

Report to Congressional Requesters

September 2002

# SSA DISABILITY DECISION MAKING

Additional Measures
Would Enhance
Agency's Ability to
Determine Whether
Racial Bias Exists



# Contents

Letter		1
	Results in Brief	3
	Background	4
	SSA's Study of Racial Disparities Was Extensive, but Methodological Weaknesses in Available Documentation	
	Preclude Conclusions	6
	SSA Has Taken Limited Steps to Address Possible Racial Bias in Its	
	Hearings Level Decision-Making Process	14
	Conclusions	18
	Recommendations	19
Appendix I	Comments from the Social Security Administration	21
	GAO Comments	26
Appendix II	GAO Contacts and Staff Acknowledgments	28
	GAO Contacts	28
	Staff Acknowledgments	28

#### **Abbreviations**

ALJ	Administrative Law Judge
DDS	Disability Determination Service
DI	Disability Insurance
EAB	Enumeration at Birth
OHA	Office of Hearings and Appeals
SSA	Social Security Administration
SSI	Supplement Security Income
SSN	Social Security Number



## United States General Accounting Office Washington, DC 20548

September 9, 2002

The Honorable Charles B. Rangel Ranking Minority Member Committee on Ways and Means House of Representatives

The Honorable Robert T. Matsui Ranking Minority Member Subcommittee on Social Security Committee on Ways and Means House of Representatives

The Honorable Gene Green House of Representatives

The Social Security Administration (SSA) is responsible for administering the Social Security Disability Insurance (DI) and the Supplement Security Income (SSI) programs—the nation's two largest disability programs. In calendar year 2001, SSA provided cash assistance through these two programs to 8.8 million working-age beneficiaries with qualifying disabilities, and about 3.4 million people applied for benefits.¹ SSA is required to administer its disability programs in a fair and unbiased manner. Nevertheless, the proportion of African American applicants allowed benefits has been historically lower than the proportion of white applicants.² These allowance rate differences have occurred with respect to disability determinations made by state Disability Determination Service (DDS) offices³ and in decisions made at the hearings level by Administrative Law Judges (ALJs). For example, in 1985, DI allowance

<sup>&</sup>lt;sup>1</sup>In calendar year 2001, about 1.5 million people applied for DI benefits and about 1.9 million people applied for SSI benefits. Our total estimate of about 3.4 million people applying for benefits is not adjusted for concurrent applications arising from individuals applying for both programs.

<sup>&</sup>lt;sup>2</sup>In 1992, GAO published data reported by SSA in the 1980s that showed that African Americans experienced a lower allowance rate than whites. U.S. General Accounting Office, *Social Security: Racial Difference in Disability Decisions Warrants Further Investigation*, GAO/HRD-92-56 (Washington, D.C.: Apr. 21,1992).

<sup>&</sup>lt;sup>3</sup>DDS offices are funded by SSA but administered by states, and make disability determinations in accordance with SSA's policies and procedures.

rates (including DDS and ALJ decisions) were 33 percent for African Americans compared with 40 percent for whites. However, differences in allowance rates alone do not necessarily point to racial bias in the system; additional analysis must be done to determine whether differences can be explained by other key factors that are considered by SSA in determining eligibility for disability benefits—including factors such as claimant impairment, age, and education. Our 1992 report analyzed racial differences in DI and SSI allowance rates and found that racial differences in allowance decisions—mostly with respect to ALJ decisions—could not be explained by these and other key factors. We recommended that SSA investigate this issue further and, if needed, take appropriate actions to correct and prevent any unwarranted racial disparities.

You asked us to investigate whether discrepancies exist within the DI and SSI programs and to examine the steps SSA has taken to address unwarranted racial and ethnic disparities. In conducting our review, we encountered potentially significant limitations with SSA's administrative data that are needed to determine whether discrepancies exist. Therefore, we are responding to your request in two parts. This report examines: (1) SSA's efforts to study potential racial disparities in ALJ decisions and (2) steps SSA has taken at the hearings level to address possible racial and ethnic bias in ALJ decision making.<sup>5</sup>

To perform our work, we evaluated the scope and statistical methods used by SSA in its study of racial disparities and interviewed outside consultants who were hired by SSA to assist in this study. We also interviewed SSA officials who have been involved in studying or otherwise addressing potential racial disparities since our 1992 report and obtained and reviewed SSA's internal working documents pertaining to these efforts. We did our work in accordance with generally accepted government auditing standards between September 2001 and July 2002.

#### <sup>4</sup>GAO/HRD-92-56.

<sup>&</sup>lt;sup>5</sup>In the next phase of our work, we plan to examine the extent to which disparities still exist within SSA's disability decision-making process.

<sup>&</sup>lt;sup>6</sup>In evaluating SSA's study, we were unable to obtain some information about the study because, due to the passage of time, much of the documentation and some of the key individuals involved in the study were not available. SSA officials told us that the agency conducted additional analyses not reflected in available documentation. However, we were not able to either review or corroborate these efforts.

#### Results in Brief

In response to our 1992 report, SSA initiated an extensive study of racial disparities in ALJ decisions, but methodological weaknesses preclude conclusions being drawn from it. The study—the results of which were not published—set out to analyze a representative sample of cases to determine whether race significantly influenced disability decisions, while simultaneously controlling for other factors. It spanned over 4 years and involved about 50 full-time technical staff and three outside consultants. SSA officials told us that, by 1998, they found no evidence that race significantly influenced ALJ decisions. However, we were unable to draw these same conclusions due to weaknesses in sampling and statistical methods evident in the limited documentation still available for our review. For example, although more than three-quarters of the case files in SSA's initial sample could not be included in the final sample—in large part because they were either in use or missing—SSA performed only limited analyses to test whether the final sample was representative of the universe of ALJ decisions. We also identified other weaknesses in SSA's methods, such as the inclusion and exclusion of certain variables in the analysis, which could lead to inaccurate or misleading results. Although SSA's study was part of a larger ongoing effort to conduct quality reviews of ALJ decisions, SSA no longer analyzes these decisions for racial bias and has no specific plans to conduct additional studies on racial disparities in the future. However, SSA continues to review ALJ decisions—the results of which are used to assess the overall accuracy of such decisions—but, as in its racial disparities study, still obtains a low percentage of files. Moreover, SSA no longer performs analyses of the sample to test its representativeness. Any future analyses of race might be hindered by this and by the fact that SSA has significantly scaled back the collection of race data.

