American College of Cardiology and American Heart Association Methodology for the Selection and Creation of Performance Measures for Quantifying the Quality of Cardiovascular Care

John A. Spertus, MD, MPH, FACC; Kim A. Eagle, MD, FACC, FAHA;
Harlan M. Krumholz, MD, FACC, FAHA; Kristi R. Mitchell, MPH;
Sharon-Lise T. Normand, PhD, MSc, FACC; for the American College of Cardiology and the
American Heart Association Task Force on Performance Measures

Abstract—The ability to quantify the quality of cardiovascular care critically depends on the translation of recommendations for high-quality care into the measurement of that care. As payers and regulatory agencies increasingly seek to quantify healthcare quality, the implications of the measurement process on practicing physicians are likely to grow. This statement describes the methodology by which the American College of Cardiology and the American Heart Association approach creating performance measures and devising techniques for quantifying those aspects of care that directly reflect the quality of cardiovascular care. Methods for defining target populations, identifying dimensions of care, synthesizing the literature, and operationalizing the process of selecting measures are proposed. It is hoped that new sets of measures will be created through the implementation of this approach, and consequently, through the use of such measurement sets in the context of quality improvement efforts, the quality of cardiovascular care will improve. (J Am Coll Cardiol 2005;45:1147–56.)

Key Words: ACC/AHA Performance Measures ■ cardiovascular diseases ■ healthcare quality ■ policy making

Medicine is experiencing an unprecedented increased focus on quantifying and improving the quality of health care. Although healthcare quality is a multidimensional construct that, as articulated by the Institute of Medicine (1), encompasses concepts of safety, equity, evidence-based medicine, timeliness of care, efficiency, and patient-centeredness, the foundation of efforts to improve care is predicated on measurement. Without the ability to quantify quality, the opportunity to identify practices that lead to higher-quality care, and the opportunity to learn how such care was delivered, quality cannot be improved. Therefore,

developing a framework to measure components of the quality of health care is of paramount importance.

The American College of Cardiology (ACC) and the American Heart Association (AHA) have developed a multifaceted strategy to facilitate the process of improving the quality of cardiovascular care. The initial phase of this effort was to create clinical practice guidelines that carefully review and synthesize the available evidence to better guide patient care. As articulated in a recent overview of the guidelines process, the creation of guidelines is but one component of the ACC's and the AHA's commitment to improving the

quality of cardiovascular care (2,3). Because guidelines are written in a spirit of suggesting diagnostic and/or therapeutic interventions for patients in most circumstances, a significant amount of judgment by clinicians is required to adapt the guidelines to the care of individual patients. Accordingly, the ACC/AHA guideline recommendations are generated with varying degrees of confidence based on the available evidence. Occasionally, the evidence supporting a particular structural aspect or process of care is so strong that failure to perform such actions reduces the likelihood that optimal patient outcomes will occur. Quantifying adherence to such aspects of care can therefore serve as a direct measure of the quality of care provided (or at least some important components of that quality) and as a foundation for quality improvement. In addition, certain outcomes may be so closely associated with the quality of care provided that they also can be used to measure healthcare quality. Creating mechanisms for measuring these opportunities to quantify healthcare quality in the course of routine practice is an important and pressing challenge.

This statement describes the methodology by which the ACC/AHA Task Force on Performance Measures develops performance measures. By clearly articulating the process by which performance measures are created, it is hoped that an understanding of the logic of these measures may be better appreciated by their users.

The applications of performance measures are designed to allow a transparent discussion of the quality of health care. Performance measures are not intended to be an end but rather a means for measuring and improving care. Specifically, punitive consequences such as restricting privileges, contracting selectively, or instituting penalties based on the performance of health systems or individual caregivers would undermine efforts to improve quality, particularly because a natural consequence of such efforts would be for clinicians and healthcare systems to manipulate the assessment process so that their performance appears better than it actually is. The intent of measuring performance is instead to allow healthcare providers to learn from one another how systems may be redesigned so that needed processes of care are applied uniformly to patients who are the most likely to benefit.

The development and implementation of ACC/AHA performance measures is a multiphase process, consisting of 3 basic phases that are inherent in building a performance measurement system: construction of the measurement set, assessment of the feasibility and reliability of data collection, and measurement of clinicians' performance. To avoid pitfalls in application, measurement, and interpretation, the Task Force identified key methodological areas associated with each phase that should be considered in developing and implementing ACC/AHA performance measurement sets.

Analytic issues associated with evaluating and/or monitoring providers via performance data are not discussed in this statement, although the Task Force recognizes these issues to be critical to any measurement system. The latter portion of this statement provides an overview of performance measure development and implementation (Table 1).

Phase I: Constructing Measurement Sets

Performance systems involve a set of measures that are targeted toward a particular patient population. From this high-priority population, a particular period of care can be identified that lends itself to measurement and improvement. Developing a set of performance measures entails 5 sequential tasks.

