

STATE-OF-THE-ART PAPER

Integrating Quality Into the Cycle of Therapeutic Development

Robert M. Califf, MD, FACC,* Eric D. Peterson, MD, MPH, FACC,*
Raymond J. Gibbons, MD, FACC,† Arthur Garson, Jr, MD, MPH, FACC,‡
Ralph G. Brindis, MD, MPH, FACC,§ George A. Beller, MD, FACC,||
Sidney C. Smith, Jr, MD, FACC¶

Durham, North Carolina; Rochester, Minnesota; Houston, Texas; San Francisco, California; Charlottesville, Virginia; and Chapel Hill, North Carolina

The quality of healthcare, particularly as reflected in current practice versus the available evidence, has become a major focus of national health policy discussions. Key components needed to provide quality care include: 1) development of quality indicators and performance measures from specific practice guidelines, 2) better ways to disseminate such guidelines and measures, and 3) development of support tools to promote standardized practice. Although rational decision-making and development of practice guidelines have relied upon results of randomized trials and outcomes studies, not all questions can be answered by randomized trials, and many treatment decisions necessarily reflect physiology, intuition, and experience when treating individuals. Debate about the role of "evidence-based medicine" also has raised questions about the value of applying trial results in practice, and some skepticism has arisen about whether advocated measures of clinical effectiveness, the basic definition of quality, truly reflect a worthwhile approach to improving medical practice. We provide a perspective on this issue by describing a model that integrates quantitative measurements of quality and performance into the development cycle of existing and future therapeutics. Such a model would serve as a basic approach to cardiovascular medicine that is necessary, but not sufficient, to those wishing to provide the best care for their patients. (J Am Coll Cardiol 2002;40: 1895-901) © 2002 by the American College of Cardiology Foundation

The quality of medical care has become a major focus of health policy discussions, especially since publication of the Institute of Medicine reports on medical errors (1) and quality gaps (2). These reports describe the wide, deep gulf separating current evidence from current practice. They also outline key components needed to provide quality care: development of quality standards based on specific practice guidelines, better ways to disseminate such guidelines and standards to the public and providers, and development of information technology and other support tools to promote standard practice. To enhance progress in this effort, practitioners must understand the quantitative and qualitative concepts involved in quality improvement.

DEFINITION OF QUALITY AND THE ROLE OF CLINICAL TRIALS

Quality in medicine can be defined as how much health services increase the likelihood of desired health outcomes

and how closely they adhere to professional knowledge (3). The Rand Institute defines quality care as "providing patients with appropriate services in a technically competent manner, with good communications, shared decision-making, and cultural sensitivity" (4). Along with safety and efficacy, the Institute of Medicine considers patient-centered activities (responsiveness to individual preferences and needs), timeliness, efficiency, and equitable resource use to be critical to quality healthcare. However, although these elements are important, medicine must above all provide safe and effective use of diagnostic and therapeutic technologies, or *clinical effectiveness*, for the rest to have meaning.

Randomized trial and outcome studies have provided a basis for informed decisions about the use of medical technologies. Randomized trials cannot answer all questions, however, and many decisions in practice must be made based on an understanding of physiology, intuition, and experience when treating individuals. Questions have also arisen about the value of applying randomized trials in clinical practice, with skepticism about whether advocated measures of clinical effectiveness truly reflect a worthwhile approach to improving medical practice.

We provide a perspective on this issue by presenting a model that integrates quantitative measurements of quality and performance into the development cycle for therapeutics. Such a model could serve as a basic approach to cardiovascular medicine that is necessary, but not sufficient,

From the *Duke Clinical Research Institute, Durham, North Carolina; †Mayo Clinic Foundation, Rochester, Minnesota; ‡Baylor College of Medicine, Houston, Texas; §Kaiser-Permanente San Francisco Medical Center, San Francisco, California; ||University of Virginia Health System, Charlottesville, Virginia; and the ¶University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. Supported by the Agency for Healthcare Research and Quality (AHRQ) Centers for Education and Research on Therapeutics (CERTs) cooperative agreement grant #U18HS10548.

Manuscript received June 6, 2002; revised manuscript received June 27, 2002, accepted August 9, 2002.