Concurrent with SSA's study of racial disparities, SSA's Office of Hearings and Appeals (OHA) took some limited steps at the hearings level to address possible racial bias in ALJ decision making. OHA instituted a mandatory diversity sensitivity training course for ALJs. All incumbent ALJs were given at least 1-1/2 days of diversity training in 1992-93, and all incoming ALJs take a 1-day diversity course as part of their orientation. Additionally, OHA increased its efforts to recruit minorities for ALJ and other legal positions by attending conferences for minority bar associations, where SSA distributed information and gave seminars on how to become an ALJ. Finally, in keeping with its commitment to provide fair and impartial hearings, SSA established a new process under the direction of OHA for the review, investigation, and resolution of claimant complaints about alleged bias or misconduct by ALJs. Although intended to address misconduct, including possible racial bias, this new complaint

process does not include a mechanism for easily identifying complaints involving racial discrimination; nor does it track the race of complainants. As a result, this new process will not help SSA's OHA staff to identify patterns of possible racial bias among the complaints filed, which may warrant further investigation or corrective action.

In order to better assess the overall accuracy of ALJ decisions, and to facilitate future analyses of racial disparities in ALJ decision making, we are recommending that SSA take steps to further test the representativeness of the final sample used for its ongoing quality assurance review of ALJ decisions and, if needed, make appropriate changes to its sampling methods. In addition, to help OHA more readily identify recurring incidences or patterns of possible racial bias in complaints against ALJs, we are recommending that SSA build special mechanisms in its ALJ complaint process to facilitate the collection and periodic analysis of these data. We further recommend that the results of these analyses be reported to the Commissioner and that SSA develop action plans, if needed.

In its written comments to our report, SSA agreed with our recommendations and agreed to take steps to implement them. SSA also provided general and technical comments with respect to our findings, which we incorporated in the report as appropriate. SSA's general comments and our response are printed in appendix I.

#### Background

DI and SSI are the two largest federal programs providing cash assistance to people with disabilities. Established in 1956, DI provides monthly payments to workers with disabilities (and their dependents or survivors) under the age of 65 who have enough work experience to be qualified for disability benefits. Created in 1972, SSI is a means-tested income assistance program that provides monthly payments to adults or children who are blind or who have other disabilities and whose income and assets fall bellow a certain level. To be considered eligible for either program as an adult, a person must be unable to perform any substantial gainful activity by reason of a medically determinable physical or mental impairment that is expected to result in death or that has lasted or can be expected to last for a continuous period of at least 12 months. Work

<sup>&</sup>lt;sup>7</sup>SSI also provides income assistance to the aged who have income and assets below a certain level.

activity is generally considered substantial and gainful if the person's earnings exceed a particular level established by statute and regulations.<sup>8</sup> In calendar year 2001, about 6.1 million working age individuals (age 18-64) received about \$59.6 billion in DI benefits, and about 3.8 million working-age individuals received about \$19 billion in SSI federal benefits.<sup>9</sup>

To obtain disability benefits, a claimant must file an application at any of SSA's offices or other designated places. If the claimant meets the nonmedical eligibility criteria, the field office staff forwards the claim to the appropriate state DDS office. DDS staff—generally a team comprised of disability examiners and medical consultants—review medical and other evidence provided by the claimant, obtaining additional evidence as needed to assess whether the claimant satisfies the program requirements, and make the initial disability determination. If the claimant is not satisfied with the DDS determination, the claimant may request a reconsideration within the same DDS. <sup>10</sup> Another DDS team will review the documentation in the case file, as well as any new evidence the claimant may submit, and determine whether the claimant meets SSA's definition of disability. In 2001, the DDSs made 2.1 million initial disability determinations and over 514,000 reconsiderations.

If the claimant is not satisfied with the reconsideration, the claimant may request a hearing by an ALJ. Within SSA's OHA, there are approximately 1,100 ALJs who are located in 140 hearing offices across the country. The ALJ conducts a new review of the claimant's file, including any additional evidence the claimant submitted since the DDS decision. The ALJ may also hear testimony from medical or vocational experts and the claimant regarding the claimant's medical condition and ability to work. The

<sup>&</sup>lt;sup>8</sup>The Social Security Commissioner has the authority to set the substantial and gainful activity level for individuals who have disabilities other than blindness. In December 2000, SSA finalized a rule calling for the annual indexing of the nonblind level to the average wage index of all employees in the United States. The current nonblind level is set at \$780 per month. The level for individuals who are blind is set by statute and is also indexed to the average wage index. Currently, the level for blind individuals is \$1,300 of countable earnings.

<sup>&</sup>lt;sup>9</sup>DI low-income beneficiaries can also receive SSI benefits. Of the 6.1 million DI beneficiaries, about 1.1 million also received SSI in 2001. Thus, there was a total of 8.8 million working-age beneficiaries in 2001, with 12.5 percent receiving both DI and SSI.

<sup>&</sup>lt;sup>10</sup>While most claimants may request a reconsideration, at the time of our study, SSA was implementing an initiative that attempts to ensure that all legitimate claims are approved as early in the process as possible. As part of this initiative, DDSs in one state in each region have eliminated the DDS reconsideration step in the appeals process.

hearings are recorded, and claimants are usually represented at these hearings. In fiscal year 2001, ALJs made over 347,000 disability decisions.<sup>11</sup>

SSA is required to administer its disability programs in a fair and unbiased manner. However, in our 1992 report, we found that, among ALJ decisions at the hearings level, the racial difference in allowance rates was larger than at the DDS level and did not appear to be related to severity or type of impairment, age or other demographic characteristics, appeal rate, or attorney representation. We recommended, and SSA agreed, to further investigate the reasons for the racial differences at the hearings level and act to correct or prevent any unwarranted disparities.

SSA's Study of Racial
Disparities Was
Extensive, but
Methodological
Weaknesses in
Available
Documentation
Preclude Conclusions

Following our report, SSA undertook an extensive effort to study racial disparities in ALJ decisions at the hearings level, but weaknesses in available documentation preclude conclusions from being drawn. The study involved 4 years of data collection, outside consultants, and many staff who collected and analyzed data from over 15,000 case files. Although the results were not published, SSA officials told us that their statistical analyses of these data revealed no evidence of racial disparities. On the basis of our review of SSA's internal working papers and other available information, we identified several weaknesses in sampling and statistical methods. Presently, SSA has no further plans to study racial disparities but, if it did, its ability to do so would likely be hampered by data limitations.

SSA's Effort to Study Racial Disparities Was Extensive, but Results Were Not Published In response to our 1992 report, SSA initiated a study of racial disparities at the ALJ level that involved several components of the agency. SSA obtained help in designing and conducting the study from staff in its Office of Quality Assurance and Performance Assessment; the Office of

<sup>&</sup>lt;sup>11</sup>If the claimant is not satisfied with the ALJ decision, the claimant may appeal the claim to the Appeals Council within SSA and, ultimately, to the federal courts.

