BiDil for Heart Failure in Black Patients: The U.S. Food and Drug Administration Perspective

Robert Temple, MD, and Norman L. Stockbridge, MD, PhD

Critics of the U.S. Food and Drug Administration (FDA) approval of the fixed combination of hydralazine hydrochloride, 37.5 mg, and isosorbide dinitrate, 20 mg, for treating heart failure in black patients have suggested that data were insufficient to distinguish treatment effects in black and white people; that distinctions based on race, rather than pathophysiology, were scientifically unreasonable; and that a “race-based” approval could be a commercial ploy to avoid a more expensive and prolonged full evaluation of a drug. The criticisms acknowledge that data supporting the approval came from a well-designed clinical trial in which self-identified black patients with heart failure who took hydralazine hydrochloride–isosorbide dinitrate with standard therapy experienced a statistically significant 43% (95% CI, 11% to 63%) reduction in mortality compared with those who took only the standard therapy. The criticisms do not always recognize that the decision to conduct the trial in only black patients reflected careful analyses of 2 previous trials in racially mixed patient populations that compared hydralazine hydrochloride–isosorbide dinitrate with placebo or with enalapril. Both trials showed little or no overall effect of hydralazine hydrochloride–isosorbide dinitrate in the mostly white patient population but hinted at a substantial effect in subsets of black patients. Perhaps most critically, the criticisms do not appreciate the urgency of strong scientific evidence of a substantial survival benefit in black patients. A serious attempt to avoid race-based approval by mandating study of a mixed population to identify a possible white patient-responder subset, particularly without a plausible hypothesis as to what that subset might be, would have required years of work, many thousands of patients, and wholly unreasonable delay in approval of a treatment whose effectiveness had been well-documented in the group for which it was intended.


For author affiliations, see end of text.

C onsiderable discussion has followed the FDA approval in June 2005 of BiDil, a fixed-dose combination of isosorbide dinitrate, 20 mg, and hydralazine hydrochloride, 37.5 mg, for use in “the treatment of heart failure as an adjunct to standard therapy in self-identified black patients to improve survival, to prolong time to hospitalization for heart failure, and to improve patient-reported functional status” (1). Approval followed the unanimous recommendation of the FDA’s Cardiovascular and Renal Drugs Advisory Committee, with 7 of 9 members supporting the race-specific label and 2 of 9 members urging a broader claim (2). While the importance of the effect shown in black patients has generally been recognized (3), critics have asked whether data were adequate to distinguish the effects of hydralazine hydrochloride–isosorbide dinitrate in black and white patients; whether commercial rather than medical considerations led the drug manufacturer to restrict the critical clinical trial of BiDil to an entirely black patient population (3, 4); whether the FDA should have allowed or encouraged such a trial; whether distinguishing drug responses in black and white patients by race rather than pathophysiology is scientifically reasonable (3–5); and whether a “race-based” drug approval can be abused, leading to suggestions of racial inferiority or stereotyping. These questions and others are worthy of discussion, but they did not cause the FDA to doubt that BiDil should be approved.

The FDA Perspective

1. Data from 3 clinical trials showed dramatic effectiveness of hydralazine hydrochloride–isosorbide dinitrate in black patients and supported a differential effect in black and white patients.

There was very strong evidence that hydralazine hydrochloride–isosorbide dinitrate was extremely effective in self-identified black patients and considerable evidence that the effects were far smaller, if present at all, in white patients. This evidence has been substantially understated by many commentators and critics (3–6).

The principal trial of BiDil was the African American Heart Failure Trial (A-HeFT) (7), a randomized comparison of hydralazine hydrochloride–isosorbide dinitrate and placebo in self-identified black patients with New York Heart Association (NYHA) class III or IV (mostly III) heart failure. Treatment was added to standard therapy (94% of patients received diuretics, 87% received β-blockers, 93% received angiotensin-converting enzyme [ACE] inhibitors or angiotensin II blockers, 62% received digitalis glycoside, and 39% received aldosterone antagonist).

