

# Validity versus feasibility for quality of care indicators: expert panel results from the MI-Plus study

ADOLFO PEÑA<sup>1,2</sup>, SANDEEP S. VIRK<sup>3</sup>, RICHARD M. SHEWCHUK<sup>4,5</sup>, JEROAN J. ALLISON<sup>6</sup>,  
O. DALE WILLIAMS<sup>5,7</sup> AND CATARINA I. KIEFE<sup>6</sup>

<sup>1</sup>University of Alabama-Veterans Affairs National Quality Scholars (VAQS) Fellowship Program, Birmingham VA Medical Center, Birmingham, AL, USA, <sup>2</sup>C-SMART Center for Surgical, Medical Acute Care Research and Transitions, Birmingham VA Medical Center, Birmingham, AL, USA, <sup>3</sup>University of Alabama-Montgomery Internal Medicine Residency Program, Birmingham, AL, USA, <sup>4</sup>Department of Health Services Administration, School of Health Related Professions, University of Alabama, Birmingham, AL, USA, <sup>5</sup>Center for Outcomes and Effectiveness Research and Education, University of Alabama, Birmingham, AL, USA, <sup>6</sup>Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA, and <sup>7</sup>Department of Medicine, University of Alabama at Birmingham School of Medicine, Birmingham, AL, USA

Address reprint requests to: Adolfo Peña, C-SMART Center for Surgical, Medical Acute Care Research and Transitions, Birmingham VA Medical Center, 151 REAP, 700 South 19th Street, Birmingham, AL 35233, USA. Tel: +1-(205)-933-8101-5301; Fax: +1-(205)-212-3998; E-mail: adolfope@uab.edu

Accepted for publication 13 March 2010

---

## Abstract

**Background.** In the choice and definition of quality of care indicators, there may be an inherent tension between feasibility, generally enhanced by simplicity, and validity, generally enhanced by accounting for clinical complexity.

**Objectives.** To study the process of developing quality indicators using an expert panel and analyze the tension between feasibility and validity.

**Design and participants.** A multidisciplinary panel of 12 expert physicians was engaged in two rounds of modified Delphi process to refine and choose a smaller subset from 36 indicators; these were developed by a research team studying the quality of care in ambulatory post-myocardial infarction patients with co-morbidities. We studied the correlation between validity/feasibility ranks provided by the expert panel. The correlation between the quality indicators ranks on validity and feasibility scale and variance of experts' responses was also individually studied.

**Results.** Ten of 36 indicators were ranked in both the highest validity and feasibility groups. The strength of association between validity and feasibility of indicators measured by Kendall tau-b was 0.65. In terms of validity, a strong negative correlation was observed between the ranks of indicators and the variability in expert panel responses (Spearman's rho,  $r = -0.85$ ). A weak correlation was found between the ranks of feasibility and the variability of expert panel responses (Spearman's rho,  $r = 0.23$ ).

**Conclusion.** There was an unexpectedly strong association between the validity and feasibility of quality indicators, with a high level of consensus among experts regarding both feasibility and validity for indicators rated highly on each of these attributes.

**Keywords:** clinical guidelines, feasibility, myocardial infarction, quality indicators, quality of health care, validity

---

## Introduction

The use of more effective therapies, based on strong evidence, can decrease morbidity and mortality from coronary heart disease and myocardial infarction (MI). For example, if all heart attack survivors receive timely beta-blocker therapy, an estimated of 1500 deaths could be avoided each year [1]. Although most physicians are aware of evidence-based clinical practice guidelines, promulgation of guidelines does not

necessarily translate into their application in clinical practice [2, 3]. Several barriers to increase the adherence toward clinical practice guidelines have been identified, for example, expectations about outcomes with guidelines and the inertia of previous practices [4]. Additionally, patients with chronic disease, such as ischemic heart disease, have other co-morbidities, which further complicate the process of care that can be applied uniformly. Also, the movement towards evidence-based medicine has resulted in an intense proliferation of clinical practice guidelines, with thousands now available.

Indicators of quality of care represent measurable aspects of care that reflect key components of a guideline. An indicator is an observable trait or variable that is assumed to point to the assessment of some other trait, usually difficult to observe directly [5]. Increasingly, quality indicators are being developed to assess different dimensions of care as adherence to clinical practice guidelines, physician performance and quality of care for specific conditions like acute MI [6], urinary tract infections [7] and epilepsy [8]. Each guideline can be used to develop multiple indicators based on that guideline.

