Journal of the American College of Cardiology
© 2002 by the American College of Cardiology Foundation
Published by Elsevier Science Inc.

Vol. 39, No. 9, 2002
ISSN 0735-1097/02/$22.00
PII S0735-1097(02)01773-4

REVIEW ARTICLES

Surrogate End Points in Heart Failure

Inder S. Anand, MD, FRCP, DPHIL (OXON), FACC,* Viorel G. Florea, MD, PHD, DSC,* Lloyd Fisher, PhD, FACC†

Minneapolis, Minnesota and Seattle, Washington

Because of the increasing number of pharmacologic strategies available for treatment of heart failure (HF), the time has come to reassess the adequacy of end points used to evaluate therapeutic efficacy. Interest in the use of surrogate end points in clinical studies is increasing. A surrogate end point is defined as a measurement that can substitute for a true end point for the purpose of comparing specific interventions or treatments in a clinical trial. A true end point is one that is of clinical importance to the patient (e.g., mortality or quality of life), whereas a surrogate end point is one biologically closer to the disease process (e.g., ejection fraction or left ventricular volume in HF). The prime motivation for the use of a surrogate end point concerns the possible reduction in sample size or trial duration. Such reductions have important cost implications and in some cases may influence trial feasibility. Another, perhaps more important, aspect of measuring surrogate end points is that they increase our understanding of the mechanism of action of drugs and thus may help physicians to take a more enlightened approach in managing their patients. In this article we have analyzed the possible potentials of the surrogate end points in clinical studies of patients with chronic HF. Other uses of possible surrogates are discussed, and the limitations in finding true surrogates are mentioned. At this time we conclude there is no well established surrogate in HF. (J Am Coll Cardiol 2002;39:1414–21) © 2002 by the American College of Cardiology Foundation

Chronic heart failure (HF) affects 2% of the population and is the fourth leading cause of adult hospitalizations in the U.S. and the most frequent cause of hospitalization in patients older than 65 years (1,2). The primary objectives in the treatment of patients with HF are to improve quality of life (QOL) and increase survival. Because the number of pharmacologic agents used in the treatment of HF is increasing and the mortality decreasing, the time has come to reassess the adequacy of end points used to evaluate therapeutic efficacy. Reduction in mortality is regarded as the gold standard. However, the large sample size that may be required to show an incremental survival advantage of a new drug is relatively prohibitive. Therefore, there has been recent increased interest in the use of surrogate end points, variables that can substitute for true end points for the purpose of comparing specific interventions in a clinical trial.

Surrogate end points have several potential advantages. Clinical trials evaluating surrogate end points require smaller sample size and can sometimes be completed in months rather than years. The ability to bring effective therapies to clinical practice quickly and inexpensively makes surrogate end points attractive in the drug-approval process. The principal disadvantage of using surrogates to assess therapies is the possibility of an incomplete, inadequate, or misleading evaluation (3–5). Drugs usually have multiple effects, and resorting to a single surrogate end point, focused exclusively on one intermediate effect, often precludes the evaluation of other intended or unintended health effects. Although even large clinical trials may not provide a complete evaluation of safety for rare adverse events, the relatively small sample size and the short duration typical in surrogate end-point trials provide little assurance of long-term safety. To use only surrogate end points is therefore to accept as empirical evidence for clinical practice a hypothesis about health benefits that has never been directly tested. The aim of this article was to analyze the possible potential of suggested surrogate end points for clinical studies of patients with chronic HF.

Most of the discussion of the use of surrogates has been in a very strong sense of the word "surrogate." For example, for a number of years the U.S. Food and Drug Administration (FDA) has approved medications that lower blood pressure with a presumed acceptable benefit/risk ratio. The presumption has been that any compound that lowers blood pressure in a hypertensive population would convey the benefits of preventing selected cardiovascular events as shown in early long-term placebo-controlled trials. The benefit is assumed even though the underlying biological mechanism of lowering blood pressure may differ. This would be one type of surrogate: a *strong surrogate*. For a well-defined end point or end points, appropriate changes in the surrogate are regarded as establishing appropriate changes in the end point(s).

There may be other potential variables not accepted as definitely establishing benefit for an end point(s) but considered so close that it is rational to use the putative surrogate variable for drug development in all but the final

From *University of Minnesota Medical School and Veterans Affairs Medical Center, Minneapolis, Minnesota, and †Department of Biostatistics, University of Washington, Seattle, Washington.

Manuscript received December 27, 2000; revised manuscript received January 31, 2002, accepted February 6, 2002.

Abbreviations and Acronyms

BEST	= Beta-blocker Evaluation of Survival Trial
BNP	= brain natriuretic peptide
CONSENSUS	= Cooperative North Scandinavian Enalapril Survival Study
EF	= ejection fraction
FDA	= Food and Drug Administration
HF	= heart failure
LV	= left ventricular/ventricle
MOXCON	= Effect of Sustained Release Moxonidine on Mortality and Morbidity in Patients with Congestive Heart Failure
NE	= norepinephrine
NYHA	= New York Heart Association
Peak V_{O_2}	= maximal oxygen intake
PRIME II	= Second Prospective Randomized Study of Ibopamine on Mortality and Efficacy
PROMISE	= Prospective Randomized Milrinone Survival Evaluation trial
QOL	= quality of life
REFLECT	= Randomized Evaluation of FLosequinan on ExerCise Tolerance
SAVE	= Survival And Ventricular Enlargement trial
Val-HeFT	= Valsartan Heart Failure Trial
V-HeFT	= Vasodilator-Heart Failure Trial

phase III trials of a new compound. Such variables might be used for dose ranging, preliminary proof of efficacy (phase I and II evaluation) and initial examination of the benefit/risk ratio. Such a variable might serve as a *developmental surrogate*, a measurement considered closely related to an end point of clinical benefit that makes it appropriate to use the developmental surrogate for dose ranging and preliminary proof of efficacy and benefit/risk ratio. Use of a developmental surrogate could result in considerable savings in time and resources, albeit with some risk for the final evaluation if the surrogate is not appropriate.