Abbreviations and Acronyms

- ACC = American College of Cardiology
- AHA = American Heart Association
- CPG = clinical practice guideline
- ESC = European Society of Cardiology
- MI = myocardial infarction

to those wishing to provide the best care for their patients. These concepts have evolved largely through ongoing efforts of the American College of Cardiology (ACC) and American Heart Association (AHA) to develop clinical practice guidelines (CPGs) and performance indicators for cardiology, but the issues raised in this review pertain to all areas of medicine.

INTEGRATION OF QUALITY IN THE DEVELOPMENT CYCLE

We present a model to integrate quality measures into the development cycle for therapeutics, adapted from a previous concept, the “great circle” (Fig. 1) (5). There are six main “stops” along the circle, each representing a chance to apply quantitative strategies to integrate quality.

To summarize the circle, first, hypotheses (*Concepts*) evolve from biological discoveries or clinical observations. These include proposed mechanisms of disease, diagnostic technologies, or therapeutics that emerge from basic and animal research. After these concepts are refined, they are tested in various phases of *Clinical Research* (randomized trials and outcomes studies). Initial clinical research produces “proof of principle” and preliminary safety data, whereas large, representative trials measure the clinical benefits and risks of technologies. The latter form of clinical

research provides the highest-level evidence for the creation of practice *Guidelines*.

Recommendations about diagnosis and treatment contained in a given guideline can be synthesized into algorithms, which then can be used as *Quality Indicators*, specifying the clinical circumstances under which to use a technology. By determining how well a provider or institution meets these quality indicators, actual *Performance Measures* can be assessed. The final stage in the cycle then links measured performance with the ultimate goal of healthcare, better *Outcomes*. For all of these elements to contribute optimally to the overall system, continuous education and feedback about findings and concepts are needed; thus, these aspects are in the center of the cycle.

As an example, suppose the preponderance of basic and clinical evidence leads a CPG to recommend that all eligible patients receive a beta-adrenergic blocking agent after acute myocardial infarction (MI). This recommendation could translate into the quality indicator “prescription for beta-blocker at discharge after MI.” The corresponding performance measure then would be “proportion of eligible patients prescribed a beta-blocker at discharge after MI.” Stated simply, the guideline generates a criterion (the quality indicator), and how well it is met by providers or institutions is the performance measure.

By studying the links between and among cycle elements, we might begin to develop ideas for building a quality system. Of note, this approach deals with only one component of quality, although this quantitative component relating to outcomes may distinguish the medical environment from other elements affecting healthcare quality. Without excellence in the subjective element, these quantitative elements are moot. With this idea of giving care to the individual patient in mind, several attributes that would

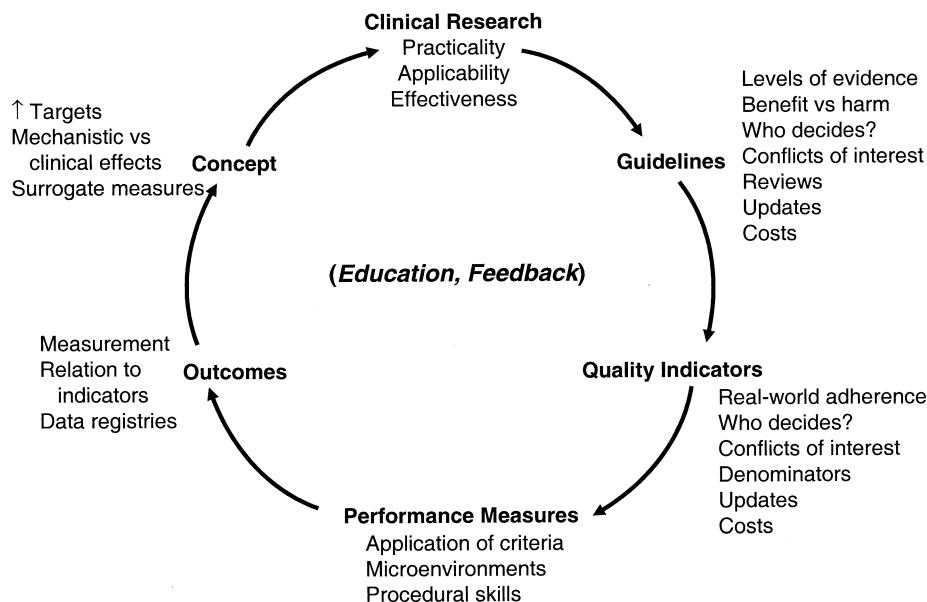


Figure 1. Model for the integration of quality into the therapeutic development cycle.

enhance each cycle element become evident, and many research questions can be posed.

