Chapter 2. Methodology

Scope of Work

AHRQ described the scope of work as a quantitative analysis and evidence report on the effectiveness of treatment of HF using ACE inhibitors and beta-blockers. The project had five key steps:

1. Identify technical experts to provide input and advice to the project.

2. Refine the research questions.

3. Perform a literature search and evaluation.

4. Systematically synthesize the evidence.

5. Produce and disseminate and evidence report.

Original Potential Key Questions

The American College of Physicians, the American Society of Internal Medicine, and the American Academy of Family Physicians nominated this topic. They submitted the following potential key questions to AHRQ.

1. What evidence exists on the effectiveness of nurse management programs? Health food supplements?

2. What evidence exists on the treatment of sleep apnea in patients with HF?

3. What is the evidence on the treatment of specific myocardial disorders; e.g., myocarditis, sarcoidosis, and amyloidosis, in patients with HF?

4. What interventions are effective for patients with diastolic dysfunction?

5. Which patients benefit from which beta-blockers?

6. What are the effects of potassium levels on HF outcomes?

7. Do angiotensin blockers improve outcomes?

8. What, if any, are the differences in treatment effectiveness associated with patient gender, race, age, and income level?

Technical Expert Panel

Project staff assembled a technical expert panel (TEP) that included leading cardiologists working in academic and nonacademic settings, researchers, clinicians, and health care managers. Panelists assisted the project with topic refinement, retrieval of unpublished data, and review of the final evidence report. The TEP members (and relevant affiliations) are listed here:

Michael Barrett American College of Physicians
Greg Fonarow UCLA Medical Center
Barry Greenberg UCSD Medical Center
Paul Heidenreich Palo Alto VA Hospital
Stanford-UCSF Evidence-based Practice Center
Tom Knabel UnitedHealthcare
Marvin Konstam New England Medical Center
Michael Rich Washington University of School of Medicine
Anthony Steimle Kaiser Permanente, Northern California
Lynne Warner Stevenson Brigham and Women's Hospital

After "congestive heart failure" was nominated as a topic, but prior to assignment of this contract to the Southern California Evidence-based Practice Center (SCEPC), the American Heart Association (AHA) and the American College of Cardiology (ACC) released practice guidelines on the management of HF. AHA/ACC graciously provided the SCEPC a draft copy for confidential review. On September 8, 2000, a conference call was held with our technical expert panel to limit the key questions to be addressed in the evidence report. The purpose of the conference call was to identify topic areas for this report that would complement but not duplicate the draft guidelines, a copy of which had been made available to each TEP member. The technical experts judged that several of the original key questions posed by the nominating organizations had been adequately answered in the AHA/ACC guidelines, major studies were underway that would answer several more of the questions, and published data would be insufficient to reach meaningful conclusions for still others. The technical experts identified three areas where they believed significant contribution could still be made:

- Assessment of the effects of age over 70, gender, race, and assisted living on treatment outcomes

- Cost-effectiveness of medication combinations

- Assessment of outcomes in patients with various comorbidities, particularly diabetes mellitus, renal dysfunction, and cognitive dysfunction.

Our TEP members determined that for clinical questions 1 and 3, only the results of placebo-controlled randomized trials (RCTs) of ACE inhibitors or beta-blockers that measured outcomes of interest to patients and policymakers (including mortality, utilization, and costs) would be accepted as evidence. The TEP judged that a formal explication of a causal pathway was not needed, because numerous randomized trials had already addressed the overarching clinical

questions of the effect of the drugs on mortality and utilization. As a starting point for our research, our experts provided us with references to eight pertinent studies and the names and acronyms of the major ACE inhibitor and beta-blocker trials.

Preliminary Search

In addition to the eight reports provided by the expert panel, we searched the following databases for articles on HF treatment for the specific populations under study.

Medline, produced by the U.S. National Library of Medicine, is widely recognized as the premier source for bibliographic coverage of biomedical literature. It encompasses information form Index Medicus, Index to Dental Literature, and the Cumulative Index to Nursing and Allied Health Literature as well as other sources of literature in the areas of allied health, biological and physical sciences, humanities and information science as they relate to medicine and health care.

Healthstar, produced by the American Hospital Association, contains over one million references covering topics in hospital administration, personnel, planning, budget, accreditation, and health care delivery.