An ideal quality of care indicator would be valid, reliable, sensitive, specific and feasible [9, 10], among other characteristics related to the importance of the construct being measured. Validity can be understood as the measurement under consideration corresponding to the true condition of the event being measured. However, because a quality indicator reflects a minimal acceptable standard of care, validity also relates to the degree of relevance of the proposed recommendations (clinical practice guidelines). Feasibility refers to the ease with which the quality indicator can be measured accurately; for example, the frequency of prescriptions of beta-blockers for patients with MI is highly feasible if there is a computerized order entry system that allows flexible search inquiries [10].

Conceptually, there is an overlap between validity and feasibility. However, increasing validity may increase complexity of measuring, and thus render the indicator more difficult to measure, thus decreasing feasibility. Conversely, easily measurable indicators may ignore clinical complexity and raise questions regarding their validity. One study showed that feasibility rankings were not associated with the overall utility rankings of quality indicators for Parkinson's disease [11]. Another study reported a negative correlation between feasibility and 'room for improvement' of quality indicators for stroke [12]. Development of quality indicators is an arduous task, and to understand the relationships among the characteristics of quality indicators is helpful. Also, the validity and feasibility of a particular guideline have been shown to predict implementation of the guideline in clinical settings [13, 14]. Thus, understanding the tension between feasibility and validity should help in finding the reasons why some clinical practice guidelines are implemented more effectively than others in clinical practice.

MI-Plus is a group-randomized trial funded by the National Heart Lung and Blood Institute (NHLBI) that developed and tested an intervention to increase physician adherence to guidelines for complex post-MI patients with co-morbidities. As the backbone of this intervention, valid quality indicators for use in feedback and performance measurements were developed. This manuscript describes this process and examines the tension between feasibility and validity in the selection of quality of care indicators.

## Methods

### Clinical guidelines review process

In 2003, before identifying appropriate quality indicators, the MI-Plus research team reviewed available guidelines for the

management of ischemic heart disease as well as for the co-morbidities that are common in post-MI patients, including hypertension, hyperlipidemia, diabetes mellitus, heart failure, renal disease, stroke, peripheral vascular disease and depression. In all, 22 guidelines and position papers from recognized entities were used (Table 1). Disease-specific guidelines along with the level of evidence for these co-morbidities were then organized into a summary statement. American College of Cardiology and the American Heart Association (ACC/AHA) guidelines were the primary source for this summary, but we used other guidelines as well. For example, to address lipid management, guidelines from the National Cholesterol Education Program (NCEP) ATP III guidelines were preferred over ACC/AHA guidelines as these guidelines covered a broader aspect of management of patients with lipid disorders.

### Development of quality indicators

Key recommendations were selected from the previously reviewed guidelines. Based on the available literature, the level of evidence was graded for each recommendation to determine its scientific soundness. Then the research team selected processes of care that exert a large impact on the population and manifest demonstrable variations of care, thus revealing potential for improvement of quality of care of post-MI patients with co-morbidities. The indicators were derived from these processes of care that respond to recommendations that are evidence based, actionable, and under the control of the physician or health-care organization. Quality indicators were constructed using an 'IF-THEN-BECAUSE' format based on the methodology used in studies by Shekelle *et al.* [15] and MacLean *et al.* [16]. IF refers to the clinical characteristics that describe a person's eligibility for the quality indicator; THEN indicates the actual process of care that should or should not be performed and BECAUSE refers to the expected health impact if the indicator is performed. For example, IF a patient has had an MI and is not on aspirin AND no contraindications are present, THEN aspirin should be prescribed at a level of 75–325 mg/day BECAUSE aspirin inhibits platelet aggregation and significantly reduces the risk of reinfarction and mortality in post-MI patients.

Thirty-six quality indicators were drafted. These indicators were presented to the panel of expert members as quality measures focusing on standard processes of care. Hence, in the numerator, the appropriate process of care was included, and in the denominator, only patients with clear indications and no contraindications for treatment or recommendations should be included. Although the research team paid attention to proper specific numerator, denominator, sampling methodology, ease interpretation, risk adjustment and economical and logistic data, it was not possible to provide specific information about reliability, sensitivity, specificity or specific auditing techniques for some indicators. Information about the evidence rating or classification (i.e. Classes I, II, III and A, B, C) for each recommendation was provided along with pertinent references. Moreover, a summary document provided all this information as well as information