<sup>&</sup>lt;sup>12</sup>At the time of our 1992 report, there was no independent measure for controlling for severity of impairment in our analysis since DDSs make the severity determinations. However, some of the effects of severity were indirectly accounted for in our analyses that included impairment type, a factor that is associated with severity. Furthermore, in addition to conducting analyses based on all applicants, we conducted analyses based on only those cases classified as severely impaired by the DDSs. This enabled us to examine whether racial difference persisted after excluding cases that the DDSs considered to be nonsevere.

Research, Evaluation and Statistics; and the Office of Hearings and Appeals. SSA also created a new division within the Office of Quality Assurance—the Division of Disability Hearings Quality—to spearhead the collection of data needed to study racial disparities and to oversee ongoing quality assurance reviews of ALJ decisions.<sup>13</sup>

Data collection for this study was a large and lengthy effort. In order to construct a representative sample of cases to determine whether race significantly influenced disability decisions, SSA selected a random sample each month from the universe of ALJ decisions, stratifying by race, region, and decisional outcome (allowance or denial). This sample of over 65,000 cases was drawn over a 4-year period—from 1992 to 1996. Then, for each ALJ decision that was selected to be in the sample, SSA requested the case file and a recording of the hearing proceedings from hearing offices and storage facilities across the country. Obtaining this documentation was complicated by the fact that files were stored in different locations, depending on whether the case involved an SSI or DI claim, and whether the ALJ decision was an allowance or denial.

In addition to obtaining files and tapes, the data collection effort included a systematic review of each case—the results of which SSA used, in part, for its analysis of racial disparities. Specifically, each case used in the analysis received three reviews: a peer review by an ALJ, a medical evaluation performed by one or more medical consultants (depending on the number and type of impairments alleged by the claimant), and a general review of the documentation and decisions by a disability examiner. In total, a panel of 10 to 12 ALJs, whose composition changed every 4 months, worked full-time to review cases. <sup>16</sup> In addition, over a

<sup>&</sup>lt;sup>13</sup>Prior to 1992, the Office of Quality Assurance conducted quality assurance reviews at the DDS level but did not have a quality assurance process for the hearing level.

<sup>&</sup>lt;sup>14</sup>The case file contains the application for benefits, appeal requests, disability information provided by the claimant, medical evidence furnished at each level of the appeal, DDS determinations, claimant's appointment of an attorney/representative (if applicable), copies of notices to the claimant and the ALJ decision. For ALJ allowance decisions, the file will also contain documentation of benefit computation and payment.

<sup>&</sup>lt;sup>15</sup>For example, in all allowances where a hearing was held, hearing offices are directed to send the recording of the hearing on a cassette to OHA's Computer Cassette Library in Falls Church, Virginia. Cassettes, thus stored separately from the case files, were requested separately by the Division of Disability Hearings Quality. Folders and cassettes were associated, if both were received, in the Division of Disability Hearings Quality.

 $<sup>^{16}</sup>$ In order to prevent conflicts of interest, ALJs did not evaluate cases from their own regions.

4-year period, 37 to 55 staff, including disability examiners, worked full-time reviewing case files that were used for this study. <sup>17</sup> Ultimately, about 15,000 cases received all three reviews necessary for inclusion in this study.

During and after the 4-year data collection effort, SSA worked with consultants to analyze the data in order to determine the effect of race on ALJ decisions. SSA used descriptive statistics to show that overall application and allowance rates of African Americans differed from whites. 18 In addition, SSA used multivariate analyses to examine the effect of race on ALJ decisions while controlling for other factors that influence decisions. 19 One of SSA's consultants—a law professor and recognized expert in disability issues—reviewed SSA's analytical approach and evaluated initial results. In his report to SSA, this consultant expressed overall approval of SSA's data collection methods, but made several recommendations on how the analysis could be improved—some of which SSA incorporated into later versions of its analysis. 20 SSA subsequently hired two consulting statisticians to review later versions of the analysis. These statisticians expressed concerns about SSA's methods and offered several suggestions. According to SSA officials, these suggestions were not incorporated into the analysis because they were perceived to be labor intensive and SSA was not sure the effort would result in more definitive conclusions.

According to SSA officials, the agency's final analysis of the data revealed no evidence of racial disparities, but the results were considered to be not definitive enough to warrant publication. Specifically, SSA officials told us that, by 1998, they found no evidence that race significantly affected ALJ decisions for any of the regions. However, these officials also told us that, due to general limitations of statistical analysis, especially as applied to such complex processes as ALJ decision making, they believed that they

<sup>&</sup>lt;sup>17</sup>This number does not include the medical consultants who reviewed cases.

<sup>&</sup>lt;sup>18</sup>Although SSA's original sample stratification and preliminary analyses examined African American claimants and "all other" claimants, SSA informed us that its concluding analyses, which were not available for our review, examined African American claimants and white claimants.

<sup>&</sup>lt;sup>19</sup>Specifically, to estimate a model of the ALJ's allowance decision, SSA used logistic regression, a common technique that is used when the dependent variable is binary.

<sup>&</sup>lt;sup>20</sup>For example, the consultant recommended that SSA include a separate variable for DI, SSI, and concurrent applications, which SSA incorporated into its analysis.

could not definitively conclude that no racial bias existed. Given the complexity of the results and the topic's sensitive nature, SSA officials told us the agency decided not to publish the conclusions of this study. $^{21}$ 

Available Documentation of Racial Disparities Study Indicated Some Methodological Weaknesses From our review of SSA's internal working papers pertaining to the study, and information provided verbally by SSA officials, we identified several weaknesses in SSA's study of racial disparities. These weaknesses include: using a potentially nonrepresentative final sample of cases in their multivariate analyses, performing only limited analyses to test the representativeness of the final sample, and using certain statistical techniques that could lead to inaccurate or misleading results.

Although SSA started with an appropriate sampling design, its final sample included only a small percentage of the case files in its initial sample in part because staff were unable to obtain many of the associated case files or hearing tapes. SSA was not able to obtain many files and tapes because they were missing (i.e., lost or misplaced) or they were in use and were not made available for the study. For example, according to SSA officials, files for cases involving appeals of ALJ decisions to SSA's Appeals Council—about half of ALJ denials—were in use and, therefore, excluded from the study.<sup>22</sup> In addition, SSA officials told us that, due to resource constraints, not all of the obtained files underwent all three reviews, which were necessary for inclusion in SSA's analysis of racial disparities.<sup>23</sup> In the end, less than one-fourth of the cases that were selected to be in the initial sample were actually included in SSA's final sample.<sup>24</sup>

With less than one-fourth of the sampled cases included in the final sample, SSA took steps to determine whether the final sample of cases

<sup>&</sup>lt;sup>21</sup>In response to a Freedom of Information Act request, SSA officials presented the results orally to the National Organization of Social Security Claimants Attorney Representatives on July 10, 2001.

