CASE 8

Data Monitoring Experience in the Metoprolol CR/XL Randomized Intervention Trial in Chronic Heart Failure: Potentially High-Risk Treatment in High-Risk Patients

Jan Feyzi
Desmond Julian
John Wikstrand
Hans Wedel

ABSTRACT

The Metoprolol CR/XL Randomized Intervention Trial in Chronic Heart Failure (MERIT-HF) was a double-blind, randomized, placebo-controlled trial in 3,991 patients with New York Heart Class II–IV heart failure and LVEF ≤0.40.[1,2] The two primary objectives were to determine the effect of metoprolol CR/XL on all-cause mortality and on the combined endpoint of all-cause mortality or all-cause hospitalizations (time to first event). There was a two-week placebo run-in period. after which patients were randomized to either metoprolol CR/XL at a dose of 12.5 mg (NYHA III–IV) or 25 mg (NYHA II) once daily or matching placebo. The randomized treatment was titrated up to 200 mg once daily or to the highest tolerated dose over an eight-week titration phase. The trial was designed to follow patients for a total mean follow-up of 2.4 years. The Data and Safety Monitoring Board (DSMB) had two tasks. The first was to review all reported Serious Adverse Events (SAEs) on a monthly basis and produce a short report to the sponsor aimed for regulatory agencies. This was done because the sponsor had received a waiver for expedited reporting of SAEs from regulatory agencies including the U.S. Food and Drug Administration (FDA). The second was to perform three prespecified interim analyses of total mortality. After the second interim analysis, at the point of observing one-half of the targeted number of deaths, the

trial was stopped early by the International Steering Committee on recommendation of the DSMB (mean follow-up time 1 year). Final results showed that all-cause mortality was lower in the metoprolol CR/XL group compared to the placebo group (145 deaths, corresponding to 7.2% per patient-year of follow-up for the metoprolol CR/XL group versus 217 deaths, 11.0% per patient-year of follow-up for the placebo group, $p = 0.0062$ adjusted for interim analyses, $p = 0.00009$ nominal).[2] The second primary endpoint of all-cause mortality combined with all-cause hospitalizations was also lower for the metoprolol CR/XL group (641 events) compared to placebo (767 events), $p = 0.00012$ nominal.[3] The procedures developed by the DSMB to implement the required intense safety follow-up will be described.

INTRODUCTION AND BACKGROUND

Chronic heart failure is a progressive clinical syndrome arising from a variety of pathological processes. The central mechanism is the heart's inability to meet the circulatory and metabolic demands of the body. The most common etiology of chronic symptomatic systolic heart failure is coronary heart disease, often complicated by hypertension and diabetes mellitus.

Heart failure is a major and growing public health problem in industrialized countries worldwide, and has a significant impact on the health care system. Estimates of the prevalence of heart failure in the general population in the Western countries range from 0.4% to 2%.[4–8] A conservative estimation indicates that about four million patients in Europe and two million patients in the U.S. have chronic heart failure, and the numbers are expected to increase substantially in the next few decades. An increased proportion of elderly in the population and improved survival after acute myocardial infarction very likely explain this. It is estimated that 90% of new cases of heart failure occur in patients above the age of 60 years.[5]

The prognosis of heart failure is, in general, poor. Approximately half of those patients diagnosed with chronic heart failure will die within four years, and of those with severe chronic heart failure, half will die within one year.[6,9] Chronic heart failure accounts for a considerable proportion of all cardiovascular related hospitalizations; about 20% of admissions and 30% of hospital days are due to this condition. The total economic burden amounts to 1% to 2% of total health care expenditure, of which hospitalization costs make up two-thirds.[7]

Therapy of chronic heart failure caused by left ventricular systolic dysfunction is mainly based on inhibition of neurohormonal stimulation secondary to pump failure. Angiotensin-converting enzyme (ACE) inhibitor treatment in combination with diuretics was initially found to improve survival and symptoms;[10] however, mortality (especially due to sudden death)

remained high.[11] Thus, there was a need for continued improvements in reducing mortality and morbidity in this patient population.

