United States General Accounting Office

GAO

Report to the Chairman, Subcommittee on Social Security, Committee on Ways and Means, House of Representatives

August 1997

SOCIAL SECURITY DISABILITY

SSA Must Hold Itself Accountable for Continued Improvement in Decision-making

GAO/HEHS-97-102

United States
General Accounting Office
Washington, D.C. 20548

Health, Education, and
Human Services Division

B-270669

August 12, 1997

The Honorable Jim Bunning
Chairman, Subcommittee on Social Security
Committee on Ways and Means
House of Representatives

Dear Mr. Chairman:

This report, prepared at your request, evaluates the Social Security Administration's decision-making process for disability determinations and assesses its efforts to improve the process.

As agreed with your office, we are sending copies of this report to the Commissioner of the Social Security Administration and the Director of the Office of Management and Budget. We will also make copies available to others on request.

If you have any questions about this report, please call me at (202) 512-7215. Other major contributors to this report are listed in appendix IV.

Sincerely yours,

Jane L. Ross

Jane L. Ross
Director, Income Security Issues

Executive Summary

Purpose

Each year, about 2.5 million people apply to the Social Security Administration (SSA) for disability benefits. Disability determination services (DDS), which are state agencies that conduct disability determinations on behalf of SSA, award benefits to about 35 out of every 100 of these initial applicants. Of the 65 denied applicants, about 43 abandon their claims, and about 22 appeal to administrative law judges (ALJ). On appeal, 14 of 22 claimants, or almost two-thirds, are subsequently awarded benefits. This rate of ALJ benefit awards raises concerns in the Congress and elsewhere about the accuracy of disability DDS and ALJ decisions, length of time claimants must wait for a decision if they appeal, and costliness of deciding cases on appeal rather than upon initial application.

In 1995 testimony before the Social Security Subcommittee of the House Committee on Ways and Means, GAO reported on the timeliness and consistency of SSA's disability decisions.[1] On the basis of that testimony, the Chairman asked GAO to report on (1) factors that contribute to differences between DDS and ALJ decisions and (2) SSA's actions to make decisions in initial and appealed cases more consistent. This report details GAO's findings, which were reported in testimony earlier this year.[2]

Background

SSA operates the Disability Insurance (DI) and Supplemental Security Income (SSI) programs—the two largest federal programs providing cash benefits to people with disabilities. The law defines disability for both programs as the inability to engage in any substantial gainful activity by reason of a severe physical or mental impairment that is medically determinable and is expected to last at least 12 months or result in death. The programs have grown in the last 10 years, and today over 7 million working-age adults are on the rolls. These and other beneficiaries receive cash benefits totaling about $61 billion a year.[3]

Disability determinations begin at the DDSs, where a disability examiner and a medical or psychological consultant, working as a team, analyze an applicant's documentation, gather additional evidence as appropriate, and make a disability determination. Denied applicants may ask the DDS to

[1]Social Security Disability: Management Action and Program Redesign Needed to Address Long-Standing Problems (GAO/T-HEHS-95-233, Aug. 3, 1995).

[2]Social Security Disability: SSA Actions to Reduce Backlogs and Achieve More Consistent Decisions Deserve High Priority (GAO/T-HEHS-97-118, Apr. 24, 1997).

[3]Included in the $61 billion of benefits are payments to all SSI blind and disabled beneficiaries regardless of age.

reconsider its finding, and if denied again, may appeal to an ALJ. The ALJ usually conducts a hearing and must consider the findings of the DDS medical consultant but is not legally bound by them. In addition, claimants may testify before an ALJ and present new evidence. Claimants whose appeals are denied may request review by SSA's Appeals Council and then may file suit in federal court. The average initial DDS decision in DI cases costs about $540, though a hearing can cost an additional $1,200. In addition, appeals can add an average of 378 days to the length of time that a claimant must wait for a final decision.

Both DDS and ALJ adjudicators use a sequential evaluation process when determining disability. Under this process, applicants are awarded benefits when their medical condition meets or equals criteria in SSA's regulations (commonly referred to as the medical listings). For those whose condition does not meet or equal the listings, the adjudicators focus on the functional consequences of applicants' medically determined impairments.

As part of its 1994 plan for redesigning the disability determination process, SSA set a goal of "making the right decision the first time." As a first step, SSA has begun an initiative, called process unification, to improve the consistency of its decisions. Under redesign, the agency expects more award decisions to be made by the DDSs, reducing the need for appeals. Meanwhile, SSA faces several other competing demands, including significant increases in continuing disability reviews and increasing SSI workloads mandated by recent legislation.[4] Over the longer term, SSA plans to improve its methods for assessing applicants' capacity to function in the workplace.

Results in Brief

ALJs made nearly 30 percent of all awards in 1996. Moreover, because two-thirds of all cases appealed to ALJs have resulted in awards, questions have arisen about the fairness, integrity, and cost of SSA disability programs. Differences in assessing applicants' functional capacity and procedural factors, as well as weaknesses in quality assurance, contribute to inconsistent decisions.

Differences in assessing functional capacity help explain the inconsistent decisions of ALJs and DDSs. ALJs are far more likely than DDSs to find claimants unable to work on the basis of their functional capacity.

[4]The Social Security Independence and Program Improvements Act of 1994 and the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 increased claims workloads for drug addicts and alcoholics, noncitizens, and children on SSI, and both significantly increased SSA's requirements to conduct continuing disability reviews.

Moreover, this outcome has occurred even when ALJ and DDS adjudicators review the same evidence for the same case. Most notably, DDS adjudicators tend to rely on medical evidence such as the results of laboratory tests; ALJs tend to rely more on symptoms such as pain and fatigue. In addition, the opinions of claimants' own physicians may more likely influence ALJs than DDSs; DDSs may give more weight to other medical evidence such as laboratory findings.

DDS and ALJ decision-making practices and procedures also contribute to inconsistent results because they limit the usefulness of DDS evaluations as bases for ALJ decisions. For instance, DDSs often do not ensure that medical consultants write adequate explanations of their opinions. SSA regulations require ALJs to consider these explanations, but this has little practical value if the explanations are not well documented. In addition, SSA procedures often lead to substantial differences between the evidentiary records examined by DDSs and ALJs. Specifically, ALJs may examine new evidence submitted by a claimant and hear a claimant testify. As a result, even with a well-explained DDS decision, ALJs could reach a different decision because the evidence in the case differs from that reviewed by the DDS.

Finally, SSA has not used its quality review systems to identify and reconcile differences in approach and procedures used by DDSs and ALJs. In fact, the quality review systems for the initial level and appeals levels of the decision-making process merely reflect the differences between the levels; they do not help produce more consistent decisions.

Although SSA has not managed the decision-making process well in the past, its current process unification initiatives, when fully implemented, could significantly help to produce more consistent decisions. Competing workload pressures at all adjudication levels could, however, jeopardize SSA's efforts. As a result, SSA, in consultation with the Congress, will need to sort through its many priorities and be more accountable for meeting its deadlines and establishing explicit measures to assess its progress in reducing inconsistency. This may include, for example, setting a goal, under the Government Performance and Results Act, to foster consistency in results, set quantitative measures, and report on its progress in shifting the proportion of cases awarded from the ALJ to the DDS level.