EmBase, the Excerpta Media database produced by Elsevier Science, is a major biomedical and pharmaceutical database indexing over 3,800 international journals. EMBASE currently contains over six million records, with more than 400,000 citations and abstracts added annually.

Ageline covers subjects that include aging, gerontology, health sciences, psychology, and sociology. References date from1978 to the present.

SciSearch is a database that contains all records published in Science Citation Index and additional records from about 1,000 journals listed in Current Contents. Every subject area within the board fields of science, technology, and biomedicine is included.

The **Cochrane DARE** (Database of Abstracts and Reviews of Effectiveness) contains structured abstracts of systematic reviews that have been critically appraised by reviewers at the Centre for Reviews and Dissemination, York, UK.

The specific search strategies are listed in Table 1.

Paul Shekelle, MD, and Colonel Sid Atkinson, MD, reviewed the list of retrieved titles. Of the 1,647 titles retrieved, 315 articles were deemed relevant to our undertaking and were ordered. An additional 88 articles found through mining reference lists were also ordered. Literature was tracked using Pro-Cite and Access software.

Additional Sources of Evidence

The TEP made us aware that reports of several recent studies were in press and thus would not be found through a search. Prepublication copies were provided to us.

In hopes of obtaining data on all ACE inhibitors and beta-blockers approved for HF by the Food and Drug Administration (FDA), we requested filings for each of these agents through the Freedom of Information (FOI) act. Approved ACE inhibitors included captopril, enalapril, fosinopril, lisinopril, quinipril, ramipril, and trandolapril. Approved beta-blockers included bucindolol, bisoprolol, carvedilol, and metoprolol. As discussed in the Results section, we eventually obtained data from the FDA for two studies.

Another TEP conference call was held on April 4, 2001. During this phone call, we reviewed the preliminary results of our literature search. The TEP advised us to attempt to obtain subgroup

data on all RCTs that had at least 12 weeks of followup. Since most published studies did not report on our special populations of interest, project staff sent letters to original authors requesting subgroup data (see Appendix A for sample letter). Nonrespondents were sent a reminder letter on May 8, 2001. In addition, expert panel members agreed to call or email selected nonrespondents.

Our yield from this process was poor. After mailing 62 letters, we netted four agreements (all from studies with relatively small sample sizes), 12 new contacts, 10 refusals, 32 nonresponses, and four responses categorized as other.

Based on this poor response, we modified our plan to seek subgroup data more intensively from the biggest studies through personal contacts by TEP members to the authors of those studies and through attempts to obtain individual patient data on any study that had been submitted to the FDA as part of the regulatory process. Our rationale was that we had enough resources to attempt these intensive methods on only a select number of studies and the biggest studies would provide us the greatest statistical power. We calculated that the seven largest ACE inhibitor studies enrolled 14,932 patients, whereas the remaining 19 ACE inhibitor trials enrolled an aggregate of 3,033 patients. Similarly, the five largest beta-blocker studies enrolled 12,726 patients, whereas the remaining 19 beta-blocker studies enrolled 2,938 patients. Therefore, by targeting our efforts at the largest studies, we were able to make the most effective use of our resources. However, this strategy assumes that the large and small studies are measuring the same effect.

With the assistance of our TEP members, we succeeded in obtaining the individual patient level data for TRACE from the principal investigator, Dr. Torp-Pederson. With the help of the Task Order Officer, we negotiated a confidentiality agreement with the FDA that gave us access to data submitted to the FDA as part of the regulatory process. In discussions with FDA staff, it was clear that within the constraints of time and resources, we could assess only data that had been submitted to the FDA in electronic form. FDA staff identified two studies (MERIT-HF and COPERNICUS) that had electronic data. Our confidentiality agreement required us to examine these data onsite; therefore, our quantitative analyst spent two days at the FDA working with the original data to calculate the subgroup results needed for our pooled analyses. The outcomes of our efforts to obtain subgroup data and the sources of data used in our pooled analysis are shown in Tables 2 and 3, respectively.

During the data extraction phase, it became apparent that few studies reported the relevant data stratified by age or nursing home residence. In addition, health care outcomes and health outcomes other than mortality were reported variably in the studies, making pooling less justified. For these reasons, we further restricted key questions 1 and 3 to assess only data stratified by gender, race, and diagnosis of diabetes, and to use all-cause mortality as the sole outcome of interest.

