Patients with heart failure and reduced ejection fraction account for about half of the 6 million patients with heart failure and the $35 billion spent annually on related medical costs. Over the past 25 years, heart failure management has evolved through randomized trials of specific therapy and the extensive clinical experience accumulated at referral centers, many of which arose initially to offer comprehensive medical care for patients in parallel with cardiac transplantation, as described by Fonarow et al in 1997.\textsuperscript{1} The recommended therapies have been the subject of extensive guidelines, from the first Agency for Health Care Policy and Research guidelines chaired by Konstam in 1994\textsuperscript{2} to the most recent American College of Cardiology (ACC)/American Heart Association (AHA) guidelines chaired by Hunt in 2005\textsuperscript{3} and updated by Jessup et al in 2009.\textsuperscript{4} Lessons from implementation of measures for acute myocardial infarction have translated into heart failure management, as described by O’Connor.\textsuperscript{5} Outcomes with heart failure have improved remarkably as these therapies have become widely implemented, and even disparities by race have begun to diminish, as underlined by Yancy et al;\textsuperscript{6} however, the remaining burdens of readmission and mortality warrant ongoing efforts to further improve care for the diverse populations with heart failure.

The Registry to Improve the Use of Evidence-Based Heart Failure Therapies in the Outpatient Setting (IMPROVE HF)\textsuperscript{7} successfully increased the number of therapies provided to patients with heart failure. Testing the impact of a strategy on practice is challenging, because the investigators cannot be blinded and are exposed to center effects, which may have played a role here. Simultaneous trends in continuing education and community practice also influence practice patterns over time. Data were collected by 34 centralized chart review specialists, whose abstractions were likely highly reliable for identifying the numerator of therapies prescribed but more limited in determination of the denominator of eligibility for those therapies. It is particularly difficult to identify a history of previous adverse effects or chronic noncardiac conditions that limit therapies and survival. However, the intensity and magnitude of effects demonstrated here provide strong evidence that the IMPROVE HF strategy was responsible for increasing the number of therapies provided to patients with heart failure in the 167 participating institutions.

The strategy encompassed multiple components: Education for physicians included algorithms for recommended therapies, and information for patients included equivalent emphasis on specific medical therapies and pacing/defibrillation devices. The increase in prescription of therapies was much higher when the same patients were compared at baseline and during follow-up than when new cohorts of patients were compared during the trial. Thus, it appears that physician practice was influenced less by the general education tools than by the monitored feedback tool that identified specific patients as needing additional therapies. The impact on prescriptions for patients deemed eligible ranged from an additional 1 in 14 patients for angiotensin-converting enzyme inhibitor (ACEI)/angiotensin receptor blockers (ARBs) and β-blockers to 1 in 4 for spironolactone and 1 in 3 for cardiac resynchronization therapy (CRT) and implantable cardioverter-defibrillators (ICDs). A strategy that can increase intervention to this degree confers a somber obligation to ensure that the additional intervention does indeed offer to enhance the quality of life or length of quality life for the patient, not just the reported quality of the physician being assessed. The more powerful the missile launched, the more important it is to aim it in the right direction.

Small Impact on Approved Performance Measures

Four of the 7 measures included as quality measures in the IMPROVE HF tools have been approved by the ACC/AHA as performance measures. Acceptance as a performance measure requires not only evidence of benefit but also sufficient simplicity of indications and transparency of contraindications that eligible patients can feasibly be identified.\textsuperscript{8} Therapies approved as performance measures generally have wide applicability across multiple subgroups, revealed either in specific trials or through extensive postapproval experience. We would also anticipate that the universality of such therapies would lead to equivalent endorsement and use in most developed nations with similar populations. As included
in the heart failure performance measure development led by Bonow et al in 2005, the measures of ACEI/ARB and β-blockers have easily cleared these bars. As anticipated, penetrance for ACEI/ARB was high (80%) at baseline in this study. The absolute increase of 6.7% in the longitudinal cohort was small and did not quite reach significance when adjusted for center effects. The 86% penetrance of β-blockers at baseline was even higher and increased by 7.4%. Anticoagulation for atrial fibrillation was only 69% at baseline, and this number did not increase despite the feedback about underuse. This discrepancy suggests that perceived contraindications for atrial fibrillation may not have been reliably captured by the chart review. Heart failure education, for which virtually all patients should be eligible, was only 62% at baseline; although the absolute increase of 9% was significant, the residual 29% gap in documented education for which virtually all patients should be eligible, was only 62% at baseline; although the absolute increase of 9% was significant, the residual 29% gap in documented heart failure education in a population with high utilization of ACEI/ARB and β-blockers suggests an unanswered deficiency either in practice or in documentation.

