Published in final edited form as:
J Law Med Ethics. 2008 ; 36(3): 449–457. doi:10.1111/j.1748-720X.2008.290.x.

Flaws in the U.S. Food and Drug Administration's Rationale for Supporting the Development and Approval of BiDil as a Treatment for Heart Failure Only in Black Patients

George T. H. Ellison, **Jay S. Kaufman**, **Rosemary F. Head**, **Paul A. Martin**, and **Jonathan D. Kahn**
, is a Professor of Health Sciences at St. George's, University of London, Cranmer Terrace, London. ***Jay S. Kaufman, Ph.D.****, is an Associate Professor of Epidemiology at the School of Public Health, University of North Carolina at Chapel Hill.* ***Rosemary F. Head, Ph.D.****, is the Senior Research Fellow in Epidemiology and Biostatistics at St. George's, University of London, Cranmer Terrace, London.* ***Paul A. Martin, Ph.D.****, is a Reader in Science and Technology Studies at the Institute for Science and Society, University of Nottingham, U.K.* ***Jonathan D. Kahn, J.D., Ph.D.****, is an Assistant Professor of Law at Hamline University School of Law in St. Paul, MN.*

Strengths and Weaknesses of the FDA's Rationale for Approving BiDil

There is likely to be widespread agreement with much of the FDA's rationale for approving BiDil (a combination of hydralazine hydrochloride and isosorbide dinitrate; H-I) as a treatment for heart failure.[1] In particular, most would agree that the evidence of effectiveness provided by the African American Heart Failure Trial (A-HeFT) is compelling.[2] Likewise, few health scientists would believe that it is either necessary or responsible to withhold therapies such as BiDil from those who might benefit until there is a full understanding of how they work.[3] And although there is substantial concern that biomedical differences between racial groups are routinely misinterpreted as evidence of innate genetic differences[4] (hence Jonathan Kahn's call for all such claims to be supported by genetic evidence),[5] most would concede that using race as a "descriptive" variable[6] can help identify differences in health and access/response to treatment that might warrant further investigation or intervention.[7]

A recent article by Robert Temple and Norman Stockbridge recounted the history and logic of the FDA's decision,[8] and it is now clear from this account[9] that the FDA's "encouragement" of A-HeFT, and its subsequent approval of BiDil, relied on *post hoc* subgroup analyses of the two earlier H-I trials (the Veterans' Administration Cooperative Vasodilator Heart Failure Trials: V-HeFT I and II).[10] At best, such analyses could only ever provide limited evidence for racial differences in response to H-I,[11] and they could not justify a "single population trial" such as A-HeFT (to examine the effect of H-I in only black patients) given the substantial changes in routine therapy for heart failure that had occurred since the V-HeFT I and II results were published (in 1986 and 1991, respectively).[12] At the same time, Temple and Stockbridge's account — of the FDA's "perspective" on the circumstances leading up to the approval of BiDil — appears to dismiss the serious ethical concerns that arise when the development of group-specific therapies invoke race as a discrete marker of innate difference, and are subject to commercial incentives rather than scientific evidence or therapeutic imperatives.[13]

Kirsten Bibbins-Domingo and Alicia Fernandez published a response to the article by Temple and Stockbridge,[14] but for the most part, they did not address these methodological or ethical concerns. Instead, their response focused on the way the FDA the need to address racial disparities in health, and went on to question whether the FDA's decision: (1) had "endorsed

a biological model of race"; (2) elevated "biological difference in medication response to an important cause of [racial] health disparities without evidence"; and (3) represented a "perverse" justification for treating poverty-related health disparities with expensive medicines. These are certainly important issues to consider when mapping out the likely consequences of the FDA's decision to approve BiDil. But given the FDA's stated desire to improve the efficiency of drug development and the hype surrounding "personalized medicine" (based on matching individual genotypes to the most appropriate therapies), it is equally important to expose the fallibility of the (*post hoc*) subgroup analyses this approach is likely to involve and the dangers posed by commercial incentives to adopt this approach. The discussion that follows therefore sets out to do the following: (1) demonstrate the weaknesses of the scientific evidence used to support the approval of BiDil only for use in black patients; (2) call for further analysis of the V-HeFT I and II data which might clarify whether responses to H-I vary by race; and (3) evaluate the consequences of commercial incentives to develop racialized medicines.