<sup>&</sup>lt;sup>22</sup>In other instances, cases were excluded from the study because a payment action was pending and the case file was being updated. In addition, a small proportion of cases were excluded because they were later identified as being cases that were not intended to be included in the sample.

<sup>&</sup>lt;sup>23</sup>According to SSA officials, a total of 22,700 cases were reviewed by at least one team. Of these, 15,000 received all three requisite reviews for inclusion in the study.

<sup>&</sup>lt;sup>24</sup>In their reports to SSA, none of the consultants indicated that they were aware that SSA was using only a small subset of the initial sample in its analysis.

was still representative of all ALJ decisions. While the investigation SSA undertook revealed no clear differences between cases that were and were not included in the final sample, we found no evidence that SSA performed certain analyses that could have provided more assurance of the sample's representativeness. For example, SSA made some basic comparisons between claimants who were included in the final sample and those that were in the initial sample but not the final sample. SSA's results indicate that these two groups were fairly similar in key characteristics such as racial composition, years of education, and years of work experience. However, we found no indication in the documentation provided to us that SSA tested whether slight differences between the two groups were or were not statistically significant. Further, we found no indication that SSA compared the allowance rates of these two groups. This is an important test because, in order to be statistically representative, claimants in the final sample should not have had significantly different allowance rates from claimants who were not included in the final sample. In addition, although children were not included in SSA's analysis of racial disparities, SSA's tests to determine the representativeness of the final sample included children in one group and excluded children from the other. By including children in one of the comparison groups, SSA could not assess whether characteristics of the adults in the two groups were similar.<sup>25</sup>

Another weakness, as documented in internal working papers available for our review, was the inclusion of certain variables in the multivariate analyses of ALJ decisions, which could lead to biased results. SSA guidelines clearly define the information that should be considered in the ALJ decision, and SSA appropriately included many variables that capture this information in its multivariate analysis. However, SSA also included several variables developed during the review process that reflected the reviewer's evaluation of the hearing proceedings. For example, SSA included a variable that assessed whether the ALJ, in the hearing decision, appropriately documented the basis for his or her decision in the case file. This variable did not influence the ALJ's decision, but evaluated the ALJ's compliance with SSA procedures and should not have been included in the multivariate analysis. This and other variables that reflected a posthearing

<sup>&</sup>lt;sup>25</sup>Specifically, 7.48 percent of the group for which files were not included in the analysis was comprised of children, while only 0.21 percent of the final sample was comprised of children. The inclusion of children in one group reduces that group's average age, education level, and years of experience. It is impossible to know what these averages would have been had children been similarly excluded from both groups.

evaluation of ALJ decisions were included in SSA's multivariate analysis. If these variables are associated with race or somehow reflect racial bias in ALJ decision making, including such variables in multivariate analysis will reduce the explanatory power of race as a variable in that analysis. For example, if a model includes a variable that may reflect racial bias—such as one that indicates the reviewer believed that the original ALJ decision was unfair or not supported—then that variable, rather than the race of the claimant, could show up as a significant factor in the model. The statisticians hired by SSA as outside consultants also expressed concern about the inclusion of these variables in SSA's analyses.

Finally, in its internal working papers, SSA used a statistical technique stepwise regression—that was not appropriate given the characteristics of its analysis. Specifically, SSA researchers first identified a set of variables for potential use in their multivariate analysis—variables drawn mostly from data developed during the case file review process. Then, to select the final set of variables, SSA used stepwise regression. Stepwise regression is an iterative computational technique that determines which variables should be included in an analysis by systematically eliminating variables from the starting variable set that are not statistically significant. Using the results from this analysis, SSA constructed a different model for each of SSA's 10 regions, which were used in SSA's multivariate analysis to test whether African Americans were treated differently than whites in each region.<sup>26</sup> Stepwise regression may be appropriate to use when there is no existing theory on which to build a model. However, social science standards hold that when there is existing theory, stepwise regression is not an appropriate way to choose variables. In the case of SSA's study, statutes, regulations, rulings and SSA guidance establish the factors that ALJs should consider in determining eligibility, and thus indicate which variables should be included in a model. By using the results of stepwise regression, SSA's regional models included variables that were statistically significant but reflected the reviewer's evaluation of the hearing proceedings—which an ALJ would not consider in a hearing—and therefore were not appropriate. As mentioned earlier, including these

<sup>&</sup>lt;sup>26</sup>In its internal working papers, SSA tested whether African Americans were treated differently than non-African Americans. Specifically, SSA created a set of "race specific models" by multiplying each variable in the regional models by a dummy variable for race. SSA used a standard statistical test, called the log-likelihood ratio test, to test whether differences in the results of the models with and without the race variables were statistically significant. SSA informed us that its concluding analyses, which were not available for our review, were based on African American claimants and white claimants and also showed no differences by race.

variables may have reduced the explanatory power of other variables—such as race; this, in conjunction with the use of stepwise regression, may explain why race did not show up as statistically significant in the regional models. Had SSA chosen the variables for its model on the basis of theory and its own guidelines, race may have been statistically significant. The statisticians hired by SSA as consultants also noted this as a concern.

According to an SSA official, the analysts directly responsible for or involved in the study conducted other analyses that were not reflected in the documentation currently available and provided to us. For example, this SSA official told us that the analysts involved in the study would have tested the statistical significance of slight differences between the cases included and not included in the final sample. This official also said that the analysts used multiple techniques in addition to stepwise regression—and ran the models with and without variables that reflected posthearing evaluations—and still found no evidence of racial bias. However, due to the lack of available documentation, we were unable to review these analyses or corroborate that they were performed.

#### Future Studies of Racial Disparities Would Be Hampered by Data Limitations

Since the conclusion of its study of racial disparities, SSA no longer analyzes race as part of its ongoing quality review of ALJ decisions, and SSA officials told us they have no plans to do so in the future. SSA still samples and reviews ALJ decisions for quality assurance purposes. However, since 1997, SSA no longer stratifies ALJ decisions by race before identifying a random sample of cases—a practice that had helped to ensure that SSA had a sufficient number of cases in each region to analyze decisions by race. Although the dataset used for SSA's ongoing quality assurance review of ALJ decisions still includes information on race, SSA no longer analyzes these data to identify patterns of racial disparities.

Even if SSA decided to resume its analysis of racial disparities in ALJ decisions, it would encounter two difficulties. First, SSA collects files for only about 50 percent of sampled cases in its ongoing review of ALJ decisions for quality assurance purposes, such that its final samples may be nonrepresentative of the universe of ALJ decisions. SSA uses this review data to produce annual and biennial reports on ALJ decision making.<sup>27</sup> Data in these reports are also used to calculate the accuracy of

<sup>&</sup>lt;sup>27</sup>SSA has issued findings from its Disability Hearings Quality Review Process on a biennial basis since 1995. In 1999, SSA also began issuing a less comprehensive report of its findings on an annual basis.