The A-HeFT investigators used a complex weighted combination of all-cause death, first hospitalization for heart failure, and change in quality of life at 6 months as the primary end point, with the individual components specified as secondary end points. The study was terminated—after 1050 patients had been randomly assigned—on...
the basis of a survival advantage (a 43% [95% CI, 11% to 63%] reduction in mortality) in the hydralazine hydrochloride–isosorbide dinitrate group. The primary composite end point showed a significant effect ($P < 0.021$), but the more critical results were mortality rate and time to first heart failure hospitalization. Both outcomes statistically significantly improved with treatment (Table 1).

The decision to conduct A-HeFT in black patients was supported by 2 earlier well-controlled studies that had strongly suggested a differential mortality effect of hydralazine hydrochloride–isosorbide dinitrate in black and white patients. Two Veterans Administration Cooperative Vasodilator Heart Failure Trials (V-HeFT) tested the drugs in patients with predominantly NYHA II and III heart failure. The first trial (V-HeFT I [8]) compared hydralazine hydrochloride–isosorbide dinitrate with placebo (and prazosin, not further discussed), while the second trial (V-HeFT II [9]) compared hydralazine hydrochloride–isosorbide dinitrate with enalapril, which was by then an established heart failure treatment. Table 2 shows the results of these studies. The first trial showed a nearly statistically significant effect of hydralazine hydrochloride–isosorbide dinitrate on survival, and the second trial showed a nearly statistically significant advantage of enalapril over hydralazine hydrochloride–isosorbide dinitrate. Thus, neither study showed an effect of the combination in the general patient population.

Post hoc subset analyses by race, however, indicated that responses were not uniform in racial groups. In V-HeFT I, there was a large reduction in mortality in the small subset of black patients ($n = 128$) that just reached nominal statistical significance (although it was not corrected for multiple end points). There was a much weaker favorable trend in the larger subset of white patients. In V-HeFT II, hydralazine hydrochloride–isosorbide dinitrate was statistically significantly inferior to enalapril in white patients ($P = 0.02$). The magnitude of the difference was not far from what one might expect in a comparison of enalapril and placebo. In contrast, the mortality rates in the enalapril and hydralazine hydrochloride–isosorbide dinitrate groups were almost identical in black patients. Although no formal noninferiority analysis was conducted, the effect size in black patients, in a setting in which enalapril was almost superior to hydralazine hydrochloride–isosorbide dinitrate in white patients, suggested that the trial had assay sensitivity: the ability to distinguish effective treatment from ineffective treatment (10).

We (at the FDA) were asked to approve hydralazine hydrochloride–isosorbide dinitrate for black patients on the basis of the post hoc racial analyses of V-HeFT I and II. Although we did not consider these analyses to be a sufficient basis for approval, the substantially different effect of hydralazine hydrochloride–isosorbide dinitrate in white and black patients in 2 studies led us to believe that there was a good case for conducting an additional study in black patients. We also believed that a study of reasonable size in a population of both black and white patients would have little chance of detecting a treatment effect in white patients, particularly if the treatment were added to current best therapy, which already reduces mortality substantially. A reasonably powered ($\beta = 80\%$) study in white patients, assuming an effect size of, say, 15%, would require about 16,000 patients if placebo mortality were similar to that observed in A-HeFT. The FDA, therefore, advised NitroMed that the demonstration of a mortality benefit in self-identified black patients could be a basis for marketing approval of hydralazine hydrochloride–isosorbide dinitrate.

The mortality benefit of BiDil in black patients is, thus, supported by 3 well-controlled studies: most convincingly in A-HeFT and V-HeFT I and II by an informal noninferiority finding. Approval of BiDil was not based on a single trial where all data came from the black patient population, as has been suggested. The FDA’s encouragement of A-HeFT, a single-population trial, arose from recognition that a larger study of black and white patients was not likely to yield any additional useful information.