Task 1: Defining the Target Population and Observational Period

Quantifying the quality of care often is centered on a specific disease or its treatment. Thus, performance measures are designed to assess the care of a cohort of patients and, often, specific subsets of patients with a given disease. Accurately defining the target population for a performance measurement system is critical to ensuring the validity of these quality measures. By being concise in defining the target population, excessive inclusion and exclusion criteria can be avoided, and implementation can be more practical. Examples may include patients discharged from a hospital with heart failure, patients receiving procedures in specific clinical settings, or the treatment of acute or chronic aspects of a disease.

Two dimensions of time are relevant to performance measurement. One dimension is the "period of care" for an individual patient, during which certain care processes would be expected to occur. The second dimension is the "period of observation," during which a provider treats a number of individual patients. The period of care has implications for the specific aspects of care that are relevant and can be measured (see task 4). Under some circumstances, restrictions may be required to collect complete data. For example, a physician practice group may be interested in assessing the quality of ambulatory care for patients with heart failure 1 year after an initial diagnosis was made. In this instance, the target population consists of patients with heart failure, and the period of care is 1 year after diagnosis. An additional restriction that patients are continuously enrolled during the observational period may be required to obtain accurate information during the entire period of care.

The period of observation for the target population is the time frame during which sufficient cases accrue to provide reasonably accurate information about quality. The window of time selected has implications for both the number of cases that are available for measurement and the specific aspects of care that are relevant. For example, observational periods may be as short as 6 months or as long as 3 years, depending on the volume of cases within a practice. As longer periods of observation are considered, changes in technology and delays in providing analyses may limit the relevance of the data collected.

Clear, concise, and implementable definitions of the target population and the observational period that will become the foundation of the performance measurement set are needed. In addition, the ongoing efforts of the ACC/AHA Task Force on Clinical Data Standards (4,5) can provide data definitions for important clinical variables that are related to performance measurement. A sample framework for defining a target population is provided in Table 2.

TABLE 1. Summary of Performance Measure Development

Task	Description
Phase I: Constructing Measurement Sets:	
Task 1: Defining the target population and observational period	Develop a clear, concise, and implementable definition of the sample (eg, adults more than 29 years of age, discharged alive with a principal diagnosis of heart failure (ICD-9: 398.91, 402.01, 402.11, 402.91, 428.0, 428.1, 428.9), with a length of stay of at least 1 day, excluding patients with an AMI in the previous month continuously enrolled for 6 months after discharge.
Task 2: Identifying dimensions of care	Explicitly define each aspect of care that should be quantified to ensure a valid assessment of the most meaningful aspects of care. Potential dimensions include diagnosis, risk stratification and patient education, treatment, self-management, and reassessment of patient's health status.
Task 3: Synthesizing and reviewing the literature	Review published literature (including guidelines and other performance measurement systems) with a team of clinicians and researchers with expertise in meta-analysis.
Task 4: Defining and operationalizing potential measures	For each measure, determine which data sources are available and define the data elements needed to construct it (including period of care).
Task 5: Selecting measures for inclusion in the performance measures set	Present information based on tasks 1–3 to writing group and other relevant individuals, and put in place a formal mechanism to decide upon the measures that will be selected for inclusion.
Phase II: Determining Measure Feasibility:	
Definition of sample	Calculate sensitivity and specificity of selection criteria whenever possible. Document sources of case attrition (eg, medical record never sent, not continually enrolled, died during period of care). Develop an algorithm to assign patients to providers (eg, primary care provider, specialist) and validate the accuracy of the algorithm.
Feasibility of measures	Report validity, reliability, and completeness of collected data. If chart abstraction is used, then interabstractor reliability needs to be measured; if patient survey is used, then item and unit nonresponse must be measured. Data lags in identifying and surveying patients need to be assessed.
Phase III: Measuring Performance:	
Determining reporting unit	Determine at what level information will be reported (eg, physician-level data will typically require longer accrual period, even if only for internal monitoring).
Determining number and range of measures	Cost constraints may dictate how many measures can be measured. For quality improvement, how many measures will be evaluated and/or whether a combined measure is necessary will need to be determined.
Evaluating Performance	*Caution*: To determine whether a provider has "improved" care over time or whether a provider is sufficiently different from others, a sample size calculation that incorporates the relevant statistical features of the "test" (within- and between-provider variability, size of test, significance of test) should be undertaken.

ICD indicates International Classification of Diseases; AMI, acute myocardial infarction.

Task 2: Identifying Dimensions of Care

Given the multiple domains of providing care that can be measured, explicit articulation of which domains are being quantified by a given performance measure set is needed. All aspects of the care process, including diagnosis, risk stratification and prognosis, treatment, compliance, and patient reassessment should be considered. As the writing group plans to develop a measurement set, the group may find it useful to consider the range of steps needed to deliver optimal care. Figure 1 illustrates an example of care dimensions for the ambulatory care of patients with heart failure.

The initial step in rendering care to a patient with heart failure is to make a proper diagnosis. The next step involves educating patients about the nature of heart failure and what to expect regarding treatment (including lifestyle interventions) and prognosis. The third phase of care is to recommend the initial treatment. It is the evidence for treatment that most often dominates the work of guidelines committees. Ensuring that treatment recommendations are followed is the next step along the path of ideal care and includes teaching patients techniques of self-management such as weight monitoring and medication compliance. Finally, serial assessments of patients' responses to treatment and a monitor of the status of their heart failure are needed to continuously optimize the other aspects of the care of patients with heart failure.