ALJ decisions—a key performance indicator used in SSA's 2000–03 performance plans pursuant to the Government Performance and Results Act. The reasons for obtaining only half of the files are the same, potentially biasing reasons as for the racial disparities study—files are either missing or not made available if the cases are in use for appeals or pending decisions. However, SSA's annual and biennial reports do not cite the number or percentage of case files not obtained for specific reasons. In addition, SSA officials told us that they do not conduct ongoing analyses to test the representativeness of samples used for quality assurance purposes, and SSA's annual and biennial reports do not address whether the final sample used for quality assurance purposes and for calculating the performance indicator for ALJ accuracy is representative of the universe of ALJ decisions. In addition to not obtaining about 50 percent of the case files, SSA officials told us that medical consultants and disability examiners only review a portion of cases for which a file was obtained due to limited resources.

Second, future analyses of racial disparities at either the DDS or hearings level is becoming increasingly problematic because, since 1990, SSA no longer systematically collects race data as part of its process in assigning Social Security Numbers (SSN).<sup>28</sup> For many years, SSA has requested information on race and ethnicity from individuals who complete a form to request a Social Security card. Although this process is still in place, since 1990 SSA has been assigning SSNs to newborns through its Enumeration at Birth (EAB) program, and SSA does not collect race data through the EAB program.<sup>29</sup> Under current procedures, SSA is unlikely to subsequently obtain information on race or ethnicity for individuals assigned SSNs at birth unless those individuals apply for a new or replacement SSN (due to change in name or lost card). As of 1998, SSA did not have data on race or ethnicity for 42 percent of SSI beneficiaries under the age of 9. As future

<sup>&</sup>lt;sup>28</sup>Because race is not a factor in determining DI or SSI eligibility, SSA does not ask claimants to provide information on their race. Race information may be present in the claimant's file to the extent it is provided by treating physicians as part of the medical documentation relating to the claimant. Because this information is not self-reported, it may differ from a claimant's view of his or her race.

<sup>&</sup>lt;sup>29</sup>The EAB program allows parents to apply for an SSN for their child as part of the hospital birth registration process. The hospital transfers data to the state Bureaus of Vital Statistics and SSA contracts with the Bureaus to electronically transfer certain fields from the birth file to SSA. The birth file for most states includes race data for parents, but only one state (Washington) includes race data for the baby. These data are not transferred to SSA.

generations obtain their SSNs through the EAB program, this number is likely to increase.<sup>30</sup>

# SSA Has Taken Limited Steps to Address Possible Racial Bias in Its Hearings Level Decision-Making Process

Concurrent with SSA's study of racial disparities, SSA's Office of Hearings and Appeals took several steps to address possible racial bias in disability decision making at the hearings level. These steps included providing diversity training, increasing recruitment efforts for minority ALJs, and administering a new complaint process for the hearings level to help ensure fair and impartial hearings. The complaint process was intended, in part, to help identify patterns of possible racial and ethnic bias and other misconduct; however, this process lacks mechanisms to help OHA easily identify patterns of possible racial or ethnic bias for further investigation or corrective action.

#### OHA Has Taken Some Steps to Address Possible Racial Bias

SSA's OHA adopted a mandatory diversity sensitivity program in 1992. All of SSA's incumbent ALJs were required to attend a 2- or 1-1/2-day course immediately after its development. In addition, the course (now 1 day in length) is included in a 3-week orientation for newly hired ALJs. The course was designed and is conducted by an outside contractor. The course addresses topics such as cultural diversity, geographic diversity, unconscious bias, and gender dynamics through a series of exercises designed to help the ALJs understand how their thought processes, beliefs, and past experiences with people influence their decision-making process.

OHA also increased its efforts to recruit minorities for ALJ and other legal positions, although the impact of these efforts on the racial/ethnic mix of SSA's ALJ workforce has been limited. According to OHA officials, OHA has attended conferences held by several minority bar organizations, to raise awareness about the opportunities available at SSA to become an ALJ.<sup>31</sup> In addition to having information booths and distributing information on legal careers at OHA, OHA presented a workshop called "How to Become an Administrative Law Judge at OHA," at each

 $<sup>^{30}</sup>$ In 2001, 69 percent of the original SSNs SSA assigned that year were processed through the EAB program, and approximately 75 percent of newborns received their SSNs through EAB

<sup>&</sup>lt;sup>31</sup>These minority organizations have included the Hispanic Bar Association, the National Bar Association, the Federal Bar Association, the Asian American Bar Association, and various Women's Bar conventions.

conference. Despite these efforts, there have not been significant changes in the racial/ethnic profile of SSA ALJs.<sup>32</sup>

In addition to these efforts, in 1993 SSA instituted a complaint process under the direction of OHA that provides claimants and their representatives with a new mechanism for voicing complaints specifically about bias or misconduct by ALJs. The ALJ complaint process supplements and is coordinated with the normal appeals process. All SSA claimants have the right to appeal the ALJ decision to the Appeals Council and, in doing so, may allege unfair treatment or misconduct. According to OHA officials, the vast majority of allegations of unfair hearings are submitted by claimants or their representatives in connection with a request for Appeals Council review. Under the 1993 process, claimants or their representatives may also file a complaint at any SSA office, send it by mail, or call it into SSA's 800 number service. Any complaints where there is a request for Appeals Council review are referred to the Appeals Council for its consideration as part of its normal review. For complaints where

<sup>&</sup>lt;sup>32</sup>In 1992, 91.8 percent of SSA's ALJs were white. Ten years later, in April 2002, 89.0 percent were white. OHA officials attributed the limited progress in minority hiring of ALJs to restrictive Office of Personnel Management requirements. For example, OHA officials noted that ALJs must be hired from the Office of Personnel Management's register, which gives preference to veterans and is traditionally underrepresented by minorities. Officials also pointed out that because of litigation involving the register of ALJ candidates, SSA has been prevented from hiring ALJs since 1999—with an exception in 2001. Finally, officials noted that they have increased the number of Regional Chief ALJs who are members of a racial minority group from 1 to 3 (out of 10) since 1992.

<sup>&</sup>lt;sup>33</sup>SSA published a notice of procedures for the new ALJ complaint process in the Federal Register on October 30, 1992. Claimants are informed of ALJ complaint procedures through instructions posted in all hearing offices and on SSA's Web site. Claimants may also learn about the process through their attorneys.