Inherent Weaknesses of the V-HeFT *Post Hoc* Subgroup Analyses

The *post hoc* subgroup analyses of V-HeFT I and II suffer from the same potential problems as those faced by all *post hoc* subgroup analyses of randomized controlled trials: a loss of statistical power and the potential for covariate imbalances.[15] Given that *post hoc* subgroup analyses require the study population to be divided into two or more subgroups, it is rare that such analyses involve sufficient participants in each subgroup to satisfy the sample sizes needed for evaluations of stratum-specific effects. As a result, multiple *post hoc* subgroup analyses increase the risk of generating statistically significant yet arbitrary group differences by chance alone (i.e., type 1 errors) while missing genuine differences between groups (i.e., type 2 errors). Likewise, because *post hoc* subgroup analyses are not part of the original trial design, it is unlikely that subgroup participants are randomly allocated to the intervention and control groups. Any subgroup effects are therefore susceptible to covariate imbalances — i.e., due to chance or systematic differences in the prevalence of confounders amongst participants allocated to the intervention or the control.

It is now clear that the FDA's "encouragement" of A-HeFT, and its subsequent approval of BiDil, relied on *post hoc* subgroup analyses of the two earlier H-I trials

Sample Size Limitations of the V-HeFT Subgroup Analyses

No sample size calculations were reported for V-HeFT II,[16] but for V-HeFT I the researchers estimated that a sample size of 720 participants would be required to provide an 84% chance of detecting a reduction in mortality with prazosin (the second of two drugs tested in this trial) or H-I of 33% below that observed with the placebo.[17] In practice, V-HeFT I failed to recruit the required number of participants for the analysis of main effects, recruiting just 642 (89.2%) of the 720 required. As a result, both of the racial subgroups fell far short of this figure with only 451 (62.6% of 720) participants classified as white and just 180 (25% of 720) classified as black. However, because the inclusion/exclusion criteria and the H-I dose were the same in both V-HeFT I and II, the V-HeFT researchers should have been able to improve their estimates of subgroup effects by combining the results of the 49 black and 132 white patients allocated to receive H-I in V-HeFT I with the 109 black and 282 white patients allocated to receive H-I in V-HeFT II. Indeed, although V-HeFT II did not have a control group (because it simply compared two drugs, H-I and the ACE-inhibitor enalapril), the combined mortality rate of participants receiving H-I in V-HeFT I and II could have been compared to that observed amongst participants receiving a placebo in V-HeFT I. Unfortunately, it is not possible to combine the published results of these two trials without access to the original data (because

the outcome measure used, "annual mortality rates," is specific to each of these analyses). Nonetheless, the combined mortality rate for black patients allocated to H-I would inevitably fall somewhere between the V-HeFT I estimate of 31% and the V-HeFT II estimate of 36%, and would therefore be closer to the mortality rate of 44% observed amongst those allocated to the placebo in V-HeFT I. Likewise, the combined mortality rate for white patients allocated to H-I would inevitably fall somewhere between the V-HeFT I estimate of 42% and the V-HeFT II estimate of 39%, and would therefore be closer to the mortality rate of 44% observed amongst those allocated to the placebo in V-HeFT I.

Given the importance of these combined estimates to the FDA's "encouragement of A-HeFT" and their subsequent approval of BiDil, it is not clear why the racial subgroup analyses or BiDil's FDA application overlooked the benefits of combining the two trials' data in this way. One possibility is that the investigators considered that the recruitment of 121 of the V-HeFT I participants who had been allocated to

> The *post hoc* subgroup analyses of V-HeFT I and II suffer from the same potential problems as those faced by all *post hoc* subgroup analyses of randomized controlled trials: a loss of statistical power and the potential for covariate imbalances.

prazosin or placebo (i.e., not to H-I) as participants in V-HeFT II[18] made it problematic to compare the combined estimates of participants allocated to H-I in V-HeFT I and II with those allocated to the placebo in V-HeFT I (some of whom would have gone on to receive H-I in V-HeFT II). Alternatively, perhaps, the researchers and the FDA were already aware that the doses of H-I administered in V-HeFT I were not bioequivalent to those administered in V-HeFT II.[19] These issues aside, it is also possible that the researchers were reluctant to analyze or report the combined estimates of black and white participants allocated to H-I in V-HeFT I and II since these combined estimates seem likely to have weakened the strength of the racial differences they had found in V-HeFT I, and would have thereby undermined their FDA application to approve BiDil for use only in black patients. Whatever the reason(s), now that BiDil has been approved, if these researchers were to combine estimates from the data they have at their disposal, it would not only strengthen the evidence for (or against) racial differences in response to H-I, but it would also help to establish whether BiDil should be approved for use in white patients and patients from other racial groups rather than solely for use in those classified as black.