Beta-Blockers in Chronic Systolic Heart Failure

For more than two decades after the first positive report was published,[12] use of beta-blockers in chronic heart failure was avoided because of concerns about adverse effects. Three survival studies were then run in parallel investigating the effect of beta-blockers in systolic heart failure: CIBIS II, MERIT-HF, and COPERNICUS. CIBIS II[13] studied bisoprolol an immediate release $beta_1$-selective beta-blocker, MERIT-HF[1] utilized controlled-release/extended-release metoprolol succinate (metoprolol CR/XL, $beta_1$-selective), and COPERNICUS[14] utilized carvedilol, a non-selective beta-blocker with a weak α_1-blocking property. The data and safety monitoring by the DSMB in MERIT-HF was conducted in this context.

PROTOCOL DESIGN

The MERIT-HF trial was designed to evaluate the effect of metoprolol CR/XL in patients with mild to moderate chronic systolic heart failure. The trial had two primary endpoints, total mortality and total mortality plus all-cause hospitalization (time to first event).[1] In MERIT-HF, a total of 3,991 patients were randomized from February 1997 through April 1998 at 313 sites in the US and 13 European countries. Eligibility criteria included patients with NYHA class II–IV heart failure, left ventricular ejection fraction of 0.40 or lower, age between 40 and 80 years, and heart rate of at least 68 beats per minute at enrollment. Patients with acute myocardial infarction or unstable angina within 28 days before randomization were excluded. In addition, patients with a supine systolic blood pressure below 100 mm Hg at enrollment were excluded. The intention of the protocol was that no more than 40% NYHA class II patients were to be randomized.

The randomization was performed according to an optimal allocation procedure which balanced the metoprolol CR/XL and placebo groups for pre-specified baseline factors. The study medication was up-titrated during eight weeks, starting with 12.5 mg (NYHA functional class III–IV) or 25 mg once daily (NYHA II). The target dose was 200 mg once daily or highest tolerated dose. Follow-up visits then occurred every three months. Data on mortality, hospitalizations, and adverse events were collected during these visits. All predefined endpoints were classified by an independent endpoint committee using available medical records.

The trial was initially designed to randomize 3,200 patients over a 14-month period. When recruitment had been ongoing for ten months, the

number of randomized patients was higher than expected. The Steering Committee then decided to continue recruitment for the planned 14-month recruiting period, thereby increasing the sample size of the trial. This was done partly in order to increase the power of the trial.

The first-draft Study Protocol defined one primary endpoint, which was total mortality analyzed on an intention-to-treat principle with an alpha-value of 0.05 and a power of more than 80%. After discussions with the U.S. FDA in September 1996, it was decided when planning the trial, to define two primary endpoints: total mortality, and a combined endpoint of total mortality or all-cause hospitalizations (time to first event).[1] The reason for this was that if the trial had failed to show a statistically significant effect on total mortality, there would be a second option for a combined endpoint when filing for registration. An alpha-value of 0.04 was set aside for all-cause mortality and 0.01 for the second primary endpoint.[1] However, the two primary endpoints are related, which means a total alpha-spending of less than 0.05 altogether. The cumulative alpha-value (0.0015) spent on interim analyses at the final analysis of total mortality should be covered by the saving of alpha caused by the correlation between the first and second primary endpoint.

THE DATA MONITORING EXPERIENCE

The DSMB monitored safety issues during the trial based on safety reports prepared by an independent statistical analysis center. The task was to meet each month (via phone conference) to monitor all reported serious adverse events (transferred electronically each month from the sponsor) and adverse events leading to discontinuation of blind study medication, and also to perform three pre-planned interim analyses of total mortality. The procedures were governed by pre-specified DSMB monitoring guidelines stating that the second primary endpoint, i.e., the combined endpoint of total mortality or all-cause hospitalizations (time to first event) should not be monitored with interim analyses during the course of the trial. The stopping rule for efficacy was based on the total number of expected deaths, analyzed based on the intention-to-treat principle.