Meta-Analysis

Our principal questions for meta-analysis, as determined by our TEP, were the following:

- What is the association between treatments (ACE inhibitors or beta-blockers) and all-cause mortality for female patients, male patients, patients with diabetes, patients without diabetes, black patients, and white patients with HF?

- Does this association vary (e.g. are there statistically significant differences) by gender (female versus male), diabetic condition (those with diabetes *versus* those without), and race (black versus white patients)?

Because individual studies did not enroll sufficient number of patients in the sub-populations of interest, meta-analysis is an appropriate technique to consider for these questions.

We first retrieved all articles that pertained to the eleven large placebo-controlled studies on ACE inhibitors and beta-blockers mentioned above. The SOLVD study consisted of two distinct trials on prevention and treatment respectively; thus, we considered a total of twelve studies. The same meta-analysis was done separately for the ACE inhibitor and beta-blocker sub-populations of studies, respectively.

Our outcome of interest was all-cause mortality. For studies for which both patient-level data and published statistics were available, we always chose the patient-level data over published statistics in the event of disagreement. Among the five studies for which we had patient-level data, three datasets disagreed with related publications. The differences were extremely small, never more than two patients in particular sub-populations; for example, the number of nondiabetic patients in the published article was two fewer than in the patient-level dataset. For the studies for which we had only published data, no two articles presented conflicting data about the same patient subgroup.

Relative Risks

All published reports that included the relevant patient subgroup data presented those data in the form of a two-by-two table of all-cause mortality by treatment or placebo group for each subgroup separately. If the patient-level data were available, we could construct this table directly. For example, the report of an ACE inhibitor study that stratified data by gender would provide the data in separate two-by-two tables, one for women and one for men. For each subgroup (e.g., women), we estimated the log mortality relative risk, which is equal to the log of the risk of dying for women who received ACE inhibitors divided by the risk of dying for women who received placebo. The extraction of data from patient-level datasets is described below.

The standard error for the log relative risk was also estimated,[5] and a 95% confidence interval was constructed. A similar log relative risk and confidence interval were calculated for men. We then back-transformed to the unlogged scale for interpretability so that our final statistic for each subgroup in each study was the relative risk with its associated confidence interval. The reason for conducting the estimation on the log scale is that the variance is more stable and the errors are more symmetric in this metric.

For subgroup comparisons for which we had data from more than two studies, we pooled the logs of the relative risks across studies using the DerSimonian and Laird random effects model,[6] and back-transformed the pooled estimate to the unlogged scale to produce a pooled relative risk (e.g., for women) across all relevant studies. We also constructed a 95% confidence interval and provide a p-value for the test of whether the pooled relative risk is different from 1. We tested for heterogeneity using a chi-squared test.[7] We note that in the case when sufficient heterogeneity across studies is not found, the DerSimonian and Laird estimate of the between-study variance is 0, and the random effects estimate is the same as a fixed effects estimate, the latter incorporating

only within-study variance. Significant heterogeneity was not observed for almost all our beta-blocker pooling situations, indicating that the studies were not heterogeneous, though we acknowledge that the chi-squared test of heterogeneity has low power to detect differences across studies, and the DerSimonian and Laird estimate is only a one-step iterative method. For ACE inhibitor studies, there was substantial heterogeneity, and the random effects analysis is designed to take this into account. This meta-analysis and the ones described below were conducted in the statistical package Stata using the "meta" and associated commands.[8] The analysis just described informed us about the association between various patient characteristics (such as gender) and mortality, when association is measured on the relative risk scale. Thus, this analysis answered our first question of interest.

To answer our second question, that is, whether the association differed between sub-populations (e.g., female versus male), we needed to test whether the relative risks of the two subgroups were statistically different. We did this by constructing a test statistic equal to the ratio of relative risks (RRR), which (for the example given) equals the female relative risk divided by the male relative risk. If this test statistic differs significantly from 1, then we infer that the two subgroup relative risks are significantly different. As before, we performed the analysis on the log scale. The log ratio of relative risks equals the log of the relative risk for women divided by the relative risk for men, and its standard error equals the square root of the sum of the variances of the two log relative risks. We constructed a confidence interval on the log scale. We then back-transformed the estimate and its confidence interval to the unlogged scale so that our final test statistic for each study was the RRR.