<sup>&</sup>lt;sup>34</sup>In addition to the complaint and appeals processes, another process is available to claimants who feel they were discriminated against. All SSA customers, including claimants, can file a discrimination complaint to SSA's Office of General Counsel, which is responsible for conducting civil rights investigations for SSA. According to SSA officials, it is rare for claimants to file complaints of bias by ALJs to the Office of General Counsel. For example, in 2001 the Office of General Counsel only received about 10 civil rights complaints involving ALJs. According to officials, complaints received by the Office of General Counsel against ALJs are generally referred directly to OHA's Special Counsel Staff; that is, the Office of General Counsel does not conduct its own investigation unless it specifically involves a civil rights allegation.

<sup>&</sup>lt;sup>35</sup>The Special Counsel Staff receive from the Appeals Council copies of all claims in which a bias/unfair hearing complaint has been made—whether the complaint was found to be supported or not. The Appeals Council sends the Chief ALJ only those claims in which the bias/unfair hearing complaint was found to be supported by the record.

the complainant did not request an Appeals Council review, the complaint is reviewed by the appropriate Regional Chief ALJ, and the findings are reported to the Chief ALJ. Regardless of how the complaint was filed or which office reviewed the complaint, OHA's Special Counsel Staff is notified of all claims and any findings from either the Appeals Council or the Regional Chief ALJ. On the basis of these findings, OHA may decide to take remedial actions against the ALJ, such as a counseling letter, additional training, mentoring or monitoring, an official reprimand, or some other adverse action. OHA's Special Counsel Staff may also decide to conduct a further investigation. Regardless of which office handles the complaint, OHA acknowledges the receipt of each complaint in writing, notifies the complainant that there will be a review or investigation (unless to do so would disrupt a pending hearing or decision), and notifies the complainant concerning the results of the investigation.

Officials from the Special Counsel Staff told us OHA receives about 700 to 1,000 complaints (out of 400,000 to 500,000 hearings) per year. About 90 percent of these are notifications from the Appeals Council that involve an allegation of bias or misconduct. Officials from the Special Counsel Staff also said that few complaints are related to race. For example, officials noted that, in response to a special request, Special Counsel Staff

<sup>&</sup>lt;sup>36</sup>Specifically, for complaints where there is no request for Appeals Council review, or the Appeals Council did not address the complaint, the complaint is forwarded to the appropriate Regional Chief ALJ to determine the facts pertinent to the case. The Regional Chief ALJ reports the findings and any recommendations to the Chief ALJ at OHA headquarters. The Chief ALJ decides whether the report supports some remedial action and forwards the report, findings, and recommendations to the Special Counsel Staff.

<sup>&</sup>lt;sup>37</sup>Other adverse actions may include a removal, suspension, reduction in grade, reduction in pay, and furlough of 30 days or less.

<sup>&</sup>lt;sup>38</sup>The Special Counsel Staff told us that, if the Appeals Council or the Chief ALJ does not have sufficient information to take an appropriate action but there is enough on the record to raise a question about a judge's conduct, the Special Counsel Staff would conduct a full field investigation of the complaint, which entails reviewing the file, listening to the hearing tapes, and interviewing the complainant and collateral witnesses. On the other hand, the Special Counsel Staff does not conduct in-depth investigations of complaints for which the Appeals Council or Chief ALJ either found no basis for misconduct or collected sufficient information to take appropriate action.

<sup>&</sup>lt;sup>39</sup>These officials also noted that, annually, about 15-20 complaints filed with the Special Counsel Staff result in full field investigations.

 $<sup>^{40} \</sup>rm{These}$  notifications are out of the approximately 115,000 reviews conducted by the Appeals Council annually.

reviewed all 372 complaints filed during the first 6 months of 2001, and found that only 19 (5.1 percent) were in some way related to race.<sup>41</sup>

#### ALJ Complaint Process Lacks Mechanisms to Identify Patterns of Racial Bias

While the ALJ complaint process provides a mechanism for claimants to allege discrimination, it lacks useful mechanisms for detecting patterns of possible racial discrimination. In SSA's public notice on the creation of this process, it was stated that SSA's Special Counsel would "collect and analyze data concerning the complaints, which will assist in the detection of recurring incidences of bias or misconduct and patterns of improper behavior which may require further review and action." However, OHA's methods of collecting, documenting, and filing complaints make this difficult to do. For example, in its instructions to the public, SSA directs complainants to describe, in their own words, how they believe they were treated unfairly. This flexible format for filing complaints may make it difficult for OHA to readily identify a claim alleging racial bias. In contrast, SSA's Office of General Counsel's complaint form specifically asks complainants to categorize their claim as being related to such factors as race or sex.

Similarly, OHA does not use a standardized internal cover-sheet to summarize key aspects of the review, such as whether the complaint involved racial or some other type of bias or misconduct, and whether the complaint had merit and what action, if any, was taken. The lack of a cover-sheet makes it difficult to quickly identify patterns of allegations involving race that have merit. In order to determine whether patterns exist, OHA staff would have to reread each complaint.

Additionally, OHA staff do not record key information about complaints—such as the nature of the complaint—in an electronic database so that patterns of bias can be easily identified. OHA's Special Counsel Staff files complaints and related documents manually, and in chronological order by hearing office. <sup>43</sup> According to OHA officials, this filing system was

 $<sup>^{\</sup>rm 41}\!\text{According}$  to the officials, these complaints were out of 170,225 hearings during this 6-month period.

<sup>&</sup>lt;sup>42</sup>57 Fed. Reg. 49186, 49187 (1992).

<sup>&</sup>lt;sup>43</sup>SSA's Office of the Chief ALJ maintains its complaint data in an electronic format. However, this database is not as complete as the Special Counsel Staff's manual files in that it does not include claims considered by the Appeals Council and found to be not supported by the record. The Chief ALJ's electronic database also does not capture information on the nature of the complaint.

developed in 1993 when the process was created and complaint workloads were much lower. Today, SSA receives and reviews 700 to 1,000 complaints a year. In order to identify patterns of bias, Special Counsel Staff must not only reread each file, it must tabulate patterns by hand—a time-consuming process that it does not perform on a routine basis.

Finally, OHA does not currently obtain demographic information (such as race, ethnicity, and sex) on complainants, which are important in identifying patterns of bias. These data are important for identifying patterns of possible racial bias because complainants—aware only of their own circumstances and lacking a basis for comparison—may not specifically allege racial bias when they file a complaint about unfair treatment. Without demographic data, it is impossible to discern whether certain types of allegations are disproportionately reported by one race (or sex) and whether further investigative or corrective action is warranted. Although SSA is currently obtaining less race data in its process of assigning SSNs, OHA staff could still obtain data on race and sex for most complainants from the agency's administrative data.