Concepts: biological insights and the treachery of surrogates. Insight into disease mechanisms is essential to develop concepts for diagnostic and therapeutic products. Given the accelerating insights from genomics and proteomics, a wealth of biological targets seems probable. Of more immediate relevance, bioengineering progress is making devices and their combinations with drugs or biologics an increasingly routine part of medicine, as evidenced in cardiology by the advent of coated stents (6), wider indications for defibrillators (7,8), and mechanical assist devices for heart failure (9). Therapies based on genomic and proteomic technology will soon follow, and diagnostic tests will eventually use analysis of genetic variations to identify patients more or less likely to respond to given therapies.

Pressure continues to increase to develop ways to assess the efficacy and safety of new technologies. Many have advocated the use of biological “surrogates,” known as biomarkers, to substitute for clinical outcomes during development. Although biomarker results should be considered when deciding which theories to pursue in trials (10,11), they provide only an entry point for medical products. Even therapies that produce substantial benefits for a respected surrogate may fail because of other safety problems (12,13). Perhaps most important, although biomarkers may identify particular benefits of therapies, they cannot reliably reflect the balance between the risks and benefits of therapies, information critical to determining their value (14).

The treachery of surrogates has caused cardiovascular specialists to require large outcomes trials as a basis for the highest level recommendations in CPGs. Arrays of biomarkers are urgently needed, however, to determine when it is reasonable to invest in such trials. Of note, single biomarkers are insufficient for this purpose. Antithrombotic drugs, for example, can affect markers of thrombin, platelet activation, and inflammation differently, precluding translation into a cohesive, quantitative estimate. Imaging methods provide a promising approach to biomarker evaluation, by integrating structure and function into a common measurement.

Clinical research: the standard of evidence. Clinical trials are preferable as the source of evidence whenever possible. From a regulatory standpoint, a definitive clinical trial must be “adequate and well controlled” and must assess the safety and efficacy of a product when used in the intended population. To be helpful in the qualitative component of quality, however, a trial also must address the issues of practicality, applicability, and effectiveness. This means that a trial should measure clinically relevant outcomes in a representative population given the treatment in practice for a clinically relevant duration. “Large, simple trials”—with minimal data collection and “harnessing” data from electronic medical records—often can accomplish this goal (15).

Table 1. General Principles to Consider in the Design of Clinical Trials (16)

| |
|--|
| Treatment effects usually modest |
| Qualitative interactions rare |
| Quantitative interactions common |
| Unanticipated effects universal |
| Combinations are unpredictable |
| Short-term effects do not reliably predict long-term effects |
| Class effect almost impossible to define |
| Most therapies produce a combination of helpful and harmful effects |
| Effective therapies do not usually save money, but they are incrementally cost-effective |

Empirical experience has now provided cardiovascular practitioners with general principles to consider when designing trials designed to inform the cycle of quantitative evidence (Table 1) (16). For example, the modest nature of most treatment effects mandates large trials, so that effects can be detected or excluded with certainty. Large trials also can and should enroll a wide variety of patients, so that quantitative and qualitative interactions can be estimated for policy reasons. Large trials also maximize the possibility of detecting unanticipated effects of therapies, alone and combined with other technologies.

Being large is not enough, however. Accurate assessments of effectiveness also require clinically relevant follow-up periods, given that short-term outcomes may not reliably predict long-term events (17). Furthermore, most major cardiovascular diseases already have at least one effective treatment. The finding that not all therapies in a general class have the same balance of risk and benefit (18,19), with inherent uncertainties in combining therapies, mandates comparative trials. Trials also must consider current CPG recommendations. For example, a trial of a new secondary prevention approach that prohibits statin use might be more likely to show an effect of the experimental agent, but it would provide no information about whether the new agent should be preferred over standard treatment (which includes statins), combined with standard treatment, or substituted for standard treatment.