Optimization can be accomplished through the serial assessment of patients' symptoms, functioning, and quality of life. Suboptimal health status (eg, symptoms, function, and quality of life) should trigger a repeated pursuit, following the same steps outlined above, of opportunities to improve a patient's condition. By creating a conceptual model of the dimensions of care, writing groups can be certain that evidence-based measures for quantifying each important aspect of care are developed. The measurement of all phases of ideal care can readily illuminate the sources of clinical inertia within current practice. Importantly, guidelines writing groups should review such models to ensure the content validity of their work (ie, that all important domains are being meaningfully quantified).

Task 3: Synthesizing and Reviewing the Literature

The goal of task 3 is to identify a set of indicators that are likely to improve quality. This is accomplished by reviewing summaries of the evidence-based literature (ie, guidelines) and existing performance measures from other organizations. The scientific foundation of clinical medicine is expanding rapidly. Because performance measures imply that adherence to these measures is a direct reflection of the quality of care provided, it is essential that a thorough review and synthesis of the medical literature be conducted. The critical issues that

TABLE 2.   Sample Framework for Defining the Target Population

Patient Characteristic Category	Definition	Acceptable Responses	Potential Data Sources	Comments
Age	Years alive	18–120, unless otherwise noted	Patient records	Because the ACC/AHA guidelines are typical for adult patients, these performance measures are, in general, meant for patients 18 years of age or older, unless otherwise noted to include pediatric patients. This field is best calculated from birth date and date care was provided if HIPAA regulations permit collection of the birth date. If not, age should be calculated and entered.
Gender	Sex	Female, male	Patient records	Because the ACC/AHA guidelines define practices that meet the needs of most patients in most circumstances, the performance measures will be for both men and women.
Principal diagnosis	Diagnosis most responsible for visit/admission	XXX.XX	Patient records, billing records	Consider including relevant ICD-9 codes that define the group of patients that are the focus of the study. For prospective implementations, clinical criteria (eg, positive troponin with clinical features consistent with MI) could be used.
Principal procedure	Primary procedure performed	XXX.X	Patient records, billing records	Consider including relevant CPT-5/ICD-9 procedure codes that define the group of patients that are the focus of the study.
Period of care	Duration of care being studied	X months, Y years, etc	Patient records, billing records	Define the time period over which quality will be assessed for the cohort of patients.
Period of observation	Duration of time during which care is measured	MM/YY–MM/YY	N/A	This is the time during which cases may accrue for the observed provider.
Other restrictions	Continuous enrollment, discharged alive, etc	XXXXX	Patient records, billing records	Define any other general restrictions that are needed (eg, if assessing outpatient care up to 6 months after a hospital discharge, then patients should be discharged alive and continuously enrolled for up to 6 months so that data are complete).

HIPAA indicates Health Insurance Portability and Accountability Act; CPT, Current Procedure Technology; other abbreviations as in Table 1.

should be considered when reviewing the literature include the following:

1. The strength of evidence (ie, multiple efficacy and effectiveness studies consistently demonstrate meaningful benefit on patient outcomes, potentially including Bayesian analyses (6), that give a strong post-test probability of benefit) that supports measure inclusion.

2. The clinical relevance of the outcome associated with adherence to the performance measures (ie, that the outcomes are meaningful to patients and society and are not surrogate markers of outcome).
3. The magnitude of the relationship between performance and outcome (ie, that "significant" improvements in patients' health will be realized with greater adherence to performance).
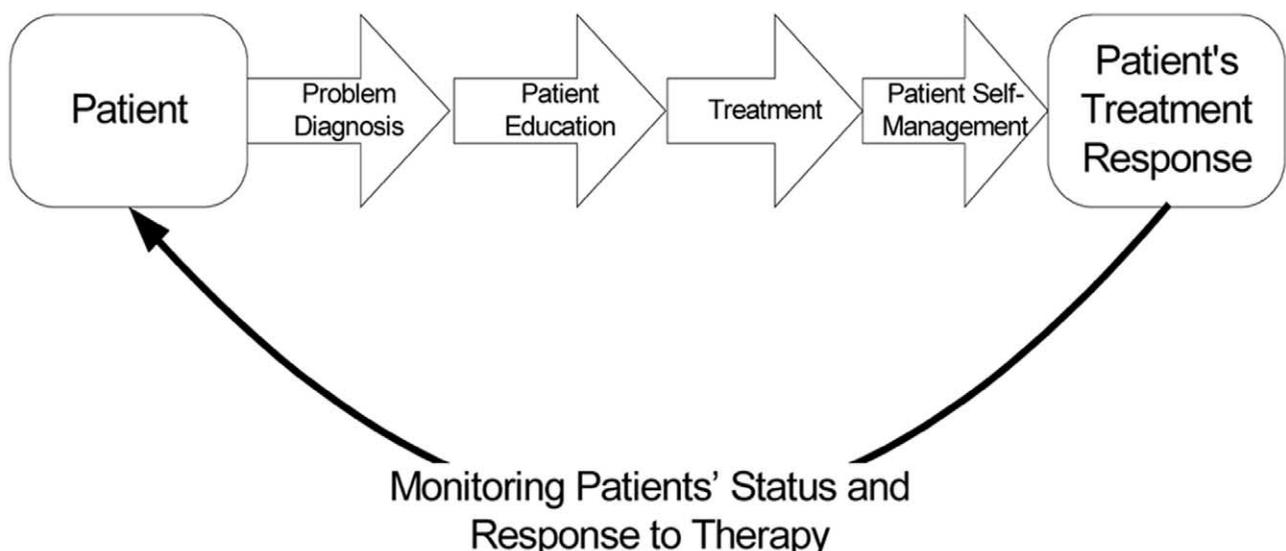