#### Conclusions

The steps SSA has taken over the last decade have not appreciably improved the agency's understanding of whether or not, or to what extent, racial bias exists in its disability decision-making process. SSA's attempt to study racial disparities was a step in the right direction, but methodological weaknesses evident in SSA's remaining working papers prevent our concluding, as SSA did, that there is no evidence of racial bias in ALJ decision making. SSA does not have an ongoing effort to demonstrate the race neutrality of its disability programs. Moreover, the continuing methodological weaknesses in SSA's ongoing quality assurance reviews of ALJ decisions hamper not only its ability to ensure the accuracy of those reviews but also its ability to conduct future studies to help ensure the race neutrality of its programs. Furthermore, in the longer term, SSA's ability to analyze racial differences in its decision making will diminish due to a lack of data on race and ethnicity. Finally, SSA's complaint process for ALJs lacks mechanisms—such as summaries of key information on each complaint, an electronic filing system, and information on the race and ethnicity of complainants—that could help identify patterns of possible bias. SSA is not legally required to collect and monitor data to identify patterns of racial disparities, although doing so would help SSA to demonstrate the race neutrality of its programs and, if a pattern of racial bias is detected, develop a plan of action.

#### Recommendations

To address shortcomings in SSA's ongoing quality assurance process for ALJs—which would improve SSA's assessment of ALJ decision-making accuracy—we recommend the agency take the following steps:

- conduct ongoing analyses to assess the representativeness of the sample
  used in its quality assurance review of ALJ decisions, including testing the
  statistical significance of differences in key characteristics of the cases
  included in the final sample with those that were not obtained;
- include the results of this analysis in SSA's annual and biennial reports on ALJ decision making; and
- use the results to make appropriate changes, if needed, to its data collection or sampling design to ensure a representative sample.

To more readily identify patterns of misconduct, including racial bias, in complaints against ALJs, we also recommend that SSA's Office of Hearing and Appeals:

- adopt a form or some other method for summarizing key information on each ALJ complaint, including type of allegation;
- use internal, administrative data, where available, to identify and document the race and/or ethnicity of complainants; and
- place the complaint information in an electronic format, periodically analyze this information and report the results to the Commissioner, and develop action plans, if needed.

# Agency Comments and Our Response

We provided a draft of this report to SSA for comment. SSA concurred fully with our recommendations and agreed to take steps to implement them. In its general comments, SSA expressed concern that the title of the report might foster the perception that its disability decision making, particularly at the OHA level, is suspect. Although we believe the draft report's title accurately reflected the report's content and recommendations, we have modified the title to ensure clarity. SSA also cited a number of reasons for the low percentage of cases included in its final sample, as well as steps it took to ensure the representativeness of its final sample. Nevertheless, we continue to believe that SSA could have performed additional analyses to provide more assurance of the sample's representativeness.

SSA also provided technical comments and clarifications, which we incorporated in the report, as appropriate. SSA's general comments and our response are printed in appendix I.

We are sending copies of this report to the Social Security Administration, appropriate congressional committees, and other interested parties. We will also make copies available to others on request. In addition, the report will be available at no charge on the GAO Web site at http://www.gao.gov.

If you or your staff have any questions concerning this report, please call me or Carol Dawn Petersen, Assistant Director, at (202) 512-7215. Staff acknowledgments are listed in appendix II.

Robert E. Robertson

Director, Education, Workforce, and Income Security Issues

# Appendix I: Comments from the Social **Security Administration**



August 13, 2002

Mr. Robert E. Robertson Director, Education, Workforce and Income Security Issues

U.S. General Accounting Office Washington, D.C. 20548

Dear Mr. Robertson:

Enclosure

Thank you for the opportunity to review and comment on the draft report, "Social Security Administration Disability Decision-Making: Additional Measures Would Enhance Agency's Understanding of Potential Racial Bias" (GAO-02-831). Our comments on the report are enclosed. If you have any questions, please have your staff contact Trudy Williams at (410) 965-0380.

SOCIAL SECURITY ADMINISTRATION BALTIMORE MD 21235-0001

COMMENTS OF THE SOCIAL SECURITY ADMINISTRATION (SSA) ON THE GENERAL ACCOUNTING OFFICE (GAO) DRAFT REPORT, "SOCIAL SECURITY ADMINISTRATION DECISION-MAKING: ADDITIONAL MEASURES WOULD ENHANCE AGENCY'S UNDERSTANDING OF POTENTIAL RACIAL BIAS" (GAO-02-831)

We are concerned about the title of the report because the current title may foster the perception that the Agency's disability decision making, particularly at the OHA, is suspect. We suggest that the GAO use a title that clearly reflects that GAO has yet to find that the Agency has discriminated against claimants based on race. Furthermore, the title suggests the report addresses all races, when in fact it looks only at two races. A suggested title may be--"Additional Measures to Enhance the Perceived Fairness of Agency Decision-Making."

#### Recommendation 1

SSA should conduct ongoing analyses to assess the representativeness of the sample used in its quality assurance review of Administrative Law Judge (ALJ) decisions, including testing the statistical significance of differences in key characteristics of the cases included in the final sample with those that were not obtained.

#### Comment

We agree. SSA's Office of Quality Assurance (OQA) will conduct ongoing analyses to assess the representativeness of the sample used in our quality assurance review of ALJ decisions, including testing the statistical significance of differences in key characteristics of the cases included in the final sample with those that were not obtained.

#### Recommendation 2

SSA should include the results of this analysis in SSA's annual and biennial reports on ALJ decision-making.

#### Comment

We agree and will include the results of this analysis in SSA's reports on ALJ decision-making.

#### Recommendation 3

SSA should use the results to make appropriate changes, if needed, to its data collection or sampling design to ensure a representative sample.

#### Comment

We agree and will use the results to make appropriate changes, if needed, to our data collection or sampling design to ensure a representative sample.

See comment 1.

2

#### Recommendation 4

The Office of Hearings and Appeals should adopt a form or some other method for summarizing key information on each ALJ complaint, including type of allegation.

#### Comment

We agree. We will design and implement a tracking form to capture key data from each complaint.

#### Recommendation 5

OHA should use internal, administrative data, where available, to identify and document the race of complainants.

#### Comment

We concur with this recommendation. We agree with GAO that such data is increasingly unavailable and, for claimants who are children, may not be available at all. We will, however, explore what data is available while developing the information tracking system mentioned above. At a minimum, we will be able to identify and quantify actual complaints based on race.

#### Recommendation 6

OHA should place the complaint information in an electronic format, periodically analyze this information and report the results to the Commissioner, and develop action plans, if needed.

#### Comment

We concur with this recommendation and will include electronic format and analysis in development of the tracking system.