Principal Findings

DDSs and ALJs Differ Mainly Over Claimants' Functional Abilities

Differences in assessing claimants' residual functional capacity (RFC) by DDSS and ALJS are the main reason for most ALJ awards, according to GAO's analysis. ALJS are much more likely than DDSS to find that claimants have severe limitations in functioning in the workplace, as indicated by an ongoing SSA study of the appeals process. For instance, in the view of awarding ALJS, 66 percent of cases merited an RFC of "less than the full range of sedentary work"—a classification that often leads to an award. In contrast, DSS reviewers found that less than 6 percent of the cases merited this classification. DDS and ALJ differences in assessments were also apparent in a 1982 SSA study that controlled for differences in evidence. This study indicated that DDS and ALJ adjudicators often reach different results even when presented with the same evidence. Specifically, DDS reviewers would have awarded benefits in 13 percent of the cases, while ALJS would have awarded benefits in 48 percent of the cases.

The use of medical experts and the application of judgment in weighing evidence seem to influence the differences in DDS and ALJ decisions. For example, at the DDSS, medical or psychological consultants assess applicants' RFC. DDSS appear to rely more on objective medical findings when assessing the impact of symptoms, such as pain and fatigue, on functional capacity. In contrast, ALJS have the sole authority to determine RFC and often rely on a claimant's testimony and treating physicians' opinions. Although ALJS may have independent medical experts testify at hearings, only about 8 percent of cases in which benefits are awarded have used such experts, according to our analysis.

SSA issued rulings in July 1996, which were written to clarify ALJS' use of DDS medical consultants' findings, treating source opinion, and assessing RFC. In addition, SSA plans to issue a regulation to provide further guidance on assessing RFC for both DDSS and ALJS, specifically clarifying when a "less-than-sedentary" classification is appropriate. SSA expects this classification to be used rarely.

DDS Evaluations of Limited Use to ALJs

Several factors at both the DDS and ALJ levels limit the usefulness of DDS evaluations as bases for ALJ decisions. Often, ALJS cannot rely on DDS evaluations because they lack the supporting evidence and explanations of

the reasons for the denial, laying a weak foundation for an ALJ decision if an applicant appeals the case. Moreover, although SSA requires ALJs to consider the DDS medical consultants' assessments of RFC, DDS procedures do not ensure that such assessments are clearly explained. Without this, an ALJ could neither effectively consider such assessments nor give them much weight.

At the ALJ level, claimants may submit additional evidence and claim new impairments. This also affects the consistency of DDS and ALJ decisions. Claimants submit additional evidence in about 75 percent of appealed cases; and, in about 27 percent of hearing allowances, additional evidence is an important factor in the decision. In about 10 percent of appealed cases, claimants switch their primary impairment from a physical to a mental one.

SSA has acknowledged the need to ensure that DDS decisions are better explained and based on a more complete record so that they are more useful if appealed. The agency plans to issue instructions and provide additional training for the DDSs on how and where in the case files to explain their decisions and on explaining the decisions. SSA also plans to issue a regulation clarifying the reliance on DDS medical consultants' opinions at the ALJ level. To deal with the possible effect of new evidence, SSA plans to return about 100,000 selected cases a year to the DDSs for further consideration when new evidence is introduced at the ALJ level. The DDSs might award benefits at this point, eliminating the need for costly and time-consuming ALJ hearings.

Quality Review Systems Neither Identify nor Reconcile Inconsistency Between DDS and ALJ Decisions

SSA has several quality review systems for disability decisions, each with its own specific purpose; none, however, is designed to identify and reconcile factors that contribute to differences between DDS and ALJ decisions. For example, although ALJs must consider as evidence medical consultants' conclusions about claimants' functional capacity, DDS quality reviews do not focus adequate attention on explaining these conclusions in the record. Moreover, SSA reviews of ALJ awards are too limited to ascertain whether ALJs appropriately consider this evidence or whether DDS explanations could be made more useful to ALJs. Feedback about both of these issues—DDSs' explanations of decisions and ALJs' consideration of them—would help improve SSA's reviews of DDS and ALJ decisions and make DDS decisions more useful to ALJs.

SSA has started to focus its quality reviews on achieving greater consistency between DDS and ALJ decisions. In late 1996, the agency started to increase its reviews of ALJ awards, setting a first-year target of 10,000 cases. In the longer term, SSA plans to unify its DDS and ALJ quality review processes, providing systematic review of decision-making. The agency hopes this will ensure that the correct decision is made at the earliest point in the process.

Recommendations

GAO supports SSA's process unification initiatives and recommends that SSA, using available systems and data collected so far, move quickly ahead to implement its quality assurance initiative to provide consistent feedback to DDS and ALJ adjudicators as soon as possible. In addition, SSA should expand its effort to return cases to DDSs for their review when new evidence is introduced on appeal.

GAO also recommends that SSA set specific goals for measuring the effectiveness of process unification in reducing inconsistent decisions.

Agency Comments

In its written comments on a draft of this report, SSA stated that the goal of process unification was the linchpin of the agency's disability redesign efforts and that GAO's findings and suggestions would help SSA achieve this goal. SSA generally agreed with GAO's conclusions and recommendations and provided specific comments and observations about areas of the report that it believed should be changed. Where appropriate, GAO has revised the report. A number of SSA's specific comments and GAO's evaluation of these comments appear in chapter 6; the full text of SSA's comments and GAO's response appear in appendix III.

Contents

Executive Summary 2

Chapter 1
Introduction 12
Claimants May Pursue Several Levels of Appeal 12
ALJ Procedures Foster Independent Decision-making 16
Differences in Decision Results Are Long-standing 17
Decision-making Process Yields Much Inconsistency Between DDSs and ALJs 19
Recent SSA Efforts to Reduce Inconsistency 20
Objectives, Scope, and Methodology 22

Chapter 2
Disability Decision-making: A Complex Process Requiring Much Judgment 23
DDSs and ALJs Use a Standard Approach, the Sequential Evaluation Process 23

Chapter 3
DDSs and ALJs Differ Most When Assessing Residual Functional Capacity 33
Most ALJ Awards Result From RFC Assessments That Differ From Those of DDSs 33
DDSs and ALJs Differ in Their Decision-making Approaches 35
Effect of Differences in Policy Documents Difficult to Assess 36
SSA Is Taking Actions to Improve Consistency of Decisions 37

Chapter 4
DDS Evaluations of Limited Use to ALJs 39
DDS Medical Consultants Often Inadequately Explain RFC Assessments 39
ALJ Awards Are Often Based on Information Not Available to DDSs 40
Effect of Other Factors Does Not Appear Major or Is More Difficult to Substantiate 42
SSA's Planned Improvements in Procedures 43

Chapter 5 Quality Reviews Do Not Focus on Inconsistency Between DDS and ALJ Decisions 45

Quality Reviews Not Designed to Address Differences 45
Current Quality Reviews Mirror Differences in Approach and Procedures 47
SSA Has Plans to Improve Quality Reviews 49

Chapter 6 Conclusions, Recommendations, Agency Comments, and Our Evaluation 51

Conclusions 51
Recommendations 52
SSA's Comments and Our Evaluation 53

Appendixes

Appendix I: DDS and ALJ Disability Decisions and Operations 56
Appendix II: SSA Studies Addressing Differences Between DDS and ALJ Decision-making 60
Appendix III: Comments From the Social Security Administration and Our Evaluation 68
Appendix IV: GAO Contacts and Staff Acknowledgments 75

Tables

Table 1.1: Levels of Appeal and Actions Taken by Disability Adjudicators 13
Table 1.2: Award Rates by DDSs and ALJs by Impairment Type 19
Table 1.3: Back Impairment Award Rates by DDS and ALJs by Claimant Age 20
Table 2.1: Five-Step Sequential Evaluation Process for Determining Disability 24
Table 2.2: Definition of Five Exertional Demand Categories 27
Table 2.3: Directed Decisions Under Medical-Vocational Rules for Applicants Aged 50 or Older Whose Exertional Ability Is Limited to Sedentary Work 28
Table 2.4: Directed Decisions Under Medical-Vocational Rules for Applicants Under Age 50 Whose Exertional Ability Is Limited to Full Range of Sedentary Work 29
Table 2.5: Factors Adjudicators Consider in Weighing Medical Opinions 31