Figure 1. Example of care dimensions for ambulatory care for patients with heart failure.

A review of the medical literature also should acknowledge the expense of implementing performance measurement. As such, it is recommended that writing groups consider pursuing the creation of performance measures for only those aspects of care with the greatest likelihood of providing meaningful benefit.

As the foundation for the scientific evidence that underpins the performance measurement set, the appropriate ACC/AHA Practice Guidelines that are relevant to the topic with which the writing group has been charged should be reviewed. In addition, the writing group should perform or have access to an environmental scan of additional national or international performance measures for the condition of interest. Finally, the writing group will benefit by familiarizing itself with any pending revisions of relevant guidelines. Ideally, performance measurement sets would be released at the same time that guidelines revisions are published.

Clinical Practice Guidelines

Clinical practice guidelines are particularly rich sources of potential performance measures. The writing group should have copies of the relevant ACC/AHA Clinical Practice Guidelines at its disposal. In the event that other guidelines have been written on the same topic, the writing group should be informed and should have an opportunity to review these as well. The following steps are recommended:

1. Identify relevant ACC/AHA and non-ACC/AHA Practice Guidelines.
2. Review recommendations for each guideline.
3. Determine relevant areas of quality to consider when developing the performance measurement set.

In general, ACC/AHA Class I and III indications for therapy identify potential dimensions of care and processes for performance measurement; however, not all Class I and III guidelines recommendations should be selected for performance measurement. Specific considerations may include the following:

● *The magnitude of evidence supporting the process of care*: In addition to randomized clinical trials in ideal patients, effectiveness data demonstrating that the process is related to improved outcomes in more diverse clinical settings are of critical importance. Furthermore, in situations in which the relevance of clinical trial data to the patients that are being considered for performance measurement is debated, a mechanism for requesting and reviewing clinical trial data relevant to the issue at hand ideally should be acquired and reviewed.

● *The relationship of adherence to the performance measure with clinically meaningful outcomes*: The scientific method is predicated on discovering the pathways by which diseases develop and progress. This process often requires a focus on surrogate markers of disease progression. For example, left ventricular ejection fraction or coronary occlusion may be used to document the progression of heart failure or coronary disease, yet these characteristics are less relevant to patients than are survival and health status. When developing performance measures, writing groups should consider only those aspects of care associated with disease outcomes that are relevant to patients and society. For clinical trials with combined end points, careful attention to which outcomes were most influenced by a given process of care and the importance of those outcomes to patients and society are critical considerations in selecting potential areas for the development of performance measures.

● *Separating statistical and clinically significant differences in outcomes*: It is not uncommon for large clinical trials to identify treatments that have shown small but statistically significant improvements in outcome. Given the expense of ultimately collecting potential performance measures, it is the responsibility of the writing group to make judgments about the magnitude of the relationship between adherence to a performance measure and improvements in clinically meaningful outcomes. Those attributes of care that are associated with greater absolute (as opposed to relative) improvements in outcome should be made a priority.

In general, reviewing clinical guidelines annually or biannually is highly recommended. This recommendation reflects the rapid pace at which knowledge is being generated (3). The role of performance measures writing groups is not to perform a primary evaluation of the medical literature; this task should be undertaken by guidelines writing groups. It is appropriate—and recommended—that performance measures writing groups work collaboratively with guidelines writing groups so that the guidelines may be written with a degree of specificity that supports performance measurement and that new knowledge can be incorporated rapidly into performance measurement.

Existing Performance Measures

An additional important source for identifying potential performance measures is existing performance measures endorsed by other groups. Therefore, a review of existing performance measures being promulgated by other professional organizations should be conducted. Where possible, synergy with existing performance measures must be created so that the burden of data collection may be minimized when reporting to the different assessors of quality. Upon completing an environmental scan of performance measures on the specific clinical topic, the writing group should create a table describing the performance measurement sets reviewed and detailing the measure specifications for specific target populations. Table 3 provides an example of a systematic method to organize information collected from tasks 1 to 3.