Clinicians feel increasingly pressed for time, and financial constraints leave little room for altruistic efforts to engage patients in discussions about trials. Questions about the professional responsibilities of physicians (duty to individual patients) versus trialists (duty to answer questions), and the constraints of institutional review boards and consent mechanisms, also make the conduct of research in clinical practice more daunting. The new Health Information Privacy, Portability, and Accountability Act, which places criminal penalties on the misuse of medical information, has exacerbated these concerns. Nevertheless, cardiovascular medicine has led the way by engaging many practices in addressing important questions that cannot be answered by a separate trials infrastructure (15). Highly organized cardiology practices have been the cornerstone of successful trials (20,21), and the federal government now reimburses phy-

Table 2. Judging the Evidence for Clinical Practice Recommendations (24)

| Measure | Description |
|-------------------------|--|
| Class of recommendation | |
| I | Intervention is useful and effective |
| IIa | Evidence conflicts/opinions differ but lean towards efficacy |
| IIb | Evidence conflicts/opinions differ but lean against efficacy |
| III | Intervention is not useful/effective and may be harmful |
| Level of evidence | |
| A | Data from many randomized clinical trials |
| B | Data from single randomized trial or nonrandomized studies |
| C | Expert consensus |

sicians for the routine costs of trials in patients covered by Medicare (22).

Finally, in many cases randomized trials cannot be performed because: 1) they would be impractical, 2) they would be unethical, or 3) the follow-up needed would exceed society's "willingness to wait." However, standards for observational comparisons have not evolved to the same level as for clinical trials (23). Although newer statistical techniques allow better control of treatment selection in observational research, unmeasured sources of bias will always be a concern. Thus, observational studies will continue to support randomized studies: hypothesis generation, testing (if randomized trials are impossible), and confirming that the results of randomized trials can be generalized to other providers and patient populations.

Clinical practice guidelines: synthesis of evidence. A hierarchy of evidence forms the basis for formulation of CPGs. The highest level of evidence reflects large trials addressing the specific question of interest or several smaller trials with consistent results. The ACC/AHA guideline for unstable angina and non-ST-segment elevation MI provides a framework for considering evidence (24). As shown in Table 2, evidence can be generated in two vectors, roughly representing the quantitative and qualitative perspectives. A level of evidence of "A" for a recommendation is derived from multiple, consistent randomized trials or a single large, definitive trial. A "C" level represents expert opinion and no definitive data, and a "B" level encompasses various intermediate-quality data. In the other vector, a class I recommendation reflects consensus that the practice should be done, whereas a class III recommendation reflects consensus that the practice should be avoided. A class IIa recommendation denotes a situation in which consensus does not exist but the practice is generally reasonable, and class IIb recommendation does not endorse the practice but does not definitively recommend against it.

A recent document from the ACC and AHA provides insight into the CPG process (24). An obvious first question is the composition of the guidelines committee. The ACC/AHA strategy is to select committee members through

consensus of the parent guideline task force and the presidents of the two organizations. Efforts are made to include representatives from the ACC, AHA, and other professional organizations and views from different regions, practices, and organizations within cardiovascular medicine. Real and potential conflicts of interest (25) are identified at the outset and reviewed periodically.

In addition to committee composition, review of the guidelines before finalization is a major issue. Ideally, all of the affected constituencies should agree with the draft CPGs. When appropriate, joint ACC/AHA guidelines typically also are reviewed by the American Academy of Family Practitioners, the American College of Physicians-American Society of Internal Medicine, and other major professional organizations caring for cardiovascular patients. A recent experimental guideline for atrial fibrillation also has attempted to reach consensus between the American organizations and the European Society of Cardiology (ESC) (26).

Updating the guidelines is another significant component of the process. In many areas of medicine, trials are being performed at such a pace that major findings relevant to clinical practice are common. For example, the Global Use of Strategies To Open occluded arteries (GUSTO)-IV trial of acute coronary syndromes (27) was reported within days of publication of both the ACC/AHA guidelines (24) and ESC guidelines on unstable angina and non-ST-segment MI (28). Within months of the publication of the ACC/AHA guidelines on heart failure (29), a randomized trial (9) showed a survival advantage for left ventricular assist devices, and initial data from another randomized study (7) appeared to expand the indications for implantable defibrillators. As other areas of clinical practice change, existing therapies must be reevaluated.