Large Impact to Increase Therapies Not Approved for Performance Measures

Spironolactone, CRT, and ICD were the therapies for which use was most increased by the IMPROVE HF strategies. Although listed as level I recommendations for specific patients with symptomatic heart failure, these have not been approved as performance measures. Generally, guideline-recommended therapies that are not performance measures present 1 or more of the following concerns: Benefit may be considered insufficient or harm excessive in an important subgroup for which statistical power was insufficient; adequate expertise and resources may not be widely available for safe and effective implementation; patient commitment and compliance for monitoring may not be reliable; or accurate determination of the denominator is not feasible. The 3 additional measures of spironolactone, CRT, and ICD included as indicators of quality improvement in this study were neither approved as performance measures in 2005 by the ACC/AHA nor endorsed by the National Quality Forum. During the exhaustive deliberations that surround level I guideline recommendations, committee members may express concern about balancing the medical risk of prescribing and the legal risk of withholding a recommended therapy due to complex considerations, particularly in the older heart failure patient who may have noncardiac limitations, including frailty. It is not possible to list all such considerations in a guideline document; it is assumed that physicians will meet their responsibility to integrate multiple medical and personal aspects of care for each patient. Concern for defensive practice has in the past been softened by the strong reassurance that guidelines are only to guide, not to dictate, practice. A clue that a level I recommendation has been the subject of dissent and compromise may be revealed in the length and lack of specificity of the qualifying phrases. This may apply to the recommendation for spironolactone, which is described in the ACC/AHA Heart Failure guidelines as “reasonable in selected patients with moderately severe to severe symptoms of [heart failure] and reduced [left ventricular ejection fraction] who can be carefully monitored for preserved renal function and normal potassium concentration. Creatinine should be ≤2.5 mg/dL in men and ≤2.0 mg/dL in women and potassium should be less than 5.0 mEq/L. Under circumstances where monitoring for hyperkalemia or renal dysfunction is not anticipated to be feasible, the risks may outweigh the benefits.” With fewer words but more global uncertainty, the ICD recommendation for primary prevention specifies patients “with expected survival with good functional capacity for >1 year.” It is not an accident that these phrases defy translation into the quantitative criteria required for the denominator of a performance measure. Defining a therapy as a quality measure does not necessarily circumvent the concerns that hinder its approval as a performance measure.

Is More Therapy Always Better Therapy? Aldosterone Antagonists

The use of aldosterone antagonists increased from 34% to 62% of patients considered eligible in the group followed up during the 24 months. Although the evidence from the Randomized Aldactone Evaluation Study (RALES) trial has shown benefit in select populations under close protocol supervision, the pilot trial for the Randomized Aldactone Evaluation Study showed 13% incidence of hyperkalemia with doses of 25 mg and 20% with doses of 50 mg, with a risk ratio of 2.9 for doses of ACEIs that were half or more of target doses. The hazard of translating trials into community practice was suggested in the study of 1.3 million adults older than 65 years of age with heart failure in Canada. An excess of 560 hyperkalemia hospitalizations and 73 hospital deaths was attributed to a small absolute increase from spironolactone prescription from 34 to 149 per 1000 patients with heart failure, for whom the anticipated benefit to decrease readmission rate was not seen. In a community study in the United States, Bozkurt et al reported that patients receiving aldosterone antagonists were less carefully selected and followed up than in the trials, with a 24% incidence of hyperkalemia, and in a Scandinavian population with a mean age of 73 years, there was a 10% incidence of potassium levels >6 mmol/L.