Potential Covariate Imbalances in the V-HeFT Subgroup Analyses

It is regrettable that these combined estimates from V-HeFT I and II have yet to be reported; it is also unfortunate that the racial subgroup analyses of V-HeFT I and II did not provide disaggregated data on the baseline characteristics of the black and white participants allocated to H-I, prazosin, enalapril, or placebo.[20] Without these data, it is not possible to assess whether the apparent differences in response to H-I by race are simply the result of racial differences in the distribution of covariates amongst participants receiving each of the different therapies.[21] And while the various papers describing the main findings and clinical subgroup analyses of V-HeFT I and V-HeFT II collectively reported on at least 36 different baseline variables (see Table 1),[22] these variables were not consistently reported and only 11 (30.6%) were included in what Temple and Stockbridge described as the "careful" racial subgroup analyses.[23] Nonetheless, it was evident from these 11 variables that there were significant sociodemographic and biomedical differences between black and white participants in both trials.[24] For example, in V-HeFT I, black participants were significantly younger than white participants and had a significantly lower prevalence of coronary artery disease, and a significantly larger cardiothoracic ratio. In V-HeFT II, there was no significant difference in the mean age of black and white participants, but the black participants again had a significantly

lower prevalence of coronary artery disease and a significantly larger cardiothoracic ratio than white participants. Black participants in V-HeFT II were also significantly more likely to have a history of hypertension, and to have lower maximal oxygen consumption and plasma norepinephrine levels than white participants. Although the authors did not specify which variables they included in their subsequent multivariable analyses, they found that after adjusting for what they called "baseline imbalances," there was no interaction between race and response to H-I in either of the V-HeFT trials.[25]

Of course, one of the principal criticisms of racial subgroup analyses, even when randomization has been stratified by race, is that race is not (and cannot be) randomly allocated to trial participants.[26] For this reason, and because race (like sex and class) is such an important determinant of the way in which health risks are unequally distributed throughout racially stratified societies, racial subgroup analyses will always be undermined by differential health risks and their myriad physiological consequences. However, if we assume that the "baseline imbalances" in the racial subgroup analyses of H-I[27] were not on the causal pathway between race and response to H-I (whether as a result of pre-existing genetic, behavioral, and/or socioeconomic differences), then the absence of an interaction between race and H-I following adjustment suggests that the unadjusted racial differences in response to H-I may have been the result of differences in baseline characteristics operating as confounders. And even if these baseline characteristics *were* on the causal pathway between race and the response to H-I, the adjusted findings would still imply that any white participants with similar baseline characteristics to average black participants would be likely to have had a similar response to H-I, and vice versa. Either way the racial subgroup analyses of V-HeFT I and V-HeFT II only provide equivocal evidence for a differential response to H-I amongst black and white patients with heart failure, and the FDA was unjustified to conclude unequivocally that this constituted "considerable evidence that the effects [of H-I] were far smaller...in white patients."[28]

Of course, one of the principal criticisms of racial subgroup analyses, even when randomization has been stratified by race, is that race is not (and cannot be) randomly allocated to trial participants.

Changes in Accepted Therapy Since V-HeFT I and II Undermine the Rationale for A-HeFT

Moreover, the FDA's subsequent "encouragement of A-HeFT" occurred after substantial changes had occurred to the accepted therapy for heart failure, which meant that A-HeFT tested H-I in combination with very different baseline therapies to those used in V-HeFT I and II: V-HeFT I and V-HeFT II examined the effect of H-I on patients who were only receiving digoxin and diuretics while A-HeFT examined patients who were also receiving "angiotensin-converting-enzyme (ACE) inhibitors, angiotensin-receptor blockers, beta-blockers...digoxin, spironolactone, and diuretics."[29] This is important, not least because there are other examples of drugs that were initially thought to be differentially effective across different sub-populations which subsequently proved to be equally effective across all patients when administered in combination with other drugs. For example, ACE-inhibitors were initially reported to have a blunted effect in black patients, but numerous reports have since shown that ACE inhibitors work equally well in black and white patients when combined with diuretics (although these later findings are routinely ignored by contemporary claims based on the original research).[30]