The trial used an asymmetric group sequential procedure to monitor total mortality.[15] A Peto-type boundary was used for monitoring a positive trend.[16] This approach favors a large critical Z-value for all interim tests before the end of the trial. The cumulative alpha for benefit was set to be 0.0012, 0.0024, and 0.0036 at the first, second, and third interim analyses to take place when 25%, 50%, and 75%, respectively, of the total number of the expected 581 deaths had occurred. It was felt that these boundaries were too conservative for harm, the cumulative probability of early stopping for harm was therefore set to be 0.005, 0.010, and 0.015 at the first, second, and third interim

analyses, respectively. The sequential boundaries for MERIT-HF are shown in Figure 1, including the mortality results for each formal interim DSMB analysis.

Safety reports were prepared by the independent statistical analysis center and consisted of a primarily graphical examination of accrual data, baseline characteristics, and adverse event data (including serious adverse events and all-cause mortality). All data were presented in a blinded manner to the DSMB (i.e. with treatments denoted as "A" and "B"). The DSMB initially remained blinded to the corresponding treatment assignment; however, they elected to unblind themselves during the February 1998 safety teleconference due to a widening difference between the two treatments in number of deaths (37 on arm A and 64 on arm B).

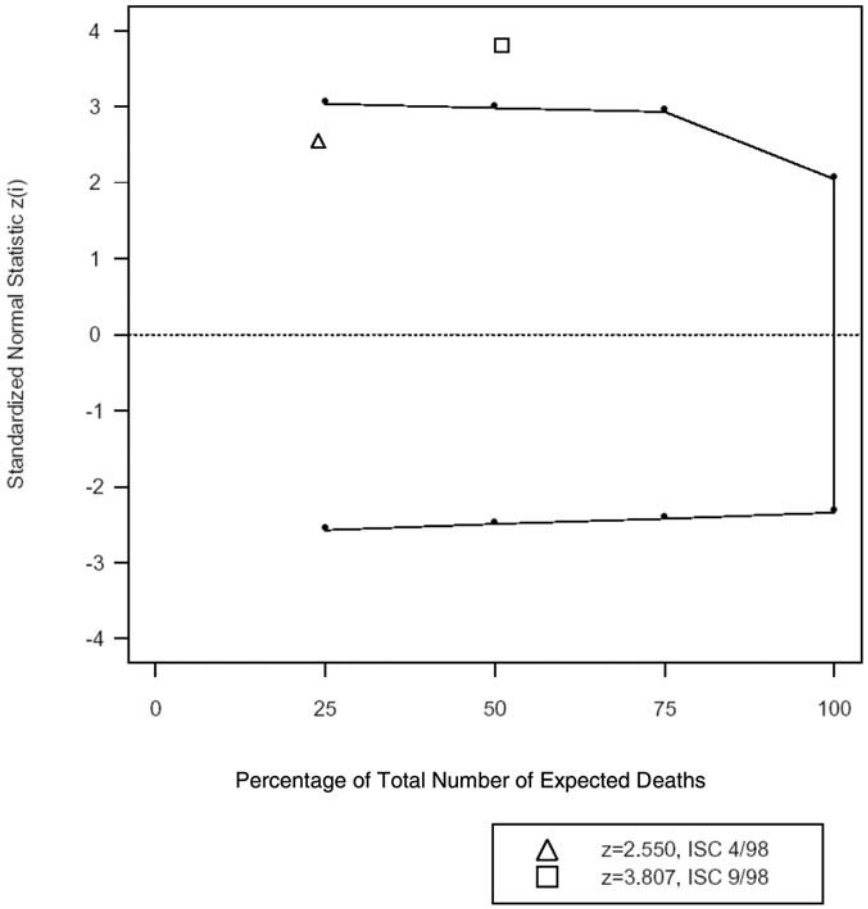


Figure 1 MERIT-HF monitoring bounds for mortality.