Chart review of eligibility for aldosterone antagonists must have been a daunting task, requiring location of both the objective and subjective criteria listed above. The complexity of the denominator for this recommendation would likely not meet criteria for a true performance measure. Only 17% of patients in the IMPROVE HF study were considered eligible for aldosterone antagonists. More than half of the 80% increase in recommended use during the 24 months actually resulted from a decline in the denominator due to newly documented contraindications or intolerance; however, we do not have any information on whether the tool may have indiscriminately increased use in the many patients who were not considered eligible because of baseline conditions or limited surveillance of renal function and potassium. It is concerning that there was no significant interaction between the likelihood of increased use and patient age, with use of aldosterone antagonists in patients older than 80 years of age increasing from 25% to 46% after the intervention to change practice.
Cardiac Resynchronization Therapy
Cardiac resynchronization has dramatically improved the quality of life and also lengthened survival for responsive patients. Unlike other therapies for symptomatic heart failure, this therapy is recommended only for patients with QRS >120 ms, as defined by the entry criteria for the trial populations. This degree of QRS prolongation defines approximately one third of patients with class III or IV symptoms of heart failure and reduced ejection fraction. However, in the pivotal trials of MIRACLE (Multicenter InSync Randomized Clinical Evaluation) and COMPANION (Comparison of Medical Therapy, Pacing, and Defibrillation in Heart Failure), the benefit was demonstrable only in the smaller group of patients with a QRS >150 ms. In the CARE-HF (Cardiac Resynchronization in Heart Failure) trial, which showed mortality benefit, >80% of patients had a QRS duration ≥150 ms. The benefit for patients with atrial fibrillation and intrinsic atrioventricular conduction remains under debate, although there is increasing evidence that patients with right bundle-branch block may actually worsen with CRT. The AHA/ACC guidelines that recommended cardiac resynchronization also emphasized the need for further information in these subgroups before benefit is assumed. With these caveats, does increased use of CRT in all symptomatic patients with QRS >120 ms represent improved care? Interestingly, more of the increase in use of this recommended therapy was in patients who became newly eligible during the 2 years than in patients who were initially eligible. Although the relative increase in CRT use was high, only 8% of patients were considered eligible at baseline and 14% at 24-month follow-up, so as with the use of aldosterone antagonists, CRT use would not serve as a broad measure of care.

Implantable Defibrillators
The measure directed to increase use of implantable defibrillators affected the largest number of patients during the IMPROVE HF study. Of the 7605 patients who underwent 24-month follow-up, the population eligible for ICD was 70% the size of the ACEI/ARB-eligible population at the beginning of the trial and 86% the size of the ACEI/ARB population at the end of the trial. Although the number of patients receiving ACEI/ARB declined by 573 in the longitudinal study of 7605 patients, the number of patients with ICDs increased by 1482, 769 of whom became newly eligible during the 24 months.

Although the portrait of the population to derive benefit from ICD remains indistinct, even current recommendations are not fully aligned. The design of the IMPROVE HF strategy selected an ejection fraction criterion of ≤35% for symptomatic heart failure, in contrast to the 30% in the AHA/ACC level 1 recommendations for symptomatic heart failure, in which ejection fraction from 30% to 35% is relegated to a level IIa recommendation. The criteria for IMPROVE HF more closely resemble the guidelines for device-based therapy. Both guidelines specify that patients considered for ICD implantation for primary prevention should have reasonable expectation of survival with good (heart failure guideline) or acceptable (device guideline) functional status for at least 1 year. Anticipation of survival and functional capacity poses a major challenge for clinicians and likely represented the most difficult challenge during the chart review in IMPROVE HF, which may explain in part why the denominator of eligible patients appeared surprisingly high compared with the overall proportion of patients taking ACEI/ARBs.