2. Not understanding the reasons for the difference in treatment effect by race did not justify withholding the treatment from those who could benefit from it.

Race or ethnicity is clearly a highly imperfect description of the genomic and other physiologic characteristics that cause people to differ, but it can be a useful proxy for those characteristics until the pathophysiologic bases for observed racial differences are better understood (11). The history of treating heart failure should give pause to anyone who imagines that developing effective therapies requires an understanding of pathophysiology or that pathophysiologic-based expectations reliably predict outcomes. Drugs that were once contraindicated in heart failure ($\beta$-blockers)

### Table 1. African American Heart Failure Trial Results*

<table>
<thead>
<tr>
<th>End Point</th>
<th>Hydralazine Hydrochloride–isosorbide Dinitrate Group (n = 518), %</th>
<th>Placebo Group (n = 532), %</th>
<th>Hazard Ratio (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>6.2</td>
<td>10.2</td>
<td>0.57 (0.37–0.89)</td>
<td>0.012</td>
</tr>
<tr>
<td>First CHF hospitalization</td>
<td>16.4</td>
<td>24.4</td>
<td>0.61 (0.46–0.80)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* CHF = congestive heart failure.
have become standard treatment and the most logical treatments for heart failure, inotropic agents to strengthen the weakened heart muscle, have neutral effects on mortality at best (12) and have more often proved to be lethal (13–17). The effectiveness requirement of the Federal Food, Drug, and Cosmetic Act asks only for evidence that a drug will have its claimed effect, without reference to why it has that effect. The idea that the FDA should have waited to approve hydralazine hydrochloride–isosorbide dinitrate until we understood why white and black patients differ in response, despite the 3 well-controlled trials suggesting or showing a mortality benefit in black patients, has led to peculiar proposals.

For example, Avorn (5), unimpressed by the 43% mortality risk reduction demonstrated in A-HeFT and apparently believing that the racial difference hypothesis was based on a post hoc analysis of a single trial, thought that “[t]his interesting observation could have been enormously important in helping us understand the pathophysiology of [heart failure] in a particularly vulnerable population.” He suggested that instead of FDA approval of BiDil, a plausible next step would have been to test the racial difference hypothesis in a controlled trial that enrolled both black and white patients to look for differences in outcomes and predictors of those differences, including “genetic markers, self-identified race, diet, and other risk factors” (4). Understanding pathophysiology is good, of course, but it ranks well behind a documented survival effect in importance, and the suggested study is more or less the same as V-HeFT I. Without any plausible hypotheses (genetic, dietary, and other risk factors) about which white patients might respond, the study would have had little chance of revealing predictors of racial differences. Even if anyone were willing to perform such a study, we believe that spending years exploring the basis for the observed black–white difference before approving hydralazine hydrochloride–isosorbide dinitrate for the population in which it dramatically reduced mortality would not have been responsible.

3. Race and other demographic characteristics have long been important to consider in analysis of trials and as a matter of equity and justice.

Since the early 1980s, there has been substantial concern about inadequate representation of women, elderly people, black people, and other groups in the drug development process. The interest was partly ethical, reflecting concerns about unequal access to potentially valuable treatments, but it was far more a concern that lack of participation of these subgroups would lead to incorrect conclusions for those groups about benefits or adverse effects of treatments. Although, on examination, participation of black people and both sexes approximately reflected the prevalence of the conditions being treated in the population (18), it was found that clinical databases in new drug applications were not examined for potential differences in response among demographic subgroups, at least through the 1980s.

By 1993, the FDA had finalized specific guidelines on assessment of elderly persons (19) and of both sexes (18) in drug development and more general guidelines on evaluation of demographic subgroups for safety, effectiveness, and dose-response in new drug applications (20). Finally,
the regulations describing new drug application submissions (21) were changed in 1998 to require analyses of safety and effectiveness in demographic subsets of the patient population. Recent changes in labeling (package insert) regulations (22) also describe the need to include pertinent demographic subset information in many sections of labeling.

