CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER
20-297/S-007

Medical Review(s)

DIVISION OF CARDIO-RENAL DRUG PRODUCTS

Clinical Review

NDA: 20-297/SE1-007

Sponsor: GlaxoSmithKline

Submission: 3 October 2001 proposed changes in label reflecting COPERNICUS.

Review date: October 15, 2001

Reviewer: N. Stockbridge, M.D., Ph.D., HFD-110

The sponsor has adequately explained differences between the primary review's and their own accounting of end point events. The reviewers included a small number of events after cardiac transplant and after withdrawal of consent, the inclusion of which had no material effect on the final results. The sponsor's numbers should appear in the label.

The sponsor argues that patient global assessment data should appear in the label, because censoring for early termination of the study was not informative, the nominal p-values are quite small, and the results of COPERNICUS are consistent with the earlier CHF development program. This reviewer concurs.

The sponsor was asked to reflect on effects on the primary end point in subgroups defined by sex, race, and age. Of these, the data are least compelling with regard to effects by race, probably because there were only about 5% Black subjects. The following proposal is made for the paragraph the sponsor adds on page 8:

> *Beneficial effects of Coreg were seen for all-cause mortality and for combined end points of mortality plus CHF, CV, or total hospitalization in subgroups based on sex or age.*

The sponsor has applied the Temple Rule to the lists of adverse events, tabulating events more common on active treatment after round to the nearest 1%. There are data to support the proposed statement that adverse event rates were pretty similar in demographic subgroups, but, once again, there are not many Black subjects, so this component of the statement should be dropped.

Various other minor changes proposed by the sponsor are all reasonable.

DIVISION OF CARDIO-RENAL DRUG PRODUCTS

Clinical Review

NDA: 20-297

Sponsor: GlaxoSmithKline

Submission: SE-0071 (28 February 2001): a request to extend the indication for carvedilol in light of a reduction in all-cause mortality in the COPERNICUS study.

Review date: June 7, 2001

Reviewer: N. Stockbridge, M.D., Ph.D., HFD-110

J Hung, Ph.D., HFD-710

Concurrence: G. Chi, Ph.D., HFD-710

Summary: This is a review of the COPERNICUS study, an evaluation of the effect of a single dosing strategy for carvedilol versus placebo on all-cause mortality in subjects with severe heart failure. The study was stopped early for an effect favoring carvedilol. The benefit extended to prospectively defined secondary end points of mortality plus hospitalization, but these appear to be driven by the mortality effect, largely attributable to a reduction in sudden death in the absence of worsening heart failure symptoms.

Distribution: NDA 20-297

HFD-110/Project Manager

Table of contents

List of Tables

List of Figures

1 COPERNICUS protocol

1.1 Name of study

MF 4477/SB 287: Carvedilol prospective randomized cumulative survival trial (COPERNICUS).

1.2 Basis of review

The description of the study protocol is based on the fully amended protocol, dated 21 August 2000, and 4 amendments.

1.3 Objectives

The purpose of the study was to determine the effect of carvedilol on all-cause mortality in subjects with severe chronic heart failure.

Secondary objectives were the evaluation of all-cause mortality plus heart-failure-specific hospitalization, all-cause mortality plus cardiovascular-specific hospitalization, all-cause mortality plus all-cause hospitalization, and subjects' global assessment.

1.4 Population

Subjects were to be males and females of low childbearing potential, over 18 years, with at least a 3-month history of heart failure, with at least a 2-month history of symptoms of fatigue or dyspnea at rest or with minimal exertion, not receiving intravenous inotropes or vasodilators within 4 days of screening. To be randomized, subjects needed a left ventricular ejection fraction <25% within 6 months or since any cardiac events or procedures. Subjects had to have been receiving a diuretic for 2 months, an ACE inhibitor or angiotensin II receptor antagonist for 2 months (unless contraindicated). They could be receiving digitalis, hydralazine, or amiodarone, but the digitalis and amiodarone must have been present for at least 2 months. Other requirements included sitting systolic pressure >85 mmHg, serum potassium 3.5-5.2 mM, serum creatinine <2.8 mg/dL (and not rising 0.5 mg/dL during screening), body weight stable within 1.5 kg during screening, no rales, ascites, or peripheral edema attributable to heart failure, and not currently hospitalized for cardiovascular reasons.

Other exclusion criteria were cardiac dysfunction not attributable to left ventricular dysfunction; potentially reversible cardiomyopathy; prior cardiac transplantation, left ventricular assist device, or left ventricular remodeling procedure; myocardial infarction, unstable angina, life-threatening ventricular arrhythmia, stroke, TIA, or cardiac surgery within 2 months; regular out-patient treatment with intravenous inotropes or vasodilators; use of beta-blockers within 2 months; need for alpha-adrenergic blockers, calcium channel blockers, or class-I antiarrhythmics; sitting heart rate <68 bpm; significant conduction defects without a pacemaker; peripheral arterial disease symptomatic at rest; history of asthma or reversible chronic obstructive pulmonary disease; unstable insulin-dependent diabetes; hepatic enzymes or bilirubin >3xULN; endocrine disorders; other illness limiting life expectancy; and alcohol or drug abuse within 1 year.

1.5 Procedures

Subjects were to be randomized evenly to either placebo or carvedilol. Study drug was initiated at 3.125 mg bid and doubled, as tolerated, at 2-week intervals to a target dose of 25 mg bid. During the maintenance phase, visits were scheduled every 2 months.

The plan was to enroll 2500 subjects and continue until there were a total of 900 deaths or until the study was terminated at the discretion of the DSMB. With the decision to terminate the study, subjects were to return for a final on-treatment assessment of study end points.

Specific advice was given to manage dose-related vasodilation, bradycardia, heart block, and worsening heart failure. Temporary discontinuation of study drug or down-titration was permitted.

Other cardiac medications could be adjusted as needed for each subject. The only specific prohibition was the use of beta-blockers.

The primary end point was all-cause mortality. The primary population was all randomized subjects. The primary analysis was to be a log-rank comparison of time-to-event for active treatment and placebo with two-sided alpha = 0.04 (after interim analysis). All subjects were to be followed until the end of the study. The assumptions in the sample-size calculation were (1) 90% power, (2) recruitment over 2.5 years, (3) follow-up of 1.25 years post-recruitment, (4) mortality on placebo of 28% per year, and (5) 20% reduction in mortality on carvedilol.

Subjects who were lost to follow-up or who underwent cardiac transplant were to be censored at those times.

The study administration included an Executive Committee and a Data and Safety Monitoring Board. An independent Biostatistical Center statistician was responsible for providing data to the DSMB. The head of the Executive Committee could be at DSMB meetings (at the latter's discretion). The DSMB could elect to view unblinded outcome data.

There was to be an Endpoint Committee, comprised of investigators, to judge whether deaths were non-cardiovascular, progression of heart failure, sudden death, myocardial infarction, or other cardiovascular causes. This assessment was based on blinded data.

There were 4 protocol amendments. The first (9 September 1997) made minor changes in inclusion and exclusion criteria, intended to insure a more stable population at baseline. This amendment was made before the first subjects were randomized. The second amendment (10 January 1998) made minor changes. The third amendment (22 June 1998) made changes to the criteria to encourage the enrollment of subjects with worse symptoms of heart failure. Also, hospitalization was defined as confinement for medical reasons, exceeding 24 hours. The fourth amendment (8 October 1999) increased the sample size from 1800 to 2500 subjects, without changing the goal of continuing until 900 deaths were observed.

2 Results

2.1 Enrollment

The first subject was enrolled 14 October 1997. The last subject was enrolled 16 March 2000. The nominal closing date was 20 March 2000 (as a result of the recommendation of the DSMB). Vital status is known for all randomized subjects as of 20 March 2000.

There were 334 centers in Europe[1] (152), US (117), Canada (35), Israel (14), Australia (7), South Africa (5), Argentina (3), and Mexico (1).

Individual centers enrolled from 1 (56 centers) to 44 subjects. The median number of subjects per center was 4.

Enrollment by country is shown in Table 1.

Table 1. Enrollment by country[2]

	N		N
US	482	Italy	37
Russia	309	South Africa	33
Poland	299	Lithuania	21
Israel	196	Austria	20
Hungary	176	Australia	20
Canada	145	Switzerland	17
Ukraine	132	Mexico	13
Czech Republic	118	Argentina	13
Great Britain	93	France	9
Netherlands	76	Portugal	9
Germany	71	Unknown	1

The disposition of subjects in the study is shown in Table 2.

Table 2. Disposition of subjects

	Placebo	Carvedilol
Screened	?	
Randomized	1133	1156
Withdrew during titration	110	100
In titration at study end	129	142
Entered maintenance	894	914
Withdrew from maintenance[3]	194	128
In maintenance at study end	700	786

Baseline characteristics of randomized subjects are shown in Table 3.

[1] There were 26 centers in Poland, 17 in Great Britain, 16 in Hungary, 16 in Russia, 14 in Germany, 12 in Czech Republic, 11 in Netherlands, 10 in Italy, 5 in Austria, 5 in France, 4 in Switzerland, 4 in Lithuania, and 2 in Portugal.

[2] Analysis confirmed by the reviewer.

[3] Counts include subjects who died.

Table 3. Baseline characteristics[4]

	Placebo N=1133	Carvedilol N=1156
Age	63±12	63±11
Sex (%male)	80	79
Race (%)		
Caucasian	90	91
Black	6	5
Ischemic heart failure (%)	67	67
LVEF (%)	20±4	20±4
CHF hospitalization with 1 year (%)	65	66
Systolic pressure	123±19	123±19
Diastolic pressure	76±11	76±11
Heart rate	83±13	83±12
Serum sodium	137±3	137±
Serum creatinine	134±36	134±37
Medications (%)		
Diuretic	99	99
ACE inhibitor	89	88
Digitalis	65	67
Spironolactone	20	19
Amiodarone	17	18
Angiotensin II antagonist	9	10

As expected for a large trial, there were small differences between the treatment groups.