In contrast to the vast experience with ACEI/ARBs and β-blockers, the subpopulations from ICD trials do not provide robust evidence of benefit. Class II but not class III patients showed benefit in the Sudden Cardiac Death in Heart Failure Trial (SCD-HeFT), whereas in the Multicenter Automatic Defibrillator Implantation Trial II (MADIT II), patients who were less sick appeared to derive less benefit. A recent meta-analysis of 5 major trials suggested that women derive no benefit from an ICD, with only 1 trial showing even a numeric trend for benefit in women.

There is growing controversy about the benefit of ICDs in older patients, as summarized by Epstein et al. Benefit for older patients was shown in the Multicenter Automatic Defibrillator Implantation Trial II trial but was diminished in the Sudden Cardiac Death in Heart Failure Trial and was not shown in secondary prevention for patients older than 75 years. In the later decades of life, the contribution of sudden death to total mortality declines. Even effective termination of ventricular tachyarrhythmias may not necessarily prolong life in the unselected community heart failure population, in which the average age is approximately 75 years. In 14 000 patients after their first heart failure hospitalization, the median survival was 2.4 years, with only 6% of the 9000 deaths being attributed to cardiac causes in patients living independently at home. More than 40% of new devices are currently implanted in patients older than 70 years, and >10% are implanted in patients older than 80 years. In the IMPROVE HF experience reported here, there was a lower rate of ICD use in the older population, but even in octogenarians, the performance strategy increased use of the ICD from 45% to 66%.

Although the benefit of sudden death prevention to improve survival may not include all populations, all face the possibility of harm. Older age increases the risk of in-hospital complications after ICD implantation. Although older patients may experience fewer inappropriate shocks from exogenous sinus tachycardia, atrial fibrillation increases with advancing age. Morbidity from defective leads remains substantial. New information suggests that neurological dysfunction may be impaired by ICD implantation in some patients, particularly after defibrillation threshold testing. With a 2.4-year median survival after the first heart failure hospitalization in the community, many patients will not survive to reach the 5- to 7-year threshold for cost-effectiveness of ICDs. A recent analysis of the Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan (EVEREST) trial from Wang et al suggests that ICDs did not prolong life once heart failure worsened to require hospitalization, but their use was associated with increased rehospitalization. No information is provided on how patients received information and participated in the decision to have ICDs implanted; other experience suggests that patients lack insight into differing modes of death,
overestimate the survival benefit of the ICD, and misunder-
stand device inactivation.28,29

Can Excess Therapy Be Tracked?
There is no information about the precision of the IMPROVE HF approach to change practice. Does increased awareness of reporting encourage the addition of therapies in patients for whom they were inappropriate or contraindicated? Previous experience with spironolactone indicates common use outside guidelines in patients at high risk for hyperkalemia.12 Current national registry results suggest an absence of documented indications in 1 of 5 primary prevention ICD implantations, for CRT in 1 in 4 patients receiving a CRT defibrillator, and absence of approved training in 14% of implanting physicians.30 For ICDs, the design of the present study mentioned the exclusion of patients whose noncardiovascular conditions limited survival to less than 1 year but did not include consideration of cardiac prognosis except when New York Heart Association class was available or imputed.

The sensitivity of the IMPROVE HF tool to increase use was clearly high for the identification and treatment of patients who were not receiving therapies that they should have been getting. It would be more challenging to determine the specificity: How many of the total increases in therapy were appropriate given the overall condition of the patients? Although any accounting of the number of prescriptions will pass over the essence of continuing care, counting therapies that should have been given requires balance from counting those that should not.

What Should Be Recommended?
The search for bases of evidence has transformed the care of heart failure and led to improvement in both quality of life and survival; however, trial design reflects incentives for broad entry criteria and statistical victory that may not always translate into meaningful benefit for the individual patient. Currently, trial results are considered generalizable unless proven otherwise. Once therapies are approved for general use, experience is gained at the population margins, where neither prescription nor proscription of a therapy is judged. However, if the recommendations metamorphose into mandates and performance measures, guideline committees will have to be more conservative in their acceptance of new therapies. They will need to defer recommendations until all relevant subpopulations have been studied in detail. Furthermore, they will need to focus more on when and how to introduce new exclusions to old recommendations; physician practice will likely demonstrate hysteresis when instructed to curtail therapies previously encouraged for wide markets. Every stage, from initial approval and reimbursement through recommendations and revisions, should be illuminated with transparent independence from market forces that gain directly if their products or services are endorsed as measures of quality and thus as implicit mandates for performance.