There has been much discussion about when engaging in only one trial is adequate for regulatory approval. In addition to evidence obtained with formal hypothesis testing and the associated p value, other evidence can be brought to support the findings. Variables known to be closely associated and predictive of a beneficial outcome can be used to support and strengthen the controlled trial data. This would be a *supportive surrogate*, a variable that strengthens the plausibility of favorable results from other controlled data.

TRUE END POINTS IN HF TRIALS

Because HF is a life-threatening and debilitating disorder, the true end points in HF should reflect patients' symptoms as well as their quality and duration of life.

Survival as a true end point. Mortality is currently regarded as the most important end point for evaluation of new HF drugs, and a reduction in mortality, or alternatively another beneficial effect with assurance of no important increase in mortality, is important for regulatory approval of the drug. Although mortality is a strong and easily measured end point, it has several limitations. The main concern of using only mortality as an end point is that it refers to the extreme manifestation of HF. Thus, most of the patients in the study do not contribute to a mortality primary end point, yet have important QOL issues. Because the current management of HF with angiotensin-converting enzyme (ACE) inhibitors and beta-blockers has reduced the event rate considerably, if mortality is the primary end point, patients with advanced diseases have to be studied to get enough events for reasonable statistical power in a reasonable period of time.

Consequently, patients in early stages of HF, in whom the disease process is most likely to be halted or possibly reversed, are not evaluated. Preventive strategy is therefore not assessed. Finally, trials using mortality as the primary end point require a large sample size to show a survival advantage of a new drug.

Symptoms and quality-of-life questionnaires. The most direct approach to the evaluation of HF is to inquire about symptoms, although to date this has not been an important aspect of HF regulatory discussions. Quality-of-life (QOL) questionnaires, which provide comprehensive information about the effects of a disease and its treatment on patients' lives, are now widely used in clinical trials. The QOL questionnaires can be general or disease-specific, and there is no general agreement as to which type of questionnaire is most appropriate. Few questionnaires have been validated in HF patients in a way that shows that the results correlate with the severity of disease.

Many ongoing studies of treatments for chronic HF incorporate the Minnesota Living with Heart Failure questionnaire as a measure of QOL (6). Statistically significant improvements in the QOL score have been observed in placebo-controlled studies of enalapril, flosequinan, pimobendan, vesnarinone and valsartan (7–11). However, flosequinan, pimobendan and vesnarinone have also been shown to have an adverse effect on survival (10,12,13), raising the issue of a trade-off between improved QOL and the risk of drug-induced death. Increased mortality with these and other agents clearly indicates that symptomatic benefit in HF does not necessarily predict improved survival (14). Because it is generally believed that a pharmacologic agent that improves symptoms in HF should have at least a relatively neutral survival effect, the pharmacologic industry has never pursued an HF indication for any of the above mentioned agents. The important issue of a trade-off between QOL and mortality has been discussed extensively by Patrick and Erickson (15).

SURROGATE END POINTS IN HF TRIALS

Chronic HF is the final common end point of several processes involved in the cardiovascular continuum that are

initiated by risk factors for cardiovascular diseases. Once initiated, cardiovascular disease progresses through structural remodeling of the heart and blood vessels. Factors that contribute to this include activation of various neurohormones, growth factors and cytokines. Markers of this biological process (e.g., left ventricular [LV] hypertrophy and enlargement) and factors that contribute to it (e.g., neurohormones) may be viewed as surrogates of the progression of the disease.

DEFINITION AND REQUIREMENTS FOR A SURROGATE END POINT

Several definitions have been suggested for a surrogate end point (16,17). For a surrogate end point to be considered an adequate substitute for the real end point, several levels of evidence must be provided (16,18–20). First, the biological relevance of the postulated relationship needs to be established by showing a strong and consistent relationship between the candidate for surrogate status (e.g., LV volume in HF patients) and outcomes in numerous studies of different populations and disease states (16). Second, changes in the surrogate end point (e.g., plasma brain natriuretic peptide [BNP] level) should predict a change in the morbid event independent of treatment. Third, a consistent proportionality between the degree of change in the surrogate end point and the true end point needs to be documented in a spectrum of studies, so that the magnitude of change in the surrogate end point can be used to predict the actual change in the true end point (e.g., change in LV volume and survival). Finally, this association needs to be replicated in a variety of different populations, in both observational studies and treatment trials, using a spectrum of therapeutic interventions. Although insistence on strict mathematical proof of these four criteria may make it impossible to validate any surrogate end point fully, they provide a set of useful benchmarks in assessing the practicality of using surrogate end points.

A more realistic expectation is that the surrogate end point accounts for a substantial portion of the treatment effect on the true clinical end point. A physiologic variable can be proposed as a surrogate end point when the variable meets two basic criteria (3). First, a statistical relationship must exist between the change in the proposed surrogate end point over time and the clinical outcome. Second, a pathophysiologic basis must exist for believing that the change in the surrogate end point is the primary determinant of the outcome.

