training physicians to provide high-value, cost-conscious care. *JAMA* 2015;314:2384–400.