The Quest for Quality
One clear lesson for practice groups is the importance of systematic and accessible documentation that potentially appropriate therapies have been considered. The original performance measures proposed by the ACC/AHA in 2005 for outpatient heart failure practice included a worksheet designed to document careful consideration of therapies and their perceived contraindications.8 This living record enhances implementation and monitoring of therapies, facilitates communications between providers, and prevents dangerous and costly duplication of efforts, including identification of adverse reactions.

The mandate to establish the electronic medical record is an unusual example of an intervention that should both enhance quality and decrease costs of health care. In the IMPROVE HF population, Walsh and colleagues31 showed more of an increase in ACEI/ARB use and heart failure education in sites that used electronic medical records; on the other hand, there was less of an increase in the use of ICDs. Perhaps coordinated access to the complete medical profile will encourage more integration of the multiple medical conditions and the patient preferences that should guide decisions about implantable defibrillators.

The relationship between performance measures and quality remains elusive. In the past, centers with higher rates of the classic performance measures tended to be those that met other criteria for centers of excellence in patient care, which led to a clear association between performance measures and outcome. A large wave of retrospective studies has provided reassurance that outcomes are better in patients who are taking ACEI/ARBs and β-blockers, and in some populations, outcomes are better for those with implantable defibrillators.32 However, these waters are rendered permanently unsafe by the rip tide of confounding indications, which cannot be adjusted away in the attempt to prove that incentivized increases in therapies improve outcome. As confirmed recently in a large population, those patients at highest risk for poor outcomes are indeed those with the highest prevalence of contraindications for recommended therapies.33

Vigorous education initiatives provided to both cardiologists and primary care practitioners have enhanced the prevalence of ACEI/ARB and β-blocker use into the 80% to 90% range.6 Residual differences around these high levels may reflect more variability of practice populations than of performance and offer diminishing targets for quality improvement. It may now be more relevant to evaluate the quality of the processes of care, particularly education, as shown by Krumholz et al34 for elderly hospitalized patients, rather than to count the therapies provided. Delineation of process for both indications and contraindications is particularly important as more heart failure management is supervised by advanced practice nurses and physicians’ assistants. Their participation in the IMPROVE HF study was linked with more heart failure education but also with almost 2-fold higher ICD use, which led to speculation that “findings from IMPROVE HF may lend support to [advanced practice nurse] and [physician assistant] staffing as a potential means to improve ICD therapy use.”35 Focus on process rather than prescription is particularly crucial when one considers the specter of proposed “composite” performance measures, which are only met when each of many therapies are prescribed.36 Composite measures based on therapies may
encourage convergence to the default of prescribing everything for everyone.

On the other hand, we can count and reward systematic documentation that all recommended therapies were appropriately considered, thus retaining in the denominator those patients with documented contraindications, who also deserve quality care. We can and reward continuing education for practitioners and also reward them for the education they provide their patients. We can count and reward interventions to improve patients’ telephone access to nurses who know them and to early follow-up appointments after hospital discharge. However, to tally and compete instead for the highest percentages of prescriptions and devices may push us close to the break-even brink beyond which we may fall into doing harm. Beyond process measures, we may not be able to legislate or reimburse the “quality” of our therapy for heart failure. It is unlikely that an increase in the quantity of therapies will serve as a valid surrogate. Albert Einstein is credited with reminding us that “Not everything that can be counted counts, and not everything that counts can be counted.”

Disclosures

Dr Stevenson has previously received research support and provided consulting for Medtronic Inc and Cardiomioms.

References


**Key Words:** Editorials ■ heart failure ■ health policy ■ quality indicators ■ guidelines ■ pay for performance