Task 4: Defining and Operationalizing Potential Measures

Explicit criteria exist for the development of performance measures so that they can accurately reflect healthcare quality, including explicit quantification of the numerator and denominator of potential measures and explicit evaluation of the interpretability, actionability, and feasibility of the proposed measure. These are critical steps to take before the quality of care can be measured. Upon determining the target population and care period and reviewing pertinent scientific evidence on the topic, the writing group should operationalize

TABLE 3.   Sample Summary Information

Dimension of Care	Recommendations and Guidelines	Potential Quality Area	Site of Care	Existing Measures
Initial recognition and treatment	*ACC/AHA Class I:* A dose of 160 to 325 mg of aspirin should be given on day 1 of acute MI and continued indefinitely.	Aspirin on arrival	Inpatient	CMS, JCAHO, VA
	ACC/AHA Class I: An ECG should be obtained and interpreted within 10 min of arrival in the emergency department for all patients with suspected acute ischemic-type chest discomfort.	ECG on arrival	Inpatient	VA
Hospital treatment	*ACC/AHA Class I:* Primary PCI: As an alternative to fibrinolytic therapy in patients with AMI and ST-segment elevation or new or presumed new LBBB who can undergo angioplasty of the infarct-related artery within 12 hours of symptom onset or beyond 12 hours, if ischemic symptoms persist and if performed in a timely fashion* by those skilled in the procedure† and supported by experienced personnel in an appropriate laboratory environment‡	Appropriate primary PCI	Inpatient	JCAHO, CMS, VA
Pharmacotherapy	*ACC/AHA Class I:* Beta-blockers: Patients without a contraindication to beta-blocker therapy who can be treated within 12 hours of onset of infarction, irrespective of administration of concomitant fibrinolytic therapy	Beta-blocker on admission	Inpatient	CMS, JCAHO, VA
Risk stratification and treatment	*ACC/AHA UA/NSTEMI Guidelines:* Biomarkers of cardiac injury should be measured in all patients who present with chest discomfort consistent with ACS. A cardiac-specific troponin is the preferred marker, and if available, it should be measured in all patients. In patients with negative cardiac markers within 6 hours of the onset of pain, another sample should be drawn in the 6- to 12-hour time frame (eg, at 9 hours after onset of symptoms).	Early risk stratification	Inpatient	
Secondary prevention	*ACC/AHA Class I:* Patients with LDL-cholesterol levels greater than 100 mg/dL should be placed on drug therapy on hospital discharge, with preference given to statins (Level of Evidence: A). Patients with LDL less than 100 mg/dL or unknown LDL should be prescribed statins on hospital discharge (Level of Evidence: B).	Lipid-lowering therapy on discharge	Inpatient	VA
	AHA Secondary Prevention Guidelines		Outpatient	
Provider–patient interaction		Patient-presented alternative treatment options	Inpatient	
Long-term treatment	*Symptom & Activity Assessment:* Regular assessment of patients' angina symptoms and levels of activity is recommended.	Patient assessment with or without risk stratification	Outpatient	ACC/AHA/AMA Physician Consortium

CMS indicates Centers for Medicare and Medicaid Services; JCAHO, Joint Commission on Accreditation of Healthcare Organizations; VA, Department of Veterans Affairs; AMI, acute myocardial infarction; LBBB, left bundle-branch block; PCI, percutaneous coronary intervention; ACS, acute coronary syndromes.

*Performance standard: balloon inflation within 90 plus or minus 30 minutes of admission.

†Individuals who perform more than 75 PCI procedures per year.

‡Centers that perform more than 200 PCI procedures per year and have cardiac surgical capability.

the areas of quality identified in task 3. This is the most time-consuming and challenging task because it involves translating recommendations to specific measures. To accomplish this task, 3 key items for constructing each measure should be defined as follows.

Defining the Period of Care

The writing group should specify the time period during which each performance measure is to be evaluated. For example, some processes of care are required to be carried out within 24 hours of admission, others before discharge, and still others within 3 months of discharge. The writing group should give due consideration to the circumstances of routine clinical practice when specifying the period of care. For

example, although aspirin should be prescribed within 24 hours of a heart attack and beta-blocker use should be started during initial heart failure hospitalization, maximizing the beta-blocker dosage may be better completed in the outpatient setting rather than at the time of hospitalizing a patient for decompensated heart failure. In the case of a heart attack, it would be appropriate to assess performance upon hospital discharge. In the case of maximizing beta-blocker dosage in patients with heart failure, however, information may not be feasibly collected until 3 months after discharge.

Specifying the Denominator

The denominator of a performance measure refers to the target population that is eligible for the assessment of each

measure. In defining the denominator, the writing group provides direction for data to be collected and identifies consistent sources for information. Occasionally, the denominator will exclude subsets of patients within the target population and the dimension of care for the performance measure. This often arises when physicians provide a rationale for not applying the performance measure or when emerging evidence dictates that an alternative treatment strategy may be appropriate but evidence is insufficient to support that treatment as satisfying the performance measure. For example, in 2003, evidence was insufficient to recommend that angiotensin-receptor blockade be used for all patients with congestive heart failure, particularly if they could tolerate angiotensin-converting enzyme (ACE) inhibitor medications. If a physician recommends angiotensin-receptor blockade, however, then treatment with an ACE inhibitor may not be necessary. In this situation, when sufficient uncertainty exists in the medical literature to support the use of angiotensin-receptor blockers as an alternative to ACE inhibitors, then patients treated with angiotensin-receptor blockers may be excluded from the denominator so that they neither count as fulfilling nor as not fulfilling the performance measure.