The first formal interim analysis occurred in April 1998, when 24.3% of the expected deaths had been observed. At that time, preliminary results from the CIBIS II trial had been presented (March 1998). This trial was closed prematurely, with positive results showing that bisoprolol reduced mortality. The MERIT-HF DSMB discussed the impact of this trial's early closing and the possible impact on the MERIT-HF trial. The DSMB noted the differences between these trials, especially that the CIBIS II trial studied patients with NYHA Class III–IV heart failure (contrasting to MERIT-HF, with approximately 40% NYHA Class II patients). The logrank Z-value for MERIT-HF at this point was 2.550, below the pre-specified monitoring bound (3.04) for benefit at this point in the trial. Hence, the DSMB recommended to continue the trial; however, it was noted that the Z-value was fairly close to the upper bound and may cross at the time of the next interim analysis if the trend was to continue. Trial randomization was to stop on April 14, 1998. The full CIBIS II results were to be presented in August, 1998, near the time of the next interim analysis.

In scheduling the second formal interim analysis, the DSMB decided to wait until the pre-specified 50% point (September 1998) in order to give the DSMB time to understand and reflect on the CIBIS II results. MERIT-HF safety reviews prior to this scheduled meeting did not show any unexpected safety concerns in the metoprolol CR/XL arm. Updated numbers for deaths and hospitalizations were given and the trend for a lower number of both deaths and hospitalizations for patients on metoprolol CR/XL continued.

The Second Interim Analysis

The second interim analysis meeting of the DSMB was held on September 21, 1998. The chairman of the DSMB had informed the chairman of the Executive Committee that the DSMB wanted to meet with the Executive Committee directly after their second interim analysis. At that time, the CIBIS II results had been already presented and confirmed the initial reports that bispropolol reduced all-cause mortality in patients with NYHA Class III–IV heart failure.[13] The mean follow-up time for patients in the MERIT-HF trial was 10.8 months at this point in the trial. There were 115 deaths on the metoprolol CR/XL arm and 181 deaths on the placebo arm, representing 51% of the expected number of deaths. The logrank Z-value was 3.807, substantially exceeding the upper monitoring bound for benefit of 2.98 prespecified in the DSMB monitoring guidelines. Although not formally tested, the results for the second primary endpoint of all-cause mortality and all-cause hospitalization were consistent with the mortality results. After discussion of these results and a thorough examination of consistency of results over protocol-specified subgroups, the DSMB unanimously voted to recommend termination of the trial.

The DSMB prepared a brief initial statement and the limited results on mortality were immediately presented to members of the MERIT-HF Executive Committee. The Executive Committee then deliberated as to whether to accept the recommendations of the DSMB, based on these limited data. The Executive Committee voted to accept the DSMB recommendations. The DSMB then fully debriefed the Executive Committee as to the overall results, including baseline data, compliance, mortality, mortality plus hospitalization, adverse events and several pre-specified subgroups.

The DSMB then issued the following statement:

> *On September 21, 1998, the Independent Safety Committee undertook the secondary interim analysis of the MERIT-HF study. The Committee found that the previously defined criteria for termination of the study for mortality reduction had been met and exceeded ($z = 3.807$ versus 2.98 as defined in the protocol). These mortality results were consistent across the predefined subgroups. The findings with regard to the second primary endpoint of death and/or hospitalization were consistent with the mortality results. Discontinuation of study medication was similar in the two groups. Serious adverse effects were commoner in the placebo group than in the metoprolol CR/XL treated patients. In view of the highly significant benefit observed, the Independent Safety Committee recommend termination of the study as soon as practicable.*

This statement was given to the Executive Committee and kept secret until the Executive Committee met with the Steering Committee of MERIT-HF two weeks later (see below). Furthermore the DSMB recommended that mortality data be published as soon as possible.