Thus, surrogate end points in HF patients should unequivocally reflect the true end points (i.e., survival and the QOL). More than 150 clinical, hemodynamic, or exercise variables have been identified as predictors of survival in patients with HF. Some have been tested in clinical trials, but not completely validated. We will focus on hemodynamic measurements, ventricular arrhythmias, autonomic nervous system markers, exercise capacity, neurohormones and variables of LV structure and function as potential surrogate end points for patients with chronic HF.

Hemodynamic measurements. During the 1980s, HF was considered primarily a hemodynamic disorder, and physicians believed that therapeutic interventions that improved pump function would predictably benefit patients. Invasive hemodynamic studies to assess cardiac output and right ventricular (RV) and LV filling pressures were viewed as crucial in development programs for new drugs. Later studies, however, have raised important concerns about the validity of acute hemodynamic changes as surrogate end points. A number of controlled clinical trials conducted in the 1990s (9,10,13,21–23) have shown that drugs that produce striking hemodynamic benefits do not necessarily produce long-term clinical benefits. These findings have discouraged the use of hemodynamic variables as surrogate markers for drug efficacy. However, the converse is not true. All the drugs approved for treatment of HF have long-term beneficial hemodynamic effects, and there are no drugs that worsen hemodynamics and improve long-term outcomes. Hemodynamic variables are still viewed as the main variables for assessing the efficacy of intravenous (IV) drugs for the treatment of acute HF (3).

Ventricular arrhythmias. Ventricular premature depolarizations, a risk factor for sudden and non-sudden cardiac death after myocardial infarction (MI) (24), have been associated with a lower ejection fraction (EF) and larger heart size in patients with chronic HF (25). Drugs that suppress ventricular premature beats were therefore expected to prolong life (26). Several clinical trials, however, have failed to show that drugs that reduce the frequency and complexity of disturbances in rhythm decrease the long-term risk of sudden death (27). These results demonstrate that ventricular premature beats cannot be considered valid surrogate markers, though they are clearly an important risk factor.

Autonomic nervous system markers. Recent compelling evidence linking the autonomic nervous system and cardiac mortality, including sudden death (28–31), suggests that parameters such as heart rate variability, baroreceptor sensitivity and ventricular repolarization characteristics (QT dispersion) may well serve as potential surrogate markers. Yee and Struthers (32) analyzed the impact of drugs on mortality and autonomic nervous system surrogate markers and found heart rate variability and QT dispersion to be the most promising surrogates for sudden death in HF. Nevertheless, the predictive value of many of the autonomic markers is at present uncertain. It is important to reiterate that a strong predictive marker may not result in the expected-as-modeled change in outcome when using a drug that also changes the predictor.

Exercise capacity. Traditionally, functional capacity in HF is classified according to a categoric scale, such as that of the New York Heart Association (NYHA), and the efficacy of a therapeutic intervention is judged using each patient as his or her own control. Although it is simple, such a classifica-

tion is not quantifiable, is subject to considerable interobserver variability and lacks adequate sensitivity to detect important changes in functional capacity. Given these difficulties, various quantitative and objective measurements of functional capacity have been developed in recent years.

The *6-min walk test* (33) was found to predict long-term mortality and HF hospitalization rates in patients with LV dysfunction of varying cause and severity (34,35). The test can be administered safely in an outpatient setting without specialized equipment and is well accepted by patients. However, the test is subjective, and its usefulness in serially monitoring therapy and progression of LV dysfunction has not been fully explored. Some recent data have not confirmed the predictive value of the distance walked on survival (36), especially in patients with mild HF and preserved exercise tolerance (37). The 6-min walk distance is therefore considered helpful in clinical descriptions of HF patients but cannot be used as a surrogate marker for assessing the survival prognosis.

Treadmill or *cycle exercise testing* has generally shown that therapeutic interventions that lessen symptoms in HF patients also improve exercise tolerance and, conversely, that symptomatically ineffective drugs produce little change in exercise capacity. Exercise tolerance, expressed as exercise time or workload achieved on an ergometer, has been recognized for several decades as an important prognostic marker in patients with heart disease (38). In recent years, there has been increased interest in directly measured maximal oxygen uptake (peak V_{O_2}) during exercise. Peak V_{O_2} is considered the best criteria of exercise capacity in patients with chronic HF (39–41). An objective measure of maximal exertion, peak V_{O_2} is used as an independent prognostic indicator in HF (25,42) and in the selection of patients for cardiac transplantation (43,44).

Some HF trials have shown that therapeutic interventions that increase exercise capacity also improve survival (45). However, an improvement in survival has not been demonstrated with each therapeutic agent that improves effort tolerance. Results from the Prospective Randomized Milrinone Survival Evaluation (PROMISE) (21) and Randomized Evaluation of FLosequinan on ExerCise Tolerance (REFLECT) (46) trials have shown that early treatment-induced improvements in exercise tolerance were unreliable predictors of actual treatment effects on survival. In light of the above, it might be useful to distinguish between the short-term benefit on exercise capacity shown in PROMISE (21) and REFLECT (46) trials and long-term effects shown in the Vasodilator-Heart Failure Trial (V-HeFT) studies (45).

Neurohormones. There is convincing evidence that neurohormonal systems play major roles in the pathogenesis and progression of HF (47). Two sets of neurohormones with opposing effects are activated in the syndrome of HF. The vasoconstrictor hormones are anti-natriuretic, antidiuretic and generally have growth-promoting properties, whereas the vasodilator hormones are natriuretic, diuretic and have anti-mitogenic effects. Norepinephrine (NE) and the natriuretic peptides are the most studied neurohormones in HF, and the strongest evidence for their pathogenetic role comes from studies showing that modulation of these neurohormones is associated with changes in clinical course and survival.