For subgroup comparisons for which we had data from more than two studies, we pooled the logs of the RRRs across studies using the DerSimonian and Laird random effects model.[6] We back-transformed the pooled result to the RRR scale for interpretation, and present the pooled ratio of relative risks, its 95% confidence interval, and a p-value for the test of whether the pooled RRR is different from 1.

We note that the ratio of the pooled relative risks may not exactly equal the pooled ratio of relative risks due to the nature of the weighting. The reason for pooling of the RRRs in order to compare the relative risks, rather than pooling the relative risks separately in each subgroup and then taking the ratio, is that comparison (i.e., taking the ratio) should be done separately within each study to control for study differences.

The directions (definitions of the numerator and denominator) of the RRRs were as follows. For the effect of gender, we compared outcomes for women (numerator) versus those for men (denominator). For the effect of diabetes we compared those who had diabetes with those who did not. For the effect of race, we compared black patients to white patients if the data were stratified appropriately. If not, we compared black patients to nonblack patients, or, if necessary, we compared nonwhite patients to white patients. We conducted a sensitivity analysis as described below to assess this hierarchical approach and to determine whether the inconsistency of race classification across studies affected our conclusions.

Hazard Ratios

Followup times for outcome assessment varied across studies, and the relative risk calculations do not take this variation, or the censoring of observations, into account. Thus, we also assessed the mortality associated with ACE inhibitors and beta-blockers on the hazard ratio

scale. The hazard ratio accounts for the variable contribution made to followup by patients who dropped out of the study for whatever reason. We followed the strategy for data extraction and pooling as described in Parmar, Torri, and Stewart.[9] The majority of the studies included in our analysis presented hazard ratios and confidence intervals, and after transforming these statistics to the log scale, we extracted the log hazard ratio and its standard error for each study.

For each patient subpopulation of interest, we estimated the log hazard ratio for each study that provided the data stratified on that dimension. We followed the same analytic strategy for the hazard ratio as for the relative risk, conducting a random-effects pooled analysis on the log scale and back-transforming to the unlogged scale. We then calculated a ratio of hazard ratios (RHR) to compare the hazard ratios in each subgroup.

Extraction of Data from Patient-Level Datasets

We obtained data directly from patient-level datasets for five studies: CONSENSUS, COPERNICUS, MERIT-HF, SOLVD, and TRACE. For CONSENSUS, SOLVD, and TRACE, the entire patient-level files were available to us directly, and we could conduct any analyses that we wished. As described above, we constructed two-by-two tables of mortality by treatment for each subgroup of interest to estimate a relative risk and constructed a Cox proportional hazard model in SAS[10] with treatment or control as the single covariate to estimate the hazard ratio for each patient subgroup of interest.

As previously mentioned, for the other two studies, COPERNICUS and MERIT-HF, we were able to analyze the patient-level data that the FDA provided. However, we were required to analyze the data at the FDA facility. The FDA allowed one of our statisticians to have access to the data at the FDA facility in Maryland. The analyst spent one day extracting and analyzing the data for both studies. The FDA provided our analyst with a computer workstation, and the data for both studies were in SAS format. The data for each study had a table of contents in a PDF file, which, along with the drug questionnaire, was used to locate the necessary variables. Once the data were compiled in a usable format for analysis, relative risks and hazard ratios were calculated for patient sub-populations. We were able to assess all-cause mortality separately from cardiac-cause mortality.

For COPERNICUS, the randomization group, gender, race, outcome status (dead or alive at the end of the trial), and time of death or dropout (i.e., censored) variables were each in separate files and had to be merged together by patient identification number. We defined the "diabetes" subgroup as any patients whose files were identified by searching the medical history text for the root "DIABET." Two subjects who were coded as "dead" but whose files did not show dates of death were dropped from the analysis.

For MERIT-HF, an analysis file with most of our variables of interest was already available. The number of days from enrollment until death or censoring had to be calculated using either the date of death or the date of last interview.