Fifty-five percent of subjects had medical history of myocardial infarction and 35% had angina.

2.2 Conduct

2.2.1 Financial disclosure

In a submission of 4 June 2001, the sponsor asserts that no investigator's compensation was affected by study outcome and that no investigator received 'significant payments' as defined by 21CFR 54.2(f).

2.2.2 DSI audit

Because this was a large multi-center study for which no center contributed a substantial fraction of the subjects, a decision was made not to perform any audits of the clinical sites.

2.2.3 Interim analyses

The available documentation from DSMB meetings is consistent with there having been 5 meetings. The first (March 1998, with 153 subjects enrolled) was an organizational meeting at which the plan for 4 interim analyses was discussed. The plan called for interim analyses with 0.2, 0.4, 0.6, and 0.8 of the information content. At the second meeting (November 1998), the DSMB elected to view unblinded safety data. The recommendation to increase the enrollment goals was made after the third DSMB meeting, based on lower-than-expected mortality in the placebo group, despite an earlier amendment that permitted enrollment of subjects still in the hospital with heart failure. Z-scores rose progressively from the first to the fourth analysis. The fourth analysis (March 2000, with 2055 subjects randomized) took place with about 0.3 of the

[4] *Sponsor's analysis.*

information available. At that time, the boundary for the fourth analysis was crossed, and the DSMB recommended termination of the study as soon as possible.

2.2.4 Protocol violations

Six subjects randomized to placebo and 3 subjects randomized to carvedilol had ejection fractions >0.25. All but two of these (one in each group) had ejection fractions <0.3.

2.2.5 Dosing

Figure 8 shows the distribution of subjects on each dose in the active treatment group. Most subjects transitioned from the initial dose of 3 mg to the target dose of 25 mg.

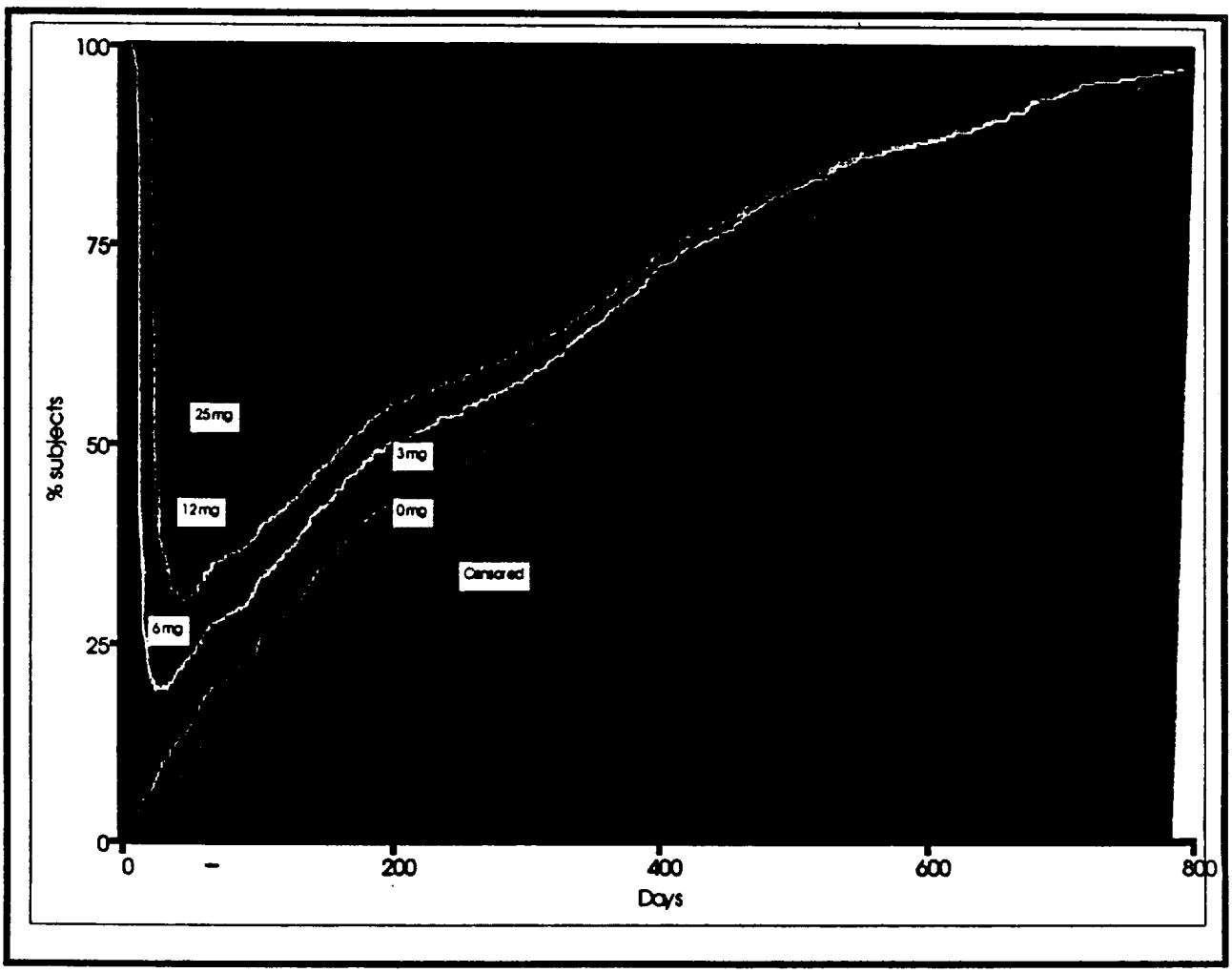

Figure 1. Distribution of doses by time in study[5]

This is a stacked bar chart in which each subject contributes to exactly one state at each point in time. The censored region includes subjects who died and subjects whose participation was limited by the premature termination of the study. Reviewer's analysis of COMP dataset.

[5] *Reviewers' analysis.*

2.3 End points

2.3.1 Primary

2.3.1.1 Mortality

2.3.1.1.1 Primary analysis

The primary analysis was a log-rank comparison of all-cause mortality for all randomized subjects. Because of the interim analyses, the nominal α for the final analysis would be 0.04, if the study had continued to its planned end. In fact, the study was stopped early (20 March 2000) on the recommendation of the DSMB. The recommendation was based on the finding of a highly significant beneficial effect of carvedilol on survival.

The sponsor's analysis of progression of the Z-score in relation to the interim analysis monitoring boundaries is shown in Figure 2. The reviewers' analysis, based on the SAS datasets, was similar.

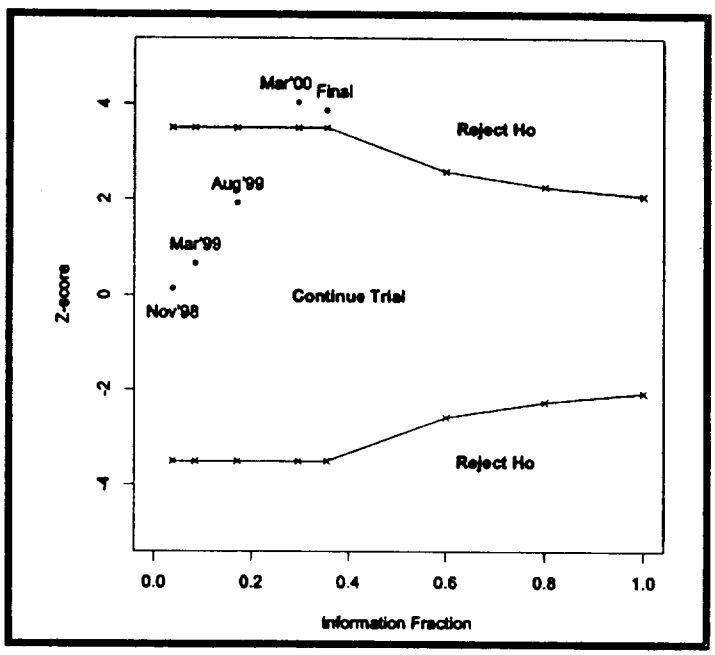

Figure 2. Interim analyses for all-cause mortality

Survival analyses for all-cause mortality were performed at 4 times during the conduct of COPERNICUS. The figure, from the sponsor's analysis[6], shows the Z-score as a function of the information fraction, in relation to the prospectively set monitoring boundaries. The trial was stopped early, shortly after the boundary was crossed by the interim analysis of March 2000.

The time to censoring appeared to be balanced between the two treatment groups, as shown in Table 4.

[6] *From page 55 of mf4477statisticalanalysis.pdf*

Table 4. Distribution of time to censoring for mortality[7]

		Placebo N=1133	Carvedilol N=1156
Deaths		191 (16.9%)	132 (11.4%)
Time to censoring (days)	Max	866	874
	99th %ile	839	844
	95th	732	747
	90th	656	671
	75th	481	487.5
	50th	313	317
	Mean	324	330
	25th	133	131
	10th	41	42
	5th	20	20
	1st	6	4
	Min	0	0

Based on the SAS data set STEND.SD2 provided by the sponsor, there were a total of 303 deaths recorded up to the study termination date (March 20, 2000). There were no subjects lost for follow-up for vital status. Two of the deaths (Subjects #03100022 and #43803373) in the carvedilol group were not reported on Endpoint Committee forms.