Most current CPGs do not emphasize costs, partly because information about the cost-effectiveness of therapies is often scant and based on models rather than empirical data. In theory, best practice is defined by the quality of evidence for the intervention rather than its price. The fact that cost does affect therapeutic choice has produced a crisis, however, as illustrated by the escalation of healthcare costs versus the increasing imperatives to implant defibrillators (7,8), coated stents (6), or left ventricular assist devices (9). These examples suggest that CPGs might need to be country-specific, depending on national resources. In the U.S., an incremental cost of \$50,000 to \$70,000 per year of life saved has become a de facto standard based on the national right to dialysis (30), but in many countries, even when a therapy clearly is beneficial, it simply is unaffordable given competing demands for financial resources.

Quality indicators. Once a CPG has been developed, its recommendations must be translated into a series of variables, or indicators, that reflect the quality of care (or lack thereof). In an ideal clinical world, for every clinical decision there would be an indicator based on a guideline based on evidence from randomized trials, such that a standard of

Table 3. Class I, Level of Evidence A, Recommendations from Selected ACC/AHA Guidelines

| Guidelines | Recommendation |
|-------------------------------------|--|
| Acute coronary syndromes (ACS) (24) | <ul style="list-style-type: none"> ● Assessment of patients with definite ACS and ST-segment elevation for immediate reperfusion therapy ● Immediate, continued antiplatelet therapy with aspirin; add clopidogrel as soon as possible, continue for ≥1 month (indefinitely if aspirin not tolerated) ● Platelet glycoprotein IIb/IIIa inhibition (plus aspirin and heparin) if continued ischemia or other high-risk features ● Subcutaneous LMWH or intravenous unfractionated heparin (plus aspirin and/or clopidogrel) ● Platelet glycoprotein IIb/IIIa antagonist (plus aspirin and heparin) if catheterization and PCI planned ● Early invasive strategy for: recurrent angina/ischemia: 1) at rest, 2) with low-level activities despite intensive anti-ischemic therapy, or 3) with HF symptoms, S3 gallop, pulmonary edema, worsening rales, or new/worse mitral regurgitation; elevated troponins; new ST-segment depression; high-risk findings on noninvasive stress testing; left ventricular systolic dysfunction (e.g., ejection fraction <40% on noninvasive test); hemodynamic instability; sustained ventricular tachycardia; PCI <6 months; prior bypass surgery ● PCI for multivessel disease with suitable anatomy, normal LV function, and no diabetes ● Bypass surgery for significant left main disease, three-vessel disease, or two-vessel disease with significant pLAD and either LV dysfunction or ischemia during noninvasive testing ● Lipid-lowering drugs (HMG-CoA reductase inhibitor) and diet if LDL level >130 mg/dl ● Hypertension control to <130/85 mm Hg ● ACEI for patients with HF, ejection fraction <40%, hypertension, or diabetes ● Diabetes is an independent risk factor |
| Heart failure (HF) (29) | <ul style="list-style-type: none"> ● Overall: revascularization if concurrent angina, anticoagulants if concurrent atrial fibrillation or thromboembolic event, control of ventricular response with beta-blocker (or amiodarone) if concurrent atrial fibrillation; beta-blocker in patients with HF to reduce risk of sudden death; implantable cardioverter-defibrillator (alone or with amiodarone) if prior cardiac arrest or severe ventricular arrhythmia ● Stage A patients: control of systolic and diastolic hypertension as per guidelines ● Stage B patients: all class I Stage A recommendations, ACEI if prior myocardial infarction, beta-blocker if recent infarction ● Stage C patients: all class I recommendations from prior stages; diuretics if fluid retention; ACEI, beta-blocker, digitalis in all patients unless contraindicated ● Stage D patients: all class I recommendations from prior stages, referral to experienced HF program |

ACEI = angiotensin-converting enzyme inhibitor; ACC = American College of Cardiology; ACS = acute coronary syndromes; AHA = American Heart Association; HF = heart failure; LDL = low-density lipoprotein; LMWH = low-molecular-weight heparin; LV = left ventricular; PCI = percutaneous coronary intervention; pLAD = proximal left anterior descending artery.

care could be defined for each situation. Such data exist for few clinical decisions, however. Table 3 lists the class I, level A recommendations (“almost always do it”) from the ACC/AHA guidelines for unstable angina/non-ST-segment elevation MI (24) and heart failure (29). Table 4 lists the class III, level A recommendations (“never do it”) from these same guidelines.

