CASE 11

The Data Monitoring Experience in the Candesartan in Heart Failure Assessment of Reduction in Mortality and Morbidity Program*

Stuart Pocock
Duolao Wang
Lars Wilhelmsen
Charles H. Hennekens

ABSTRACT

The Candesartan in Heart Failure Assessment of Reduction in Mortality and morbidity (CHARM) program was designed as three separate randomized trials comparing candesartan with placebo in patients with chronic heart failure (CHF) who (1) were intolerant to angiotensin converting enzyme (ACE)-inhibitor and had left ventricular ejection fraction (LVEF) ≤ 40%, (2) were on ACE-inhibitor and had LVEF ≤ 40% or (3) had LVEF > 40%. CHARM provides an interesting example of the challenges faced by a Data and Safety Monitoring Committee (DSMC).

While the primary efficacy endpoint for each component trial was cardiovascular (CV) death or hospitalization for CHF, the primary outcome for the overall program was all-cause mortality. The DSMC received monthly safety reports and also met every six months (seven times in all) to review interim reports. Statistical stopping guidelines were predefined for all-cause mortality in the overall program. The overarching principle of the DSMC was proof beyond a reasonable doubt that would be likely to influence clinical practice.

There were significant treatment differences in all-cause mortality at several interim analyses, and the statistical stopping guideline was reached on one occasion. The DSMC consistently recommended that the program

* This paper first published in the Am Heart J 2005; 149:939–943.
continue as planned. The final published results for all-cause death over a median 3.1 years were a 9% reduction in hazard (95% CI 0% to 17%, \( p = 0.055 \)), whereas for CV death or hospitalization for CHF there was a 16% reduction in hazard (95% CI 9% to 23%, \( p < 0.0001 \)). Subsequent exploratory analyses suggest that the hazard reduction in CV death was more marked in the first year after randomization, and that, if real, this apparent treatment-time interaction offers a plausible explanation for why the interim mortality data showed statistically more extreme findings than the overall final results.

The DSMC experience in the CHARM program illustrates the importance of continuing a trial to its scheduled completion unless there is proof beyond reasonable doubt that would influence clinical practice rather than strict reliance on a statistical stopping guideline.

**INTRODUCTION AND BACKGROUND**

Angiotensin-receptor blockers such as candesartan offer the potential to improve clinical outcomes in heart failure patients as alternatives or adjuncts to those seen with angiotensin-converting enzyme (ACE) inhibitors. Accordingly the CHARM program\(^1\) was designed as three independent randomized double blind trials comparing candesartan with placebo in three populations of patients with symptomatic heart failure:

1. **CHARM—Alternative** patients (N = 2,028) had a left ventricular ejection fraction (LVEF) \( \leq 40\% \) and were not on ACE inhibitor because of previous intolerance.\(^2\)
2. **CHARM—Added** patients (N = 2,548) also had LVEF \( \leq 40\% \) and were being treated with an ACE inhibitor.\(^3\)
3. **CHARM—Preserved** patients (N = 3,023) had LVEF >40%.\(^4\)

The primary endpoint for each trial was CV death or hospitalization for worsening CHF and each required sample size was based on power calculations for this endpoint. The overall program was designed to evaluate all-cause mortality in the broad spectrum of symptomatic heart failure patients, with the overall sample size (N = 6,500) equal to the sum of all three trials.\(^1\) With an estimated overall annual mortality in the placebo group of 8% the program had over 85% power to detect a 14% reduction in mortality at two-sided 5% significance based on a logrank test.

All three trials were done at the same 618 sites in 26 countries. The CHARM program exceeded its recruitment goal of 6,500 by enrolling 7,599 patients between March 1999 to March 2001, who were followed for a minimum of two years. Hence, all follow-up was concluded on March 31, 2003, resulting in a median duration of 3.14 years. The final results were