The reviewers' analysis used all deaths reported in the STEND dataset. Carvedilol significantly decreased the mortality risk by 34% with 95% CI (18%, 47%) and a nominal p = 0.0002. According to the study report, the three previous interim analyses were performed after 34, 76, and 154 deaths prior to transplant had occurred. Using this information, the adjusted p-value for the final interim analysis is p=0.0014.

Table 5. All-cause mortality[8]

	Placebo N=1133	Carvedilol N=1156	Hazard ratio (95% CI)	Nominal p-value[9]	Adjusted p-value[10]
Deaths	191 (16.9%)	132 (11.4%)	0.66 (0.53, 0.82)	0.0002	0.0014

The sponsor's final life table analysis of all-cause mortality is shown in Figure 3.

APPEARS THIS WAY
ON ORIGINAL

[7] *Reviewers' analysis*

[8] *Reviewers' analysis*

[9] *Unadjusted for interim analyses.*

[10] *Adjusted for three prior interim analyses.*

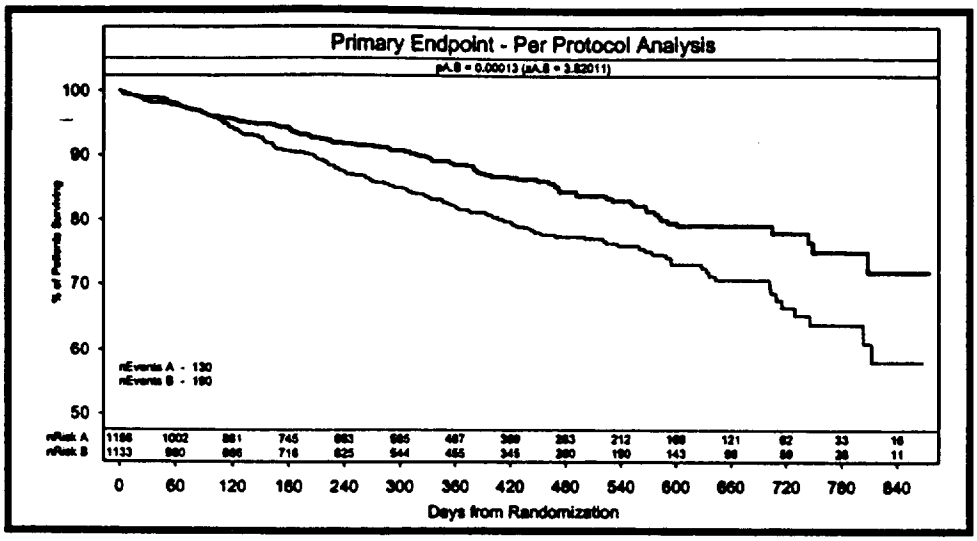

Figure 3. All-cause mortality

Survival curves for all-cause mortality as analyzed by the sponsor[11].

Survival curves for placebo and carvedilol separate after about 3 months and diverge for the remaining 2.5 years of follow-up.

Had the study gone to completion with subsequent deaths being equal in the two treatment groups, the reviewers estimate that the final counts would have been about 480 on placebo and 420 on carvedilol. This difference would have still been statistically significant (p=0.004).

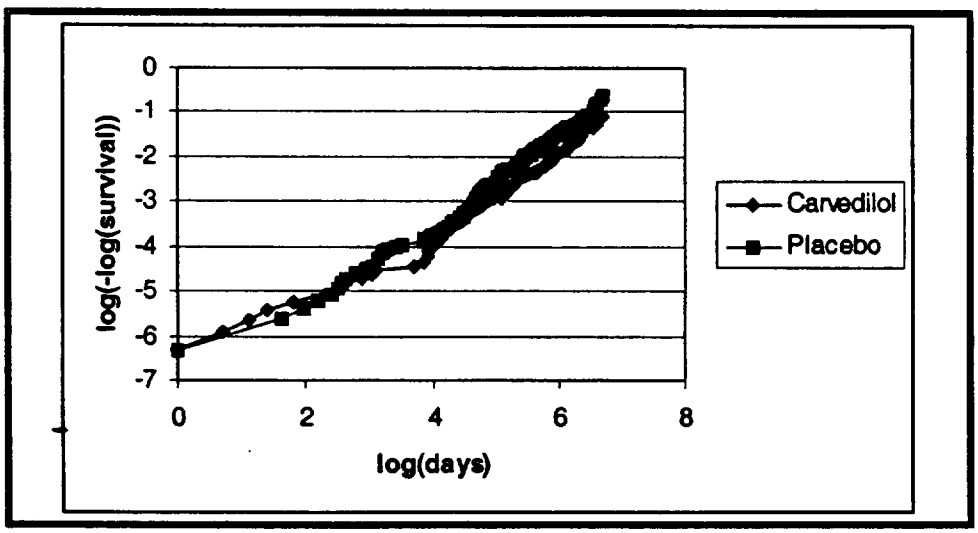

Figure 4. Log(-log(survival) vs. log(days)[12]

A parallel time course is indicative of a constant hazard ratio as a function of time.

As shown in Figure 4, the log(-log(survival)) curves of the two treatment groups appeared to be approximately parallel, once the difference between the groups becomes manifest. In addition, based on the Cox regression analysis using the model containing treatment by log(days) interaction, there is no evidence to indicate that the hazard ratio increases or decreases over time during the study duration (p = 0.52).

[11] *Page 47 of mf4477statisticalanalysis.pdf*

[12] *Reviewers' analysis*

2.3.1.1.2 Subgroups

The sponsor performed numerous subgroup analyses for all-cause mortality, as shown in Figure 5.

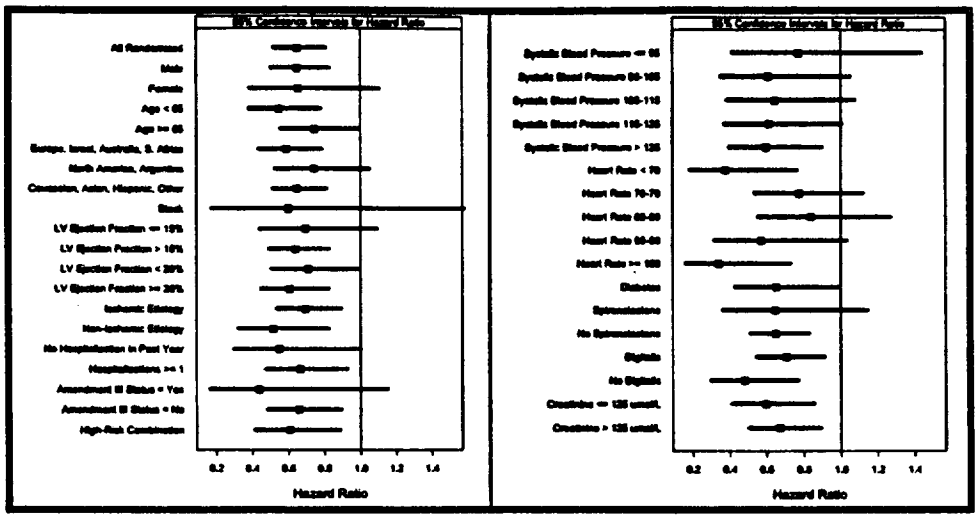

Figure 5. Subgroup analyses for all-cause mortality

Hazard ratio for carvedilol versus placebo for all-cause mortality in various sponsor-selected subgroups[13].

Only about 5% of subjects were Black, so the confidence limits for that subgroup are large, but the point estimate is a greater benefit in this group than among Caucasians.

As part of this review, the crude death rates on carvedilol and placebo were compared for each distinct percentile for ejection fraction for which there was at least one death on carvedilol and placebo. These results are shown in Figure 6.

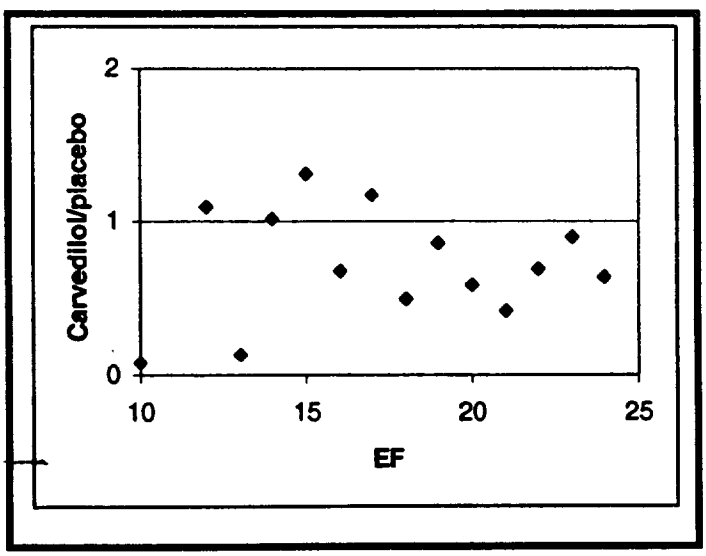

Figure 6. Crude death rate comparison by ejection fraction.

LVEF values were reported to the nearest percent. For values of LVEF for which there was at least one death on placebo and carvedilol, the crude death rates were computed for each treatment group and the figure shows the rate on carvedilol divided by the rate on placebo. Reviewer's analysis of STEND and LVEF datasets.

The results do not suggest an effect of ejection fraction on the treatment effect for all-cause mortality (p=0.99 for treatment-by-ejection fraction interaction, based on a Cox regression analysis).

As part of this review, an analysis of crude mortality ratios by country was performed. The results are shown in Figure 7.