Early Stopping

The International Steering Committee of MERIT-HF met October 2, 1998, and decided to close the trial on October 31, 1998, on the recommendation made by the DSMB. However, for regulatory reasons and as previously decided by the Steering Committee, the blind study medication code could not be broken until all data were in, and clean file had been declared, which would take many months after trial closure. The solution was a controlled down-titration of blind study medication in parallel with an optimal up-titration of open label metoprolol CR/XL according to the recommendation made by the Steering Committee of the trial. Since it would take some time to declare clean file at the sponsor it was agreed to base the publication on mortality results on analyses performed by the independent statistical analysis center.[2]

The Executive Committee and sponsor recommended that the DSMB remain functional throughout the close-out period of the trial in order to ensure patient safety. This would also allow the Executive Committee and

sponsor to remain blinded for individual patient assignments during the final data collection period. The DSMB continued to monitor patient safety until the database was locked (June 1999). The final published mortality[2] and mortality plus hospitalization results[3] are summarized in Table 1. Kaplan–Meier plots of the time to death and time to death plus hospitalization can be found in Figures 2a,b. Baseline characteristics are summarized in Table 2.

LESSONS LEARNED

The MERIT-HF monitoring experience provided a number of lessons for the monitoring of patient safety in future trials. First, because of the waiver for expedited reporting of SAEs, the DSMB had to meet more often than usual in order to provide reports to regulatory agencies. In order to comply with this request, the independent statistical analysis center provided the DSMB with monthly safety reports with subsequent discussion by the DSMB via teleconference. Scheduling of such meetings could be potentially problematic; however, a time convenient to all DSMB members was established at the onset and remained predictable throughout the trial to encourage consistent participation. Any early trends, both positive or negative, could be addressed with such monitoring.

In addition, because of the lack of data for long-term exposure to beta-blockers in this patient population, asymmetric monitoring bounds were established with a conservative upper bound for benefit and a less conservative lower bound for harm. This allowed for less stringent statistical criteria to be met in the case of a negative trend. However, the pre-defined DSMB

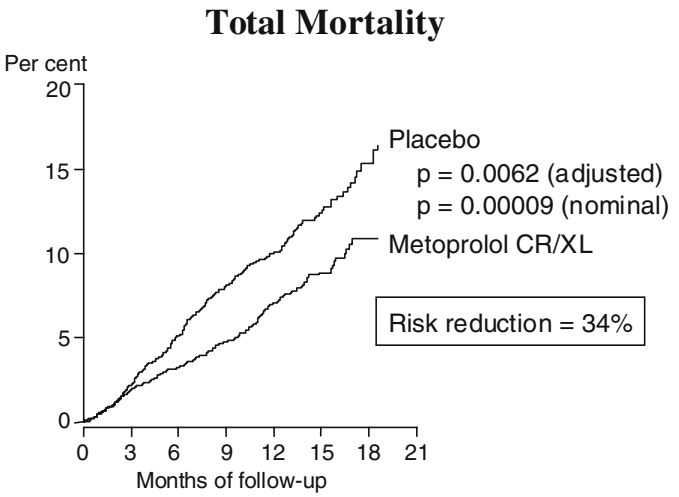
Table 1 Mortality and Mortality Plus Hospitalization Results

Endpoints (N, %)*	Metoprolol CR/XL (n = 1990)	Placebo (n = 2001)	Relative risk (95% CI); p-value†
Total mortality	145 (7.2%)	217 (11.0%)	0.66 (0.53–0.81); p = 0.0062 (adj)
Cardiovascular mortality	128 (6.4%)	203 (10.3%)	0.62 (0.50–0.78); p = 0.000022
Sudden death	79 (3.9%)	132 (6.7%)	0.59 (0.45–0.78); p = 0.0002
Total mortality or all-cause hospitalization	641 (38.8%)	767 (48.0%)	0.81 (0.73–0.90); p = 0.0001
Total mortality or hospitalization for CHF	311 (16.5%)	431 (23.5%)	0.69 (0.60–0.80) p = 0.0000008

* Percentage per patient year of follow-up (2,004 vs. 1,977; 1,651 vs. 1,599; and 1,882 vs. 1,837 patient years for the different endpoints, respectively).

† For total mortality, p-value adjusted for interim analysis is given; otherwise the nominal p-value is given.