**Table 1** Guidelines used in identifying quality indicators

- 
- 1 ACC/AHA Guideline update for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction, 2002
  - 2 ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult, 2001
  - 3 ACC/AHA guidelines for the management of patients with acute myocardial infarction. 1999 updated version, p. 141
  - 4 ACC/AHA/ACP-ASIM guidelines for the management of patients with chronic stable angina, 1999
  - 5 ACC/AHA/NHLBI clinical advisory on the use and safety of statins. *J Am Coll Cardiol* 2002;40(3).
  - 6 AHA scientific statement: secondary prevention of coronary heart disease in the elderly. *Circulation* 2002
  - 7 AHA/ACC guidelines for preventing heart attack and death in patients with atherosclerotic cardiovascular disease: 2001 update.
  - 8 American Diabetes Association: clinical practice recommendations. *Diabetes Care* 2002: S51
  - 9 Cardiac rehabilitation. A national clinical guideline, Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network (SIGN), 2002, p. 32
  - 10 Congestive heart failure in adults. Institute for clinical systems improvement, 2002
  - 11 Executive summary of the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 2001
  - 12 Institute of clinical systems improvement guidelines; treatment of acute myocardial infarction, 2001
  - 13 JNC VI: risk stratification and treatment of recommendations. 1999 Clinical advisory update: treatment of hypertension and diabetes.
  - 14 NIH/NHLBI clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults, 1998
  - 15 Secondary prevention of coronary heart disease following myocardial infarction. SIGN publication 41, 2000
  - 16 SIGN publication No. 55: Management of diabetes. Management of diabetic cardiovascular disease, 2001, p. 14
  - 17 The pharmacological management of chronic heart failure. VA National Clinical Practice Guidelines Group, 2001
  - 18 The sixth report of the JNC on prevention, detection, evaluation and treatment of high blood pressure, 1997
  - 19 The Thrombosis Interest Group of Canada-Practice Treatment guidelines Post-MI, 2002
  - 20 Treating tobacco use and dependence. A clinical practice guideline. U.S. Department of Health and Human Services, Public Health Service, 2000
  - 21 University of Newcastle upon Tyne: prophylaxis for patients who have experienced a myocardial infarction: drug treatment, cardiac rehabilitation and dietary manipulation, 2001.
  - 22 AACE/ACE position statement on the prevention, diagnosis and treatment of obesity. 1998 Revision. *Endocrine Practice* 1998:P297
- 

about the methodology used to create each indicator. These materials are available upon request. Table 2 shows the list of abbreviated quality indicators (in terms of processes of care) that were proposed for review by the expert panel.

### Expert panel review

The expert panel was composed of 12 practicing physicians. The 12-member panel included experts in cardiovascular disease, community practitioners and academic physicians from different geographical areas of the USA. The primary purpose of the expert panel was to provide external validation of the study team's selection of quality indicators for use in our MI-Plus study's performance measurement system. The expert panel had three primary tasks (i) To review a set of quality indicators; (ii) to rate the indicators on a set of validity and feasibility criteria and identify a final set of indicators; and (iii) to assist the study team in operationalizing the final panel of quality indicators for use in performance measurement.

### Panel process

During summer and fall of 2003, two rounds of the modified Delphi process were used to obtain the validity and

feasibility of the proposed indicators using the RAND/UCLA appropriateness method [17]. Round 1 involved rating the indicators on validity criteria alone and also collecting information on barriers to implementation. The validity of quality indicators was rated using a 9-point scale with higher numbers indicating greater validity. The validity criteria proposed to the expert panel was defined by: (i) adequate scientific evidence or professional consensus supported a link between the performance of that indicator and overall positive outcome for post-MI patients; (ii) a physician with higher rates of adherence to that indicator would be considered a higher quality provider; and (iii) the physician or health plan influenced most factors that determine adherence to the indicator [11, 15, 18].

Round 2 focused on areas of disagreement about the validity of indicators, refining the indicators to improve specificity, and ranking (not rating) the indicators on feasibility criteria. We considered an indicator to be feasible if (i) information on adherence to the indicator is expected to be available and documented in the medical record; OR (ii) information on the indicator is expected to be available from patient or proxy surveys or interviews and is likely to be accurate; AND (iii) information collection is of reasonable cost and does not impose an inappropriate burden on