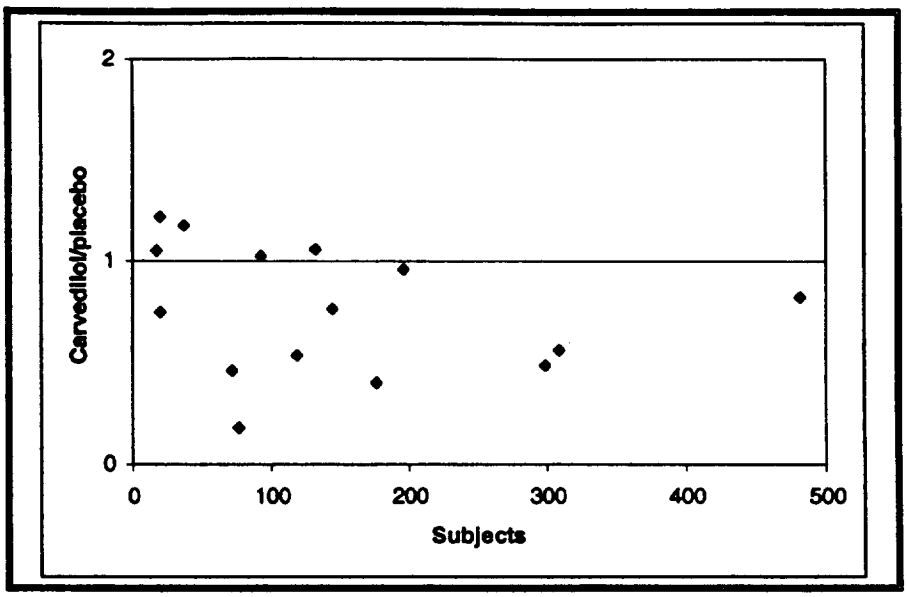

Figure 7. Crude death rate comparison by country

For each country with at least one death on placebo and carvedilol, the ratio of crude mortality rates was plotted as a function of the number of subjects randomized in that country. The US result is the rightmost point.

This is a classic appearance to a "funnel plot", showing progressively less variation in countries that randomized greater numbers of subjects. The US does not appear to be an outlier. However, compared with non-US regions as a whole, the effect in the US was somewhat smaller (risk ratio of 0.80 in the US vs. 0.60 in the rest of the world).

2.3.1.1.3 Causes of death

The distribution of causes of death, as adjudicated by the end points committee, is shown in Table 6.

Table 6. Causes of death[14]

		Placebo N=1133	Carvedilol N=1156
Left ventricular dysfunction	Sudden death without worsening HF	88	48
	Sudden death with worsening HF	13	13
	Pump failure	49	43
Other cardiovascular	Cerebrovascular disease	6	5
	Myocardial infarction	5	2
	Cardiac procedure	1	0
	Aortic, mesenteric, renal, or peripheral vascular disease	1	0
	Pulmonary embolus	1	1
	Other	2	4
Non-cardiovascular		13	10
Unknown[15]		12	6
Total		191	132

The largest difference was in the category with the largest number of deaths—sudden death without worsening heart failure. This category accounted for 43% of the deaths and 67% of the difference between the treatment groups.

2.3.2 Secondary

2.3.2.1 Death plus hospitalization

The reviewers' analyses of secondary end points involving death and hospitalization are summarized in Table 7.

Table 7. Death plus hospitalization

	Placebo N=1133	Carvedilol N=1156	P-value	Hazard ratio (95% CI)
Death	191	132	0.0002	0.66 (0.53, 0.82)
Death + CHF-hospitalization	363	274	0.0001	0.69 (0.59, 0.81)
Death + CV-hospitalization	403	317	0.0001	0.73 (0.63, 0.84)
Death + any hospitalization	510	437	0.0003	0.79 (0.70, 0.90)

Δ = 59
Δ = 89
Δ = 86
Δ = 73

The difference between the groups in the numbers of subjects died or hospitalized for worsening heart failure is about 50% wider than the difference in deaths. This is suggestive of benefit in reducing hospitalization, as well as a reduction in mortality.

The Packer manuscript[16] contains numerous other analyses of the effects of carvedilol on hospitalization, few of which were specified in the protocol. These were (1) the number of randomized subjects hospitalized for worsening heart failure, (2) the number of randomized subjects hospitalized for cardiovascular reasons, (3) the number of randomized subjects hospitalized for any reason, (4) the number of randomized subjects hospitalized more than once, (5) the number of hospitalizations for worsening heart failure, (6) the number of hospitalizations for worsening heart failure per randomized subject, (7) the number of hospitalizations for cardiovascular reasons, (7) the number of cardiovascular hospitalizations per randomized subject, (8) the number of hospitalizations for atrial tachyarrhythmias, (9) the number of hospitalizations for

[14] *Reviewers' analysis.*

[15] *Includes three subjects not evaluated by the Endpoint Committee.*

[16] *Mf4477summary.pdf*

ventricular tachyarrhythmias, (10) the number of hospitalizations for symptomatic bradycardia, (11) the number of hospitalizations for symptomatic heart block, (12) the number of hospitalizations for myocardial infarction, (13) the number of hospitalizations for unstable angina, (14) the number of hospitalizations for other cardiovascular reasons, (15) the number of hospitalizations for noncardiovascular reasons, (16) the number of hospitalizations for any reason, the number of hospitalizations for any reason per randomized subject, (17) the number of randomized subjects requiring intravenous diuretics during hospitalization, (18) the number of randomized subjects requiring intravenous vasodilators during hospitalization, (19) the number of randomized subjects requiring intravenous positive inotropes during hospitalization, (20) the number of randomized subjects requiring echocardiography during hospitalization, (21) the number of days spent in the hospital for any reason, (22) the number of days spent in the hospital for any reason per randomized subject, (23) the number of days spent in the hospital for any reason per hospitalized subject, (24) the number of days spent in the hospital for any reason per admission for each randomized subject, and (25) the number of days spent in the hospital for any reason per admission for each hospitalized subject. For 15 of these comparisons, there is a nominal p-value displayed. All 25 of these analyses of hospitalization alone suffer from the problem of the competing risk of mortality.

The sponsor's time-to-first-event analyses of death plus hospitalization obviously do not take into consideration the number and duration of hospitalizations. As part of this review, the number of days alive or alive and out of the hospital were analyzed[17]. This was not a pre-specified analysis, but unlike the 25 analyses of hospitalization alone, this analysis does not have a problem with competing risk; it accounts for every subject on every day of study. The results are shown in Table 8.

Table 8. Time alive and unhospitalized

		Placebo N=1133	Carvedilol N=1156
Days at risk	Total	414495	420740
	Per subject	365.8	364.0
Alive	Total	364583	384082
	Per subject	321.8	332.3
Fraction of time alive		0.880	0.913
Alive and unhospitalized	Total	354819	376624
	Per subject	313.2	325.8
Fraction alive and unhospitalized		0.856	0.895

The results are consistent with the preservation of some benefit when all of the time spent dead or hospitalized is counted, but the difference between the groups in days alive and unhospitalized (+12.6) is almost entirely attributable to the difference in days alive (+10.5).

In fact, the magnitude of the effect on "days hospitalized" is in the same direction and is about the same magnitude as the effect on "days dead"; it is simply that the number of "days dead" is about 5 times as large as the number of "days hospitalized". This is simply another way to say that the "death plus hospitalization" end point is mostly driven by, or representative of, effects on mortality.

[17] *Reviewer's analysis of STEND, HOSPEC, and DEMOG datasets. The number of days at risk was calculated as the number of days from randomization to 20 March 2000. Calendar days of hospitalization or death were counted similarly. The Packer manuscript (mf4477summary.pdf) clearly counts days of hospitalization differently, perhaps excluding hospitalizations less than 24 hours.*

In contrast to the US-vs.-non-US comparison for mortality alone, the US had a slightly lower risk ratio for death plus hospitalization, as shown in Table 9.

Table 9. Comparison of US and non-US death plus hospitalization[18]

	US			Non-US		
	Placebo N=233	Carvedilol N=249	Risk ratio	Placebo N=900	Carvedilol N=907	Risk ratio
Death	50	44	0.80	141	88	0.60
Death + CHF-hospitalization	104	77	0.62	259	197	0.72
Death + CV-hospitalization	119	88	0.61	284	229	0.77
Death + any hospitalization	145	122	0.69	365	315	0.83

2.3.2.2 Subjects' global assessment

Subjects responded to the question "In general terms, how do you feel as compared to how you felt at the start of the study (before you began taking the new medication)?" categorically as no change, or slightly, moderately, or markedly worse or improved. For various reasons, including the early termination of the study, only 1633 (71%; similar in both groups) of subjects had any global assessment. The distribution of scores is shown in Figure 8.