Clarity of the denominator is needed so that the selected performance measures are clinically relevant. A tension exists between specificity and inclusivity of the denominator. When considering the most appropriate denominator, the writing group should entertain issues of the population's magnitude (ie, the larger the number of eligible patients, the more important the performance measurement set), variability in care, and the association with outcome of greater adherence to the potential performance measure.

Specifying the Numerator

The numerator of a performance measure indicates the subset of the denominator that has had the performance measure met. Patients from the denominator enter the numerator if documentation that the performance measure has been executed is available. Alternatively, if the quality measure is continuous (eg, blood pressure), then the performance measure can be either a mean (or median or other summary) across the patients who are eligible for the measure or dichotomized as meeting a prespecified desirable goal. Table 4 provides an example of the definition of a performance measure for a target population of adults discharged alive with a principal diagnosis of acute myocardial infarction.

Task 5: Selecting Measures for Inclusion in the Performance Measurement Set

On the basis of the information collected, the writing group will be able to choose from a range of measures. Selecting which potential measures to endorse should involve considering the interpretability, actionability, and feasibility of implementing each measure. Interpretability reflects the degree with which a practitioner is likely to understand what the results mean and can take action if necessary. Actionability represents an assessment of the degree to which a practitioner can influence the quality of the care being delivered by the health system. Because the

TABLE 4. Example of Measure Specifications for Daily Aspirin Use for Adults Discharged Alive With Acute Myocardial Infarction

Performance Measure	Daily Aspirin Use
Description of the measure	75 to 325 mg daily should be used routinely by all patients with acute and chronic ischemic heart disease with or without manifest symptoms in the absence of contraindications
Clinical recommendation	Aspirin is effective secondary prevention for survivors of myocardial infarction against reinfarction and death
Period of care	At the time of hospital discharge and continued throughout the period of care being assessed
Denominator	Patients 18 years of age or older discharged alive with a principal diagnosis of acute myocardial infarction AND who do not have a contraindication to aspirin use, including *x*, *y*, and *z*
Numerator*	Number of patients in the denominator whose medical record documents a prescription for daily aspirin at discharge
Data sources	Inpatient medical record; flow sheet
Methods of reporting	Proportion of a physician's eligible patients recorded as receiving aspirin (standard error)

* Mean or median if measure is continuous.

purpose of quality assessment is to improve care, it is important that the performance measure be under the locus of control of the entity being assessed. Finally, the feasibility of collecting the data required for the performance measure must be assessed. Feasibility addresses whether the required data can be typically abstracted from patient charts through easily implemented prospective or retrospective data collection systems or from national registries/databases that are readily available.

To assist in the selection process, it is recommended that the writing group pursue a formal strategy of evaluating potential measures. A systematic determination of the usefulness, specification, and likely feasibility of implementation will focus discussion on the advantages and disadvantages of each measure. Determining measure feasibility, a critical component of the ACC/AHA's multiphasic approach to building a performance-measurement system, is described in phase II. Such a determination can be assessed through a survey of writing group members and the parent committee of the ACC/AHA Task Force on Performance Measures. In addition, when resources permit, extension of the survey to practitioners and healthcare systems would be an excellent strategy for assessing the feasibility of a measurement set before its initial publication. A sample survey form (Figure 2) and a guide for its completion (Figure 3) are presented here. Rules for selecting performance measures should be decided upon before the survey is completed.

After digesting and integrating the feedback from these initial surveys, a final proposed measurement set is developed. At this point, a broader review of the performance measurement set occurs. This parallels the approach used in the review of proposed guidelines (2), whereby disease experts, other organizations, representatives from the ACC Board of Trustees and the AHA Scientific Advisory Committee, and the public are invited to review the proposed measurement set during an established

Name of Measure:					
Clinical Rationale:					
Numerator:					
Denominator:					
Measure:					
Rate this measure on the following criteria	**Disagree**		**Moderate Agreement**		**Agree**
	1	2	3	4	5
Useful in Improving Patient Outcomes					
1. **Evidence-based:** The scientific basis of the measure is well established.	1	2	3	4	5
2. **Interpretable:** The results of the measure are interpretable by practitioners.	1	2	3	4	5
3. **Actionable:** The measure addresses an area that is under the practitioner's control.	1	2	3	4	5
Measure Design					
1. **Denominator:** The patient group to whom this measure applies (denominator) is clinically meaningful.	1	2	3	4	5
2. **Numerator:** The definition of conformance for this measure is clinically meaningful.	1	2	3	4	5
3. **Validity:** The measure appears to measure what it is intended to (face validity).	1	2	3	4	5
a. The measure captures most meaningful aspects of care (content validity).	1	2	3	4	5
b. The measure correlates well with other measures of the same aspect of care (construct validity).	1	2	3	4	5
4. **Reliability:** The measure is likely to be reproducible across organizations and delivery settings.	1	2	3	4	5
Measure Implementation					
1. **Feasibility:** The data required for the measure is likely to be obtained with reasonable effort.	1	2	3	4	5
a. The data required for the measure is likely to be obtained at reasonable cost.	1	2	3	4	5
b. The data required for the measure is likely to be obtained within the period allowed for data collection.	1	2	3	4	5
Overall Assessment					
Considering your assessment of this measure on all dimensions above, rate this measure overall for inclusion into the ACC/AHA Heart Failure Performance Measurement Set.	**Do Not Include**		**Could Include**		**Must Include**
	1	2	3	4	5

Figure 2. Sample rating form. Adapted from the Q-SPAN-CD collaboration.[7]

comment period. The writing group then responds to all comments and completes the performance measurement set (ie, phase I).