(a)



(b)

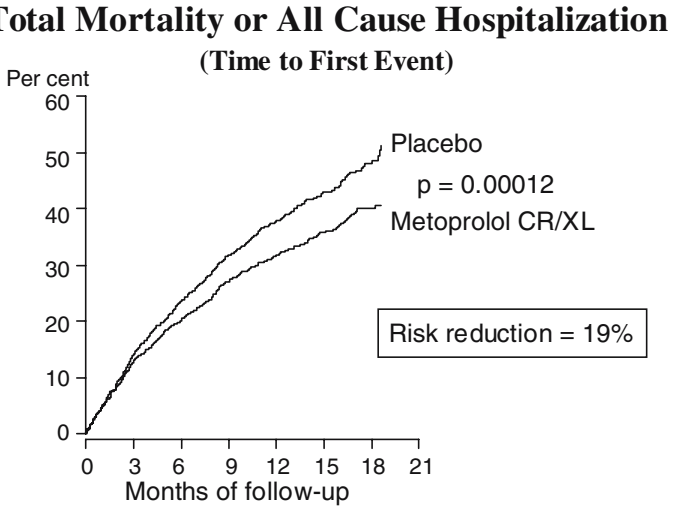


Figure 2 Kaplan–Meier estimates of the first primary endpoint of total mortality (a), and of the second primary endpoint of total mortality or all-cause hospitalization (time to first event; b). From references 2 and 3, with permission.

Table 2 Baseline Characteristics

Baseline characteristics	Metoprolol CR/XL (n = 1,990)	Placebo (n = 2,001)
Age, mean, yr	64	64
Sex, % female	23	22
White, %	94	94
Ischemic etiology of heart failure	65	66
NYHA class, %		
II	41	41
III	56	55
IV	3.4	3.8
Ejection fraction, mean	0.28	0.28
Previous myocardial infarction, %	48	49
Time since last myocardial infarction <1 yr, %	8	7
Hypertension, %	44	44
Diabetes mellitus, %	25	24
Medications, %		
Diuretics	91	90
ACE inhibitor	89	90
A-II-blocker	7	6
ACE inhibitor or A-II-blocker	95	96
Digitalis	63	64
Spironolactone	7	8

Revised from MERIT-HF (2000) with permission from JAMA.

monitoring guidelines stated that in the event that a negative mortality trend would emerge during the course of the MERIT-HF study, the DSMB should proceed until a definitive result had been obtained. Although the negative trend may be sufficient to rule out any possible positive benefit, the DSMB should continue the trial until a harmful effect could be distinguished from neutrality. The rationale for this was that being able to distinguish between a harmful mortality effect and a neutral mortality effect was important in this patient population since metoprolol may be used for other beneficial effects than mortality.

The release of the results from other similar trials can both simplify and complicate the decision-making process. The CIBIS II results provided confidence that the effect of beta-blockers on survival in patients with heart failure is a real phenomenon. However, there was a real concern that the early release of results of CIBIS II to the medical community could have adversely affected completion of the MERIT-HF trial. Luckily, randomization was near complete and would not have been compromised. However, had the patient populations in the two trials been more alike regarding their heart failure profiles, an ethical dilemma as to whether to continue the trial in light of the CIBIS II results could have arisen. The results of the COPERNICUS trial (of carvedilol in patients with severe heart failure) were released

shortly after the closure of the MERIT-HF trial and further established the efficacy of beta-blockers in the treatment of patients with heart failure. Thus, external consistency of the effects of beta-blockers on both mortality and mortality plus hospitalization is reassuring.

The method used by the DSMB to reveal the results to the Executive Committee is of interest. This process was discussed in closed-session by the DSMB after the unanimous vote to recommend termination of the trial. It was decided that the first information to be given to the Executive Committee was that the DSMB had recommended that the trial be terminated and to ask whether the Executive Committee would like to be unblinded to the trial results. Members of the Executive Committee who were present met in closed-session to further discuss this issue. They decided that they indeed wanted more information. The DSMB immediately informed them of additional results.

Finally, in order to speed up the publication of the mortality results,[2] the independent statistical analysis center generated all the analyses for the MERIT-HF Steering Committee while the sponsor was still working on the clean file process.