Table 3.1: DDS and ALJ Differences in RFC Assessment Classifications for Physical Impairment Awards 34
Table 4.1: SSA Reviews Differ by Organization Reviewed and Purpose 45
Table I.1: DDS Decisions—September 1992 Through April 1995 56
Table I.2: ALJ Decisions—September 1992 Through April 1995 57
Table I.3: DDS and ALJ Operations, FY 1986 Through First Quarter, FY 1997 58
Table II.1: Design of SSA's Bellmon Study 61
Table II.2: Three-Tier DHQRP 65

Figures

Figure 1.1: Disability Appeals Process and Outcomes 15
Figure 1.2: ALJ Awards as a Proportion of All Awards, FYs 1986-96 18

Abbreviations

ALJ	administrative law judge
APA	Administrative Procedure Act
CDR	continuing disability review
DDS	disability determination service
DHQRP	Disability Hearings Quality Review Process
DI	Disability Insurance
HALLEX	Hearings, Appeals, and Law Litigation Manual
IG	Office of the Inspector General
OHA	Office of Hearings and Appeals
OPIR	Office of Program and Integrity Reviews
PER	pre-effectuation review
POMS	Program Operations Manual System
RFC	residual functional capacity
SSA	Social Security Administration
SSI	Supplemental Security Income
SSR	Social Security Ruling

Introduction

The Social Security Disability Insurance (DI) and Supplemental Security Income (SSI) programs are the two largest federal programs providing cash payments to people with long-term disabilities. The DI program, authorized in 1956 under title II of the Social Security Act, provides monthly cash insurance benefits to insured, severely disabled workers. The SSI program, authorized in 1972 under title XVI, provides monthly cash payments to aged, blind, or disabled people whose income and resources fall below a certain threshold. About 2.5 million people apply to the Social Security Administration (SSA) each year for disability benefits.

Between 1985 and 1995, the number of DI beneficiaries increased about 53 percent to about 5.0 million, and the number of working-age SSI recipients increased 81 percent to 2.4 million. In 1995, SSA distributed about $61 billion to these and other disability beneficiaries and spent $3 billion on program administration, which accounted for more than half of SSA's total administrative expenses.[5]

Both the DI and SSI programs are administered by SSA and state disability determination services (DDS), which determine benefit eligibility. DDSs award benefits to about 35 percent of applicants.[6] Denied applicants may appeal to an administrative law judge (ALJ) in SSA's Office of Hearings and Appeals (OHA). About a third of all applicants found not disabled by DDSs appeal to an ALJ, and almost two-thirds of claimants who appeal to an ALJ are subsequently found disabled.

Cases appealed to ALJs add considerably to SSA's administrative expense and increase the time claimants must wait for a decision. The average initial DDS decision in DI cases costs about $540, while a hearing can cost an additional $1,200. In addition, appeals can add an average of 378 days to the length of time that an applicant must wait for a final decision. Moreover, because ALJs award a high percentage of appealed cases that have already been denied twice by the DDS, the integrity of the process is called into question.

Claimants May Pursue Several Levels of Appeal

Claimants apply for DI and SSI disability benefits in SSA field offices, which forward these applications, along with any supporting medical evidence, to the appropriate state DDS. A DDS adjudication team, consisting of a disability examiner and a medical or psychological consultant, makes the

[5]Included in the $61 billion of benefits are payments to all SSI blind and disabled beneficiaries regardless of age.

[6]DDSs are funded by SSA and make decisions in accordance with SSA's policies and procedures.

initial decision on each claim. If the DDS denies a claim, the claimant may ask for reconsideration. For the reconsideration review, a new team of DDS adjudicators makes an independent decision on the basis of its own evaluation of all the evidence, including any new evidence the claimant might submit.

If, after reconsideration, a DDS denies benefits, the claimant may pursue several levels of appeal (see table 1.1) and may introduce new evidence at almost every level. First, the claimant has the right to request a hearing before an ALJ. Before the hearing, the ALJ may obtain further medical evidence, for example, from the claimant's own physician or by hiring a consultative physician to examine the claimant. The hearing before the ALJ is the first time that a claimant has an opportunity for a face-to-face meeting with an adjudicator. SSA hearings are informal and nonadversarial; SSA does not challenge a claimant's case.

Table 1.1: Levels of Appeal and Actions Taken by Disability Adjudicators

Adjudicative action	Adjudicative decisionmakers
State DDS	
Make initial decision	Medical consultant and disability examiner team
Reconsider decision to deny benefits	Different medical consultant and disability examiner team
SSA	
Review appealed DDS denial	ALJ
Review ALJ denial	Appeals Council members
Federal courts	
Review final agency decision (by ALJ or Appeals Council) to deny benefits	Federal courts

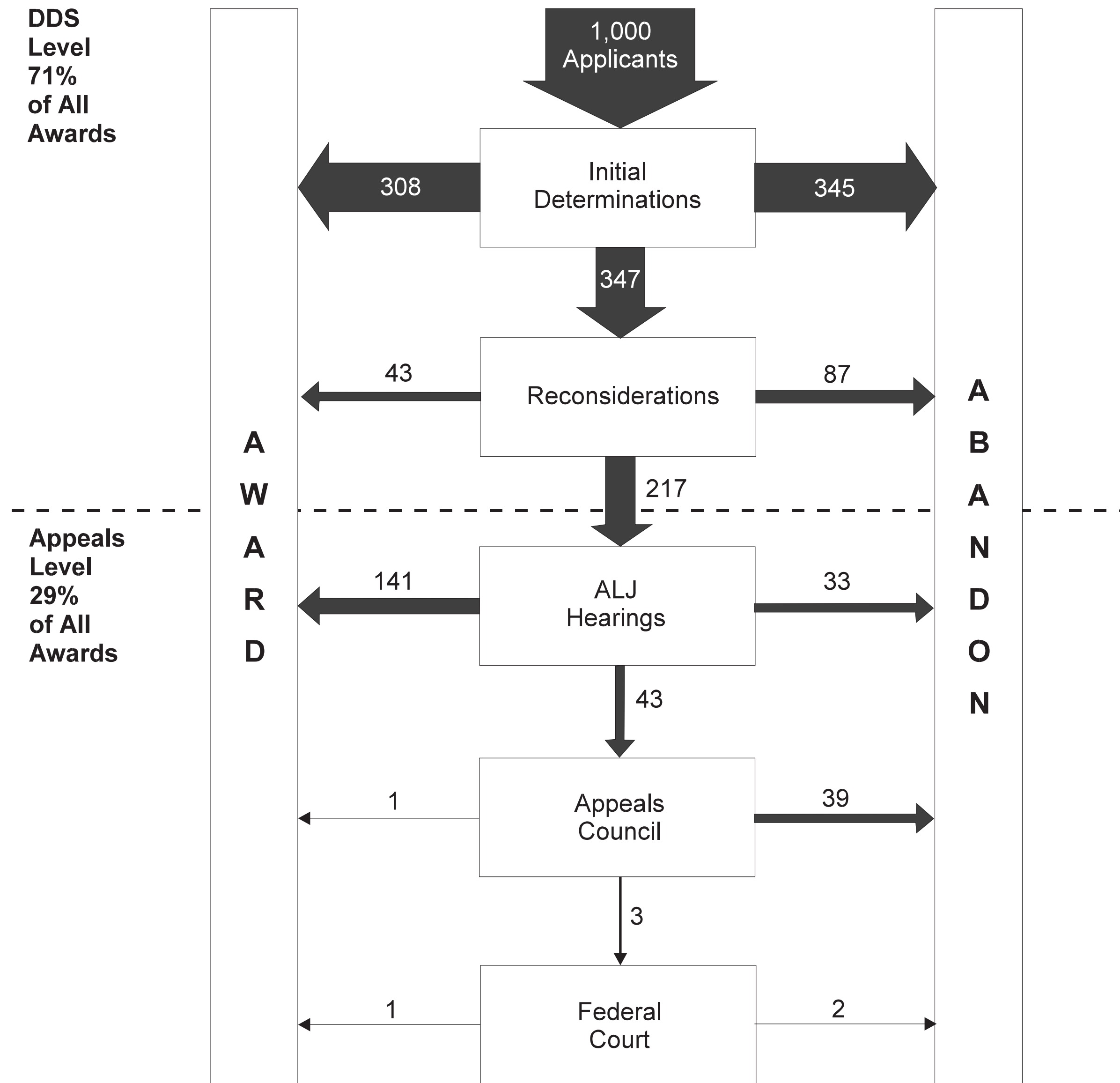
The claimant and witnesses—who may include medical or vocational experts—testify at the hearing. The ALJ asks about the issues, receives relevant documents into evidence, and allows the claimant or the claimant's representative to present arguments and examine witnesses. If necessary, the ALJ may further update the evidence after the hearing. When this is completed, the ALJ assesses the effects of the claimant's medical impairment on capacity to function at work. The ALJ then issues a decision based on his or her assessment of the evidence in the case and is generally authorized to do so without seeking input from a medical professional.