Publication Bias

We assessed the possibility of publication bias for the studies corresponding to each drug and patient comparison subgroup by graphically evaluating a funnel plot of the individual study log relative risk and hazard ratio for symmetry resulting from the nonpublication of small, negative

studies. Because graphical evaluation can be subjective, we also conducted an adjusted rank correlation test[11] and a regression asymmetry test[12] as formal statistical tests for publication bias. We found no evidence of publication bias in any of the study subpopulations assessed.

Sensitivity Analyses

As described above, studies varied in their definitions of racial groups. For the black patient versus white patient comparison, if the researchers reported data separately for blacks and whites, we utilized those data. If such data were not available, we used data reported for black versus nonblack patients, or, as a last resort, data comparing nonwhite with white patients. For those studies that provided the data for more than one of these comparisons, we compared the relative risk and hazard ratio statistics. The results of this sensitivity analysis did not differ markedly from the results of our primary hierarchical approach. We acknowledge that this sensitivity analysis cannot assess whether the potentially different race definitions (e.g., inclusion of Hispanic black patients in the Hispanic subgroup versus the black group) had an effect. However, the sensitivity analysis did permit us to evaluate some of the effects of different race definitions and stratifications across studies.

Cost-effectiveness Analysis

At the April 4, 2001, teleconference, Paul Heidenreich, MD, proposed to the TEP that based on his analysis of the data that were suitable for cost-effectiveness modeling, the most feasible cost-effectiveness analysis would be that of the use of ACE inhibitors for asymptomatic left ventricular systolic dysfunction, rather than an analysis of the cost-effectiveness of combinations of medications, as was originally proposed. This plan was accepted by the TEP and approved by the Task Order Officer. Later, based on the findings of this analysis, a further cost-effectiveness analysis that assessed screening for left ventricular dysfunction was proposed and approved by the Task Order Officer.

Assessing Treatment of Asymptomatic Left Ventricular Dysfunction

Decision Model

We developed a decision model using EXCEL (Version 5.0, Microsoft Corporation, Redmond, WA) and DATA (Version 3.0, TreeAge Software, Boston, MA) software. Using two treatment strategies, we modeled the lifetime health and economic outcomes for a hypothetical cohort of 55-year-old asymptomatic patients with ejection fraction of 35% or less but no history of HF (Figure 1). In the first strategy, asymptomatic patients are treated with ACE inhibitors. In the second strategy, patients are not treated with ACE inhibitors until they develop HF.

Each time period (month), patients with no history of HF can remain asymptomatic, develop HF, or die. Of those patients who developed HF, we assumed 33% would be hospitalized during their initial episode.[3] Once patients develop HF, they can remain in stable HF, be hospitalized, or die during each time period. The model follows patients until each has died (or to age 120).

Health Outcomes

Published data from the SOLVD prevention trial were used to calculate rates for the development of HF and death for asymptomatic patients with and without ACE inhibitor treatment.[3] We used actual event rates during the four years of reported followup. To model outcome after four years, we used an average of the yearly event rates weighted by the number of subjects still enrolled during each year of followup. Using this method, we estimated that the yearly rate of progression to symptomatic HF would be 6.5% for patients treated with ACE inhibitors and 9.8% for those not treated. We used a similar method to determine the yearly relative risk of death (compared to the general population) for patients with asymptomatic left ventricular dysfunction who are treated (2.9) and those not treated (3.3) with ACE inhibitors.

We used data from the SOLVD treatment trial to estimate hospitalization and death rates for patients with HF treated with ACE inhibitors.[3] The data consisted of actual event rates during the four years of reported followup for the SOLVD treatment trial. To model outcome following four years of living with HF, we used an average of the annual event rates weighted by the number of subjects participating during each year of the trial. This method estimated that the yearly relative risk of death (compared to the general population) for patients with symptomatic left ventricular dysfunction was 6.5 when treated with ACE inhibitors.

To determine quality-adjusted survival, we assigned a utility value of 0.71 to each year of life for patients living with HF, based on prior studies using the time-tradeoff utility of patient preferences in HF.[13] Patients with asymptomatic left ventricular dysfunction were assumed to have a utility value of 0.87.[13] We varied these quality-of-life assumptions in sensitivity analysis (range 0.5 to 1).