Measurements of plasma NE were performed in the Vasodilator-Heart Failure Trial II (V-HeFT II) to examine the effects of therapy on neuroendocrine activation and the responses to therapy among patients with different degrees of activation. The baseline plasma NE data were grouped into three relatively homogeneous strata: plasma NE $<$600 pg/ml, 600 to 900 pg/ml, and $>$900 pg/ml (48). Cumulative mortality was found to differ significantly between strata: NE values $<$600 pg/ml were associated with the lowest risk, values between 600 and 900 pg/ml were associated with an intermediate risk, and values $>$900 pg/ml identified a group at exceedingly high risk. The group treated with enalapril had a significantly lower mortality than the group treated with hydralazine-isosorbide dinitrate, and this benefit was most evident in patients with NE values $>$900 pg/ml (48). Similarly, in the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS), significant reduction in mortality seen with enalapril was confined to patients with baseline NE levels above the median (49).

Other studies have raised important concerns about the validity of plasma NE as a surrogate marker in HF treatment trials. In the Australia–New Zealand Carvedilol Heart Failure Trial, high baseline NE levels did not predict additional survival benefit with carvedilol, which significantly reduced HF admissions only in those patients with NE levels below the median (50). The most worrisome examples of disagreement between survival data and plasma NE values come from studies with ibopamine (51,52) and moxonidine (53). The PRIME II (Second Prospective Randomized Study of Ibopamine on Mortality and Efficacy) (52) and MOXCON (Effect of Sustained Release Moxonidine on Mortality and Morbidity in Patients with Congestive Heart Failure) (53) trials were terminated prematurely because of the adverse effects of ibopamine and moxonidine on mortality despite significant reductions in plasma NE. For the association of change in NE over time and prognosis, a nonlinear effect was observed in the recently completed BEST study (Beta-blocker Evaluation of Survival Trial) (54). In this study, compared with the intermediate or no-change group, those with a large decrease or a large increase in NE had a higher mortality (54). Moreover, the largest decrease in NE was seen in the African American population, who had a higher mortality. These results limit the use of plasma NE as a surrogate marker for HF trials.

Plasma natriuretic peptides are being recognized as important prognostic markers in patients with HF (49,55). Brain natriuretic peptide (BNP), particularly its aminoterminal portion (N-BNP), appears to be one of the most powerful neurohormonal predictors of LV function and

prognosis in chronic HF (56–59). It has also emerged as an important diagnostic marker (60,61). In the Australia–New Zealand Carvedilol Heart Failure Trial, carvedilol reduced mortality rates and HF admissions in those patients with higher baseline BNP levels but lower plasma NE (50). Moreover, N-BNP–guided treatment of HF was associated with a significant reduction of total cardiovascular events, including cardiovascular death and delay to first cardiovascular event, compared with clinically guided treatment (62).

In a recent study on patients admitted with decompensated HF, Cheng et al. (63) found that patients who had good outcomes were characterized by decreases in both their NYHA functional class and BNP levels during hospitalization, whereas patients who were re-admitted within 30 days of discharge had only minimal decreases in their BNP levels during hospitalization, despite improvement in NYHA classification. Finally, subjects who died in the hospital had increasing BNP levels and little change in symptoms. These data demonstrate a consistent relationship between the BNP as a surrogate marker and outcomes (death and change in symptom status). Changes in plasma BNP level were associated with changes in the morbid events, suggesting that plasma BNP may serve as a surrogate marker for HF trials.

LV dimensions and ejection fraction. Baseline LV dimensions and ejection fraction (EF) have been shown to be one of the most powerful predictors of survival after acute MI (64–66), in chronic HF (25,67–69) and among people free of overt cardiovascular disease (70–72). Many of the beneficial effects of angiotensin-converting enzyme (ACE) inhibitors and beta-blockers in HF appear to be related to the ability of these agents to inhibit or to reverse cardiac remodeling (73–78). In the Survival And Ventricular Enlargement (SAVE) trial, treatment with captopril was associated with attenuation of LV enlargement over time and improved clinical outcome (73,74). Serial assessments of LV dimensions and function in the SOLVD, Australia–New Zealand Carvedilol Heart Failure and CAPRICORN echocardiographic substudies, and in the MERIT-HF magnetic resonance imaging substudy, have also found an association between favorable changes in LV dimensions and EF and improved clinical outcome produced by enalapril, carvedilol or metoprolol (75–78). Treatment with xamoterol was shown to increase LV end-systolic and end-diastolic volumes (79) and to increase mortality (80).

Cintron et al. (81) assessed the prognostic value of serial changes in LVEF in patients with chronic HF and found a significant and proportionate relationship between the direction and magnitude of changes in EF over time and the one-year mortality. Changes in LVEF >5% from baseline at 6 months (V-HeFT I) and 12 months (V-HeFT II) were the strongest predictors of mortality and remained significant even after adjustment for therapy and baseline LVEF (81). This is perhaps the only study that addressed and found a consistent proportionality between the degree of change in the surrogate end point and the subsequent change in mortality. Thus, LV dimensions and their derivative EF seem to fulfill most of the criteria for surrogate end points: baseline LV dimensions and EF are significantly related to prognosis; changes in these measurements reflect changes in mortality; and both the direction and the magnitude of change in these variables cause a proportional change in survival.