If an ALJ denies an appealed claim, the claimant may request that SSA's Appeals Council review the case. The Appeals Council may deny or

dismiss the request, or it may grant the request and either remand the case to the ALJ for further action or issue a new decision. The Appeals Council's decision, or the decision of the ALJ if the Appeals Council denies or dismisses the request for review, becomes SSA's final decision. After a claimant has exhausted all SSA administrative remedies, the claimant has further appeal rights within the federal court system, up to and including the Supreme Court.

Overall, about 49 percent of all applicants receive benefits, most (71 percent) from initial or reconsideration decisions made at the DDS level. About 22 percent of all applicants appeal their cases to ALJS; about two-thirds of all claimants whose claims are denied at the DDS reconsideration level appeal to an ALJ. Overall, about 29 percent of all claims in 1996 were awarded on appeal. Figure 1.1 shows an overview of the disability decision-making appeals process.

Figure 1.1: Disability Appeals Process and Outcomes



Source: GAO analysis based on 1996 SSA data.

ALJ Procedures Foster Independent Decision-making

ALJs at SSA conduct de novo (or "afresh") hearings; in other words, they may consider or develop new evidence, and they are not bound by DDS decisions. In addition, the Administrative Procedure Act (APA) protects ALJs' independence by exempting them from certain management controls.

Although ALJs are SSA employees and generally subject to the civil service laws, the APA protects these staffs' independence by restricting the extent to which management controls them. For example, ALJ pay is determined by the Office of Personnel Management independently of SSA recommendations or ratings, and ALJs are not subject to statutory performance appraisal requirements. Such safeguards help ensure that ALJ judgments are independent and that ALJs would not be paid, promoted, or discharged arbitrarily or for political reasons by an agency.

ALJs operate under rules that differ from those of appellate courts. After a DDS denial is appealed, an ALJ at SSA holds a de novo hearing, entitling the claimant to have all factual issues determined anew by the ALJ. In contrast, appellate courts generally review the findings of lower courts and only consider whether those courts made errors of law or procedure.

Under the ALJ de novo process, the claimant receives a full in-person hearing from an adjudicator who is fully authorized to hear every aspect of the case.[7] The ALJ hearing is the first time a new claimant is guaranteed the right to testify before an adjudicator.

As SSA employees, ALJs make decisions for the Commissioner and are subject to agency rules and regulations that they must apply in holding hearings and making decisions. Review by the Appeals Council ensures that ALJ decisions follow SSA regulations and rulings. If the Council concludes that the ALJ has not followed agency rules and regulations, the Council can reverse the ALJ decision on its own or send the case back to the ALJ for further action.

Although the ALJ's review and analysis of an appealed denial must include the case file materials developed by the DDS, the ALJ makes new factual determinations. For example, even though a DDS concludes that an individual can perform work, the ALJ is free to conclude that the individual cannot.

[7]Not every appealed case involves a hearing; some are decided on the basis of the case record.

Differences in Decision Results Are Long-standing

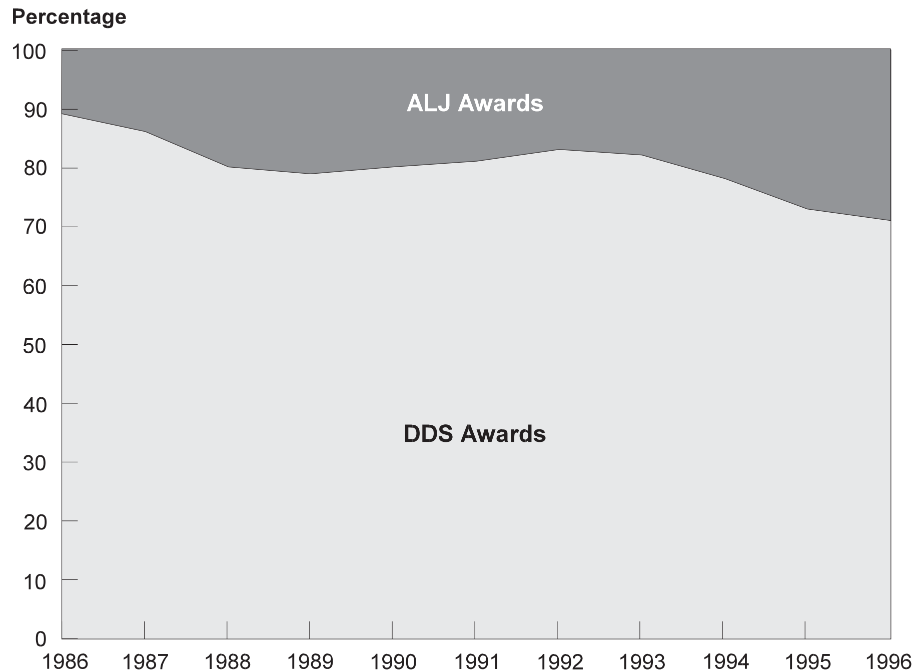
The differences between DDS and ALJ results are a long-standing problem contributing to the growth in OHA backlogs and increased case-processing time, according to our 1996 report on SSA's efforts to reduce backlogs in appealed decisions.[8] Our review of over 40 internal and external studies of the disability determination and appeals process, several of which were completed more than 20 years ago, led us to this conclusion. In the early 1990s, as part of its efforts to develop a number of strategic priority goals, SSA reviewed many of the same studies and identified inconsistent decisions as a critical issue affecting SSA's ability to improve its service to the public.

Inconsistent decisions have been evident in program data for many years. For example, since 1986, DDS award rates have ranged from 31 to 43 percent, whereas ALJ award rates have ranged from 60 to 75 percent. As shown in figure 1.2, ALJ awards, as a percentage of total awards, have ranged from 17 percent in 1986 to 29 percent in 1996.[9]

[8]Social Security Disability: Backlog Reduction Efforts Under Way; Significant Challenges Remain (GAO/HEHS-96-87, July 11, 1996).

[9]More recently, in the first quarter of fiscal year 1997, the proportion dropped to 24 percent.

Figure 1.2: ALJ Awards as a Proportion of All Awards, FYs 1986-96



Source: SSA data.