**Table 2** Quality indicators rating and ranking based on expert panel review

QI	Abbreviated quality indicators	Validity rating (1–9)	Rank on validity scale (1–36)	Feasibility average rank	Rank on feasibility scale (1–36)	Level of evidence weighting of evidence <sup>a</sup>
01	Antiplatelet therapy (aspirin)	8.92	1	3.67	1	A/I
25	ACE for CHF	8.75	2	5.42	3	A/I
15	ACE or ARBs w/impaired renal fx (proteinuria)	8.58	3	11.5	9	A/I
10	Smoking status documented	8.50	4	7.25	4	B
16	BP control for diabetes with hypertension	8.50	5	8.50	5	A
05	Beta-blockers post-MI	8.42	6	5.00	2	A/I
06	ACE w/low EF, CHF or diabetes	8.42	7	9.00	7	A/I
23	Complete lipid profile ordered (post-MI)	8.17	8	9.83	8	A
13	Monitor HbA1C in diabetics	8.08	9	12.00	10	B
11	Smoking counseling	8.08	10	16.08	12	B
27	Beta-blockers for CHF	7.92	11	16.67	14	A/I
08	Statin therapy for LDL	7.83	12	8.83	6	A/I
30	Diuretics in CHF w/edema	7.83	13	14.58	11	A
14	Glycemic control intervention for diabetes	7.75	14	17.50	15	B
24	Annual lipid profiles (follow-up)	7.75	15	18.75	16	A
19	Annual eye examination in diabetics	7.75	16	19.50	18	Not specified
36	No HRT prescribed	7.75	17	28.67	35	IIa
34	Weight loss consultation	7.67	18	25.25	26	B
33	Assess for risk of depression	7.67	19	26.92	30	Not specified
21	BB and ACE as first-line therapy in HTN	7.42	20	21.75	23	B
18	Annual influenza vaccinations in diabetics	7.33	21	16.58	13	Not specified
35	Risk of HRT discussed	7.33	22	29.25	36	Not specified
20	Annual foot examination in diabetics	7.25	23	19.42	17	Not specified
28	Calcium channel blockers in CHF	7.17	24	27.08	31	Not specified
02	Antiplatelet therapy (Clopidogrel)	7.08	25	20.33	21	A/I
32	Monitor serum K <sup>+</sup> w/dig/diuretics in combination	7.00	26	22.33	24	Not specified
31	Spirolactone for Classes III/IV CHF	7.00	27	23.58	25	Not specified
12	Nicotine replacement therapy in smokers	6.92	28	27.17	32	Not specified
04	Antiplatelet therapy (Warfarin and antiplatelet drugs)	6.67	29	25.42	28	Not specified
03	Antiplatelet therapy (Warfarin)	6.5	30	20.00	19	I, IIa, IIb
26	ARBs for CHF when ACE contraindicated	6.33	31	20.50	22	Not specified
09	Monitor CK in combination lipid lowering therapy	6.25	32	25.83	29	Not specified
07	ACE without low EF, CHF or diabetes	6.17	33	20.25	20	A
17	Monitor serum creatinine q 6mos in HTN	6.00	34	25.25	27	Not specified
22	Assess pt adherence in uncontrolled HTN	6.00	35	28.08	33	Not specified
29	Monitor dig levels in renal failure	5.92	36	28.25	34	Not specified

<sup>a</sup>In abstracting recommendations from MI-Plus guidelines, we considered the evidence rankings (A, B, C) and the ACC/AHA classifications (classes I, II, III) when assigned or provided by the guideline.

health-care systems [11, 18, 19]. Each expert was asked to divide the 36 indicators into 3 groups (12 in each) based on their perceived feasibility in clinical practice, i.e. most feasible, moderately feasible and least feasible. The panel was then asked to further rank the quality indicators in these three groups, from highly to least feasible, thus conferring a rank between 1 and 36 to the feasibility of each indicator. The results from round 2 were summarized and fed back to the panel again.

**Scoring validity and feasibility**

After round 2 feedback, as above, final values of average validity rating and feasibility ranks provided by the expert panel were tabulated. Determining the validity scoring took two steps. The first step was to average the validity rating of each indicator. The second step was to rank each average validity rating from 1 to 36. Determining the feasibility scoring took only one step: to determinate the average rank for each indicator and the rank order list.

**Analysis**

Kendall’s tau-b coefficient was used to estimate the strength of association of expert panel responses on the validity and feasibility rank scale. Since the quality indicators were ranked on the validity and feasibility scale, we used Spearman’s rho to estimate the correlation between the quality indicators ranks on validity and feasibility scale and variance in expert panel responses. To understand the expert panel preferences for quality indicators, validity average rating and feasibility average ranking were individually plotted against the variance in expert panel responses for each indicator.

**Results**

Final validity and feasibility average ratings and their corresponding ranks are shown in Table 2. Table 3 displays the cross-tabulation of final validity and feasibility categorized into tertiles of rank. The upper right-hand corner displays the 10 indicators that were ranked in both the highest validity and feasibility groups. An additional three indicators were ranked having high validity and medium feasibility. These 13 indicators were identified as most relevant for the care of post-MI patients with multiple co-morbidities (description of indicators in Table 4).