STATISTICAL APPROACH TO TESTING THE VALIDITY OF USING A SURROGATE END POINT FOR FUTURE HF TRIALS

The relatively easy part of evaluating potential surrogates in HF trials is the evaluation of the predictive value for the end point, usually survival or symptomatic status. For survival modeling, the time to an end point is a well developed area of statistics. For most purposes, proportional hazards survival models are adequate. However, when symptoms are the object of investigation over an extended time period, the expectation of many deaths is often a complicating factor. An approach with much intuitive appeal is to consider time when a patient is dead to be the "worst case," and then to take a time average (over the maximum possible exposure time for a patient including the time after death as being in the worst case status). In some cases, alternatively, the time symptom-free may be used where time after death is considered time with symptoms. It is more difficult to model the effect of therapy for several reasons, for example: 1) most surrogates will change with time, so that to model the effect on survival a time-dependent approach will need to be taken; 2) if the risk does not depend upon only the current value, but also the history of the change, then a large number of potential models may be used, and the potential to very much over-fit the data is of concern; and 3) when a surrogate is measured only occasionally and death represents informative censoring the analysis is a very complex issue, and much effort must be spent on sensitivity analyses to see whether the purported effect holds up under scrutiny. At this time, statistical theory is relatively weak in this area.

To examine the effect of change over time, both the baseline and follow-up measurements would be used in the model, and there should be statistical significance associated with the follow-up measurement beyond any predictive power of the baseline variable. If a semi-parametric model (e.g., a proportional hazards or Cox model) or parametric model holds, the model should imply an increasing treatment effect as the surrogate change increases (at least over most of the range of observed changes).

Conclusions. In the design of clinical trials, choice of the most appropriate outcome measures, referred to as the "primary end points," is crucial. The most controversial aspect of this choice often relates to the reliability of using surrogate end points as true measures of clinical efficacy. However, the use of biological markers in the progression of HF has important potential. Use of such markers as surrogate end points allows all patients in the trial to contribute

to the study end point, resulting in smaller sample size, shorter trial duration and reduced cost. Patients with even mild disease could be studied, and preventive strategies could therefore be evaluated. Moreover, surrogate end points could allow investigators to make more rapid evaluations of promising interventions than could be made using primary end points. More importantly, surrogate end points can aid our understanding of the biologic processes underlying the disease and the mechanisms of therapy. Nevertheless, the reliability of efficacy assessments based on surrogate end points can be seriously challenged, and we should be suspicious of calculated sample sizes that appear too small or follow-up periods that appear too short. The issue, therefore, is not so much the sample sizes for the surrogate points, but whether one really is ready to accept the surrogate as being adequately important.

At present, the one perfect surrogate marker for mortality and QOL in assessing patients with HF remains elusive. Chronic HF is a complex syndrome; to expect any single parameter to be universally predictive of drug effects on mortality and QOL may be too simplistic. A composite of surrogate end points may be more appropriate than any single one. The most promising surrogate end points seem to be exercise capacity assessed by oxygen consumption, plasma natriuretic peptides, and measurements of LV dimensions and EF. It is therefore important to carry out mortality trials with surrogate markers incorporated into the study. The Valsartan Heart Failure Trial (Val-HeFT), for example, evaluated the effect of valsartan on HF morbidity and mortality and measured several biological markers, such as LV dimensions, EF, neurohormones and QOL in all patients, and Holter monitoring, exercise tolerance and cardiac remodeling assessed by magnetic resonance imaging in a subset of patients (11,82). The results of some of these substudies and similarly designed trials in the future should help establish the validity of these markers as surrogates for clinical events.

In this review we have mentioned some statistical methods useful in the evaluation of potential surrogates, but both the biological knowledge and the understanding are paramount, and no amount of statistical methodology can ensure the appropriateness of a putative surrogate, although it may rule out candidates that do not perform well.

The availability of the newer drugs for HF makes it even more important to have valid surrogate end points. As more and more drugs are added to the current standard regimen, a time will come when physicians will have to consider withdrawing or replacing one of the standard drugs, because not all drugs have similar beneficial effects on all patients. At present there is no way of knowing whether an agent is having a beneficial effect on a particular patient. Availability of valid surrogate end points would help solve the ethical dilemma concerning withdrawing a standard drug in a patient who is still symptomatic on that drug. A less significant p value for the difference in a true end point may be adequate for therapeutic approval when robust favorable data on established surrogate markers are available to support the result.

Reprint requests and correspondence: Professor Inder S. Anand, Director of Heart Failure Program, Veterans Affairs Medical Center, One Veterans Drive 111-C, Minneapolis, Minnesota 55417. E-mail: anand001@tc.umn.edu.

