Using a Clinical Practice Guideline to Measure Physician Practice: Translating a Guideline for the Management of Heart Failure

Paul A. James, MD, Timothy M. Cowan, MSPH, Robin P. Graham, PhD, MPH, Barbara A. Majeroni, MD, Chester H. Fox, MD, and Carlos R. Jaén, MD, PhD

Background: **Effective clinical practice guidelines should improve clinical outcomes, and measures of physician use of clinical practice guidelines should correlate with improved outcomes. This study translates a clinical practice guideline on heart failure into review criteria to measure physician performance and the effectiveness of the clinical practice guideline.**

Methods: **A panel of 11 family physicians and 1 cardiologist systematically reviewed the clinical practice guideline for its clinical importance, educational relevance, and evaluative appropriateness. Then a subset of 4 family physicians rigorously applied each recommendation to established criteria for measurability and developed an evaluation tool useful in medical record review.**

Results: **The heart failure clinical practice guideline was found to be an excellent educational tool. Using it to measure physician performance, however, was limited to diagnostic tests and drug prescribing. Of 45 recommendations, 5 fulfilled criteria for measurability; 1 recommendation had A-level evidence, whereas 2 recommendations had B-level and 2 had C-level evidence.**

Conclusion: **This study illustrates the logistic issues and challenges in developing a measure of physician adherence to clinical practice guidelines. Medical record review is inadequate to measure many recommendations. Physician use of this clinical practice guideline must be evaluated as an intermediate step to measuring the effectiveness of clinical practice guidelines based on patient outcomes.**
(J Am Board Fam Pract 1997;10:206-12.)

Left-ventricular systolic dysfunction, or congestive heart failure (CHF), is a health problem that affects more than 2 million Americans at a cost of more than $10 billion annually. The 5-year mortality rate is about 50 percent for the almost 400,000 new cases diagnosed each year. The mortality rate for CHF patients on angiotensin-converting enzyme (ACE) inhibitors[1] and for those with angina who undergo coronary artery bypass graft surgery[2] has decreased considerably in the last 10 years.

Family physicians care for a great percentage of CHF patients and are increasingly under scrutiny for the quality of care rendered to these patients. A survey of practice styles among family physicians, internists, and cardiologists found wide variability and major cost differences.[3] Patients with CHF discharged from the hospital too early were found to have increased mortality at 90 days,[4] and in less than one half of the medical records reviewed in one study was there physician documentation of a low-salt diet.[5] In response to such concerns, the Agency for Health Care Policy and Research (AHCPR) developed its 11th clinical practice guideline, *Heart Failure: Evaluation and Care of Patients with Left-Ventricular Systolic Dysfunction*, for practicing clinicians.[6,7]

Analyses of clinical practice guidelines in the family practice literature have focused on the evidence and process of guideline development.[8-11] This study will describe an evaluation of the AHCPR heart failure clinical practice guideline for its ability to be a tool to measure effective clinical practice. Measurement of physician adherence to a clinical practice guideline is a necessary first step to evaluate effectiveness of a guideline on clinical outcomes. This measurement has important policy implications. For example, one state

Submitted, revised, 15 September 1996.

From the Office of Rural Health (PAJ, TMC), and the Department of Family Medicine (PAJ, TMC, RPG, BAM, CHF, CRJ), State University of New York at Buffalo. Address reprint requests to Paul A. James, MD, SUNY Office of Rural Health, Erie County Medical Center, 462 Grider Street, Buffalo, NY 14215.

This project was supported by the New York State Department of Health, Office of Quality Improvement.

has enacted tort reform to assist physicians who comply with a clinical practice guideline, but the success of this program has been limited by the lack of measurability of the guidelines.[12] This paper will describe the translation of the clinical practice guideline into review criteria and will discuss the use of a clinical practice guideline in measuring the quality of clinical practice.