Concerns about comparatively high ALJ award rates are not new. Although many hypotheses for inconsistent decisions have been discussed, explanations for the high rate of ALJ awards have been inadequate or unavailable. In early 1979, congressional hearings focused on high ALJ award rates, and, in 1980, the Congress passed legislation aimed at promoting greater consistency and accuracy of ALJ decision-making. This legislation required SSA to establish a system of reviewing ALJ decisions to ensure that they comply with laws, regulations, and SSA rulings. In January 1982, SSA submitted to the Congress the results of a study on progress made in reviewing ALJ decisions, including the possible causes for ALJ reversals.[10]

Soon after SSA started to perform the quality reviews required by legislation, the Association of ALJs filed suit in federal court. The lawsuit challenged SSA's plans to target these reviews to judges with high award

[10]Implementation of Section 304 (g) of Public Law 96-265, Social Security Disability Amendments of 1980 (the Bellmon Report), Secretary of Health and Human Services (Washington D.C.: Jan. 1982).

rates on the grounds that such reviews threatened ALJs' decision-making independence. The court never ruled on this issue because SSA decided to rescind targeted reviews. ALJ award rates fell temporarily from 62 percent in 1981 to 55 percent in 1983 when SSA was performing its targeted reviews, although other factors could explain the decline. When targeted reviews ended in 1984, however, ALJ award rates started to increase again and have remained at high levels ever since.

Decision-making Process Yields Much Inconsistency Between DDSs and ALJs

Not only do award rates between DDSs and ALJs differ, but the rates also differ by impairment type and other factors. For example, although DDS award rates vary by impairment, ALJ award rates are high regardless of the type of impairment. As shown in table 1.2, DDS award rates ranged from 11 percent for back impairments to 54 percent for mental retardation. In contrast, ALJ award rates averaged 77 percent for all impairment types with a smaller variation among impairment types.

Table 1.2: Award Rates by DDSs and ALJs by Impairment Type

	DDS award rates (percent)	ALJ award rates (percent)
Physical	29	74
Musculoskeletal	16	75
Back cases	11	75
Other musculoskeletal	23	76
Other physical	36	74
Mental	42	87
Illness	39	87
Retardation	54	84
All impairments	30	77

Source: GAO analysis based on SSA data from Sept. 1, 1992, through Apr. 30, 1995.

When age is considered in addition to impairment type, decisions can vary even more widely. Table 1.3 illustrates, for example, how widely DDSs and ALJs can diverge when age is considered in back impairment cases.

Table 1.3: Back Impairment Award Rates by DDS and ALJs by Claimant Age

Age of claimant	Award rates (percentage) DDS	ALJ
All ages	11	75
Under 50	2	68
50 and older	22	83

Source: GAO analysis based on SSA data for Sept. 1, 1992, through Apr. 30, 1995.

Recent SSA Efforts to Reduce Inconsistency

SSA has long known about its inconsistent decisions and the problems they pose for the disability programs and the agency. SSA has studied the problem and taken several steps to address factors known to contribute to inconsistency between DDS and ALJ adjudicators. In May 1992, SSA's Commissioner approved a study of the appeals process, later called the Disability Hearings Quality Review Process (DHQRP).[11] This study analyzed the reasons for high ALJ award rates. SSA has issued two reports based on this study, which is ongoing.[12]

Realizing that the inconsistency between DDS and ALJ decisions and the length and complexity of the decision-making process compromised the integrity of disability determinations, SSA began redesigning the process in 1993. In late 1994, it released its Plan for a New Disability Claim Process—commonly referred to as the "redesign plan"—which represents the agency's long-term strategy for addressing the systemic problems contributing to inefficiencies in its disability processes. To direct the redesign effort, SSA created a management team assisted by top SSA management, various task teams, and state and federal employees involved with disability determinations.

To address inconsistent decisions as a part of redesign, the agency established a process unification task team. This team included a diverse group of 29 SSA and DDS employees who, in addition to their own expertise, sought information from other sources and reviewed data from SSA's DHQRP study of the appeals process. In November 1995, the task team issued its final report. SSA established an intercomponent group to develop specific actions to support consistent disability decisions and a senior executive

[11] SSA's decision to begin this review of ALJ decisions was prompted in part by our 1992 report about racial disparities in ALJ allowance decisions. See Social Security: Racial Difference in Disability Decisions Warrants Further Investigation (GAO/HRD-92-56, Apr. 21, 1992).

[12] Findings of the Disability Hearings Quality Review Process, SSA, Office of Program and Integrity Reviews (Washington, D.C.: Sept. 1994 and Mar. 1995).

group to enforce needed changes. In July 1996, the SSA Commissioner approved the group's recommendations for several initiatives designed to reduce inconsistent decisions by DDSs and ALJs.

SSA Faces Several Competing Workloads

In addition to SSA's recent efforts to address inconsistent DDS and ALJ decisions, the agency faces significantly increasing workloads at all levels of adjudication. In particular, several congressional mandates will compete for time and resources with process unification efforts. For example, the Social Security Independence and Program Improvements Act of 1994 and the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 require hundreds of thousands more continuing disability reviews (CDR) to ensure that beneficiaries are still eligible for benefits. By law, SSA must conduct CDRs for at least 100,000 more SSI beneficiaries annually through fiscal year 1998. In 1996, the Congress increased CDR requirements for children on SSI, requiring CDRs at least every 3 years for children under age 18 who are likely to improve and for all low birth weight babies in the first year of life. In addition, SSA is required to redetermine, using criteria for adults, the eligibility of all 18-year-olds on SSI beginning on their 18th birthdays and to readjudicate 332,000 childhood disability cases by August 1997. Finally, thousands of noncitizens and drug addicts and alcoholics could appeal their benefit terminations, further increasing SSA's workload.

SSA Includes Performance Goals for Disability in Its Government Performance and Results Act Plan

The Government Performance and Results Act (the Results Act) of 1993 requires federal agencies to be more accountable for the results of their efforts and their stewardship of taxpayer dollars. The Results Act shifts the focus of federal agencies from traditional concerns, such as staffing and activity levels, to results. Specifically, the act directs agencies to consult with the Congress and obtain the views of other stakeholders and to clearly define their missions. It also requires them to establish long-term strategic goals as well as annual goals linked to the strategic goals. Agencies must then measure their performance toward these goals and report to the President and the Congress on their progress.[13]

The Results Act's initial implementation involves about 70 pilot tests during fiscal years 1994 through 1996 to provide agencies with experience in meeting its requirements before governmentwide implementation in the fall of 1997. As a pilot agency, SSA submitted its fiscal year 1996 annual

[13]For further details, see Executive Guide: Effectively Implementing the Government Performance and Results Act (GAO/GGD-96-118, June 1996).

performance plan to the Office of Management and Budget in May 1995. Specifically, the plan includes the strategic goals of (1) rebuilding confidence in Social Security, (2) providing world-class service, and (3) creating a supportive environment for SSA employees. It also includes a broad range of measures for disability and appeals-related performance outputs and outcomes.

Objectives, Scope, and Methodology

In 1995 testimony before the Subcommittee on Social Security, House Committee on Ways and Means, we reported on the timeliness and consistency of DDS and ALJ disability determinations.[14] After our testimony, the Chairman asked us to examine the differences between DDS and ALJ decisions in more detail. Specifically, we agreed to (1) ascertain the factors contributing to inconsistent decisions by DDSs and ALJs and (2) identify SSA's efforts to address inconsistent decisions. We reported our preliminary findings in testimony earlier this year.[15]

To respond to the first objective, we divided the possible contributing factors into three types: (1) factors related to differences in RFC assessments made by DDSs and ALJs, (2) procedural factors that contribute to differences in decisions, and (3) use of quality reviews to manage the process.