REFERENCES

1. Krumholz HM, Chen YT, Wang Y, Vaccarino V, Radford MJ, Horwitz RI. Predictors of readmission among elderly survivors of admission with heart failure. Am Heart J 2000;139:72–7.
2. Vinson JM, Rich MW, Sperry JC, Shah AS, McNamara T. Early readmission of elderly patients with congestive heart failure. J Am Geriatr Soc 1990;38:1290–5.
3. Lipicky RJ, Packer M. Role of surrogate end points in the evaluation of drugs for heart failure. J Am Coll Cardiol 1993;22:179A–84.
4. Sobel BE, Furberg CD. Surrogates, semantics, and sensible public policy. Circulation 1997;95:1661–3.
5. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996;125:605–13.
6. Rector TS, Kubo SH, Cohn JN. Patients' self-assessment of their congestive heart failure: content, reliability and validity of a new measure, the Minnesota Living with Heart Failure questionnaire. Heart Fail 1987;3:198–209.
7. Rector TS, Kubo SH, Cohn JN. Validity of the Minnesota Living with Heart Failure questionnaire as a measure of therapeutic response to enalapril or placebo. Am J Cardiol 1993;71:1106–7.
8. Massie BM, Berk MR, Brozena SC, et al. Can further benefit be achieved by adding flosequinan to patients with congestive heart failure who remain symptomatic on diuretic, digoxin, and an angiotensin-converting enzyme inhibitor? Results of the Flosequinan–ACE inhibitor Trial (FACET). Circulation 1993;88:492–501.
9. Kubo SH, Gollub S, Bourge R, et al. Beneficial effects of pimobendan on exercise tolerance and quality of life in patients with heart failure. Results of a multicenter trial. The Pimobendan Multicenter Research Group. Circulation 1992;85:942–9.
10. Cohn JN, Goldstein SO, Greenberg BH, et al. A dose-dependent increase in mortality with vesnarinone among patients with severe heart failure. N Engl J Med 1998;339:1810–6.
11. Cohn JN, Tognoni G. A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. N Engl J Med 2001;345:1667–75.
12. Rector TS, Cohn JN. Assessment of patient outcome with the Minnesota Living with Heart Failure questionnaire: reliability and validity during a randomized, double-blind, placebo-controlled trial of pimobendan. Pimobendan Multicenter Research Group. Am Heart J 1992;124:1017–25.
13. Packer M, Rouleau J, Swedberg K. Effect of flosequinan on survival in chronic heart failure: preliminary results of the PROFILE study (abstr). Circulation 1993;88 Suppl 1:301.
14. Rector TS, Tschumperlin LK, Kubo SH, et al. Use of the Living With Heart Failure questionnaire to ascertain patients' perspectives on improvement in quality of life versus risk of drug-induced death. J Card Fail 1995;1:201–6.
15. Patrick DL, Erickson P. Concepts of Health-Related Quality of Life. Health Status and Health Policy. Quality of Life in Health Care Evaluation and Resource Allocation. Oxford: Oxford University Press, 1993;76:112.
16. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989;8:431–40.
17. Temple RJ. A regulatory authority's opinion about surrogate endpoints. In: Nimmo WS, Tucker GT, editors. Clinical Measurement in Drug Evaluation. New York, NY: Wiley, 1995:322.
18. Wittes J, Lakatos E, Probstfield J. Surrogate endpoints in clinical trials: cardiovascular diseases. Stat Med 1989;8:415–25.
19. Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. Stat Med 1992;11:167–78.