#### Other Comments

Diversity in the ALJ Corps—On page 15, sentence 1, GAO notes that, "Despite these efforts, OHA has not achieved significant changes in the racial/ethnic profile of SSA ALJs." Only in footnote 27 on page 15 does GAO mention that "OHA officials attributed the limited progress in minority hiring to restrictive Office of Personnel Management requirements." In addition, GAO failed to mention the impact of the litigation in *Azdell v. OPM*, which has prevented hiring of ALJs in the Federal government since 1999 (with one limited exception in 2001). Finally, we note that our own intensified efforts have resulted in 3 of our 10 Regional Chief ALJs (30 percent) being of a racial minority. This increase is significant, since in 1992 only one Regional Chief ALJ was a racial minority.

See comment 2.

3

#### See comment 3.

On pages 2, 7 and 9 of the draft report, GAO discusses the Disability Hearings Quality Review Process' (DHQRP) target sample selection and actual sample realization rates. This discussion specifically implies that the low sample realization rate (approximately 25 percent) is primarily due to a failure to obtain case folders and hearing cassette tapes (which for hearing allowances are stored in separate locations). Such nonreceipt of folders and tapes is a factor in GAO questioning whether the realized sample is representative of the universe. However, the low sample realization rate is actually due to several factors that developed over time.

For example, during the planning, design and initial implementation of the DHQRP, a sample target of 15,000 cases was planned in order to produce statistically reliable data stratified by race and region.

- Due to various systems limitations, the automated sample selection process
  was unable to precisely discern the types of hearing cases that were to
  undergo this review. In fact, the DHQRP employs at least 20 manual
  screening criteria which would cause a sample case to be excluded from
  review after selection; and
- 2) Based on advice OQA received from SSA's Office of the General Counsel, the DHQRP process did not review "live" hearing denials while they were on appeal to the Appeals Council (AC). Moreover, the AC's processing time meant that we were not being able to obtain case files for 1 2 years after sample selection. Appealed denials continue to be included in our sample (about half the hearing denials are appealed).

When the DHQRP began, OQA did not anticipate any of the folder and tape retrieval difficulties that impacted us. Upon experiencing those retrieval problems after start-up, OQA found it necessary to maintain a sample selection process that intentionally selects more cases for DHQRP review than the Division of Disability Hearings Quality (DDHQ) has the planned capacity to review. The sample selection rate was never expected to be realized once the DHQRP was underway, but was found to be necessary in order to ensure that the sample needed for statistically reliable biennial reporting of data at the regional level could be achieved.

Ultimately, the best way to guard against nonrepresentativeness in the sample is to establish procedures, and monitor sample selection, to ensure that the sample is random. By the use of holdout samples and cross modeling in our analysis, OQA was satisfied that the group of cases sampled for this study was essentially free of sampling bias.

However, due to logistic problems in retrieving cases, not all cases selected for the sample were actually reviewed. To address the question as it applies to the body of cases reviewed, we used key characteristics of the two groups of cases (reviewed and not reviewed, or responders and non-responders) to see whether they contain different proportions of cases with these characteristics.

See comment 4.

# Appendix I: Comments from the Social Security Administration

	4
See comment 5.	In this connection, the important question is not whether these differences are statistically significant (because with the sample sizes in question, any difference larger than a percentage point is statistically significant), but whether they are large enough to have any practical effect on the variables in the main analysis. This is a question that OQA cannot answer conclusively, but in the judgment of OQA statistical staff, who are knowledgeable in this area and have examined the model input, the cases that were included in the analysis are a substantially accurate reflection of the entire universe.

#### **GAO Comments**

- 1. Although we believe the draft report's title accurately reflected the content of our report and recommendations, we have modified the title to ensure clarity. This report and its recommendations are not restricted to a discussion of only two races. Although we referred to race and ethnicity in the second objective and the conclusion section of the draft reviewed by Social Security Administration (SSA), we added the word "ethnicity" to the recommendations and the body of the report to further clarify this issue.
- 2. We added language to a note in the report regarding the litigation SSA mentions and that SSA has increased the number of Regional Chief Administrative Law Judges (ALJs) who are members of a racial minority group from 1 to 3 since 1992.
- 3. We agree with SSA that the low proportion of cases included in the final sample is due to several factors. In our report, we cited several reasons for cases not being included in the final sample that are significant in terms of the number of affected cases and that we believe have the potential for being nonrandom in nature. On the basis of a subsequent discussion with SSA officials, we added a note in our report that a small proportion of cases were excluded because they were later identified as being cases that were not intended to be included in the sample.
- 4. Although SSA noted that it used "holdout samples and cross modeling" to ensure that the group of cases sampled for this study was essentially free of sampling bias, SSA officials explained to us that these techniques were not used to test for the representativeness of the final sample.
- 5. We agree with SSA that, with large sample sizes, even small differences generally are statistically significant, and that such statistically significant differences are not always substantively significant. We do not believe, however, that a large sample is sufficient reason to forego significance tests. Moreover, our report cited additional analyses that SSA could have performed to provide more assurance of the sample's representativeness. Another approach that we do not cite in the report—but which SSA may wish to consider—is multivariate analysis of nonresponse. SSA performed bivariate comparisons of samples to determine whether they contained different proportions of cases with various characteristics. However, two samples can have very similar percentages of, for example, women and African Americans, but be very different with respect to the percentage of African American women. In contrast, multivariate analysis would allow SSA to look

Appendix I: Comments from the Social Security Administration		
systematic simultaneo	cally and rigorously at different characteristics ously.	

# Appendix II: GAO Contacts and Staff Acknowledgments

#### **GAO Contacts**

Robert E. Robertson (202) 512-7215 Carol Dawn Petersen (202) 512-7215

### Staff Acknowledgments

In addition to those named above, the following individuals made significant contributions to this report: Erin Godtland, Michele Grgich, Stephen Langley, and Ann T. Walker, Education, Workforce and Income Security Issues; Doug Sloane and Shana Wallace, Applied Research and Methods; and Jessica Botsford and Richard Burkard, General Counsel.

#### **GAO's Mission**

The General Accounting Office, the investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

# Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through the Internet. GAO's Web site (www.gao.gov) contains abstracts and full-text files of current reports and testimony and an expanding archive of older products. The Web site features a search engine to help you locate documents using key words and phrases. You can print these documents in their entirety, including charts and other graphics.

Each day, GAO issues a list of newly released reports, testimony, and correspondence. GAO posts this list, known as "Today's Reports," on its Web site daily. The list contains links to the full-text document files. To have GAO e-mail this list to you every afternoon, go to <a href="https://www.gao.gov">www.gao.gov</a> and select "Subscribe to daily E-mail alert for newly released products" under the GAO Reports heading.

#### Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. General Accounting Office 441 G Street NW, Room LM Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000

TDD: (202) 512-2537 Fax: (202) 512-6061

## To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

#### **Public Affairs**

Jeff Nelligan, managing director, NelliganJ@gao.gov (202) 512-4800 U.S. General Accounting Office, 441 G Street NW, Room 7149 Washington, D.C. 20548