There can be no doubt that the treatment of heart failure has dramatically improved since V-HeFT I and II. This is evident from the much higher mortality rates of black participants in both of these two trials (44% on digoxin and diuretic therapy in V-HeFT I, and 37% on a combination of digoxin, diuretics, and the ACE inhibitor, enalapril, in V-HeFT II) as compared

to that amongst black participants in A-HeFT receiving contemporary accepted therapy (10.2%). Under these circumstances, A-HeFT examined H-I as part of a completely different therapeutic cocktail, and it was therefore unable to establish what effects this cocktail might have had on white patients and those from other racial groups. Indeed, the FDA's calculation that a trial of around 16,000 white patients would have been required for a reasonably powered (ß = 80%) trial of H-I with an effect size of 15% assumed that: (1) "placebo mortality" would be the same as that in V-HeFT I and (2) the response of white patients to H-I in combination with contemporary heart failure therapy would be the same as that observed in V-HeFT I or V-HeFT II in combination with diuretics alone.[31] These assumptions are difficult to justify, not least because so little evidence supports a differential response to H-I amongst black and white participants in V-HeFT I and II, and there is therefore little reason to suppose that it was important to consider the race of participants in any subsequent trials. Moreover, with the benefit of hind-sight, we now know that the mortality rate amongst self-identified black patients receiving H-I together with contemporary heart failure therapies in A-HeFT was just 6.2% while the comparable statistic for black patients receiving H-I with digoxin and diuretics was 31% in V-HeFT I and 36% in V-HeFT II. Thus, even if we adopt the FDA's perspective and assume that the V-HeFT racial subgroup analyses are informative, it does not make sense to have encouraged a subsequent single population trial (i.e., with only black patients) of H-I in combination with contemporary therapies that include ACE inhibitors. This is because the subgroup analysis of V-HeFT II suggested that a combination of digoxin, diuretics, and H-I performed no better than a combination of digoxin, diuretics, and an ACE inhibitor (enalapril). In fact, it seems extraordinary that the original V-HeFT II trial was itself approved given that Temple and Stockbridge[32] suggest that enalapril "was by then an established heart failure treatment." Although there have been substantial changes in regulatory policy since V-HeFT II took place, the FDA's argument that "a larger study of black and white patients was not likely to yield any additional information" about racial variation in the benefits of H-I[33] is not supported by what they knew at the time when they offered "encouragement" to A-HeFT.

Ethical Concerns Generated by the Development of Race-Based Therapies

Notwithstanding the questionable validity of the *post hoc* V-HeFT racial subgroup analyses, there are substantial ethical issues which the FDA seems to have overlooked, ignored, or dismissed. These include growing concerns about the misuse of socially constructed racial identities as if these were natural genetic categories,[34] and the possibility that the development of A-HeFT and the subsequent approval of BiDil were motivated by commercial rather than by scientific or therapeutic considerations.[35]

> The FDA's claim that their support for A-HeFT and the subsequent approval of BiDil as a race-based therapy were justified by 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" does not stand up to closer scrutiny.

The Reification of Racial Identities as Natural Genetic Categories

Few researchers would disagree that the reification of racial identities as genetic categories serves to undermine efforts to tackle racial stigmatization, discrimination, and disadvantage. This operates in two ways: first by distracting attention away from alternative behavioral and socioeconomic determinants of racial variation in biology and health, and second, by suggesting that racial differences in biology and health are inherent, immutable, intractable, and somehow "natural." This is evident in Temple and Stockbridge's account of the FDA's approval for BiDil, when they acknowledge that race is "a highly imperfect description of the genomic and other physiologic characteristics that cause people to differ."[36] While we would

agree with their critical view of race as a useful marker of "genomic characteristics," their argument extends to "physiologic characteristics" and thereby assumes that these are intrinsic *causes* of population disparities in health rather than the extrinsic *consequences* of the ways in which these populations differ on genetic, behavioral, *and* socioeconomic grounds as a result of the way in which they are classified (which differentially allocates a modest amount genetic variation therein) and subsequently treated (which differentially constrains and conditions their behaviour and socioeconomic status). Clearly, concerns about the reification of racial identities as natural genetic categories make it difficult to discuss, expose, and explore racial variation in biology and health without alluding to, or creating data that support, naïve innate and/or genetic explanations for racial difference. And although it is clearly important to monitor racial inequalities in health (to identify inequitable exposure to health risks and inequitable access to health care), because racial identities are strongly associated with notions of innate genetic difference, such research inexorably runs the risk of stigmatizing the most disadvantaged races as innately unhealthy. For this reason, some researchers have argued that racial categories should only be used in biomedical research when their benefits outweigh their disadvantages.[37]