APPEARS THIS WAY
ON ORIGINAL

[18] *Reviewers' analysis.*

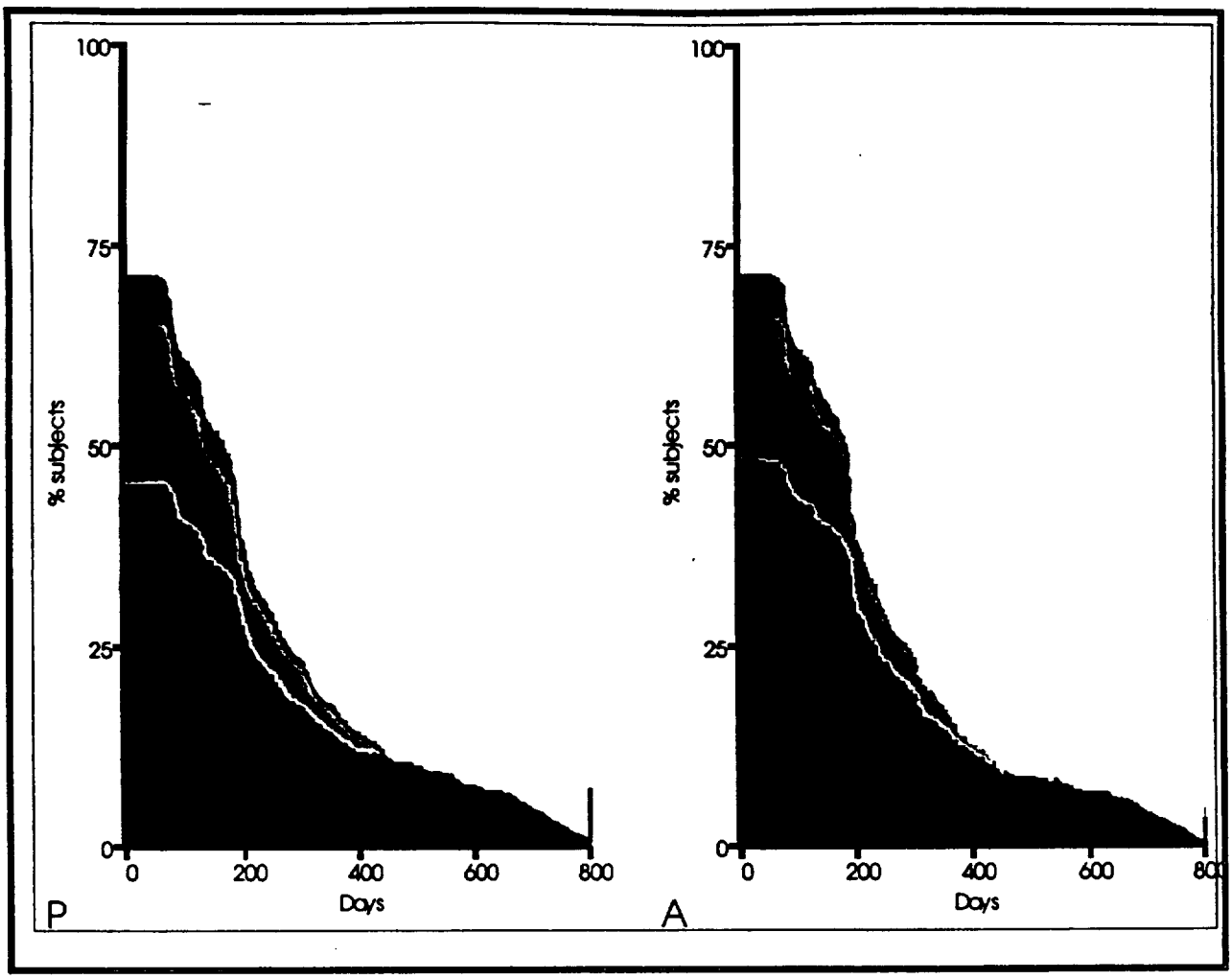

Figure 8. Distribution of global assessment scores by time in study

Panes show the distribution of global assessment scores, as a stacked bar chart, by time in study. P=placebo. A = carvedilol. Subjects with at least one global assessment contributed to the analysis, but the demoninator is always the number of randomized subjects. The score on any given assessment contributes to the distribution on that day and prior days up to the previous assessment, if any. The coarse hatched area closest to the x-axis corresponds to "markedly improved". Successive areas show "moderately improved", "slightly improved", "no change", "slightly worse", "moderately worse", and "markedly worse". Reviewer's analysis of ASSESS dataset.

Because the study was discontinued, the number of subjects with follow-up decreases progressively with time. Because of the differences in mortality, the censoring for global assessment is different in the placebo and active treatment arms. Figure 9 shows a second analysis comparing the distributions of scores among subjects with data at each point in time.

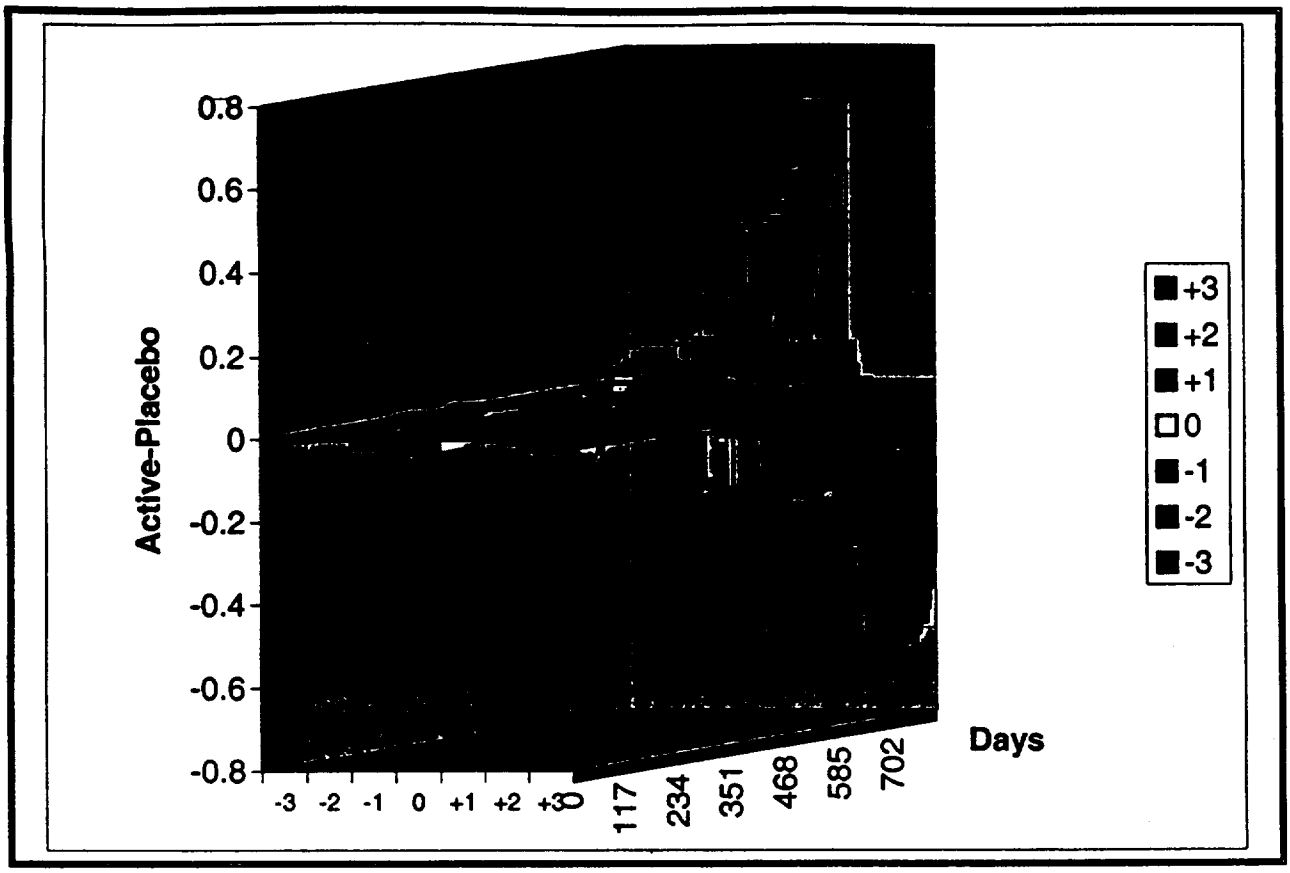

Figure 9. Change in global assessment score distributions with time

Data are the same as in Figure 8. The "+3" curve shows the proportion of carvedilol subjects with data reporting "markedly improved" minus the proportion of placebo subjects with data reporting "markedly improved". The "0" curve is the difference in proportions of subjects with data reporting "no difference". The "-3" curve is the difference in the proportions of subjects reporting "markedly worsened".

The analysis suggest that a small benefit develops (a few percent greater proportion of carvedilol subjects shifting into improved categories) immediately, but there is no consistent pattern after the first few months.

A further analysis of global assessment was based on the SAS dataset ASSESS. From all the available data in the ASSESS, but without imputation, Table 10 was generated.

Table 10. Subject global assessment at 2, 4, and 6 months of maintenance treatment.

		2 months		4 months		6 months	
	N=	Placebo 1133	Carv 1156	Placebo 1133	Carv 1156	Placebo 1133	Carv 1156
Subjects analyzed (N)		754	761	621	630	530	569
Subjects missing visit or score (%)		33	34	45	46	53	51
Percentage of subjects with score	Markedly improved	9	11	8	10	7	10
	Moderately improved	15	16	13	15	12	14
	Slightly improved	17	171	14	12	11	10
	No change	20	18	15	14	13	12
	Slightly worse	4	3	3	2	2	2
	Moderately worse	1	1	1	1	1	1
	Markedly worse	1	0	1	1	1	0
P-value by Mann-Whitney U test		0.076		0.006		0.006	

The results are different from the sponsor's (Table 7 of study report) but show a very similar pattern as theirs. The percentage of subjects with no assessment score was above 30% at 2 months of maintenance period, 45% at 4 months, 50% at 6 months. Hence the p-values, though small, are very difficult to interpret. The reviewers conclude that global assessment improvement should not be included in the label.

2.3.3 Unplanned analyses

2.3.3.1 Worsening heart failure

To further explore how the results in the COPERNICUS study relate to the earlier US development program, an exploratory analysis was undertaken to examine worsening heart failure in COPERNICUS. A time-to-first-event analysis was done on a composite end point similar to the one used in study 240. The criteria for worsening was death, end point committee-adjudicated cardiovascular hospitalization, need for new drug to treat heart failure, or a 50% increase in the dose of a heart failure medication present at baseline. Drugs considered of interest in this analysis are listed in Table 11.