Process of Guideline Development

The AHCPR contracted with RAND, a nonprofit public policy and research corporation, to develop the heart failure clinical practice guideline. With the approval of AHCPR, RAND selected a panel of experts (predominantly academic specialists, but with representation by generalists) and consumers that used an explicit approach to sort out and rate the evidence for each recommendation. The panel first met in February 1992, and drafts were later evaluated by outside persons and organizations. The final clinical practice guideline was published in June 1994.[6]

Methods

For this study the AHCPR clinical practice guideline on heart failure was reviewed by an advisory panel of 11 family physicians and an academic cardiologist. The panel comprised family physicians from diverse environments including rural, suburban, urban, academic, nonacademic, group, solo, managed care, and fee-for-service practices. Each clinical practice guideline recommendation was assessed for its educational appropriateness, clinical importance, and measurement feasibility.

Recommendations were considered specific and intended guides if printed in boldface type and assigned an evidence rating in the text.[6] Evidence ratings were based on interpretation of supporting documentation found through literature review. An A rating indicated good evidence from well-conducted randomized controlled clinical trials or cohort studies. A B rating indicated fair evidence supported by case-control studies and poorly controlled studies or those recommendations for which there was conflicting evidence from well-designed studies. C-level evidence supported recommendations based on expert opinion.

Forty-five recommendations were reviewed. They were divided among the four major areas: prevention, patient evaluation, patient management, and the need for revascularization. Four family physicians (PAJ, BAM, CRJ, CHF) experienced in quality management assessed the clinical practice guideline for measurement validity by adapting a method described by Palmer et al.[13] Criteria for this assessment are listed in Table 1.

Results

Advisory panel satisfaction with the clinical practice guideline was high. Recommendations for clinical practice were considered valid and useful as an educational resource for family physicians and their patients. The panel members felt that the clinical practice guideline represented appropriate practice for family physicians.

The panel perceived extreme difficulty in measuring physician performance based on the clinical practice guideline, however. Because the medical record is vital as a witness to medical decisions and actions, the panel believed that valid measures of physician performance would require documentation. Diagnostic test ordering and therapeutic management were believed to be most reliably recorded in the medical record. Patient education and activities requiring increased patient responsibility are less likely to be documented.

Clinical Importance and Utility

The clinical practice guideline is presented in a fashion consistent with clinical practice. Prevention of CHF in asymptomatic patients is discussed first, followed by the initial evaluation and a discussion of alternative diagnoses and management issues. Specific management includes pharmacotherapy as well as patient education. Finally, clinicians are encouraged to examine the need for revascularization in all CHF patients.

Table 1. Criteria for Determining Measurable Recommendations.

Guideline recommendation is an action
Strength of epidemiological evidence (as stated in guideline)
Strength of evidence based on physician's personal experience or experimental evidence
Condition can be detected
Data source has data element for measuring
Can make yes-or-no decision about criterion compliance
Information necessary for yes-or-no decision found in data sources
Acceptable alternatives and exclusions specified clearly
Time frame defined for observation of criterion compliance
Recommendation reflects intent of guideline

Table 2. Reasons Clinical Practice Guideline Recommendations for Heart Failure Were Not Measurable.

Reasons	Number
Recommendations informational only	8
Could not adequately recognize patients for whom the recommendation was written using existing medical record data	7
Data elements considered not available in medical records	19
No alternatives or exclusions specified for recommendation	2
A yes-or-no decision could not be made for the recommendation	3
Did not reflect intent of guideline	1

Prescribing an ACE inhibitor for asymptomatic patients with reduced left ventricular ejection fraction (LVEF) is emphasized. Family physicians will find this information useful only in the event that an asymptomatic patient will be tested for reduced LVEF. The dilemma for physicians is determining which patients are asymptomatic. The text supporting this recommendation provides a more meaningful suggestion that clinicians consider testing LVEF in patients following specific types of myocardial infarction, but no evidence rating is provided, and this is not a formal recommendation of the clinical practice guideline.

The examination of the symptomatic patient and diagnostic considerations follow. The importance of assessing volume overload is emphasized, as is its treatment with a diuretic. Patient management using ACE inhibitors, diuretics, digoxin, and hydralazine-isosorbide is discussed. Diet, exercise, and advance directives are important patient education areas to document. Additionally, physicians are encouraged to monitor weight and educate patients about self-monitoring. Though the nonpharmacologic recommendations are based on C-level evidence, the panel agreed with their importance. Nevertheless, several panel members suggested that discussing advance directives at the initial office visit might be premature; they preferred to establish the patient-physician relationship first.