The strength of association between validity and feasibility ranks measured by Kendall tau B was 0.65 ( $P < 0.001$ ), suggesting a good association between overall validity and feasibility of quality indicators. In terms of validity, a strong negative correlation was observed between the ranks of quality indicators and the variability in expert panel responses (Spearman’s rho,  $r = -0.85$ ,  $P < 0.001$ ) (Fig. 1). By contrast, the correlation between the ranks of feasibility and variability of expert panel response was weak (Spearman’s rho,  $r = 0.23$ ,  $P > 0.05$ ). Figure 2 shows an inverse ‘U’-shaped like curve suggesting the divergence of opinion was smallest on the ends of the feasibility spectrum and greatest in the middle.

**Discussion**

We describe the process of finalizing quality indicators for improving the care of post-MI patients. There was a surprisingly strong correlation between the overall validity and feasibility. Furthermore, there was an increased variance in the expert panel responses as the rating of quality indicators decreased, in terms of validity. Regarding the feasibility of

**Table 3** Cross-tabulation of validity and feasibility tertiles based on ranking of quality indicators

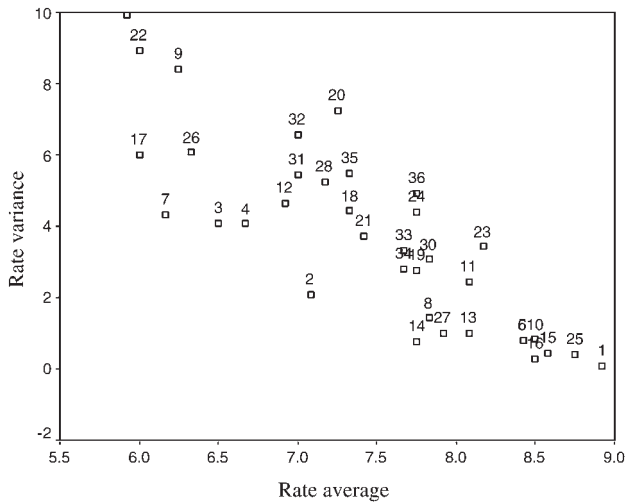
				Q101 Q106 Q110 Q115 Q123	Q105 Q108 Q113 Q116 Q125	1	Feasibility (1,-high; 3,-low)
Q102 Q107 Q131	Q103 Q126 Q132	Q114 Q119 Q121	Q118 Q120 Q124	Q111 Q130	Q127	2	
Q104 Q112 Q122	Q109 Q117 Q129	Q128 Q134	Q133 Q135 Q136			3	
	3		2		1		
Validity (1,-high; 3,-low)							

**Table 4** Quality indicators for post-MI patients rated/ranked higher on validity and feasibility criteria based on expert panel review