At the completion of phase I, the initial work of the performance measures writing group draws to a close; however, additional work, as described in phases II and III, is necessary. At the conclusion of phases II and III, it is expected that the writing group will reconvene to review initial results, troubleshoot observed difficulties, and, if necessary, refine the measures. The writing group also will convene to update those measures when ACC/AHA guidelines are updated. These steps necessitate that the writing groups for performance measures serve as "living committees" so that reviews and revisions of both the specifications of existing measures and the introduction of new measures can occur in a timely manner, again mirroring the evolution of current guidelines committees (3).

Phase II: Determining Measure Feasibility

After potential measures are selected, formal evaluations of the feasibility of assessing performance with each measure should be pursued. Within the target population, the writing group must consider 2 levels of assessment: (1) how well they can identify their sample and (2) how well they can measure the data items for each member of the sample. Identifying a test population depends on the design (eg, prospective data collection or retrospective data collection, inpatient or ambulatory-based cohorts) and the intended implementation of the performance measures in clinical practice. During this evaluation process, explicit efforts to define the sensitivity and specificity of the sample identification procedure should be determined. For example, if administrative data will be used to initially identify the sample, then medical records data or direct patient assessments may be used to validate the diagnosis in patients identified as having the disease and its absence in a population of patients not identified as having the target condition.

Once the sample is identified and a provider or providers associated with each patient determined, the Task Force recommends that the validity, reliability, and completeness of each data item be assessed. The methodology for assessing feasibility depends on the available data sources. For example, if medical records data are used, then the frequency of missing patient records should be recorded. In addition, the reliability of medical record chart abstraction should be studied and assessed. If items cannot be abstracted with sufficient reliability, then dropping the measure from the measurement set should be considered.

Alternatively, if patient survey data are used (eg, to quantify patients' health status and compliance with recommendations such as exercise or smoking cessation), then the frequency and distribution of patient nonresponse should be assessed. The time

Attribute of Performance	Considerations
Useful in Improving Patient Outcomes	
1. **Evidence-based:** The scientific basis of the measure is well established.	This can be confirmed by explicit reference to a published clinical practice guideline.
2. **Interpretable:** The results of the measure are interpretable by practitioners.	This is your assessment of the degree with which a provider can clearly understand what the results mean and can take action if necessary.
3. **Actionable:** The measure addresses an area that is under the practitioner's control.	This is your assessment of the degree with which a provider is empowered and can influence the activities of the health care system toward improvement.
Measure Design	
1. **Denominator:** The patient group to whom this measure applies (denominator) is clinically meaningful.	Depending upon intended use of the measure, the data source, any inclusion or exclusion criteria, and sampling frames are explicit. These criteria used must be clinically meaningful. An algorithm for determining the denominator may be present.
2. **Numerator:** The definition of conformance for this measure is clinically meaningful.	The numerator may be specified using either explicit or implicit criteria. These criteria used must be clinically meaningful. An algorithm for determining the numerator may be present.
3. **Validity:** a. The measure appears to measure what it is intended to (face validity). b. The measure captures most meaningful aspects of care (content validity). c. The measure correlates well with other measures of the same aspect of care (construct validity).	This can be confirmed by your judgment of the clarity and comprehensiveness of the measure. For those measures that have been actually tested for validity, you may see indications of specific testing such as comparisons with the results of other methods, criterion or gold standard validity testing, and criterion validity testing. There may also be documentation that the health care construct underlying the measure is associated with important health care processes/outcomes.
4. **Reliability:** The measure is likely to be reproducible across organizations and delivery settings.	This can be confirmed by specific tests undertaken by the measure developers. For those measures that have been actually tested for reliability, you may see indications of types of reliability testing such as test-retest reliability, inter-rater reliability, data accuracy checks, and internal consistency analyses. If the measure has not been used in practice, indicate the degree of likelihood that it is reproducible.
Measure Implementation	
1. **Feasibility:** a. The data required for the measure is likely to be obtained with reasonable effort. b. The data required for the measure is likely to be obtained at reasonable cost. c. The data required for the measure is likely to be obtained within the period allowed for data collection.	From your perspective, the required data can be typically abstracted from patient charts or there are national registries or other databases readily available. For those measures actually being used, there is information on the data collection approach and the system required to support the measure.
Overall Assessment	
Considering your assessment of this measure on all dimensions above, rate this measure inclusion in the ACC/AHA AMI Performance Measurement Set.	Consider a balance in the continuum of care. Consider overall purpose of the measurement set and the intended user.