Given the small number of definitive recommendations,

it would seem simple to develop quality indicators. The issue is much more complex, however. For each recommendation, there are exceptions and areas of uncertainty. Given unlimited space to define exceptions and elaborate on nuances, CPGs could provide a panoramic translation of evidence into recommendations. By their nature, however, quality indicators translate recommendations into measures that can be defined and quantified in the context of

Table 4. Class III, Level of Evidence A, Recommendations from Selected ACC/AHA Guidelines

| Guideline | Recommendation |
|-------------------------------|--|
| Acute coronary syndromes (24) | <ul style="list-style-type: none"> ● Immediate-release dihydropyridine calcium antagonist without beta-blockade ● Thrombolytic therapy without ST-segment elevation, left bundle-branch block, or true posterior infarction ● Abciximab if PCI is not planned |
| Heart failure (29) | <ul style="list-style-type: none"> ● Overall: class I or III antiarrhythmic drugs (except amiodarone) to prevent or treat asymptomatic ventricular arrhythmia, ambulatory electrocardiography to detect asymptomatic ventricular arrhythmia ● Stage C patients: use of an angiotensin-receptor blocker before beta-blocker in patients taking ACEI |

ACC/AHA = American College of Cardiology/American Heart Association; ACEI = angiotensin-converting enzyme inhibitor; PCI = percutaneous coronary intervention.

delivering healthcare. This means that their measurement cannot impede healthcare delivery or require such complex data collection that it becomes too costly.

Several organizations are implementing quality indicators and performance measures for cardiovascular care. Many of the same issues that pertain to construction of CPGs also apply to development of the resulting quality indicators. Who should be on the committees and how are conflicts of interest managed? What level of evidence in a guideline should merit a quality indicator? How should those who devise guidelines react when a quality indicator is advocated that is inconsistent with existing guidelines? The ACC and AHA have a task force considering this very issue, and a report is expected very soon.

A particularly vexing problem for quality indicators emerges when attempting to define which patients qualify for a particular indicator. For example, when the Cooperative Cardiovascular Project investigators measured the use of fibrinolytic therapy in patients covered by Medicare, less than half of the patients with acute MI actually qualified for measurement after a long list of potential exclusions was applied (31). Care among such filtered subgroups may not reflect providers' general practices.

Performance measures. If the quality indicator is the variable, the performance measure is the threshold. The concept is straightforward: in its simplest form, a quality indicator reflects either a class I, level A recommendation (use beta-blockers after MI) or a class III, level A recommendation (do not perform routine angioplasty of the infarct-related artery immediately after fibrinolysis). Each encounter with a patient who meets the circumstances of such recommendations provides evidence to assess the performance of providers or systems. The proportion of eligible patients with MI who receive beta-blockers, say, would be compared against some threshold level, or performance measure, in this case, perhaps 95%.

This raises the logical question of how thresholds are set. One attractive approach has been to develop "achievable benchmarks of care" (32). Instead of attempting to define purified populations for whom process indicators should approach 100%, one simply compares a provider's performance to that seen among the top 10% of practitioners, hospitals, or practices. Thus, if leading centers can prescribe beta-blockers at discharge for 90% of their patients with MI, then this would be a reasonable and achievable performance goal for the rest of the nation. The Achievable Benchmark method provides one reasonable approach and avoids striving for unrealistic goals, which could provoke practitioners into inappropriate treatment for patients who do not meet criteria or lead them to become cynical about efforts to measure quality.

A related issue concerns the number of quality indicators. Specifically, as the number of quality indicators increases, the number of performance measures also will rise, creating a signal-to-noise issue when reviewing these results. Fortunately, process performance is likely to correlate, and centers

that adhere closely to national guidelines on one measure might tend to do similarly well in other care areas. One also could develop overall "composite quality indicators" for given conditions, which combine and average provider performance on individual measures. Taking into account multiple measures per patient, composite quality indicators also increase the power for a given sample size and, thus, provide more stable estimates of performance.