The section "Need for Revascularization" highlights the importance of a primary care physician. Within the algorithm of this clinical practice guideline, there are three decision nodes that emphasize counseling and discussing options for care with the patient. Although this section adequately covers the decision analysis involving CHF patients, the tasks of quantifying and explaining the risks of various options can involve major value judgments by the physician.

Measurability of the Clinical Practice Guideline

Forty-five recommendations were evaluated for measurability based on the criteria in Table 1. A summary of recommendations that were not measurable and why is listed in Table 2. Forty of the 45 recommendations were not measurable, leaving only 5 measurable recommendations. Limitations included recommendations that are informational only (8), involve actions for patients whose condition is difficult to detect in clinical practice (7), require unavailable data (19), provide no alternative or exclusionary criteria (2), and require the reviewer to make a yes-or-no decision that is clinically difficult in an external review (3). Examples of these limitations follow.

Informational recommendations demand no action on the part of clinicians, and thus compliance is not measurable. Eight of the recommendations are informational. An example of an informational recommendation is, "Digoxin can prevent clinical deterioration in patients with heart failure due to left-ventricular systolic dysfunction and improve patients' symptoms." This A-level-evidence recommendation appears to endorse the use of digoxin in patients with CHF; however, the action recommendation for the use of digoxin has C-level evidence.

A clinician or evaluator could not distinguish the patient for whom the recommendation is intended in seven instances. An example of this is, "Asymptomatic patients who are found to have moderately or severely reduced left-ventricular systolic function should be treated with an ACE inhibitor to reduce the chance of developing clinical heart failure." This recommendation has A-level evidence and yet cannot be applied to a population of patients because there is no recommendation for how to recognize asymptomatic CHF patients.

Nineteen recommendations are not measurable because information about patient compliance would be lacking in the available data sources. These recommendations are predominantly related to patient education, counseling, and discernment of patient preferences, which are not reliably recorded in the medical record. An example

Table 3. Final Review Criteria Pertinent for Congestive Heart Failure.

1. Physicians should order a chest radiograph; electrocardiogram; complete blood count; serum electrolyte, serum creatinine, and serum albumin measurements; liver function tests; and urinalysis for all patients with suspected or clinically evident heart failure. Thyroxine and thyroid-stimulating hormone levels should also be measured in all patients older than 65 years with heart failure from no obvious cause and in patients who have atrial fibrillation or other signs or symptoms of thyroid disease [C evidence]
2. Patients with suspected heart failure should undergo echocardiography or radionuclide ventriculography to measure ejection fraction if information about left-ventricular function is not available from previous tests [B evidence]
3. Patients with heart failure and signs of severe volume overload should be started immediately on a diuretic. Patients with mild volume overload can be managed adequately on thiazide diuretics, whereas those with more severe volume overload should be started on a loop diuretic [C evidence]
4. All patients with congestive heart failure should undergo a trial of angiotensin-converting enzyme (ACE) inhibitors unless contraindicated by any of the following: [A evidence]
   a. A history of intolerance or adverse reactions
   b. Serum potassium greater than 5.5 mEq/L
   c. Symptomatic hypotension

   Patients with renal insufficiency (creatinine >3.0 mg/dL) should be given only a one-half dose of ACE inhibitors [B evidence]

Note: Review criterion 4 incorporates two measurable recommendations

of such a recommendation is, "All patients should be encouraged to complete advance directives regarding their health care preferences." Even if physicians encourage patients to complete advance directives, it is unlikely that available data would indicate physician encouragement.

Three recommendations are not amenable to yes-or-no decisions. One example is, "The panel recommends against the routine use of invasive or noninvasive tests, such as echocardiography or maximal exercise testing, for monitoring patients with heart failure." Although the data would be available to judge whether the test was done, the judgment of whether the test was ordered routinely is a complex decision.

Two recommendations list no alternatives or exclusion criteria. "Patients with a history of syncope possibly due to ventricular arrhythmias should be referred immediately to a practitioner with expertise in arrhythmias," is an example. Although no alternatives or exclusion criteria are listed for this recommendation, panel members offered several scenarios whereby they would not refer immediately.