Temple and Stockbridge's article suggests that the FDA mistakes the potential utility of racial categories in subgroup analyses of clinical trials (which use race alongside sociodemographic and biomedical characteristics to identify groups at greater risk of side effects or poor therapeutic response), as well as the regular use of racial categories in epidemiological studies, as evidence that using these categories is unproblematic. Indeed, they claim that "adverse consequences of these observations have not been identified."[38] Yet there is a growing body of research which suggests otherwise,[39] and it is not difficult to imagine the entirely justifiable outcry that might have ensued had BiDil been approved for use in white patients alone. In fact, it is precisely because there are substantive concerns about the underrepresentation of non-white populations in biomedical research (as well as other disadvantaged social groups, most notably women and the elderly) and the lack of evidence for effective treatment in such populations that Congress passed the 1993 NIH Revitalization Act.[40] This Act obliges the National Institutes of Health (NIH) to ensure that sufficient numbers of women and minority groups are included in phase 3 trials to permit valid comparisons of intervention effect, and the FDA claims that it has also "encouraged broad inclusion of patients in trials" along these lines.[41] It therefore seems extraordinary that the FDA should have offered encouragement to a trial (A-HeFT) that deliberately excluded patients from all but one racial group. The FDA's argument for doing so was based on what they felt to be compelling evidence that H-I worked well in black patients but not in white patients and, as we have already seen, this evidence remains ambiguous.

Commercial Incentives to Promote Race-Specific Therapies

Temple and Stockbridge acknowledge that one of the criticisms levelled at the FDA's encouragement of A-HeFT and their subsequent approval of BiDil for use in black patients alone is that it allows drug manufacturers to avoid the costs of developing drugs for the population as a whole by seeking out "narrow niche populations that are easy to study."[42] This is particularly worrisome because permitting drug manufacturers to focus their attention on the patient groups most likely to benefit from treatment might easily lead to a focus on patient groups most able to afford lucrative treatments. Likewise, allowing drug manufacturers to reformulate existing drugs as therapies for specific subpopulations enables them to extend their patents and maintain higher prices for their existing products. It certainly seems likely that the development of BiDil as a race-specific therapy had as much to do with extending the patent on the combination of H-I as it had with the potential effectiveness of H-I as a treatment for heart failure.[43] Although, when this article was first written, the jury was still out on whether BiDil would prove to be a commercial success, the parent company (Nitromed) "discontinued

active promotional efforts related to BiDil" in January 2008 following disappointing sales.[44] While, on the face of it, it seems entirely reasonable for Temple and Stockbridge to argue that such concerns are beyond the official remit of the FDA, which "does not regulate the economics of drug development," they nonetheless concede that the FDA *is* concerned with improving the efficiency of drug development, being keen "not to stifle innovation and efficient studies and deprive the community of valuable treatments."[45] As such, is it not more accurate to recognize that the FDA plays a significant role in the "economics of drug development" given that the "efficiency of drug development" is not only about the speed with which drugs make it to market but also about the cost involved? Either way, the FDA dismissed claims that commercial concerns had played a role in the premature approval of BiDil for a "niche population" by once more citing the questionable racial subgroup analyses of V-HeFT I and II to explain why they felt it had already been adequately tested in the broader population of patients with heart failure — a remarkable argument given the participants in V-HeFT I and II were all male.

Meanwhile, the approval of BiDil should also be placed within the broader context of moves to develop pharmacogenetics — so called "personalised medicine."[46] These new technologies aim to use genetic tests to target new or already licensed drugs to subpopulations defined by their genotype. Many important regulatory and policy questions raised by pharmacogenetics are relevant to the discussion of racially targeted drugs such as BiDil. In particular, does a new targeted drug need to be clinically tested on the population as a whole or only on the sub-group at which it is aimed? Some might argue that whole population testing might still be required for such drugs to assess the risk of adverse events were the drug to be given to other sub-populations as a result of poor prescription management or off-label use. Likewise, how is the restricted use of such drugs in specified subpopulations to be controlled, and what criteria will this entail to define the sub-populations involved? In the case of racially targeted drugs, will some form of biomarker be required to avoid the use of crude and potentially unreliable indicators such as skin color or self-identification?