Table 11. Drugs identified as related to treatment of heart failure.[19]

ALACEPRIL	ENALAPRIL	METOPROLOL SUCCINATE
AMILORIDE	ENALAPRIL MALEATE	METOPROLOL TARTRATE
AMILORIDE HYDROCHLORIDE	ENALAPRILAT	MILRINONE
ATENOLOL	EPITIZIDE	MOEXIPRIL
BENAZEPRIL	ETACRYNATE SODIUM	NAFTIDROFURYL
BENAZEPRIL HYDROCHLORIDE	ETACRYNIC ACID	NAFTIDROFURYL OXALATE
BENDROFLUMETHIAZIDE	FOSINOPRIL	NICERGOLINE
BETA-ACETYLDIGOXIN	FOSINOPRIL SODIUM	NICORANDIL
BETAXOLOL HYDROCHLORIDE	FUROSEMIDE	NITROPRUSSIDE SODIUM
BISOPROLOL	GLYCERYL TRINITRATE	PENTOXIFYLLINE
BISOPROLOL FUMARATE	HYDRALAZINE	PERINDOPRIL
BUMETANIDE	HYDRALAZINE HYDROCHLORIDE	PERINDOPRIL ERBUMINE
CANDESARTAN	HYDROCHLOROTHIAZIDE	PIRETANIDE
CANDESARTAN CILEXETIL	HYPERICUM EXTRACT	POTASSIUM CANRENOATE
CAPTOPRIL	IBOPAMINE	PROPRANOLOL
CARVEDILOL	IBOPAMINE HYDROCHLORIDE	QUINAPRIL
CHLOROTHIAZIDE	INDAPAMIDE	QUINAPRIL HYDROCHLORIDE
CHLORTALIDONE	IRBESARTAN	RAMIPRIL
CILAZAPRIL	ISOSORBIDE	RAMIPRILAT
CLOPAMIDE	ISOSORBIDE DINITRATE	SOTALOL
CONVALLARIA GLYCOSIDES	ISOSORBIDE MONONITRATE	SPIRONOLACTONE
DESLANOSIDE	LANATOSIDES	TELMISARTAN
DIGITALIS	LISINOPRIL	TERAZOSIN
DIGITOXIN	LISINOPRIL DIHYDRATE	THEOPHYLLINE
DIGOXIN	LOSARTAN	TORASEMIDE
DIHYDRALAZINE	LOSARTAN POTASSIUM	TRANDOLAPRIL
DIHYDRALAZINE SULFATE	MEFRUSIDE	TRIAMTERENE
DISOPYRAMIDE PHOSPHATE	METILDIGOXIN	TRIMETHYLHYDRAZINOPROPANOATE DIHYDRATE
DOBUTAMINE	METIPRANOLOL	VALSARTAN
DOBUTAMINE HYDROCHLORIDE	METOLAZONE	XANTINOL NICOTINATE
DOPAMINE	METOPROLOL	XIPAMIDE
DOPAMINE HYDROCHLORIDE		

[19] *Table derived from query "qryHFDrugNames":*

The distribution of censoring times for this analysis is shown in Table 12. The distributions are similar.

Table 12. Distribution of time to censoring for worsening heart failure[20]

		Placebo N=1133	Carvedilol N=1156
Time to censoring (days)	99th %ile	767	802
	95th	620	615
	90th	497	529
	75th	326	343
	50th	138	139
	Mean	206	214
	25th	44	44
	10th	21	21
	5th	14	13
	1st	7	5
	Min		

The results of the analysis for time to first worsening heart failure event are shown in Figure 10.

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

[20] *Reviewers' analysis*

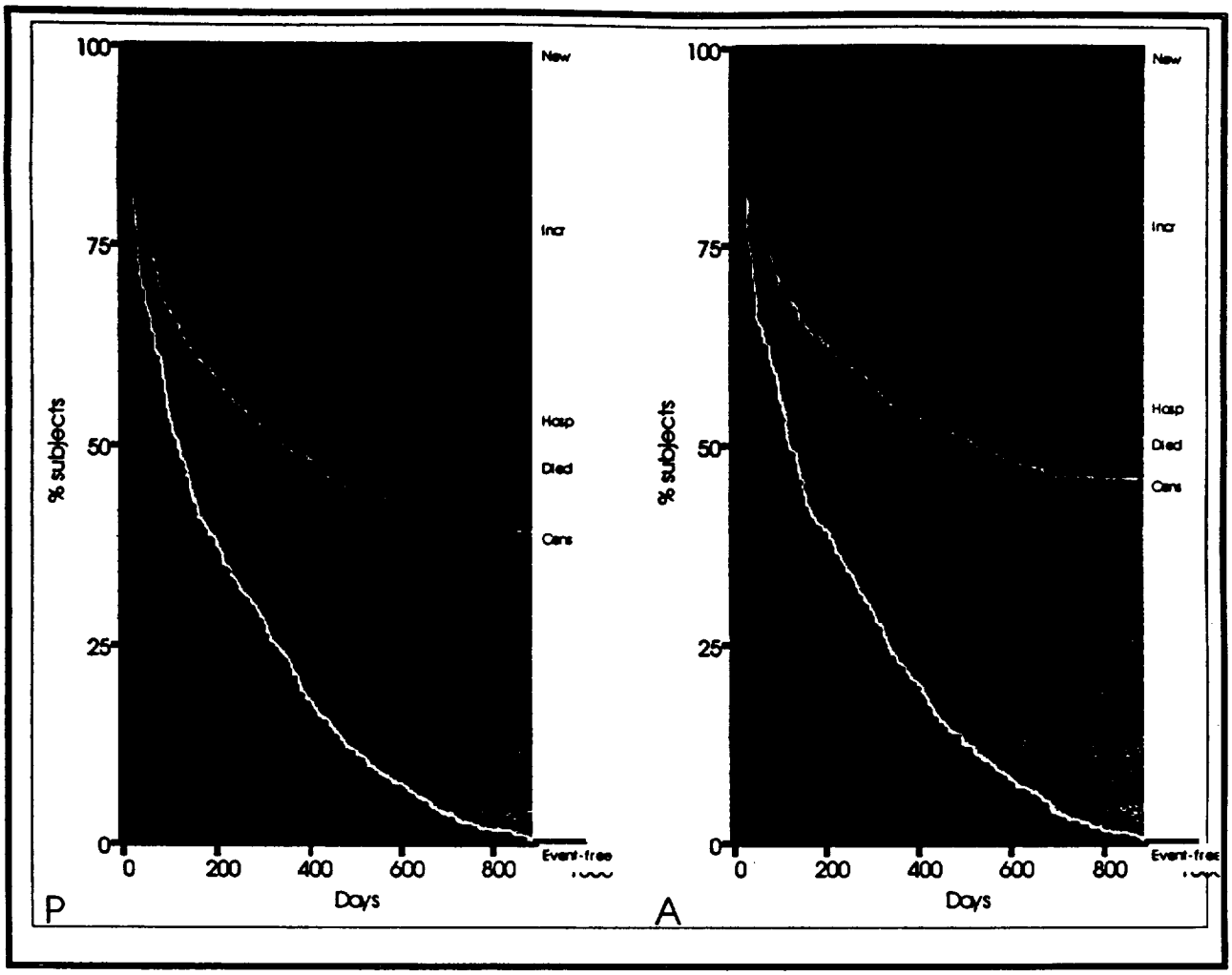

Figure 10. Worsening heart failure

P = placebo. A = carvedilol. All subjects are counted on each day as one of event-free (lowest, darkest region), censored, dead, hospitalized, having an increase in medications, or receiving new heart failure medications. Early termination resulted in many subjects being censored before an event occurred.

The results suggest that the largest difference between treatment groups arose in mortality without antecedent hospitalization or changes in medication. There was little difference in the time to changes (additions or increases) in heart failure medications.

A conventional life table plot of these data is shown in Figure 11.

APPEARS THIS WAY
ON ORIGINAL

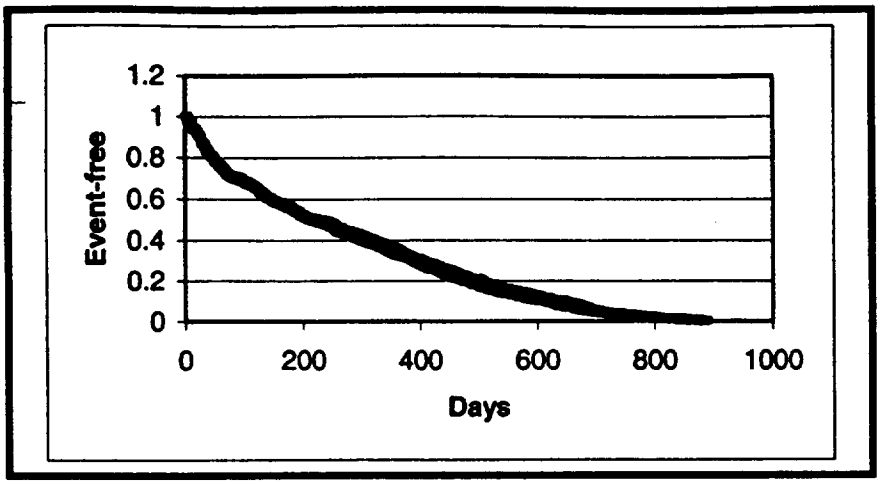

Figure 11. Event-free survival for worsening heart failure

Data are the same as in previous figure. This figure shows time to the first event of death, cardiovascular hospitalization, or change in heart failure medications, censoring at the last available observation. The upper curve is carvedilol. The lower curve is placebo.