Costs

We achieved a health care system perspective by using all direct costs of medical care (Table 4) including medical costs incurred due to increased survival. Because HF survivors will incur additional costs for care not associated with their HF diagnosis, we assigned all patients a yearly cost of medical care based on age-adjusted medical expenditures for residents of the United States.[14] In addition, we included the costs of hospitalization for HF, ACE inhibitor treatment, and other outpatient HF care. We adjusted all costs to 2001 dollars using the medical component of the Consumer Price Index.[15] We determined costs for hospitalization using Medicare reimbursement for DRG 127, costs for ACE inhibitor treatment using average wholesale price,[16] and outpatient HF care using prior published estimates updated to year 2001.[17] Costs and benefits were discounted at 3% per year.[18]

Assessing Screening for Reduced Left Ventricular Ejection Fraction

Screening Strategies

We modeled the expected costs of six screening strategies (Figure 2):

1. Echocardiography for all patients. Patients with an ejection fraction less than 35% are treated (ACE inhibitors) to prevent development of HF.

2. Electrocardiogram (ECG) first, and if abnormal, echocardiography.

3. Blood test for B-type Natriuretic Peptide (BNP) first and, if abnormal, echocardiography.

4. ECG only, with treatment based on the results.

5. BNP only, with treatment based on the results.

6. No screening for depressed left ventricular function.

Each test can provide one of four possible results (true positive, false positive, true negative, false negative). Only persons who are true or false positives are referred for treatment. True-positive patients have a higher quality-adjusted survival than false negatives, who are treated only when HF develops. True-negative patients have a normal age-specific life expectancy. False-positive patients receive a small decrement in quality-adjusted survival to account for potential side effects of treatment.

Decision Model

A decision model was developed using EXCEL (Version 5.0, Microsoft Corporation, Redmond, WA) and DATA (Version 3.0, TreeAge Software, Boston, MA) software. We obtained the lifetime health and economic outcomes for hypothetical cohorts of 55-year-old patients with (1) depressed ejection fraction (35% or less) but no history of HF treated with ACE inhibitors, (2) depressed ejection fraction but no history of HF and no treatment until HF developed, and (3) patients with heart failure but without depressed ejection fraction.

During each time period (month), patients with a low ejection fraction and without a history of HF can remain asymptomatic, develop HF, or die. Of those patients who developed HF, we assumed 33% would be hospitalized during their initial episode.[3] Once patients develop HF, they can remain in stable HF, be hospitalized, or die during each time period. The model follows each patient until death (or until age 120). Patients without depressed ejection fraction are assumed to have an average age-specific mortality based on U.S. life table data.[19]

Test Characteristics

The sensitivity and specificity of BNP and ECG for detecting depressed left ventricular ejection fraction based on echocardiography were obtained from recently published population studies as part of the MONICA heart disease project (Table 5).[20,21] The sensitivity and specificity were used for a population at least 55 years of age with a BNP threshold of 17.9 pg/ml. For the study estimating the test characteristics of the ECG (using the MONICA population) a 12-lead tracing was considered abnormal if there were pathological Q waves, left bundle-branch block, ST-segment depression, T-wave abnormalities, left ventricular hypertrophy, atrial fibrillation, or atrial flutter per the Minnesota coding system. The age-specific prevalence of depressed ejection fraction was obtained from the same population (Table 5). Although echocardiography was the gold standard used in the above studies, the SOLVD prevention trial (for which the benefit of ACE inhibitor is based) used nuclear angiography to measure ejection fraction. The accuracy of

angiographic and echocardiographic imaging are similar;[22,23] nevertheless, we assumed that nuclear angiography was the gold standard and that echocardiography would be slightly less accurate (sensitivity of 92% and a specificity of 96%) when compared to this standard.[22]

Health Outcomes

Rates for the development of HF and death for asymptomatic patients with and without ACE inhibitor treatment were based on using published data from the SOLVD prevention trial.[3] We used actual event rates during the four years of reported followup. To model outcome after four years, we used an average of the yearly event rates weighted by the number of subjects still enrolled at each year of followup. Using this method, we estimated that the yearly rate of progression to symptomatic HF would be 6.5% for patients treated with ACE inhibitors and 9.8% for those not treated. We used a similar method to determine the yearly relative risk of death (compared to the general population) for patients with asymptomatic left ventricular dysfunction who are treated (2.9) or not treated (3.3) with ACE inhibitors.