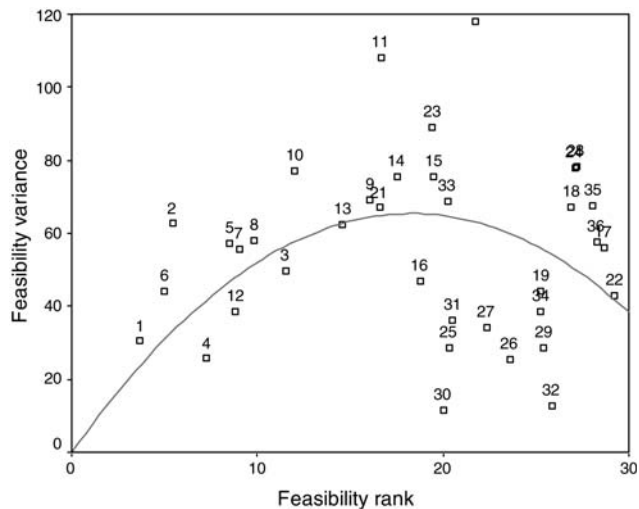
QI01	Aspirin	IF a patient has had an MI and is not on ASA therapy AND there are NO contraindications THEN aspirin therapy (75–325 mg/day) should be prescribed BECAUSE aspirin inhibits platelet aggregation and reduces risk of reinfarction and mortality in post-MI patients significantly.
QI05	Beta-blockers	IF a patient has had an MI AND there are NO contraindications THEN beta-blockers should be prescribed BECAUSE the use of beta-blockers significantly reduces post-MI mortality.
QI06	ACE inhibitors/ARBs	IF a patient has had an MI and has co-morbidities characterized by ejection fraction <40%, symptomatic/asymptomatic CHF, diabetes mellitus AND there are NO contraindications THEN ACE inhibitors or ARBs should be prescribed BECAUSE the use of ACE inhibitors or ARBs reduces mortality and the risk of non-fatal and fatal vascular events in patients with a history of MI.
QI08	Statins	IF a patient has had an MI AND, LDL cholesterol is >100 mg/dL AND there are NO contraindications THEN statin therapy should be prescribed BECAUSE patients with elevated LDL cholesterol do not have good outcomes post-MI and the use of statin therapy reduces mortality in post-MI patients.
QI10	Smoking status documentation	IF a patient has had an MI THEN smoking status should be documented BECAUSE smoking is a major risk factor in CHD and continued smoking increases the risk of a mortality and future MI.
QI11	Smoking counseling	IF a patient is post-MI AND is a current smoker THEN counseling regarding smoking cessation should be given BECAUSE counseling is effective in increasing smoking cessation and smoking is an independent risk factor for reinfarction.
QI13	A1C measure	IF a post-MI patient has diabetes THEN glycosylated hemoglobin (HgbA1c) should be measured at least every 6 months BECAUSE maintaining A1c at a desirable level reduces complications of diabetes.
QI15	ACE inhibitors/ARBs in diabetes with renal failure	IF a post-MI patient with diabetes has impaired renal function characterized by proteinuria, either microscopic or macroscopic AND there are no contraindications THEN ACE inhibitors or ARBs should be prescribed BECAUSE these drugs decrease progression to renal failure.
QI16	Hypertension control	IF a post-MI patient with diabetes has hypertension THEN therapy to lower blood pressure to goal (<130/80 mmHg) should be prescribed BECAUSE patients with better hypertension control have reduced mortality and fewer complications.
QI23	Lipid profile	IF a post-MI patient has not had a complete lipid profile (total cholesterol, LDL, HDL) since hospital discharge for the index MI (i.e. 6–12 weeks post-MI) THEN a complete lipid profile should be ordered to assess cholesterol levels BECAUSE high cholesterol increases the risk of future coronary events and mortality in post-MI patients.
QI25	ACE inhibitors/ARBs in CHF	IF a post-MI patient has CHF AND there are NO contraindications THEN an ACE inhibitor or ARBs should be prescribed BECAUSE the use of ACE Inhibitors reduces progression to severe heart failure and reduces mortality and the risk of non-fatal and fatal vascular event in patients with MI.
QI27	Beta-blockers in CHF	IF a post-MI patient has CHF and there are no contraindications THEN a beta-blocker (e.g. metoprolol, carvedilol, bisoprolol) should be used BECAUSE beta-blockers significantly reduce post-MI mortality.
QI30	Diuretics	IF a post-MI patient with CHF has fluid retention characterized by pulmonary edema, peripheral edema AND there are no contraindications THEN diuretics should be prescribed BECAUSE diuretics increase functional status, exercise tolerance and improvement in symptoms.

indicators, there was more variance in the middle of the ranking scale, signifying more consensus in the expert panel response for the quality indicators on both ends of ranking scale.

Through a consensus method, a modified Delphi process, our expert panel identified 13 quality indicators as valid and feasible for the care of post-MI patients. The 13 quality



**Figure 1** Validity rating and variance. Rate average: validity average ratings of quality indicators. Rate variance: variability of expert panel responses.



**Figure 2** Feasibility average ranking and variance. Feasibility rank: feasibility average rank of quality indicators. Feasibility variance: variability of expert panel responses.

indicators addressed different domains of post-MI care and helped prioritize the quality of improvement efforts. Several studies [6–8, 11–12, 15–16, 20, 21] have used a similar approach in working with expert panels to examine both the validity and feasibility, but very few studies have attempted to analyze the association between the attributes of quality indicators. Holloway *et al.* [12] found that the feasibility ratings of quality indicators were not associated with their overall utility rating. In a more recent study, a normative criterion was used to choose indicators. Each panel member was asked to rate each indicator on a 3-point scale: ‘do not include,’ ‘could include’ and ‘must include’ [6]; however, the relationship between this criterion and validity or feasibility was not analyzed. To the best of our knowledge, this is the first study

that evaluates the association between validity and feasibility of quality indicators.

The surprisingly high level of association between validity and feasibility appears counter the postulated tension between these two attributes. Our findings suggest that, for the members of our expert panel, both constructs were not completely bi-dimensional or independent, even when specific instructions for validity and feasibility were offered. In fact, notwithstanding our instructions and information to the panel, there may have been confusion in panel members’ minds between validity and feasibility, with a tendency to equate a certain lack of definition in the less valid indicators with less feasibility. For example, the assessment of patient adherence in uncontrolled hypertension (QI22, Table 2) had final validity and feasibility ranks of 34 and 33, respectively. Yet, it would appear to be a clearly important, and thus a valid measure of the quality of care to perform such assessments. The known difficulty in assessing patient adherence (low feasibility) might have been conducive to its very low rating on the validity scale as well.