2.3.3.2 Concomitant antiarrhythmics

The use of antiarrhythmic drugs was explored because the effect on mortality was mostly attributable to prevention of sudden death. The use of antiarrhythmic drugs was well balanced at baseline, at about 18% of subjects. During the study, there were 257 placebo subjects (23%) and 234 carvedilol subjects (20%) who received antiarrhythmic drugs. Few subjects received an antiarrhythmic drug other than amiodarone. The crude mortality rate in the subgroup receiving antiarrhythmic drugs was 17% on placebo and 16% on carvedilol.

APPEARS THIS WAY
ON ORIGINAL

3 Summary and recommendations

With COPERNICUS, the evidence increases that carvedilol improves outcome in (some) patients with CHF, but specification of the benefit is made more difficult.

There are two labeling decisions that need to be made, regarding mortality and progression of heart failure. In each case, what to say depends upon how similar one believes the populations are in COPERNICUS and the earlier US carvedilol development program. This is because, to the extent that the populations are perceived as similar, one will need to address discrepant findings in the two settings, to decide what is true or likely to be a reliably reproducible effect of carvedilol in heart failure.

3.1 Population in context

3.1.1 Intent

The population studied in COPERNICUS was to have ischemic or non-ischemic heart failure, and a recent ejection fraction (not measured at baseline) <25%. They could be receiving digitalis, diuretics, ACE inhibitor, vasodilators, and nitrates. The expected functional class of these subjects is not specified, but they were expected to have symptoms at rest or with minimal exertion and that they be not hospitalized at the time of enrollment. Exercise capacity was not assessed.

This was clearly expected to produce a population with more advanced heart failure than that in the US portion of the development program that resulted in the approval of carvedilol for heart failure. The previous US program[21] included 4 multicenter studies targeting subjects with ischemic or non-ischemic heart failure, NYHA class II-IV or III-IV (1 study), with ejection fraction ≤35%, receiving diuretics, digitalis, ACE inhibitors, hydralazine, or nitrates, and able to walk 150 m at baseline.

It should not be assumed that the COPERNICUS requirement for symptoms at rest or with minimal exertion precluded being able to walk 150 m; NYHA III-IV subjects were part of the earlier US development program.

3.1.2 Comparison on the basis of data collected

In fact, to the extent they can be compared, these populations were less different than was intended. COPERNICUS did not assess NYHA class or exercise capacity, so one is left with the following potential points of comparison.

3.1.2.1 Concomitant medications

Usage of digitalis, diuretics, and ACE inhibitor were similarly high in both populations.

3.1.2.2 Ejection fraction

The mean historical ejection fraction in COPERNICUS was 20±4%, whereas it was 21 to 23% at baseline in the various US studies[22].

3.1.2.3 Mortality

The placebo-group mortality was between 9 and 10% at 180 days in COPERNICUS (interpolation on Figure 3), but it is harder specify the appropriate comparison from the earlier US program. In study 240, whose primary end point now constitutes the indication for carvedilol in heart failure, the placebo group's 180-day mortality was about 3%; in studies 220 and 221 (with the same enrollment criteria), 180-day mortality was 14% and 7-8%, respectively. The total earlier US program mortality was

---

[21] *The Australia-New Zealand study enrolled a less severe heart failure population, so it is not germane to this discussion.*

[22] *It might also be interesting to compare median ejection fractions, but the data are not available for the earlier program.*

about 7% (including a minor contribution from the sicker target population 239).

The difference in placebo group mortality rates among the US studies is partl attributable to differences in baseline NYHA class. All three studies enrolled s NYHA classes II-IV. The lowest placebo group mortality rate was in study 240, only 15% of subjects were described as NYHA III-IV at baseline. However, the h placebo group mortality rate was in study 220, with 54% of subjects NYHA clas while study 221, with 60% of subjects NYHA III-IV, had intermediate mortality.

With respect to mortality on placebo, COPERNICUS was more like studies 220 a than it was like study 240.

3.1.3 Comparison with MERIT-HF

In comparison, the MERIT-HF (metoprolol) population was probably less severely impaired than that of even the earlier US carvedilol program. MERIT-HF enrolled subjects NYHA II-IV with ejection fraction <40%. Most subjects were NYHA III, but t mean ejection fraction was 28%. Mortality in the placebo group was lower, too, abou 11% at one year.

3.2 Findings in relation to earlier program

3.2.1 Mortality

If, on the basis of considerations described in section 3.1, one believes the population studied in the earlier US development program was clearly and importantly different, one can skip consideration of the mortality effect in the earlier program (it does not apply), and go straight to consideration of the COPERNICUS data in isolation.

3.2.1.1 Mortality findings in the US development program

A decision has already been made that a mortality benefit could not be ascribed to carvedilol on the basis of the earlier development program alone, but, on the whole, mortality was lower on carvedilol. However, "on the whole" was largely driven by the effect in one study, with no other study contributing much (including an intended replicate study). The question is whether the best indication of the effect of carvelilol on mortality in the earlier program is represented by the observations in any one study or the program as a whole.

There was no nominally statistically significant effect on mortality in the Australia-New Zealand study of the least impaired heart failure population. US studies were generally too small to provide meaningful mortality results. Study 220 appeared to show a mortality benefit for carvedilol, but a replicate study (221; with a similar distribution of subjects in various NYHA classes) failed to provide compelling corroboration of the results. Nevertheless, the US program was halted by its program-wide DSMB, because of the apparent mortality benefit. The mortality results from the previous program are shown in Figure 12 (US) and Figure 13 (Australia-New Zealand).

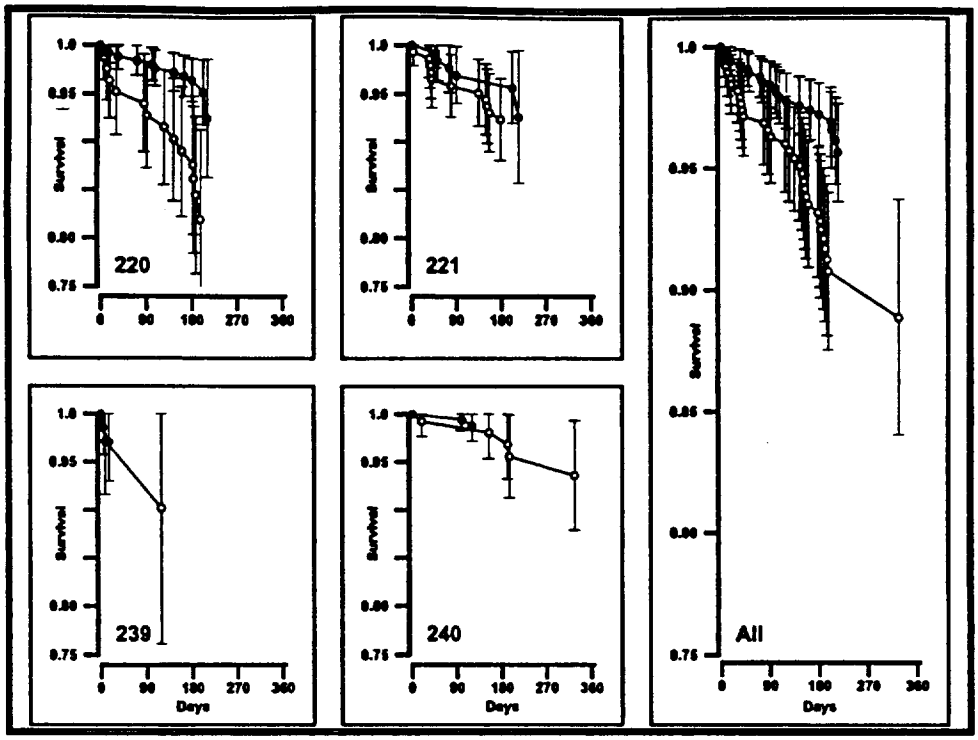

Figure 12. All-cause mortality in original carvedilol in heart failure

Studies 220, 221, and 240 were conducted in subjects with NYHA class II-III heart failure. Study 239 was conducted in subjects NYHA class III. Subjects in these studies had mean EF 21-23%.

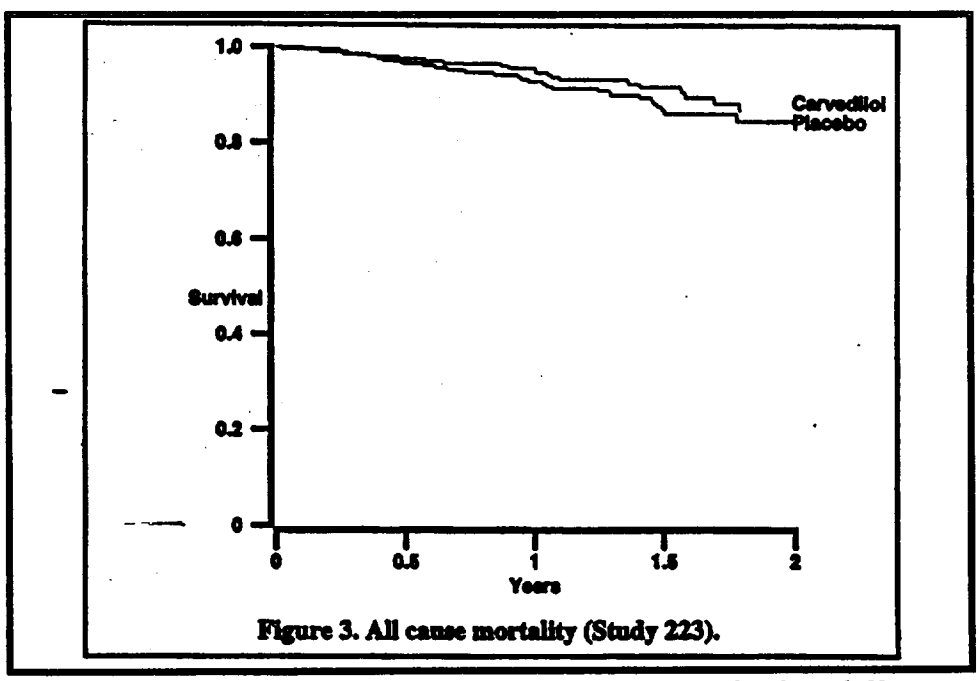

Figure 13. All-cause mortality in Australia-New Zealand study[23]

The Australia-New Zealand study was conducted among subjects with NYHA class I-III (mostly class II) heart failure. The mean EF was 28%.