As such, a major concern raised by BiDil is the extent to which it sets a precedent in answering these questions. If retrospective sub-group analyses can be used to identify new therapeutic sub-populations (be they defined by race or genotype), then the pharmaceutical industry will have a major incentive to relaunch drugs that are close to patent expiry as new "personalised medicines." Apart from being scientifically questionable, this would ensure that higher prices could be maintained in treating the very patients who might benefit most from the use of established generic products. The manner in which BiDil has been approved therefore casts serious doubts about important aspects relating to the integrity of the regulation of medicines.

Conclusion and Recommendations

The FDA's claim that their support for A-HeFT and the subsequent approval of BiDil as a race-based therapy were justified by 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"[47] does not stand up to closer scrutiny. These claims are based on ambiguous *post hoc* subgroup analyses in which the only statistically significant difference observed between black and white patients was found without any adjustment for potential confounders in under-powered subsamples that were unlikely to be adequately balanced. They also failed to consider the ethical consequences of recognizing racial categories as valid markers of innate biological difference, as well as the impact of allowing drug companies to develop or re-patent therapies for niche populations in a way that fails to address the needs of the population as a whole and increases the cost of medicines.

The FDA should re-examine its support of A-HeFT given the fact that there have been substantial changes in baseline therapy for heart failure since V-HeFT I and II, and the doses of H-I administered in V-HeFT I were not bioequivalent to those administered in V-HeFT II.[48] The FDA should also encourage the researchers involved in V-HeFT I and II to combine their analyses of black and white patients allocated to H-I in order to strengthen the evidence available for racial variation in response. If the effect of race on response to H-I is found to be lower in these combined analyses, as the available data suggest it will, then the U.S. Patent and Trademark Office (PTO) should revoke their patent of BiDil as a race-based therapy, and the FDA should approve H-I as a therapy for heart failure patients regardless of their race. Finally, the FDA and PTO should revise the procedures they use to examine applications for race-based therapies to ensure that they are based on robust claims and do not undermine the aims of the 1993 NIH Revitalization Act.

Acknowledgements

The authors thank Yu-Kang Tu for biostatistical advice, Elmien Wolvaardt for editorial comments, and the organizers and participants of two meetings: "Controversies in Biomedical Research" organized by the Royal Statistical Society in London on March 28, 2006, and "Race, Pharmaceuticals and Medical Technology" held by the Center for the Study of Diversity in Science, Technology and Medicine (CSD) at Massachusetts Institute of Technology on April 7-8, 2006. George T. H. Ellison is supported by grants 057182 and 073524/2/03/2/AW/HH from the Wellcome Trust's Biomedical Ethics Research Programme, the latter with Paul Martin from the University of Nottingham, Andrew Smart from Bath Spa University, Richard Ashcroft from Queen Mary's, University of London, and Richard Tutton from Lancaster University. See <www.raceandgenetics.org>. Jay S. Kaufman was supported by the Ethical, Legal, and Social Implications Research Program, National Human Genome Research Institute (Grant number P50-HG004488). Jonathan Kahn was supported by the Ethical, Legal and Social Implications Research Program, National Human Genome Institute (Grant number R03-HG004034-01A1).

References

1. Temple R, Stockbridge NL. BiDil for Heart Failure in Black Patients: The U.S. Food and Drug Administration Perspective. Annals of Internal Medicine 2007;146(1):57–62. [PubMed: 17200223]
2. Taylor AL, Ziesche S, Yancy C, Carson P, D'Agostino R, Ferdinand K. Combination of Isosorbide Dinitrate and Hydralazine in Blacks with Heart Failure. New England Journal of Medicine 2004;351 (20):2049–2057. [PubMed: 15533851]
3. Bibbins-Domingo K, Fernandez A. BiDil for Heart Failure in Black Patients: Implications of the U.S. Food and Drug Administration Approval. Annals of Internal Medicine 2007;146(1):52–56. [PubMed: 17200222]
4. Id.
5. Kahn J. How a Drug Becomes 'Ethnic': Law, Commerce, and the Production of Racial Categories in Medicine. Yale Journal of Health Policy, Law & Ethics 2004;4(1):1–46.
6. Ellison GTH, Smart A, Tutton R, Outram SM, Ashcroft R, Martin P. Racial Categories in Medicine: A Failure of Evidence-Based Practice? Public Library of Science Medicine 2007;4(9):1434–1436.
7. Ellison GTH. Population Profiling and Public Health Risk: When and How Should We Use Race/ Ethnicity? Critical Public Health 2005;15(1):65–74.
8. See Temple and Stockbridge, *supra* note 1.
9. Id.
10. Carson P, Ziesche S, Johnson G, Cohn J. Racial Differences in Response to Therapy for Heart Failure: Analysis of the Vasodilator-Heart Failure Trials. Journal of Cardiac Failure 1999;5(3):178–187. [PubMed: 10496190]
11. See Bibbins-Domingo and Fernandez, *supra* note 3.
12. Cohn JN, Archibald DG, Ziesche S, Cobb F, Franciosa JA, Harston WE, et al. Effect of Vasodilator Therapy on Mortality in Chronic Congestive Heart Failure: Results of a Veterans Administration Cooperative Study. New England Journal of Medicine 1986;314(24):1547–1552. [PubMed: 3520315]Cohn JN, Johnson G, Ziesche S, Cobb F, Francis G, Tristani F, et al. A Comparison of