We used SOLVD treatment trial data to estimate hospitalization and death rates for patients with HF treated with ACE inhibitors.[24] These data were actual event rates during the four years of reported followup for the SOLVD treatment trial. To model outcome following four years of living with HF, we used an average of the yearly event rates weighted by the number of subjects participating during each year of the trial. This method estimated that the yearly relative risk of death (compared to the general population) for patients with symptomatic left ventricular dysfunction was 6.5 when treated with ACE inhibitors.

To determine quality adjusted survival we assigned a utility value of 0.71 to each year of life for patients living with HF based on prior studies using the time-tradeoff utility of patient preferences in HF.[13] Asymptomatic patients were assumed to have a utility value of 0.87.[13] We varied these quality assumptions in sensitivity analysis (range 0.5 to 1).

Costs

We achieved a societal perspective by considering all costs of medical care (Table 5), including medical costs incurred due to increased survival.[18] Because HF survivors will incur additional costs for non-HF treatments, we assigned all patients a yearly age-specific cost of medical care based on medical expenditures for residents of the United States.[14] To this baseline cost, we added the costs of hospitalization for HF, ACE inhibitor treatment, and other outpatient HF care. We adjusted all costs to 2001 dollars using the medical component of the Consumer Price Index.[15] We determined costs for hospitalization using Medicare reimbursement for DRG 127, costs for ACE inhibitor treatment using average wholesale price,[16] and outpatient HF care using prior published estimates updated to year 2001.[17] Costs and benefits were discounted at 3% per year.[18] Costs of ECG and two-dimensional echocardiography were obtained from Medicare reimbursement for 2001. We assumed that Doppler and Color Doppler studies would not be performed as part of the screening echocardiogram. Because a BNP-specific reimbursement was not available, we used the commercial price of $29 per test (BioSite Inc.).

Strategy Comparisons

Because of multiple strategies, a large number of comparisons were possible. For each analysis, we first ranked the strategies by increasing effectiveness. We then compared the cost-effectiveness between the most effective strategy and the strategy that had the next-highest effectiveness. Strategies that provided less effectiveness at a higher cost were eliminated (dominance). Strategies could also be eliminated by extended dominance if a combination of two other strategies provided greater outcomes at lower costs. For example, assume the order of effectiveness of strategies is no screening< ECG screening < BNP screening.[18] If the cost-effectiveness ratio of electrocardiogram versus No Screening was greater than the cost-effectiveness ratio of BNP versus electrocardiogram, then electrocardiogram was eliminated by extended dominance. In our reporting, we excluded strategies that have been eliminated by dominance or extended dominance.

Peer Review

Identification of Peer Reviewers

At the beginning of the project, we requested nominations from several organizations for technical experts to join a panel that would advise staff throughout the project. A total of eight nominations were received for the Technical Expert Panel (TEP). In addition, experts in systematic reviews and meta-analysis were selected from a pool of experts associated with the Southern California Evidence-Based Practice Center but not involved with this project. The Project Staff, in consultation with the Task Order Officer, and Dr. Michael Rich, chairman of the TEP, suggested additional prominent cardiologists to review the report.

Peer Review Process

A copy of the draft evidence report was mailed to each peer reviewer, along with an instruction sheet for reviewing the draft evidence report (sample letter and instruction sheet included in Appendix C). The Peer Reviewers were asked to respond within three weeks. The eight of the ten peer reviewers who responded are listed below:

Stephen Gottlieb University of Maryland Medical Center
Mariell Jessup Hospital of the University of Pennsylvania
Carl Leier The Ohio State University Medical Center
Robert McNamara Johns Hopkins University
Eric Peterson Duke Clinical Research Institute
Illeana Pina University Hospitals of Cleveland
Todd Seto The Queen's Medical Center
James Young Cleveland Clinic Foundation, Kaufman Center for Heart Failure

A copy of the draft evidence report was also mailed to the members of the Technical Expert Panel and all technical experts responded with comments. Upon receipt of all responses from the peer reviewers and technical experts, the project staff compiled a summary of the comments and

changes, and revised the draft evidence report. We forwarded all comments to the Task Order Officer for review. The peer reviewers' and technical experts' comments are included in Appendix D, together with the corresponding responses or actions taken by project staff.