The FDA was hardly alone in this interest in possible differences among demographic subsets of the population. In 1993, Congress passed the National Institutes of Health (NIH) Revitalization Act (23), asking the NIH to ensure that inclusion of women and minorities in phase 3 trials was sufficient to allow for valid analyses of differences in intervention effect. A conspicuous example of the effect of the law is the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) (24), a 40,000-patient trial of blood pressure–lowering and lipid-lowering interventions in which the goal patient population was to be at least 55% black.

Finally, despite legitimate concerns about overinterpreting subset analyses (25), published reports of outcome studies in recent years almost always show forest plots of effects in a wide variety of subsets, invariably including demographic subsets (such as sex, age, and race) and other subsets (for example, disease severity and concomitant illness). These presentations illustrate the growing interest in possible subset differences in response, even where these cannot be pathophysiologically explained. Also, epidemiologic studies regularly examine racial aspects of disease prevalence and outcome. Bloche (3) notes, for example, that “[t]here is wide agreement that blacks die from heart failure at rates disproportionate to those among whites.” One might ask why, if etiology and prevalence of diseases can be racially linked, a difference in treatment response would be surprising, even if we did not understand the reason for it. Given the long history of urgent interest in searching for racial and other demographic differences, which surely accepted the possibility that such differences might be discovered, it seems surprising that there would be so much discomfort when one was found.

4. Regulatory and other concerns associated with drug approval for narrow patient populations did not justify withholding BiDil from those who could benefit from it.

The growing interest in targeting therapy to specific subgroups (“individualization”) raises legitimate questions about implications for the broader population, drug manufacturer incentives, and possible stigmatization of a target group. Not all the questions are yet answerable, but none has justified denying the benefits of hydralazine hydrochloride–isosorbide dinitrate to the subpopulation in which it was shown to be effective.

Approval of a drug for a specific subgroup means that its use in other groups would be considered “off-label.” Although off-label use is not barred by law, third-party payers may not pay for such use and, of course, manufacturer promotion must be directed at the approved popula-

tion. These limitations would represent a loss if, in fact, the drug were actually more broadly effective. The FDA has, therefore, encouraged broad inclusion of patients in trials. In that case, if treatment is ultimately directed toward a group that benefits most, this is done with the knowledge that the treatment performs less well in other groups. We believe that this describes the situation with hydralazine hydrochloride–isosorbide dinitrate. But there could be cases in which little information was available for the “other” group, a growing possibility in an era of “targeted” therapy. When a therapy is shown to be effective for a responsive subgroup, critical questions include how much data should be expected on the drug’s effects in other groups; how small an effect needs to be detected or excluded in those groups; when should the data on those other groups be expected (before or after approval and how long after approval); and to what extent the FDA can insist on the conduct of further studies. These questions are unanswered because we are still in the early days of individualization, but they are being actively considered.

We believe that, in most cases, interest in individualized therapy need not conflict with the desire for broad experience and information about the effects of drugs in a range of patients. Even if therapy is directed at a responsive subgroup, providing adequate directions for clinical use in the drug label—a requirement of law—will generally call for a reasonable amount of data on excluded patient populations or clear evidence that the treatment cannot work in those people. In the case where a drug has shown an important effect in a particular group, however, it seems hard to argue that the group can ethically be denied the therapy while investigators seek to determine whether the effect applies more broadly. In that situation, data on the effect of the drug in other populations would often come from phase 4 (postmarketing) studies. What one hopes, of course, is that the pathophysiologic basis for a differential response can be defined in future studies, allowing rational patient selection.

Fortunately, in the case of hydralazine hydrochloride–isosorbide dinitrate, we did not need to face the most difficult aspects of this question because data indicating a much smaller effect in white patients were available at the time of marketing. Whether there is any effect in white patients, and exactly who the responsive white population might be, remains to be determined. Given the lack of any plausible predictor of such a responsive subset and the massive study needed to explore effectiveness in an unselected white population, the FDA did not seek specific commitments from NitroMed to conduct studies in white patients, although the company has expressed interest in finding a pharmacologic response predictor.