Valid and available measures of one recommendation with A-level evidence were found, yet the panel felt this recommendation did not address the intent of the guideline. "Patients with heart failure and angina who will not or cannot undergo revascularization should be treated with nitrates and aspirin," is a recommendation to treat angina and is peripheral to CHF.

The final review criteria for compliance to the AHCPR clinical practice guideline for heart failure are listed in Table 3. Only 5 recommendations are acceptable as review criteria, with 2 combined to form 1 review criterion. Each of the recommendations that form the basis for review criteria is either a patient evaluation regimen or pharmaceutical prescribing. Although not listed in Table 3, a systematic weakness of the heart failure clinical practice guideline is its lack of establishing time frames for compliance with recommendations. This area is poorly studied, and the recommendation would be dependent upon expert opinion. The panel of reviewers estimated time frames to establish final review criteria.

Figure 1 illustrates the number of measurable and nonmeasurable recommendations in the clinical practice guideline and their evidence ratings. The evidence rating of a recommendation appears to have little or no relation to its measurability. Of five recommendations having A-level evidence, only one proved to be an effective review criterion. Two of 11 recommendations with B-level evidence and 2 of 29 recommendations based on C-level evidence were acceptable for review criteria.

Discussion

Special Features of This Analysis

This analysis differs from previous approaches to assess clinical practice guidelines. We adapted the methodology of Palmer because it allowed an explicit and rigorous translation of the clinical practice guideline into medical review criteria. Our

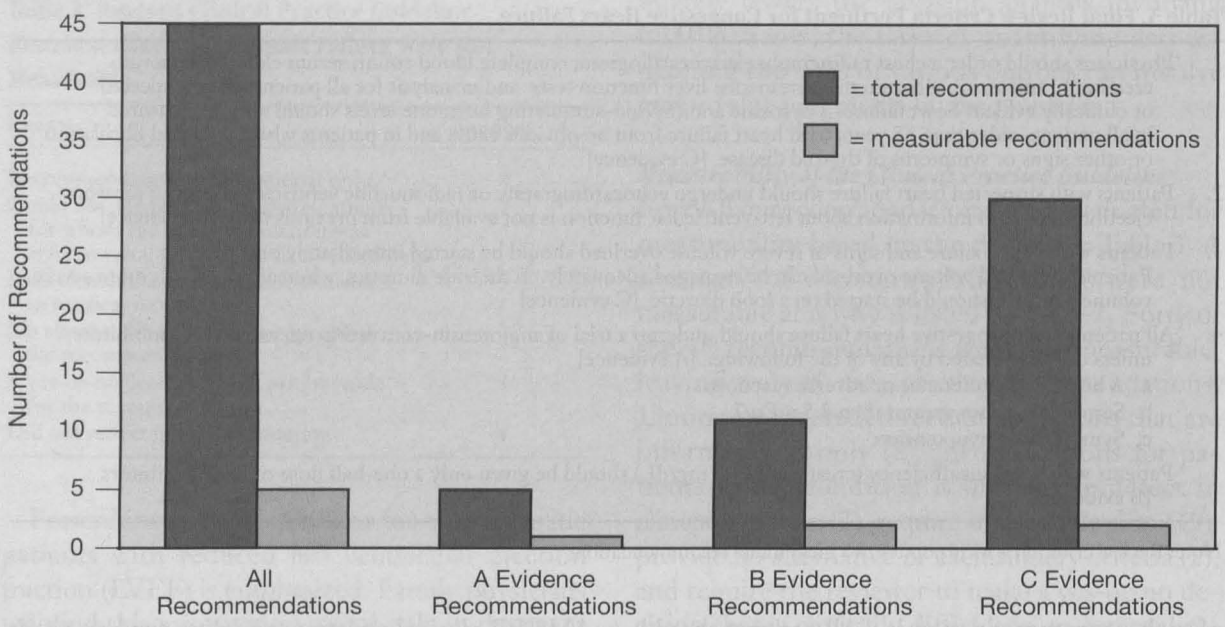

Figure 1. Comparison of evidence-based recommendations and their adaptation into measurable review criteria.

reviewers were family physicians and did not represent other specialties. The purpose of this study, however, was to analyze the practices of family physicians who care for CHF patients, and thus peer review by family physicians was most appropriate.