In conducting our review, we examined existing studies, SSA's regulations and program operations memoranda, and court cases related to the disability programs. We also obtained and analyzed program and statistical data; see appendix I for details. In addition, we interviewed DDS and SSA officials, including ALJs and OHA staff. We also attended SSA's nationwide process unification training.

We performed our review at SSA headquarters in Baltimore, Maryland; OHA headquarters in Falls Church, Virginia; and at SSA and DDS offices in Atlanta, Boston, and Denver. We conducted our review between October 1995 and June 1997 in accordance with generally accepted government auditing standards except that we did not verify agency data.

[14]Social Security Disability: Management Action and Program Redesign Needed to Address Long-Standing Problems (GAO/T-HEHS-95-233, Aug. 3, 1995).

[15]Social Security Disability: SSA Actions to Reduce Backlogs and Achieve More Consistent Decisions Deserve High Priority (GAO/T-HEHS-97-118, Apr. 24, 1997).

Disability Decision-making: A Complex Process Requiring Much Judgment

SSA requires that DDS and ALJ adjudicators follow a standard approach—called the sequential evaluation process—for making disability determinations. Although standard, the process requires adjudicators to make several complex judgments. For example, if adjudicators cannot allow the claim on the basis of medical evidence only, they must make judgments on whether claimants can perform prior or other work available in the national economy despite their disabling conditions. Such determinations may involve not only residual functional capacity (RFC) assessments, but consideration of these assessments along with the claimant's age, education, and skill levels.

To reduce the amount of judgment involved, SSA has developed medical-vocational rules. In general, the older, less educated, and less skilled the claimant, the more likely these rules will direct the adjudicator to award benefits. For claimants with functional and vocational profiles that do not fit the rules, however, adjudicator decision-making is less prescribed. In addition, before making any decision, adjudicators must decide how much weight to give to various sources of evidence and evaluate the reasonableness and consistency of any allegations the claimant makes about pain or other symptoms.

DDSs and ALJs Use a Standard Approach, the Sequential Evaluation Process

To determine whether applicants meet the Social Security Act's definition of disability, SSA regulations provide DDS and ALJ adjudicators with a sequential evaluation process (see table 2.1). Although the process provides a standard approach, determining disability requires a number of complex judgments.

For people 18 or older, the act defines disability under the DI and SSI programs as the inability to engage in substantial gainful activity by reason of a severe physical or mental impairment that is medically determinable and has lasted or is expected to last at least 1 year or result in death.[16] Moreover, the impairment must be of such severity that a person not only is unable to do past relevant work, but, considering age, education, and work experience, is also unable to engage in any substantial work available in the national economy.

[16]Regulations currently define substantial gainful activity as employment that produces countable earnings of more than $500 a month for disabled people and $1,000 a month for blind people.

Table 2.1: Five-Step Sequential Evaluation Process for Determining Disability

Step	Questions asked in the sequential process	Action or decision taken if answer to question is	
		Yes	No
1	Is the applicant engaging in substantial gainful activity?[a]	Stop—applicant is not disabled	Go to step 2
2	Does the applicant have an impairment that has more than a minimal effect on the applicant's ability to perform basic work tasks?[b]	Go to step 3	Stop—applicant is not disabled
3	Does the applicant's impairment meet or equal the medical criteria for an impairment in SSA's Listing of Impairments?[b]	Stop—applicant is disabled	Go to step 4
4	Comparing the applicant's RFC with the physical and mental demands of the applicant's past work, can the applicant perform his or her past work?	Stop—applicant is not disabled	Go to step 5
5	On the basis of the applicant's RFC and any limitations that may be imposed by the applicant's age, education, and skill level, can the applicant do work other than his or her past work?[b]	Applicant is not disabled	Applicant is disabled

[a]Under the sequential evaluation process, SSA's field offices determine whether the applicant is engaged in substantial gainful activity.

[b]In addition, the criteria require that the impairment last 12 months or be expected to result in death.

Applicants are denied benefits at step 1 if they are engaged in substantial gainful activity. At step 2, adjudicators further screen applicants by assessing whether they have a severe impairment, defined by the regulations as an impairment that has more than a minimal effect on the applicant's ability to perform basic work tasks. For those whose impairments have more than a minimal effect on ability to work, adjudicators then begin determining whether the applicant's impairments are severe enough to qualify for disability benefits.

Does the Applicant Qualify Under SSA's Listing of Impairments?

In step 3 of the sequential evaluation process, adjudicators compare the applicant's medical condition with medical criteria found in SSA's Listing of Impairments—referred to as "the medical listings"—which are published in SSA's regulations. The listings delineate over 150 categories of medical conditions (physical and mental) that, according to SSA, are presumed to

be severe enough to ordinarily prevent an individual from engaging in any gainful activity. For example, corrected vision of 20/200 or less, amputation of both hands, or an intelligence quotient of 59 or less would ordinarily qualify an individual for benefits.

An applicant may automatically qualify for benefits if the adjudicator concludes that the laboratory findings, medical signs, and symptoms of one of the applicant's impairments meet the specific criteria for medical severity cited in the listings for that impairment and the applicant is not engaging in substantial gainful activity. If an applicant's medical condition does not meet the listed criteria or if the impairment is not listed, then the adjudicator must determine whether the applicant's impairment is the medical equivalent of one in the listings.

The medical severity criteria for listed mental impairments are generally more subjective than those for physical impairments. For most mental impairments in the listings, many of the severity criteria are defined by functional limitations. Determining whether a mental impairment meets or equals the listed criteria often requires subjective evaluations about (1) restrictions of daily activities; (2) difficulties in maintaining social functioning; (3) deficiencies in concentration, persistence, or pace that result in failure to complete tasks in a timely manner; and (4) episodes of deterioration in work settings that cause the individual to withdraw or have exacerbated signs and symptoms. For example, adjudicators must decide whether the impairment has any impact at all on activities of daily living or on social functioning, and, if so, rate the impact as slight, moderate, marked, or extreme.

By contrast, the listed criteria for physical impairments generally are more objective, relating to medical diagnosis and prognosis, rather than the assessment of functional limitations in the mental listings. Determining whether the medical findings for a physical impairment meet or equal these criteria is a matter of documentation and is often more a question of medical fact than opinion. In some instances, however, the criteria for physical impairments also require that adjudicators assess functional limitations. For example, for applicants with human immunodeficiency virus, adjudicators assess their symptoms or signs, such as fatigue, fever, malaise, weight loss, pain, and night sweats as well as their subsequent effect on activities of daily living and social functioning. For musculoskeletal and other impairments, adjudicators assess the importance of pain in causing functional loss when it is associated with relevant abnormal signs and laboratory findings. Adjudicators must also

carefully determine that the reported examination findings are consistent with the applicant's daily activities.

Can the Applicant Perform Past Relevant Work?

When medical evidence does not show that an applicant's condition meets or equals the severity criteria in the listings, adjudicators must determine whether the applicant can perform past work. To do this, adjudicators use judgment when they assess an applicant's RFC—that is, what an applicant can still do, despite physical and mental limitations, in a regular full-time work setting.

To assess RFC, adjudicators must consider all relevant medical and nonmedical evidence, such as statements of lay witnesses about an individual's symptoms. In considering medical evidence, adjudicators must evaluate medical source opinions and judge the weight to be given to each opinion. Adjudicators also often evaluate issues involving pain or other symptoms and judge whether the applicant's impairment could reasonably be expected to produce the applicant's symptoms.