Given the large number of decisions that providers and systems make that are not subject to advanced levels of evidence, it is tempting to define performance in terms of higher level, complex decision-making. Indeed, most would agree that healthcare systems must be able to provide integrated care for patients with multiple comorbidities requiring complicated procedures and regimens to be considered "high-quality" systems (i.e., excellent performance). The argument is equally strong, however, that practitioners and systems who do not adhere to relatively simple guidelines based on clear evidence cannot claim to have even basal levels of quality (i.e., poor performance).

In developing performance measures, it can be useful to focus on the specific environment (microenvironment) for healthcare delivery (31,33). The inpatient cardiovascular arena, for example, contains the various microenvironments of the emergency department; cardiac care unit; inpatient service; cardiac catheterization, interventional, and electrophysiology laboratories; surgical suite and postoperative intensive care unit; and noninvasive testing and imaging systems. The cardiovascular specialist's and primary care practitioner's microenvironments include the practice structure, office staff, nonphysician practitioners, and systems. The outpatient arena is even more vast, encompassing the patient's home, workplace, various healthcare providers, social outlets, religious community, and insurance coverage. By measuring actual outcomes in populations, deviations from expected results should spur scientific insights, refined trial designs, and development of more appropriate quality indicators and performance measures for various settings.

Outcomes. From the perspective of the quality cycle, the ultimate goal is the best possible outcomes. Fortunately for cardiovascular practitioners, there is consensus about the outcome domains that are generally most important to the field, and these have been validated in many trials and observational studies. For most cardiovascular problems, survival, freedom from major cardiovascular events (stroke, MI, major arrhythmias, heart failure), and improved symptoms are the cornerstones of outcomes measurement. Much research also has focused on measurement of functional outcomes and quality of life in cardiovascular patients. Ideally, outcome measures would assess both the acute success of an episode of medical care and its long-term effects.

Outcomes measurement also faces particular challenges, however. These include how to adjust for patient risk (disease severity, comorbidity, educational, and financial status) when comparing outcomes among providers and the

instability of outcome measures at the provider level. Because of these limitations, the quality-assessment field has moved from direct measurement of outcomes to measurement of performance in most situations. As mentioned, because performance measures are essentially surrogates, the quality cycle calls for studies of the broader measurement of outcome as a function of performance in populations of patients or practitioners, to validate that the performance measures are important and that greater adherence to them improves outcome.

Recommendations. Clinical research networks and practice databases provide a convenient mechanism to tie together the quality cycle (Fig. 1). After a concept has been developed and undergone basic testing, a network could conduct clinical trials and measure incorporation of the findings (in the form of recommendations) into practice. Multiple practice registries can provide feedback about performance for individual practices while also validating the relation between greater adherence to guidelines (in the form of performance measures) and improved patient outcomes in the registry as a whole. In the process, new concepts would be developed that can then be tested, beginning the cycle anew. Finally, education and feedback, though inadequate to improve processes and outcomes, remain a necessary foundation for all elements of the cycle.

Reprint requests and correspondence: Dr. Robert M. Califf, Duke Clinical Research Institute, P.O. Box 17969, Durham, North Carolina 27710. E-mail: califf001@mc.duke.edu.