Figure 3. Sample rating form guide.

between an index-defining event and surveying a patient should be assessed to determine whether it is feasible for patients to recall needed information. Individual items within the survey should be examined in terms of completeness and clinical logic, reliability, and responsiveness so that the results are a valid reflection of patient outcomes.

If administrative data are used, then the lag time between patient events and recording the events in the files should be assessed. When data are missing, especially those based on diagnostic tests, analytic methods based on realistic scientific assumptions should be used to make inferences.

Phase III: Measuring Performance

Because the choice of a performance measurement system ultimately depends on its intended use, the Task Force recommends that researchers decide a priori both the reporting unit and the number and range of measures (many measures, a composite measure, or both) to be reported. Although all measurements will be made at the patient level, it is important to determine whether the reporting unit will be at the individual physician level, the group level, the health plan level, and so forth. The ACC/AHA Task Force on Performance Measures intends for its

measurement sets to be used by physicians to improve performance at the physician level. It is recognized, however, that for accurate estimates of performance to be obtained, a sufficient number and broad array of cases will be required to prove that providers have indeed "improved." Furthermore, many interventions needed to improve performance will be system-level interventions, and aggregating individual provider data will be needed both to assess the performance of systems of care and to monitor changes in performance over time. Assistance from individuals trained in statistics is critical for the successful aggregation of such data.

Conclusion

Quantifying clinical performance is a necessary step for improving the quality of health care. Although many entities are involved in creating methods for quantifying healthcare quality, the ACC and the AHA have joined forces to advance the field of quality assessment through the creation of performance measurement sets. (Table 1 summarizes the steps described to achieve this.)

An important consideration, although not discussed in this statement, in implementing a performance measurement relates to the frequency of measurement. This is a particularly challenging

issue and reflects a tension between the desire to provide rapid feedback on the one hand and a need to have accurate data on the other. The "accuracy" of data is a design issue that is affected by the volume of eligible cases, the variability in clinician performance, and the anticipated changes over time. Consequently, the timing of data reporting is likely to be greatly influenced by the intended purpose of such reporting. If reporting is for the sole use of the practitioner, then more frequent reporting intervals are appropriate, under the presumption that every case is an opportunity to improve the quality of care; however, if other credentialing, purchasing, or regulating entities are to review such reports, then greater statistical accuracy is needed and longer intervals between reporting periods are indicated.

It is hoped and anticipated that through the implementation of this methodological framework, new sets of performance measures will be created for cardiovascular care, and that through their use, the quality of cardiovascular care will improve (7).

References

1. Institute of Medicine Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press, 2001.
2. Gibbons RJ, Smith S, Antman E. American College of Cardiology/ American Heart Association clinical practice guidelines, part I: where do they come from? Circulation 2003;107:2979–86.
3. Gibbons RJ, Smith SC Jr., Antman E. American College of Cardiology/ American Heart Association clinical practice guidelines, part II: evolutionary changes in a continuous quality improvement project. Circulation 2003;107:3101–7.
4. Cannon CP, Battler A, Brindis RG, et al. American College of Cardiology key data elements and definitions for measuring the clinical management and outcomes of patients with acute coronary syndromes. A report of the American College of Cardiology Task Force on Clinical Data Standards (Acute Coronary Syndromes Writing Committee). J Am Coll Cardiol 2001;38:2114–30.
5. McNamara RL, Brass LM, Drozda JP Jr., et al. ACC/AHA key data elements and definitions for measuring the clinical management and outcomes of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Data Standards on Atrial Fibrillation). J Am Coll Cardiol 2004;44:475–95.
6. Diamond GA, Kaul S. Prior convictions: Bayesian approaches to the analysis and interpretation of clinical megatrials. J Am Coll Cardiol 2004;43:1929–39.
7. Normand SL, McNeil BJ, Peterson LE, Palmer RH. Eliciting expert opinion using the Delphi technique: identifying performance indicators for cardiovascular disease. Int J Qual Health Care 1998;10: 247–60.

Appendix: Author Disclosure Information

Author	Research Grant	Speakers Bureau/Honoraria	Stock Ownership	Consultant/Advisory Board
Dr Kim A. Eagle	Aventis, Biosite, Blue Cross/Blue Shield, Cardiac Science, Mardigian Foundation, Pfizer, University of Michigan, Varbedian Fund for Aortic Research, Walter & Esther Hewlett Foundation	None	None	National Heart, Lung, & Blood Institute; Sanofi-Synthelabo
Dr Harlan M. Krumholz	None	None	None	CV Therapeutics
Kristi R. Mitchell	None	None	None	None
Dr Sharon-Lise T. Normand	Harvard Medical School interfaculty/Merck grant	None	None	Institute for Clinical Evaluative Sciences
Dr John A. Spertus	CV Therapeutics	None	None	Amgen, CV Therapeutics, Orqis Medical, Otsuka, World Heart Federation

This table represents the relationships of writing group members that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all members of the writing group were required to complete and submit.