Some critics have suggested that the FDA approval of drugs in narrow subgroups will allow drug manufacturers to “get off cheap,” encouraging them to seek out narrow niche populations that are easy to study and suggest novelty, and that other populations will be deprived of the
opportunity for benefit. The FDA does not regulate the economics of drug development and generally does not evaluate drug manufacturers’ motives, but the issues are complex. There is great interest in efficient drug development, and one step toward efficiency is studying drugs in patient populations with high event rates or greater responsiveness (26). This thought is not new. The ability to enrich populations to be studied has, for example, facilitated development of such critical treatments as ACE inhibitors for heart failure. The first mortality study of enalapril (Cooperative North Scandinavian Enalapril Survival Study [CONSENSUS]) enrolled an extremely sick population with high mortality, allowing a successful study in just 253 people (27). When these results were added to the labeling for enalapril, we did not know the effect of the drug in less ill patients. Much larger studies in such patients were performed later, but the importance of the CONSENSUS finding can hardly be overstated. In some cases, there may be no reason to expect anyone to study the broader patient population until an effect can be shown in some patient population, providing “proof of concept.” While we share the community’s interest in broadly developing important new therapies, we also would not want to stifle innovation and efficient studies and deprive the community of valuable treatments. Again, this issue did not arise for hydralazine hydrochloride–isosorbide dinitrate because the broader patient population had been studied and the white patients clearly had a small response at best. There was little likelihood that BiDil’s developer would have been willing to conduct the 16 000-patient study that was needed to show a small effect in white patients, and there was no evidence of interest in such a study by any other private or government sponsor before A-HeFT.

While the FDA cannot free a manufacturer from the obligation to provide legally sufficient data to support drug approval, judgments are involved in determining what constitutes sufficient data. The striking effects in black patients in A-HeFT and V-HeFT I, the need for heart failure treatments for black patients, and the substantial delay involved in conducting an all-race study—if it could have been done at all and the near certainty that it would not be done—all argued strongly for the path that the FDA took.

Finally, the possibility that drug approval for a particular subgroup could lead to stigmatization of that subgroup has led Kahn (4) to suggest that evidence of a genetic basis for a racial distinction must clearly be shown before a race-specific approval, even in the face of compelling benefit in one race. Aside from the dubious ethics of that proposal, the concern itself seems far-fetched. Differences in drug response by sex (for example, women’s greater susceptibility to torsade de pointes), race (for example, reduced blood pressure response to ACE inhibitors and angiotensin II antagonists and greater susceptibility to ACE inhibitor–induced angioedema in black people [28]), and age (too many to count) are reasonably well-established, and demographic differences in prevalence and outcome of disease whose biological or other basis we do not understand are countless. Adverse consequences of these observations have not been identified. Indeed, some of these observations have led to appropriate efforts to improve health care delivery.

**Conclusions**

The FDA approval of a fixed combination of hydralazine hydrochloride–isosorbide dinitrate to treat heart failure in self-identified black patients was a scientifically reasonable, data-based decision, one that provided a major benefit in a group that is particularly burdened by congestive heart failure. The evidence of benefit in black patients is very strong, and the evidence that white patients have less, if any, benefit, is also strong. We hope that further research elucidates the genetic or other factors that predict the usefulness of hydralazine hydrochloride–isosorbide dinitrate. Until then, we are pleased that one defined group has access to a dramatically life-prolonging therapy.

From the U.S. Food and Drug Administration, Silver Spring, Maryland.

**Potential Financial Conflicts of Interest:** None disclosed.

**Requests for Single Reprints:** Robert Temple, MD, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993-0002; e-mail, robert.temple@fda.hhs.gov.

Current author addresses are available at www.annals.org.

**References**


21. 21 CFR §314.50.


Current Author Addresses: Drs. Temple and Stockbridge: U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993-0002.