Hadorn et al[14] published the RAND group's translation of this clinical practice guideline, and their findings are comparable to ours. Differences do exist, however. They found 34 recommendations instead of 45, but did not state their criteria for selection. Rather than using the methodology of Palmer, their assessment used 2 criteria: "importance to quality of care" and "feasibility of monitoring."[14] Only 4 of 8 final review criteria were selected unanimously by their panel, and these criteria approximate our review criteria. The other 4 review criteria either did not describe a recommendation or did not fulfill our assessment criteria. Clearly, the use of an established method for translation of clinical practice guidelines into review criteria is important to reduce interpretive differences.

This method appears to have considerable face validity. The larger panel of physicians that discussed the clinical practice guideline recommendations predicted the findings of our analysis: the clinical practice guideline would measure diagnostic tests and treatments because they are more reliably documented in the medical record. Chart reviews by our staff have demonstrated the adequacy of these review criteria. The additional measures established by the RAND group required substantially more resources and professional interpretive skill.

Education and care of chronically ill patients are dependent on patient-physician communication.[15] The heart failure clinical practice guideline reflects the importance of communication, but these recommendations cannot be validly measured. We believe the inadequacy of measurable recommendations is somewhat a reflection of clinical practice and the inadequacy of the medical record.

The purpose of this clinical practice guideline is to educate physicians and patients. Members of the RAND advisory panel expressed concern about the use of this clinical practice guideline as a tool for assessing practice patterns.[14] One might conclude that this clinical practice guideline was not efficacious as a foundation for a measurement tool and that it was inappropriately applied. On the contrary, the ultimate goal of the clinical practice guideline must be improved clinical outcomes, and an evidence-based guideline is the logical foundation for this effort. Future studies must determine whether the review criteria judged here are valid measures of quality care.

Implications of This Study

The AHCPR clinical practice guideline for heart failure is an effective resource and educational tool for primary care clinicians. Although this guideline has limitations, it provides an excellent starting point for improving clinical practice. Its format is consistent with clinical reasoning and provides flexibility. Nevertheless, the characteristics that make the clinical practice guideline an excellent educational resource hinder its effectiveness as a measurement tool. The clinical practice guideline flexibility allows wide adaptation for primary care patients, but reduces its specificity for measurement of guideline compliance.

To measure the effectiveness of a clinical practice guideline, recommendations must be translated into actions to improve specific patient outcomes. If physicians are to comply with a clinical practice guideline, they must be able to distinguish for which patient the clinical practice guideline is intended, get agreement from the patient for the course of action, implement the action, and document the action so that it is measurable.

As measurement tools are developed to monitor quality of care, we must study the patient's role in physician adherence to clinical practice guidelines. For patients with chronic diseases, the importance of patient compliance is vital to improve clinical outcomes.[15] Monane et al[16] found patient compliance among the elderly with CHF to be a serious problem in regard to use of medications and patient management. Patient-specific variables can be important considerations in explaining clinical outcomes. A weakness of this methodology is that actions physicians can control, such as prescription writing, are measured, but actions requiring mutual patient-physician agreement are not. This AHCPR clinical practice guideline on heart failure addresses the problem of patient compliance; however, effective measurement of those recommendations is limited.

Medical records are inadequate in relating patient and physician decisions. Physicians are not trained to document the variance between their recommendations and patient preferences. The patient is the only data source for this information, and to have an effective measure of physician adherence to clinical practice guidelines, access to this information must be found. Wennberg[17] suggests that the true use of clinical practice guidelines is in assisting physicians as they explain the choices patients must consider. Family physicians should look for an effective means to document these decisions in the medical record.

The development of an evaluative instrument to measure physician adherence to the AHCPR clinical practice guideline on heart failure was limited by a lack of explicit time frames for the implementation of the guideline. Our review criteria were developed using best guesses for effective time frames. Outpatient care and care for patients with chronic disease are predicated on the notion that time is an ally and that a slow and repeating management plan is often necessary.[18] Thus, more so than for hospital and acute care, the measurement of clinical practice guideline adherence in primary care must carefully establish time frames that are not overly restrictive.