20. Fleming TR, Prentice RL, Pepe MS, Glidden D. Surrogate and auxiliary endpoints in clinical trials, with potential applications in cancer and AIDS research. Stat Med 1994;13:955–68.
21. Packer M, Carver JR, Rodeheffer RJ, et al. Effect of oral milrinone on mortality in severe chronic heart failure. The PROMISE Study. N Engl J Med 1991;325:1468–75.
22. Elkayam U, Amin J, Mehra A, Vasquez J, Weber L, Rahimtoola SH. A prospective, randomized, double-blind, crossover study to compare the efficacy and safety of chronic nifedipine therapy with that of isosorbide dinitrate and their combination in the treatment of chronic congestive heart failure. Circulation 1990;182:1954–61.
23. Califf RM, Adams KF, McKenna WJ, et al. A randomized controlled trial of epoprostenol therapy for severe congestive heart failure: the Flolan International Randomized Survival Trial (FIRST). Am Heart J 1997;134:44–54.
24. Ruberman W, Weinblatt E, Goldberg JD, Frank CW, Shapiro S. Ventricular premature beats and mortality after myocardial infarction. N Engl J Med 1977;297:750–7.
25. Cohn JN, Johnson GR, Shabetai R, et al. Ejection fraction, peak exercise oxygen consumption, cardiothoracic ratio, ventricular arrhythmias, and plasma norepinephrine as determinants of prognosis in heart failure. Circulation 1993;87:VI5–6.
26. Graboys TB, Lown B, Podrid PJ, DeSilva R. Long-term survival of patients with malignant ventricular arrhythmia treated with antiarrhythmic drugs. Am J Cardiol 1982;50:437–43.
27. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med 1991;324:781–8.
28. Barr CS, Naas A, Freeman M, Lang CC, Struthers AD. QT dispersion and sudden unexpected death in chronic heart failure. Lancet 1994;343:327–9.
29. Barron HV, Lesh MD. Autonomic nervous system and sudden cardiac death. J Am Coll Cardiol 1996;27:1053–60.
30. Lown B, Verrier RL. Neural activity and ventricular fibrillation. N Engl J Med 1976;294:1165–70.
31. Schwartz PJ, Vanoli E, Stramba-Badiale M, De Ferrari GM, Billman GE , Foreman RD. Autonomic mechanisms and sudden death. New insights from analysis of baroreceptor reflexes in conscious dogs with and without a myocardial infarction. Circulation 1988;78:969–79.
32. Yee KM, Struthers AD. Can drug effects on mortality in heart failure be predicted by any surrogate measure? Eur Heart J 1997;18:1860–4.
33. Guyatt GH, Sullivan MJ, Thompson PJ, et al. The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. Can Med Assoc J 1985;132:919–23.
34. Bittner V, Weiner DH, Yusuf S, et al. Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. JAMA 1993;270:1702–7.
35. Cahalin LP, Mathier MA, Semigran MJ, Dec GW, DiSalvo TG. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. Chest 1996;110:325–32.
36. Lucas C, Stevenson LW, Johnson W, et al. The 6-min walk and peak oxygen consumption in advanced heart failure: aerobic capacity and survival. Am Heart J 1999;138:618–24.
37. Roul G, Germain P, Bareiss P. Does the 6-minute walk test predict the prognosis in patients with NYHA class II or III chronic heart failure? Am Heart J 1998;136:449–57.
38. Morris CK, Ueshima K, Kawaguchi T, Hideg A, Froelicher VF. The prognostic value of exercise capacity: a review of the literature. Am Heart J 1991;122:1423–31.
39. Weber KT, Kinasewitz GT, Janicki JS, Fishman AP. Oxygen utilization and ventilation during exercise in patients with chronic cardiac failure. Circulation 198;265:1213–23.
40. Lipkin DP, Perrins J, Poole-Wilson PA. Respiratory gas exchange in the assessment of patients with impaired ventricular function. Br Heart J 1985;54:321–8.
41. Franciosa JA, Ziesche S, Wilen M. Functional capacity of patients with chronic left ventricular failure. Relationship of bicycle exercise performance to clinical and hemodynamic characterization. Am J Med 1979;67:460–6.
42. Florea VG, Henein MY, Anker SD, et al. Prognostic value of changes over time in exercise capacity and echocardiographic measurements in patients with chronic heart failure. Eur Heart J 2000;21:146–53.
43. Mudge GH, Goldstein S, Addonizio LJ, et al. 24th Bethesda conference: Cardiac transplantation. Task Force 3: Recipient guidelines/prioritization. J Am Coll Cardiol 1993;22:21–31.
44. Costanzo MR, Augustine S, Bourge R, et al. Selection and treatment of candidates for heart transplantation. A statement for health professionals from the Committee on Heart Failure and Cardiac Transplantation of the Council on Clinical Cardiology, American Heart Association. Circulation 1995;92:3593–612.
45. Ziesche S, Cobb FR, Cohn JN, Johnson G, Tristani F. Hydralazine and isosorbide dinitrate combination improves exercise tolerance in heart failure. Results from V-HeFT I and V-HeFT II. The V-HeFT VA Cooperative Studies Group. Circulation 1993;87:VI56–V164.
46. Packer M, Narahara KA, Elkayam U, et al. Double-blind placebo-controlled study of the efficacy of flosequinan in patients with chronic heart failure. REFLECT study. J Am Coll Cardiol 1993;22:65–72.
47. Packer M. The neurohormonal hypothesis: a theory to explain the mechanism of disease progression in heart failure. J Am Coll Cardiol 1992;20:248–54.
48. Francis GS, Cohn JN, Johnson G, Rector TS, Goldman S, Simon A. Plasma norepinephrine, plasma renin activity, and congestive heart failure. Relations to survival and the effects of therapy in V-HeFT II. The V-HeFT VA Cooperative Studies Group. Circulation 1993;87: VI40–8.
49. Swedberg K, Eneroth P, Kjekshus J, Wilhelmsen L. Hormones regulating cardiovascular function in patients with severe congestive heart failure and their relation to mortality. CONSENSUS Trial Study Group. Circulation 1990;82:1730–6.
50. Richards AM, Doughty R, Nicholls MG, et al. Neurohumoral prediction of benefit from carvedilol in ischemic left ventricular dysfunction. Australia–New Zealand Heart Failure Group. Circulation 1999;99:786–92.
51. Rousseau MF, Konstam MA, Benedict CR, et al. Progression of left ventricular dysfunction secondary to coronary artery disease, sustained neurohormonal activation and effects of ibopamine therapy during long-term therapy with angiotensin-converting enzyme inhibitor. Am J Cardiol 1994;73:488–93.
52. Hampton JR, van Veldhuisen DJ, Kleber FX, et al. Randomised study of effect of ibopamine on survival in patients with advanced severe heart failure. Second Prospective Randomised Study of Ibopamine on Mortality and Efficacy (PRIME II) investigators. Lancet 1997;349: 971–7.
53. Cohn J. Heart Failure—MOXCON (Moxonidine for Congestive Heart Failure). Circulation 1999;100:e128.
54. Bristow MR, Zelis R. Baseline and three-month change in systemic venous norepinephrine as predictors of clinical outcomes in the BEST trial investigators (abstr). J Am Coll Cardiol 2001;37:183A.
55. Gottlieb SS, Kukin ML, Ahern D, Packer M. Prognostic importance of atrial natriuretic peptide in patients with chronic heart failure. J Am Coll Cardiol 1989;13:1534–9.
56. Hunt PJ, Espiner EA, Nicholls MG, Richards AM, Yandle TG. The role of the circulation in processing pro-brain natriuretic peptide (proBNP) to amino-terminal BNP and BNP-32. Peptides 1997;18: 1475–81.
57. Richards AM, Nicholls MG, Yandle TG, et al. Plasma N-terminal pro-brain natriuretic peptide and adrenomedullin: new neurohormonal predictors of left ventricular function and prognosis after myocardial infarction. Circulation 1997;97:1921–9.
58. Richards AM, Nicholls MG, Yandle TG, et al. Neuroendocrine prediction of left ventricular function and heart failure after acute myocardial infarction. The Christchurch Cardioendocrine Research Group. Heart 1999;81:114–20.
59. Yasue H, Yoshimura M, Sumida H, et al. Localization and mechanism of secretion of B-type natriuretic peptide in comparison with those of A-type natriuretic peptide in normal subjects and patients with heart failure. Circulation 1994;90:195–203.
60. Maisel A. B-type natriuretic peptide in the diagnosis and management of congestive heart failure. Cardiol Clin 2001;19:557–71.
61. Maisel A. B-type natriuretic peptide levels: a potential novel "white count" for congestive heart failure. J Card Fail 2001;7:183–93.
62. Troughton RW, Frampton CM, Yandle TG, Espiner EA, Nicholls MG, Richards AM. Treatment of heart failure guided by plasma aminoterminal brain natriuretic peptide (N-BNP) concentrations. Lancet 2000;355:1126–30.