There was high level of agreement among our experts in considering an indicator as highly valid but less agreement for indicators perceived as having lower validity. There was a great deal of consensus for validity among the panel for the quality indicators that were higher on the rating scale and considerable variability in their responses for the variables that were lower in the rating order (Fig. 1). This relationship could be attributed to the level of scientific evidence showing the efficacy of the processes of care and recommendations used to derive the indicators. This may support the intuitive notion that validity is an attribute more dependent on clinical practice guidelines and levels of clinical evidence. As we can see in Table 2, the level of evidence for the majority of the 13 indicators that ranked highly is A; however, for the low-ranking indicators, the level of evidence was not clearly stated.

The expert panel evaluation of feasibility showed a consensus on the highly feasible and least feasible quality indicators, suggesting that agreement is easily reached on the clinical decisions that are simple to do or simple to ignore. However, the indicators ranking in the middle of the feasibility scale were widely dispersed with respect to the expert panel response, indicating greater uncertainty. Besides, the ‘U-like’ shape of Fig. 2 suggests that there is not a monotonic correlation. The overall correlation between the level of agreement (i.e. the variability in the expert panel answers) and feasibility was weak (Spearman’s rho,  $r = 0.23$ ), additionally, we were unable to prove that it was statistically significant ( $P > 0.05$ ). These findings may suggest that feasibility is a construct difficult to define and assess even for a panel of experts. We speculated that this might be partially due to the fact that feasibility of quality indicators are not scientifically measured, and therefore, most of the experts’ rankings may be based on views, beliefs and common sense. Furthermore, it is plausible that feasibility is an attribute more dependent on operational issues related to the specific quality indicators derived from the clinical practice guidelines and, as such, may be more context specific. This implies important

consequences: if the level of agreement of an expert panel assessing the feasibility of quality indicators is low, future studies must pay more attention to the feasibility construct by elaborating more detailed criteria and involving professional panels with great experience in collecting data and evaluating the implementation of quality indicators. Furthermore, in the future, it will be necessary to empirically test, after implementation, the feasibility of the selected indicators.

## Limitations

The selection of quality indicators occurred several years ago during the implementation of the MI-Plus project; nevertheless, new evidence has been generated since then, thereby potentially causing changes in the choices for our final set of indicators. For example, a new guideline for the detection and management of post-MI depression has been recently proposed [22]. One of the main recommendations in this guideline indicates that MI patients should be screened for depression using a standardized depression symptom check list at regular intervals during the post-MI period; this recommendation is based on level A scientific evidence. However, in our study, the quality indicator intended to assess the risk of depression was not ranked high enough to be considered for our final set of indicators. With the new scientific evidence mentioned above, experts could plausibly rank this indicator for depression higher, which would change our final choices for the set of indicators. In terms of creating a quality indicator, while challenging the continuous flow of information and the production of scientific evidence, our set of indicators cannot adjust well to the dynamic nature of these measures. Additionally, although the members of the expert panel in this study included not only specialists but also generalists, all members were physicians. Experts in collecting data might provide different views regarding the feasibility of the indicators.

Almost all the selected quality indicators by the research team were rate-based indicators, which can be expressed as proportions of desired events; only one indicator (use of calcium channel blockers in congestive heart failure) was of sentinel type, indicators which identify undesirable actions [9]. What the tension is between validity and feasibility in the case of sentinel indicators is something this study cannot answer.

The methodology in this study was used to identify relevant processes of care, which affect outcomes based on scientific evidence. However, a quality indicator is not an exact synonymous with a process of care; therefore, the further elaboration of quality indicator is recommended: to write the measure specifications (unit of analysis, data collection specifications, and define them mathematically). How this can affect an expert's understanding of feasibility is not completely addressed by this study.

Methods to use similar data regarding validity and feasibility to differentiate dominant and weak quality indicators for medical conditions others than MI need further research. It is possible to select a small group of indicators if we

clearly know the validity, feasibility, utility, sensitivity, specificity and reliability of quality indicators. We could save time and resources if we knew whether the number of indicators affects validity. Finally, in addition to asking a panel of experts, more knowledge about these properties can be gained by testing the indicators in specific settings and with explicit methodologies. Empirical research to evaluate indicators can provide information regarding feasibility and can help to improve them.

## Conclusion

The development and the use of quality indicators pose a number of interesting methodological problems. Acquiring a better understanding of the relationship among the different attributes of quality indicators is urgent because ever more quality indicators are created to assess physician performance and quality of care, the measurement of which can become very resource intensive. We demonstrated a strong association between the validity and feasibility of quality indicators, which countered our intuitive expectation of the underlying tension in achieving a perfect quality indicator that ranked equally on both the validity and feasibility scale. This study shows increase variance in the expert panel rating low-value quality indicators and in the middle of ranking scale for feasibility. Future research should focus on further exploring the importance of correlation between validity and feasibility of indicators and developing the indicators based not only on the validity of guidelines but also on their feasibility in translation.