Assessing physical RFC requires adjudicators to judge individuals' ability to physically exert themselves in activities such as sitting, standing, walking, lifting, carrying, pushing, and pulling. Adjudicators also assess the effect of the individual's physical impairment on manipulative or postural functions such as reaching, handling, stooping, or crouching. Assessing mental RFC requires adjudicators to judge the individual's functional abilities such as understanding, remembering, carrying out instructions, and responding appropriately to supervision, coworkers, and work pressures.

After assessing an applicant's RFC, the adjudicator compares it with the demands of the applicant's prior work. The adjudicator either concludes that the applicant can perform his or her prior work and denies the claim or proceeds to the last step (step 5) in the sequential evaluation process.

Can the Applicant Perform Other Work in the National Economy?

At step 5, adjudicators evaluate whether applicants unable to perform their previous work can do other jobs that exist in significant numbers in the national economy. If the adjudicator concludes that an applicant can perform other work, the claim is denied. Again, adjudicators must apply judgment to determine whether an applicant can perform other work in the national economy, depending on whether the applicant's limitations are exertional or nonexertional.

An applicant has exertional limitations when his or her impairment limits the ability to perform the physical strength demands of work. For this evaluation, SSA places a claimant into one of five categories of physical exertion—sedentary, light, medium, heavy, and very heavy—with sedentary work requiring the least physical exertion of the five levels (see table 2.2).[17] On the basis of an applicant's RFC, adjudicators must judge which of the five exertional categories is the most physically demanding work the individual can perform. For an applicant whose maximum physical ability matches one of the five exertional categories of work, SSA provides medical-vocational rules that direct the adjudicator's decision on the basis of the claimant's age, education, and skill levels of prior work experience.

Table 2.2: Definition of Five Exertional (Strength) Demand Categories

Exertional demand category	Strength requirements	
	Requirement for lifting	**Other strength demands**
Sedentary	Requires lifting no more than 10 pounds at a time and occasionally lifting or carrying articles like docket files, ledgers, and small tools	Involves sitting; walking and standing may be required occasionally
Light	Requires lifting no more than 20 pounds at a time or carrying objects weighing up to 10 pounds	Requires a good deal of walking or standing or involves sitting most of the time with some pushing and pulling of arm or leg controls
Medium	Requires lifting no more than 50 pounds at a time with frequent lifting or carrying of objects weighing up to 25 pounds	Requires unlimited sitting, walking, and standing ability
Heavy	Requires lifting no more than 100 pounds at a time with frequent lifting or carrying of objects weighing up to 50 pounds	Requires unlimited sitting, walking, and standing ability
Very heavy	Requires lifting objects weighing more than 100 pounds at a time with frequent lifting or carrying of objects weighing 50 pounds or more	Requires unlimited sitting, walking, and standing ability

Table 2.3 shows how the medical-vocational rules direct decisions for people aged 50 or older who are limited to sedentary work.

[17]The Department of Labor developed this classification system, which is in its Dictionary of Occupational Titles.

Table 2.3: Directed Decisions Under Medical-Vocational Rules for Applicants Aged 50 or Older Whose Exertional Ability Is Limited to Sedentary Work

Education	Previous work experience	Directed decision
Limited (grades 7-11 or less)	Unskilled or none	Disabled
Limited (grades 7-11 or less)	Skilled or semiskilled—skills not transferable	Disabled
Limited (grades 7-11) or less	Skilled or semiskilled—skills transferable	Not disabled
High school graduate or more—does not provide for direct entry into skilled work	Unskilled or none	Disabled
High school graduate or more—provides for direct entry into skilled work	Unskilled or none	Not disabled
High school graduate or more—does not provide for direct entry into skilled work	Skilled or semiskilled—skills not transferable	Disabled
High school graduate or more—does not provide for direct entry into skilled work	Skilled or semiskilled—skills transferable	Not disabled
High school graduate or more—provides for direct entry into skilled work	Skilled or semiskilled—skills not transferable	Not disabled

In general, the older a person is, the more likely SSA's medical-vocational rules direct adjudicators to award benefits. For example, under the rules for those whose maximum physical capacity limits them to performing sedentary work, applicants aged 50 or older qualify for benefits under four of the scenarios shown in table 2.3. Those aged 45 through 49, however, qualify under only one scenario; applicants aged 18 through 44 qualify under no scenario (see table 2.4).

Table 2.4: Directed Decisions Under Medical-Vocational Rules for Applicants Under Age 50 Whose Exertional Ability Is Limited to Full Range of Sedentary Work

Education	Previous work experience	Directed decision	
		Aged 45-49	Aged 18-44
Illiterate or unable to communicate in English	Unskilled or none	Disabled	Not disabled
Limited (grades 7-11 or less)—at least literate and able to communicate in English	Unskilled or none	Not disabled	Not disabled
Limited (grades 7-11 or less)	Skilled or semiskilled—skills not transferable	Not disabled	Not disabled
Limited (grades 7-11 or less)	Skilled or semiskilled—skills transferable	Not disabled	Not disabled
High school graduate or more	Unskilled or none	Does not apply	Not disabled
High school graduate or more	Skilled or semiskilled—skills not transferable	Not disabled	Not disabled
High school graduate or more	Skilled or semiskilled—skills transferable	Not disabled	Not disabled

Although SSA's medical-vocational rules reduce the degree of judgment that adjudicators must use in many cases, SSA has no rules to direct adjudicators' decisions for other cases. These include cases in which (1) the applicant's maximum strength capability does not match any of the five exertional levels or (2) the applicant's primary limitations are nonexertional (or unrelated to the physical strength demands required for sitting, standing, walking, lifting, carrying, pushing, and pulling). In such cases, the medical-vocational rules can provide a guide for evaluating an applicant's ability to do other work, but the regulations instruct adjudicators to base their decisions on the principles in the appropriate sections of the regulations, giving consideration to the medical-vocational rules for specific case situations. For example, an applicant may be restricted to unskilled sedentary jobs because of a severe cardiovascular impairment. If a permanent injury of the right hand also limits the applicant to only those sedentary jobs that do not require bilateral manual dexterity, then the applicant's work capacity is limited to less than the full range of sedentary work. The ability to do less than the full range of sedentary work is not one of the five exertional levels defined in SSA's regulations; therefore, no medical-vocational rules would direct the adjudicator's decision.

On the basis of Department of Labor data, SSA estimates that approximately 200 unskilled occupations exist, each representing many jobs that can be performed by people whose limitations restrict them to the full range of sedentary work. But, if an applicant is limited to less than the full range of sedentary work, the adjudicator must determine the extent to which the exertional and nonexertional limitations reduce the occupational base of jobs, considering the applicant's age, education, and work experience, including any transferable skills or education providing for direct entry into skilled work. The mere inability to perform all sedentary unskilled jobs is not sufficient basis for a finding of disability. The applicant still may be able to do a wide range of unskilled sedentary work.

Adjudicators Must Weigh Evidence and Reasonableness of Symptom Allegations

Before making any decision, an adjudicator must assess the amount of weight to give to the various sources of evidence and evaluate the reasonableness and consistency of any allegations from applicants about pain or other symptoms.

To provide a basis for determining disability, the adjudicator must gather existing medical evidence, which includes (1) opinions of physicians or psychologists who have had an ongoing treatment relationship with the applicant and (2) hospitals, clinics, and other medical sources that have treated or evaluated the applicant but not on an ongoing basis. In addition, adjudicators may develop new medical evidence obtained from consulting sources. Medical evidence includes (1) medical history; (2) clinical findings, such as the results of physical or mental status examinations; (3) laboratory findings, such as blood pressure and X rays; (4) statement of the diagnosis of the disease or injury based on its signs and symptoms; and (5) treatment prescribed and prognosis. Medical evidence also includes statements from treating physicians or other medical sources describing work-related activities, such as sitting, standing, walking, and lifting, that the applicant can still do despite his or her impairments. In the case of mental impairments, statements should describe the applicant's ability to understand, carry out, and remember instructions and respond appropriately to supervision, coworkers, and work pressures. In making a decision, an adjudicator must assess how much weight to give to each medical source's statement of opinion. Table 2.5 describes the factors to be considered in weighing opinions.