Subsequent to the publication of the MERIT-HF trial, regulatory review raised a question regarding the consistency of results across geographic areas.[17] In particular, for mortality, the hazard ratio for the U.S. patients was near 1.0 in contrast to the non-U.S. (European) results of 0.55. For mortality plus hospitalization, the results were in fact consistent.[3] The FDA asked whether or not the trial could have been terminated early in the non-U.S. sites and allowed the U.S. sites to continue with blinded treatment. While the DSMB did not in fact deliberate on this issue, in retrospect, some members of the DSMB have conjectured that they do not believe that MERIT-HF could have been continued in the U.S. alone, given the striking overall results of MERIT-HF as well as the results of other beta-blocker studies, which did not have this anomaly. It was concluded that the best estimate of the treatment effect on total mortality for any subgroup is the estimate of the hazard ratio for the overall trial.[17]

REFERENCES

1. The International Steering Committee on behalf of the MERIT-HF Study Group. 1997. Rationale, design, and organization of the metoprolol CR/XL randomized trial in heart failure (MERIT-HF). *Am J Cardiol* 80:54–58J.
2. MERIT-HF Study Group. 1999. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). *Lancet* 353:2001–2007.
3. Hjalmarson A, Goldstein S, Fagerberg B, Wedel H, Waagstein F, Kjekshus J, et al. 2000. Effects of controlled-release metoprolol on total mortality, hospitalizations, and well-being in patients with heart failure: The Metoprolol CR/XL Randomized Intervention Trial in Congestive Heart Failure (MERIT-HF). *JAMA* 283:1295–2007.

4. Cohn, JN. *1996.* The management of chronic heart failure. *N Engl J Med* 335: 490–498
5. MattsonJack EpiOnline database accessed on 21.03.02 at http://www.epidb.com/epiOnline/Phase1/fmeLayout.asp
6. Senni M, Tribouilloy CM, Rodeheffer RJ, et al. 1999. Congestive heart failure in the community. Trends in incidence and survival in a 10-year period. *Arch Intern Med* 159:29–34.
7. McMurray J, Hart W, Rhodes G. 1993. An evaluation of the economic cost of heart failure to the National Health Service in the United Kingdom. *Br J Med Econ* 6:99–110
8. Task Force on Heart Failure of the European Society of Cardiology. 1995. Guidelines for the diagnosis of heart failure. *Eur Heart J* 16:741–751.
9. The CONSENSUS Trial Study Group. 1987. Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). *N Engl J Med* 316:1429–1435.
10. Garg R, Yusuf S. Collaborative group on ACE Inhibitor Trials. 1995. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. *JAMA* 273:1450–1456
11. Goldstein, S. 1997. Clinical studies on beta-blockers and heart failure preceding the MERIT-HF trial. *Am J Cardiol* 80(9B):50J–53J.
12. Waagstein F, Hjalmarson Å, Varnauskas, Wallentin I. 1975. Effect of chronic beta-adrenergic receptor blockade in congestive cardiomyopathy. *BMJ* 37:1022–1036.
13. CIBIS-II Investigators and Committees. 1999. The cardiac insufficiency bisoprolol study II (CIBIS II), *Lancet* 353:9–13.
14. Packer M, Coats AJS, Fowler MB, et al. 2001. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med* 344:1651–1658.
15. Lan KKG, DeMets DL. 1983. Discrete sequential boundaries for clinical trials. *Biometrika* 70:649–653.
16. Peto R, Pike MC, Armitage P, et al. 1976. Design and analysis of randomized clinical trials requiring prolonged observations of each patient. I. Introduction and design. *Br J Cancer* 34:585–612.
17. Wedel H, DeMets D, Deedwania P, Fagerberg B, Goldstein S, Gottlieb S, Hjalmarson Å, Kjekshus J, Waagstein F, Wikstrand J on behalf of the MERIT-HF Study Group: 2001 Challenges of subgroup analyses in multinational clinical trials: Experiences from the MERIT-HF trial. *Am Heart J* 143:502–511.