Enalapril with Hydralazine-Isosorbide Dinitrate in the Treatment of Chronic Congestive Heart Failure. New England Journal of Medicine 1991;325(5):303–310. [PubMed: 2057035]

13. See Kahn, *supra* note 5.
14. See Bibbins-Domingo and Fernandez, *supra* note 3.
15. Ellison GTH. Medicine in Black and White: BiDil, Race and the Limits of Evidence-Based Medicine. Significance 2006;3(3):118–121.Kent D, Hayward R. When Averages Hide Individual Differences in Clinical Trials. American Scientist 2007;95(1):60–68.Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in Medicine: Reporting of Subgroup Analyses in Clinical Trials. New England Journal of Medicine 2007;357(21):2189–2194. [PubMed: 18032770]
16. See Cohn, Johnson, Ziesche, Cobb, Francis, and Tristani et al., *supra* note 12.
17. See Cohn, Archibald, Ziesche, Cobb, Franciosa, and Harston et al., *supra* note 12.
18. See Cohn, Johnson, Ziesche, Cobb, Francis, and Tristani et al., *supra* note 12.
19. See Kent and Hayward, *supra* note 15.
20. See Carson, Ziesche, Johnson, and Cohn, *supra* note 10.
21. Id.
22. See Cohn, Archibald, Ziesche, Cobb, Franciosa, and Harston et al., *supra* note 12; Cohn, Johnson, Ziesche, Cobb, Francis, and Tristani et al., *supra* note 12;Cohn JN, Archibald DG, Francis GS, Ziesche S, Franciosa JA, Harston WE, et al. Veterans Administration Cooperative Study on Vasodilator Therapy of Heart Failure: Influence of Pre-Randomization Variables on the Reduction of Mortality by Treatment with Hydralazine and Isosorbide Dinitrate. Circulation 1987;75(5):IV49–IV54.IV54 [PubMed: 3552302]Johnson G, Carson P, Francis GS, Cohn JN. Influence of Pre-Randomisation (Baseline) Variables on Mortality and on the Reduction of Mortality by Enalapril: Veterans Affairs Cooperative Study on Vasodilator Therapy of Heart Failure (V-HeFT II). Circulation 1993;87(6):VI32–VI39.VI39 [PubMed: 8500237]
23. See Temple and Stockbridge, *supra* note 1; Carson, Ziesche, Johnson, and Cohn, *supra* note 10.
24. *Id.* (Carson et al.)
25. Id.
26. Kaufman JS, Cooper RS. Considerations for Use of Racial/Ethnic Classification in a Etiologic Research. American Journal of Epidemiology 2001;514(4):291–298. [PubMed: 11495850]
27. See Carson, Ziesche, Johnson, and Cohn, *supra* note 10.
28. See Temple and Stockbridge, *supra* note 1, at 57.
29. See Taylor, Ziesche, Yancy, Carson, D'Agostino, and Ferdinand, *supra* note 2.
30. Khan JM, Beevers DG. Management of Hypertension in Ethnic Minorities. Heart 2005;91(8):1105–1109. [PubMed: 16020613]
31. See Temple and Stockbridge, *supra* note 1.
32. Id.
33. Id.
34. See Ellison, *supra* note 7;Census, Race and Science. Nature Genetics 2000;24(2):97–98.98 Editorial. [PubMed: 10655044]
35. See Kahn, *supra* note 5.
36. See Temple and Stockbridge, *supra* note 1, at 58.
37. See Ellison, *supra* note 7.
38. See Temple and Stockbridge, *supra* note 1, at 61.
39. See alsoLa Parra Casado D, Perez AM. Scientifically Correct Racism: Health Studies' Unintentional Effects Against Minority Groups. Language and Intercultural Communication 2007;7(2):152–162.162
40. NIH Revitalization Act. 1993. Pub. L. No. 103-43
41. See Temple and Stockbridge, *supra* note 1, at 60.
42. *Id.*, at 60.
43. See Kahn, *supra* note 5.