## Acknowledgements

The authors thank all the physicians, members of the MI-Plus expert panel, who made this study possible. We also thank Dr. Joshua S. Richman assistant professor of Medicine and Biostatistics, Division of Preventive Medicine, University of Alabama at Birmingham, and Dr. Ellen Funkhouser, associate professor, Department of Epidemiology, University of Alabama at Birmingham for helpful comments and suggestions on an earlier draft of our manuscript.

At the time the MI-Plus study and this analysis were conducted, Drs. Kiefe and Allison were affiliated to the Center for Outcomes and Effectiveness Research and Education and the Department of Medicine at the University of Alabama at Birmingham.

## Funding

This project was funded in part by grant number R01 HL70786 from the National Heart, Lung and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung and Blood Institute or the National Institutes of Health.



## References

1. Phillips KA, Shlipak MG, Coxson P *et al.* Health and economic benefits of increased beta-blocker use following myocardial infarction. *JAMA* 2000;**284**:2748–54.
2. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the Twenty-first Century*. Washington, DC: National Academies Press, 2001.
3. Liang L. The gap between evidence and practice. *Health Aff* 2007;**26**:w119–w121.
4. Timmermans S, Mauck A. The promises and pitfalls of evidence based medicine. *Health Aff* 2005;**24**:18–28.
5. Bunge M. Development indicators. *Soc Indicators Res* 1981;**9**:369–81.
6. Tu JV, Khalid L, Donovan LR *et al.* Indicators of quality of care for patients with acute myocardial infarction. *CMAJ* 2008;**179**:909–15.
7. Hermanides HS, Hulscher ME, Schouten JA *et al.* Development of quality indicators for the antibiotic treatment of complicated urinary tract infections: a first step to measure and improve care. *Clin Infect Dis* 2008;**46**:703–11.
8. Pugh MJ, Berlowitz DR, Montouris G *et al.* What constitutes high quality of care for adults with epilepsy? *Neurology* 2007;**69**:2020–7.
9. Mainz J. Defining and classifying clinical indicators for quality improvement. *Int J Qual Health Care* 2003;**15**:523–30.
10. Rubin H, Pronovost P, Diette G. From a process of care to a measure: the development and testing of a quality indicator. *Int J Qual Health Care* 2001;**13**:489–96.
11. Cheng EM, Siderowf A, Swartztrauber K *et al.* Development of quality of care indicators for Parkinson's disease. *Mov Disord* 2004;**19**:136–50.
12. Holloway RG, Vickrey BG, Benesch CG *et al.* Development of performance measures for acute ischemic stroke. *Stroke* 2001;**32**:2058–74.
13. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Med Care* 2001;**39**:II46–II54.
14. Campbell SM, Roland MO, Quayle JA *et al.* Quality indicators for general practice: which ones can general practitioners and health authority managers agree are important and how useful are they? *J Public Health Med* 1998;**20**:414–21.
15. Shekelle PG, MacLean CH, Morton SC *et al.* Assessing care of vulnerable elders: methods for developing quality indicators. *Ann Intern Med* 2001;**135**:647–52.
16. MacLean CH, Saag KG, Solomon DH *et al.* Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. *Arthritis Rheum* 2004;**51**:193–202.
17. Fitch K, Bernstein SJ, Aguilar MD *et al.* *The RAND/UCLA Appropriateness Method User's Manual*, Santa Monica, CA: RAND Corporation, 2001.
18. Campbell SM, Braspenning J, Hutchinson A *et al.* Improving the quality of health care: research methods used in developing and applying quality indicators in primary care. *BMJ* 2003;**326**:816–9.
19. *Desirable Attributes of Performance Measures. A Consensus Document from The American Medical Association, The Joint Commission on Accreditation of Healthcare Organizations, and The National Committee for Quality Assurance*, 1999.
20. Mikuls TR, McLean CH, Olivieri J *et al.* Quality of care indicators for gout management. *Arthritis Rheum* 2004;**50**:937–43.
21. Saliba D, Solomon D, Rubenstein L *et al.* Quality indicators for the management of medical conditions in nursing home residents. *J Am Med Dir Assoc* 2004;**5**:297–309.
22. Post-Myocardial Infarction Depression Clinical Practice Guideline Panel. AAFP guideline for the detection and management of post-myocardial infarction depression. *Ann Fam Med* 2009;**7**:71–9.