[23] *From page 5 of the Medical-Statistical review dated 6 February 1997.*

If there is an effect on mortality, study 220 suggests it appears early, and, to the extent that a mortality difference develops in study 221, the effect appears early there, too.

3.2.1.2 Mortality findings in COPERNICUS

Mortality in the carvedilol treatment group of COPERNICUS was reduced by 34% compared with placebo, a finding sufficiently statistically significant to induce the DSMB (again) to halt the study. Although there are some aspects of this finding worthy of discussion, the p-value was low enough to be considered compelling in the absence of other supporting information.

The mortality effect seen in COPERNICUS appears only after 3-4 months, after which the curves diverge for at least 6 or 8 months.

Mortality is mostly related to reduction in the number of sudden deaths, without antecedent worsening heart failure.

Mortality in COPERNICUS is robust for various subgroups. This was observed during unblinded interim analyses of the DSMB.

3.2.1.3 Comparison of mortality in COPERNICUS and the US program

Again, whether any comparison is useful depends somewhat on one's expectations about how similar the findings should have been, and that, in turn, depends upon how similar one thinks the populations are.

The earlier US experience at least suggested a mortality benefit with carvedilol. However, the US experience suggests that the benefit develops sooner after the onset of treatment than in COPERNICUS.

The principal effect on mortality in COPERNICUS was a reduction in sudden deaths. Most of the deaths in the earlier US development program were sudden deaths, but there were too few deaths overall to make a meaningful statement about what types of death were reduced.

3.2.2 Progression of heart failure

Various measures of progression of heart failure can be imagined. The original development program used measurement of the time to the first event of death, hospitalization for heart failure, or significant increase in medications to treat heart failure, so a similar analysis was implemented by the reviewers of COPERNICUS. The "harder" two-thirds of this measure, death plus hospitalization, is worth consideration as well, although it has much less than two-thirds of the power of the former measure.

Again, the decision process bifurcates, depending on one's judgement about how much one can expect the COPERNICUS findings to replicate what was seen in the earlier US development program.

3.2.2.1 Progression in the earlier US development program

Numerous analyses of progression were undertaken with the earlier US development program. This review does not attempt to summarize them or the decision process that led to a regulatory action specifying the current indication for caredilol. The following is the description of the effects on progression from the approved label:

> *Slowing Progression of Heart Failure: One U.S. multicenter study (366 subjects) had as its primary end-point the sum of cardiovascular mortality, cardiovascular hospitalization, and sustained increase in heart failure medications. Heart failure progression was reduced, during an average follow-up of 7 months, by 48% (p=0.008).*
>
> ...

In the Australia-New Zealand study, death and total hospitalization were reduced by about 25% over 18 to 24 months. In the three largest U.S. studies, death and total hospitalization were reduced by 19%, 39%, and 49%, nominally statistically significant in the last two studies. The Australia-New Zealand results were statistically borderline.

3.2.2.2 Progression in COPERNICUS

3.2.2.2.1 Mortality plus hospitalization

The apparent mortality effect carries over into pre-specified secondary analyses of mortality plus hospitalization, most clearly in prolonging the time to the first hospitalizations related to worsening heart failure. However, an unplanned analysis of days alive and out of the hospital shows little incremental benefit compared to days alive, because the number of "days dead" was very much larger than the number of "days hospitalized".

The unplanned analysis has several advantages. It utilizes information about the full period of follow-up. It is not subject to competing risks from treatment group differences in the mortality component; every subject contributes to every time point in the full period of follow-up, as dead, alive and out of the hospital, or alive and in the hospital. Finally, this end point is a better reflection of a benefit appreciable to a patient; time-to-first-hospitalization is only a surrogate for the general need for hospitalization.

The reviewers concur that COPERNICUS showed greater time to mortality or hospitalization, but, despite the unplanned nature of the analysis, the reviewers believe this finding represents weak evidence of an additional clinical benefit.

3.2.2.2.2 Mortality, hospitalization, or need for change in medication

A further unplanned analysis of time to worsening heart failure, similar to one that led to the current indication, reveals little evidence of benefit.

3.2.2.3 Comparison of progression in COPERNICUS and the US program

The assessments of progression of heart failure were not quite the same in Study 240 and the reviewers' analysis of COPERNICUS. Progression in the original program was based on cardiovascular mortality, cardiovascular hospitalization, and sustained 50% increase in heart failure medications. Progression in COPERNICUS was based on all-cause mortality, cardiovascular hospitalization, and any 50% increase over baseline in heart failure medications.

As a follow-up to discussions of the progression of heart failure, the sponsor asserts, in the submission of 1 June 2001, that differences in protocol design prevent one from making a meaningful comparison between COPERNICUS and Study 240. Specifically, they assert [italics added] "investigators were strongly encouraged to proactively *increase* the dose of background medications (particularly diuretics) during the up-titration period of the Copernicus Study; in contrast they were strongly discouraged from doing so in Study 240."

Here are excerpts from the COPERNICUS protocol addressing the adjustment of concomitant medications during up-titration [emphasis added]:

"If the vasodilatory symptoms are moderate in severity, ...the dose of background medications should be reduced to allow the patient to tolerate a higher dose of the study medication at a later time. The investigator may elect to reduce the dose of ACE inhibitor or the dose of diuretic, or may elect to do both.

"If the vasodilatory symptoms are severe (fainting, presyncope, or syncope), the study medication may be temporarily reduced (or stopped) and at the same time the dose of background medications should be reduced. The investigator may elect to reduce the dose of ACE inhibitor or the dose of diuretic, or may elect to do both."

In contrast, here are excerpts from the COPERNICUS protocol addressing the adjustment of concomitant medications for the management of worsening heart failure [emphasis added]:

"If the degree of worsening is mild, the dose of study medication should not be advanced but instead the patient should continue to take the study medication at the dose that was previously well tolerated. The dose of background medications should be increased, including the dose of diuretic or ACE inhibitor or both.

"If the degree of worsening is moderate, the dose of study medication should not be advanced but instead the patient should continue to take the study medication at a reduced dose (lower than that which the patient had previously tolerated). The dose of background medications should be increased, including the dose of diuretic or ACE inhibitor or both.

"If the degree of worsening is severe (pulmonary edema or cardiogenic shock), the study medication should be stopped, and the dose of background medications should be increased, including the dose of diuretic or ACE inhibitor or both.

...

"If a patient experiences an increase in weight (i.e., > 1.5 kg over a period of 2 weeks) without a change in symptoms, the investigator should manage the patient as if they had worsening of heart failure that was mild in severity (using the algorithm provided above). The dose of diuretic drug should be increased until the baseline weight is restored."

Thus, the COPERNICUS protocol called for decreases in the dose of diuretics as necessary to make carvedilol more tolerable, and increases in dose as needed to treat worsening heart failure. Since the reviewers' analysis of changes in heart failure medications was compared with baseline, the analysis is biased in favor of carvedilol group, because this group was more likely to have doses adjusted downward to compensate for vasodilatory effects.

The population in COPERNICUS was less like study 240 than other parts of the earlier US development program, but some of the confidence in the findings of study 240 came from consistencies with more severely ill subjects of the other studies. To the extent that one finds overlap among the various populations, the lack of an effect on progression in COPERNICUS is disturbing.

3.2.3 Summary

There are similarities in populations studied in COPERNICUS and the earlier development program. These similarities in population lead one to seek corroboration for the effects that various parties discerned among the earlier data. The search is not altogether successful. There is a mortality difference in both sets of data, but the time course is different. There is an effect on delaying time to the first hospitalization in COPERNICUS, but it does not confer any increment in "quality time" alive and out of

the hospital. Finally, the effect of carvedilol on progression of heart failure is not sustained in the much larger and longer COPERNICUS study; this failure materially undermines confidence that slowing of progression represents a true treatment effect of carvedilol in patients with heart failure.

3.3 Recommendations

The reviewers' proposed changes to the label are summarized as follows:

The reviewers conclude that the mortality effects, disparate as they are, have been confirmed and that the indications for carvedilol should acknowledge this action.

The description of the COPERNICUS effects on mortality plus hospitalization should be described in the results of the study, but with language that says that delaying hospitalization did not result in more days alive and unhospitalized. Effects on hospitalization should not be part of the indications for the use of carvedilol in heart failure.

Failure of COPERNICUS to demonstrate even a trend for carvedilol to increase the time to worsening heart failure should result in the elimination of this existing claim.

APPEARS THIS WAY
ON ORIGINAL

4 Labeling

This section outlines areas in which either the sponsor or the reviewer have proposed changes to the approved label. The sponsor's proposed erasures are struck through in black. The sponsor's proposed insertions are underlined in black. The reviewers proposed changes are noted in the right margin (if short) or interleaved in indented paragraphs.

DRAFT

...

CLINICAL TRIALS

Congestive Heart Failure

14 pages redacted from this section of the approval package consisted of draft labeling

pp. 31-44

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

Norman Stockbridge
6/7/01 12:01:45 PM
MEDICAL OFFICER

Jim--Add G Chi if necessary

James Hung
6/7/01 12:26:01 PM
BIOMETRICS

George Chi
6/12/01 12:43:33 PM
BIOMETRICS