44. SeeNitroMed Reports Financial Results for First Fiscal Quarter 2008. May 20;2008 availaible at <http://phx.corporate-ir.net/phoenix.zhtml?c=130535&p=irol-newsArticle&ID=1141115&highlight=> (last visited.

45. See Temple and Stockbridge, *supra* note 1, at 61.

46. Martin, PA.; Lewis, G.; Smart, A.; Webster, A. False Positive? Prospects for the Clinical and Commercial Development of Pharmacogenetics. University of Nottingham and University of York; Nottingham and York: 2006.

47. See Temple and Stockbridge, *supra* note 1, at 61.

48. Center for Drug Evaluation and Research. Approval Package for: Application Number NDA 20-727. Administrative/Correspondence Reviews. N.d. May 19. 2008 available at <http://www.fda.gov/cder/foi/nda/2005/020727_S000_Bidil_AdminCorres.pdf> (last visited

Table 1

Baseline Variables Reported in Each of Papers Describing the Main Findings and Clinical Subgroup Analyses of V-HeFT I and V-HeFT II

	Main Findings		Subgroup Analyses		
	V-HeFT I	**V-HeFT II**	**Clinical V-HeFT I**	**Clinical V-HeFT II**	**Racial V-HeFT I & II**
Baseline Variable	**Cohn et al. 1986[8]**	**Cohn et al. 1991[9]**	**Cohn et al. 1987[14]**	**Johnson et al. 1993[15]**	**Carson et al. 1999[7]**
Age	+	+	+	+	+
Heart failure symptoms	+	+	-	-	-
Angina in last month	+	-	-	-	-
Coronary artery disease	+	+	+	+	+
Previous myocardial infarction	+	+	-	-	-
Alcohol excess	+	+	+	+	-
Hypertension	+	+	+	+	+
Diabetes	+	+	-	-	-
Coronary bypass surgery	+	+	+	-	-
Valve replacement surgery	+	-	-	-	-
Vasodilator therapy	+	+	+	+	-
Antiarrhythmic therapy	+	+	+	+	-
Sublingual nitroglycerine	+	+	-	-	-
Anticoagulant therapy	+	+	-	+	-
Symptom score[1]	+	-	-	-	-
Arterial pressure	+	+	-	-	-
Heart rate	+	+	-	+	+[2]
Cardiothoracic ratio	+	+	+	+	+
Ejection fraction	+	+	+	+	+
Left ventricular internal diastolic diameter	+	+	+	+	-
Exercise duration	+	-	-	-	-
Maximum O_2 consumption	+	+	+	+	+
New York Heart Association class	-	+	-	+	-
Race	-	+	-	-	-

	Main Findings		Subgroup Analyses		
	V-HeFT I	V-HeFT II	Clinical V-HeFT I	Clinical V-HeFT II	Racial V-HeFT I & II
Baseline Variable	**Cohn et al. 1986[8]**	**Cohn et al. 1991[9]**	**Cohn et al. 1987[14]**	**Johnson et al. 1993[15]**	**Carson et al. 1999[7]**
Cerebrovascular accident	-	+	-	-	-
Tobacco use	-	+	-	-	-
Atrial fibrillation	-	+	-	-	-
S_3 gallop	-	+	-	-	-
Plasma norepinephrine	-	-	-	+	+[3]
Plasma renin activity	-	-	-	+	+[3]
Cardiac hospitalization	-	-	-	+	-
Arrhythmia severity	-	-	-	+	-
Weight	-	-	-	+	-
Quality-of-life score	-	-	-	+	-
Systolic blood pressure	-	-	-	+	+
Diastolic blood pressure	-	-	-	+	+

[1] Sum of scores for dyspnea, fatigue, orthopnea, and paroxysmal nocturnal dyspnea

[2] Only included in the racial subgroup analysis of V-HeFT I

[3] Only included in the racial subgroup analysis of V-HeFT II