63. Cheng V, Kazanagra R, Garcia A, et al. A rapid bedside test for B-type peptide predicts treatment outcomes in patients admitted for decompensated heart failure: a pilot study. J Am Coll Cardiol 2001;37:368–91.
64. Hammermeister KE, DeRouen TA, Dodge HT. Variables predictive of survival in patients with coronary disease. Selection by univariate and multivariate analyses from the clinical, electrocardiographic, exercise, arteriographic, and quantitative angiographic evaluations. Circulation 1979;59:421–30.
65. Pfeffer MA, Pfeffer JM. Ventricular enlargement and reduced survival after myocardial infarction. Circulation 1987;75:IV93–7.
66. White HD, Norris RM, Brown MA, Brandt PW, Whitlock RM, Wild CJ. Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. Circulation 1987;76:44–51.
67. Unverferth DV, Magorien RD, Moeschberger ML, Baker PB, Fetters JK, Leier CV. Factors influencing the one-year mortality of dilated cardiomyopathy. Am J Cardiol 1984;54:147–52.
68. Lee TH, Hamilton MA, Stevenson LW, et al. Impact of left ventricular cavity size on survival in advanced heart failure. Am J Cardiol 1993;72:672–6.
69. Wong M, Johnson G, Shabetai R, et al. Echocardiographic variables as prognostic indicators and therapeutic monitors in chronic congestive heart failure. Veterans Affairs cooperative studies V-HeFT I and II. Circulation 1993;87:VI65–70.
70. Lauer MS, Evans JC, Levy D. Prognostic implications of subclinical left ventricular dilatation and systolic dysfunction in men free of overt cardiovascular disease (the Framingham Heart Study). Am J Cardiol 1992;70:1180–4.
71. Sandvik L, Erikssen J, Thaulow E, Erikssen G, Mundal R, Aakhus T. Heart volume and cardiovascular mortality. A 16-year follow-up study of 1984 healthy middle-aged men. Eur Heart J 1993;14:592–6.
72. Vasan RS, Larson MG, Benjamin EJ, Evans JC, Levy D. Left ventricular dilatation and the risk of congestive heart failure in people without myocardial infarction. N Engl J Med 1997;336:1350–5.
73. St. John Sutton M, Pfeffer MA, Plappert T, et al. Quantitative two-dimensional echocardiographic measurements are major predictors of adverse cardiovascular events after acute myocardial infarction. The protective effects of captopril. Circulation 1994;89:68–75.
74. St. John Sutton M, Pfeffer MA, Moye L, et al. Cardiovascular death and left ventricular remodeling two years after myocardial infarction: baseline predictors and impact of long-term use of captopril: information from the Survival And Ventricular Enlargement (SAVE) trial. Circulation 1997;96:3294–9.
75. Greenberg B, Quinones MA, Koilpillai C, et al. Effects of long-term enalapril therapy on cardiac structure and function in patients with left ventricular dysfunction. Results of the SOLVD echocardiography substudy. Circulation 1995;91:2573–81.
76. Doughty RN, Whalley GA, Gamble G, MacMahon S, Sharpe N. Left ventricular remodeling with carvedilol in patients with congestive heart failure due to ischemic heart disease. Australia–New Zealand Heart Failure Research Collaborative Group. J Am Coll Cardiol 1997;29:1060–6.
77. Groenning BA, Nilsson JC, Sondergaard L, Fritz-Hansen T, Larsson HB, Hildebrandt PR. Antiremodeling effects on the left ventricle during beta-blockade with metoprolol in the treatment of chronic heart failure. J Am Coll Cardiol 2000;36:2072–80.
78. Doughty RN, Whalley GA, Walsh H, Gamble G, Sharpe N. Effects of carvedilol on left ventricular remodeling in patients following acute myocardial infarction: the CAPRICORN echo substudy. Circulation 2001;104 Suppl:II-517.
79. Darasz KH, Bayliss J, Underwood SR, Keegan J, Poole-Wilson PA, Sutton GC. Left ventricular volume in thrombolysed patients with acute anterior myocardial infarction: the effect of captopril and xamoterol. Int J Cardiol 1995;51:137–42.
80. The Xamoterol in Severe Heart Failure study group. Xamoterol in severe heart failure. Lancet 1990;336:1–6.
81. Cintron G, Johnson G, Francis G, Cobb F, Cohn JN. Prognostic significance of serial changes in left ventricular ejection fraction in patients with congestive heart failure. Circulation 1993;87:VI17–23.
82. Cohn JN, Tognoni G, Glazer RD, Spormann D, Hester A. Rationale and design of the Valsartan Heart Failure Trial: a large multinational trial to assess the effects of valsartan, an angiotensin-receptor blocker, on morbidity and mortality in chronic congestive heart failure. J Card Fail 1999;5:155–60.