An interesting finding from this analysis was that strong scientific evidence from randomized controlled trials had only a small influence on the evaluation instrument for measuring adherence among family physicians. Of the 45 recommendations reviewed, 5 are given A-level evidence (an indication of good scientific evidence) by the AHCPR panel. Only one of these recommendations provides a useful measure for evaluation. This finding underscores the dilemma of applying results from large, highly controlled clinical trials to individual patients.[19] The application and implementation of scientific knowledge can be as complex as the acquisition of the new knowledge itself.

This study raises important questions regarding clinical practice guidelines as tools for measuring adherence. Clinical practice guidelines have been evaluated based on their developmental process, generalizability, applicability, and validity as educational tools. Our research used the clinical practice guideline as the foundation for a measurement tool applying established methodology. Further work in primary care practices measuring compliance of physicians and patients is needed. Ultimately, clinical practice guidelines must be judged by their effect on clinical practice and patient outcomes.

References

1. Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). The CONSENSUS Trial Study Group. N Engl J Med 1987;316:1429-35.

2. Bounous EP, Mark DB, Pollock BG, Hlatky MA, Harrell FE Jr, Lee KL, et al. Surgical survival benefits for coronary disease patients with left-ventricular dysfunction. Circulation 1988;78(3 Pt 2):I-151-7.
3. Fleg JL, Hinton PC, Lakatta EG, Marcus FI, Smith TW, Strauss HC, et al. Physician utilization of laboratory procedures to monitor outpatients with congestive heart failure. Arch Intern Med 1989;149: 393-6.
4. Kosecoff J, Kanouse DE, Rogers WH, McCloskey L, Winslow CM, Brook RH. Effects of the National Institutes of Health Consensus Development Program on physician practice. JAMA 1987;258:2708-13.
5. Retchin SM, Brown B. Elderly patients with congestive heart failure under prepaid care. Am J Med 1991;90:236-42.
6. Konstam MA, Dracup K, Baker DW, et al. Heart failure: evaluation and care of patients with left-ventricular systolic dysfunction. Clinical practice guideline no. 11. Rockville, Md: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, 1994. (AHCPR publication no. 94-0612.)
7. Brooks NH. Five points from the AHCPR guideline on heart failure. Am Fam Physician 1994;50:531-2.
8. Culpepper L, Froom J. Otitis media with effusion in young children: treatment in search of a problem? J Am Board Fam Pract 1995;8:305-16.
9. Stevens NG. Guidelines for the diagnosis and treatment of major depression. J Am Board Fam Pract 1994;7:49-59.
10. Stevens NG, Lyle S. Guidelines for adolescent preventive services: a critical review. J Am Board Fam Pract 1994;7:421-30.
11. Berg AO, Moy JG. Clinical guidelines and primary care guidelines for the diagnosis and management of asthma. J Am Board Fam Pract 1992;5:629-34.
12. Questions raised about impact of guidelines in Maine project. Am Med News 1996;39:21, 68.
13. Palmer RH, Banks NJ, Spath P. Checklist for developing guideline-derived evaluation instruments. In: Using clinical practice guidelines to evaluate quality of care. Vol. 2. Methods. Rockville, Md: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, 1995:73-88. (AHCPR publication no. 95-0046.)
14. Hadorn DC, Baker DW, Kamberg CJ, Brook RH. Phase II of the AHCPR-sponsored heart failure guideline: translating practice recommendations into review criteria. J Qual Improv 1996;22:265-76.
15. Kaplan SH, Greenfield S, Ware JE Jr. Assessing the effects of physician-patient interactions on the outcomes of chronic disease. Med Care 1989;27:S110-27.
16. Monane M, Bohn RL, Gurwitz JH, Glynn RJ, Avorn J. Noncompliance with congestive heart failure therapy in the elderly. Arch Intern Med 1994; 154:433-7.
17. Wennberg JE. Unwanted variations in the rules of practice. JAMA 1991;265:1306-7.
18. Rosenthal TC, Riemenschneider TA, Feather J. Preserving the patient referral process in the managed care environment. Am J Med 1996;100:338-43.
19. Asch D, Hershey JC. Why some health policies don't make sense at the bedside. Ann Intern Med 1995;122:846-50.