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS editors. *To Err Is Human: Building a Safer Health System*. Washington, DC: Institute of Medicine/National Academy Press, 2000.
2. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: Institute of Medicine/National Academy Press, 2001.
3. Lohr KN, Schroeder SA. A strategy for quality assurance in medicare. *N Engl J Med* 1990;322:707-12.
4. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? *Milbank Q* 1998;76:517-63.
5. Garson A. President's page: a target for better patient care. *J Am Coll Cardiol* 1999;34:294-5.
6. Fajadet J, Perin E, Hayashi B, et al. 210-Day Follow-up of the RAVEL Study: a randomized study with the sirolimus-eluting Bx velocity balloon-expandable stent in the treatment of patients with de novo native coronary artery lesions. *J Am Coll Cardiol* 2001;39 Suppl:20A.
7. Moss AJ, Zareba W, Hall J, et al. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *N Engl J Med* 2002;346:877-83.
8. Buxton AE, Lee KL, Fisher JD, et al. A randomized study of the prevention of sudden death in patients with coronary artery disease. *N Engl J Med* 1999;341:1882-90.
9. Rose EA, Gelijns AC, Moskowitz AJ, et al. Long-term use of a left ventricular assist device for end-stage heart failure. *N Engl J Med* 2001;345:1435-43.
10. NIH Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions in conceptual framework. *Clin Pharmacol Ther* 2001;69:89-95.
11. De Gruttola VG, Clax P, DeMets DL, et al. Considerations in the evaluation of surrogate endpoints in clinical trials: summary of a National Institutes of Health workshop. *Control Clin Trials* 2001;22:485-502.
12. U.S. Food and Drug Administration. Rezulin to be withdrawn from the market. March 22, 2000. Available at <http://www.fda.gov/bbs/topics/NEWS/NEW00721.html>.
13. Hoffman-LaRoche, Inc.. Roche announces voluntary withdrawal of Posicor. June 8, 1998. Available at <http://www.fda.gov/medwatch/safety/1998/mibefr.htm>.
14. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605-13.
15. Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? *Stat Med* 1984;3:409-20.
16. Califf RM, DeMets DL. Principles from clinical trials relevant to clinical practice. *Circulation* 2002;106:1015-21, 1172-5.
17. Lichtman JH, Roumanis SA, Radford MJ, et al. Can practice guidelines be transported effectively to different settings? Results from a multicenter interventional study. *Joint Comm J Qual Improv* 2001;27:42-53.
18. McAlister FA, Laupacis A, Wells GA, Sackett DL. Users' guides to the medical literature: XIX. Applying clinical trial results. B. Guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA* 1999;282:1371-7.
19. Furberg CD. Natural statins and stroke risk. *Circulation* 1999;99:185-8.
20. Califf RM, DeMets DL. Lessons learned from recent cardiovascular clinical trials: part 1. *Circulation* 2002;106:746-51.
21. Califf RM, DeMets DL. Lessons learned from recent cardiovascular clinical trials: part 2. *Circulation* 2002;106:880-6.
22. Health Care Financing Administration. Medicare coverage policy: clinical trials: final national coverage decision. September 19, 2000. Available at <http://www.hcfa.gov/coverage/8d2.htm>.
23. Hlatky MA, Califf RM, Harrell FE Jr, et al. Clinical judgment and therapeutic decision making. *J Am Coll Cardiol* 1990;15:1-14.
24. Braunwald E, Antman EM, Beasley JW, et al. ACC/AHA guideline update for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction. Available at http://www.acc.org/clinical/guidelines/unstable/update/pdf/UA_update.pdf.
25. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002;287:612-7.
26. Fuster V, Rydén LE, Asinger RW, et al. ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation executive summary. A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines and Policy Conferences (Committee to Develop Guidelines for the Management of Patients With Atrial Fibrillation): developed in Collaboration With the North American Society of Pacing and Electrophysiology. *J Am Coll Cardiol* 2001;38:1231-66.
27. The GUSTO IV-ACS Investigators. Effect of glycoprotein IIb/IIIa receptor blocker abciximab on outcome in patients with acute coronary syndromes without early coronary revascularisation: the GUSTO IV-ACS randomised trial. *Lancet* 2001;357:1915-24.
28. Bertrand ME, for the Task Force of the European Society of Cardiology. Management of acute coronary syndromes: acute coronary syndromes without persistent ST segment elevation. *Eur Heart J* 2000;21:1406-32.
29. Hunt SA, for the Task Force on Practice Guidelines. ACC/AHA practice guidelines for evaluation and management of chronic heart failure in the adult. *J Am Coll Cardiol* 2001;38:2101-13.
30. Mark DB, Hlatky MA. Medical economics and the assessment of value in cardiovascular medicine: part I. *Circulation*. 2002;106:516-20.
31. Marciniak TA, Ellerbeck EF, Radford MJ, et al. Improving the quality of care for medicare patients with acute myocardial infarction: results from the Cooperative Cardiovascular Project. *JAMA* 1998;279:1351-7.
32. Kiefe CI, Allison JJ, Williams OD, et al. Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. *JAMA* 2001;285:2871-9.
33. Ferlie EB, Shortell SM. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *Milbank Q* 2001;79:281-6.