Table 2.5: Factors Adjudicators Consider in Weighing Medical Opinions

Factor	General effect of each factor on the weight given to a medical source opinion
Examining relationship	More weight is given to a source who examined the applicant than to a source who did not.
Treatment relationship	More weight is given to the applicant's treating physicians because they can provide a detailed, longitudinal picture of the impairments and bring a unique perspective not available from objective medical findings alone or from single examinations or brief hospitalizations. If a treating physician's medical opinion is well supported and is not inconsistent with other medical evidence in the case file, then adjudicators must give it "controlling" weight.
Length, nature, and extent of treatment relationship	Weight is commensurate with (1) the length of time a source has treated the applicant, (2) the number of times the source has seen the applicant, and (3) the source's knowledge based on the kinds and extent of examinations and testing the source has performed or ordered from specialists and independent laboratories.
Supportability	Weight is commensurate with the extent to which the medical source (1) supports the opinion with relevant evidence, such as medical signs and laboratory findings, and (2) provides an explanation for the opinion.
Consistency	The more consistent an opinion is with the record as a whole, the more weight adjudicators must give that opinion.
Specialization	All other factors being equal, more weight is given to the opinion of a specialist on medical issues in his or her area of specialty than to the opinion of a source who is not a specialist.
Other factors	Adjudicators must consider any factors that the applicant or others bring to their attention that tend to support or contradict the opinion.

Adjudicators also must evaluate whether an applicant's impairment could reasonably be expected to produce the reported symptoms—such as pain, fatigue, shortness of breath, weakness, and nervousness. This requires the adjudicator to assess the extent to which an individual's symptoms are consistent with (1) the objective medical evidence (medical signs and laboratory findings); (2) evidence, such as statements from the applicant, medical sources, family, friends, or employers about the applicant's medical history, diagnosis, prescribed treatment, activities of daily living, and efforts to work; (3) information from social welfare agencies, nonmedical sources, and other practitioners, such as chiropractors and audiologists; and (4) any other evidence of the applicant's impairment's effect on his or her ability to work.

If the adjudicator concludes that the impairment could reasonably be expected to produce the reported symptoms, the adjudicator must then evaluate the intensity and persistence of the symptoms to determine how the symptoms limit the applicant's ability to work. In making such an evaluation, adjudicators look for objective medical evidence obtained through clinical and laboratory diagnostic techniques, such as evidence of reduced joint motion, muscle spasm, sensory deficit, or motor disruption. However, adjudicators cannot reject an applicant's statements about the intensity and persistence of pain or other symptoms or about the effect of these symptoms on the ability to work solely because the available objective medical evidence does not substantiate the applicant's statements. Because symptoms reported by the applicant sometimes suggest a more severe impairment than can be shown by objective medical evidence alone, adjudicators must carefully consider any other information provided by the applicant, treating sources, or other people about the applicant's pain or other symptoms. Following are the factors that adjudicators must consider in assessing pain and other symptoms:

- activities of daily living;
- location, direction, frequency, and intensity of the pain or other symptoms;
- precipitating and aggravating factors;
- type, dosage, effectiveness, and side effects of any medication the applicant takes or has taken to alleviate pain/symptoms;
- treatment, other than medication, the applicant is receiving or has received for relief of pain or other symptoms;
- any measures the applicant uses or has used to relieve pain or other symptoms, such as lying flat on back, standing for 15 or 20 minutes every hour, and sleeping on a board; and
- other factors concerning the applicant's functional limitations and restrictions due to pain or other symptoms.

DDSs and ALJs Differ Most When Assessing Residual Functional Capacity

SSA studies show that DDS and ALJ decisions most often differ because adjudicators make different conclusions about applicants' ability to function in the workplace. At the DDS and ALJ levels, two different types of professional staff perform residual functional capacity (RFC) assessments. At the DDS, medical staff perform the assessments; at the ALJ level, the ALJ performs them. ALJs may seek the advice of medical experts, but they do so infrequently. Study results also suggest that DDSs and ALJs differ in their assessments of the opinions of applicants' own physicians.

SSA has conducted studies of the differences between DDS and ALJ decisions and has identified key issues. To improve consistency of decisions, the agency has recently published policy clarifications, conducted training for all disability adjudicators, and is now starting to evaluate the impact of this training. SSA also plans to develop a single presentation of policy to be used by both DDSs and ALJs.

Most ALJ Awards Result From RFC Assessments That Differ From Those of DDSs

Differing DDS and ALJ assessments of a claimant's capacity to function in the workplace are the primary reason for most ALJ awards. Under the sequential evaluation process, almost all DDS denial decisions appealed to ALJs include an RFC assessment. On appeal, ALJs also follow the same sequential evaluation process and assess the claimant's functional ability in most awards they make. Both the ongoing Disability Hearings Quality Review Process (DHQRP) study and a study conducted by SSA in 1982 note the importance of differences in assessing RFC.[18] (See app. II for more details on these studies' results.)

Decisions in cases involving physical impairments clearly reflected differences in assessing RFC. Table 3.1 presents data from SSA's DHQRP study on physical impairment cases in which ALJs made awards on the basis of RFC assessments. The table compares the ALJ decisions with those of reviewers who used the DDS approach and examined the written evidence available to the ALJ. These data indicate that ALJs are significantly more likely than DDS medical consultants to find that applicants have very limited work capacity.

[18]Findings of the Disability Hearings Quality Review Process, SSA, Office of Program and Integrity Reviews (Washington, D.C.: Sept. 1994) and Implementation of Section 304(g) of Public Law 96-265, Social Security Disability Amendments of 1980 (the Bellmon Report), Secretary of Health and Human Services (Washington, D.C.: Jan. 1982).

Table 3.1: DDS and ALJ Differences in RFC Assessment Classifications for Physical Impairment Awards

Level of physical exertion determined by an adjudicator or reviewer	Quality reviewers using DDS approach (percentage of awards)	Original awarding ALJs (percentage of awards)
Heavy work (or no limiting effect on physical effort)	0	0
Medium work	22	1
Light work	56	8
Sedentary work	15	25
Less than the full range of sedentary work	6	66

Source: GAO analysis based on SSA data for ALJ awards made from Sept. 1992 through Apr. 1995.

In the view of awarding ALJs, 66 percent of the cases merited a "less than the full range of sedentary work" assessment—a classification that often leads to an award. In contrast, the medical consultants who performed the RFC assessment using the DDS approach found that less than 6 percent of cases merited this classification. The DDS and ALJ adjudicators also differed in the other classifications.

In addition, high ALJ award rates for claimants with mental impairments often reflect different assessments of functional limitations. Even ALJ mental impairment awards based on the listings reflect these differences because most such listings require adjudicators to assess functional limitations in addition to determining the claimant's medical condition.

A study known as the Bellmon Report, which controlled for differences in evidence, also found that differing RFCs played a role in differing DDS and ALJ decisions. This study found that DDS and ALJ adjudicators reached different results even when presented with the same evidence. As part of the study, two groups of reviewers looked at selected cases. One group reviewed the cases as ALJs would, and the other reviewed the cases as DDSs would. Reviewers using the ALJ approach concluded that 48 percent of the cases should have received awards; reviewers using the DDS approach concluded that only 13 percent of those same cases should have received awards.

DDSs and ALJs Differ in Their Decision-making Approaches

We identified specific differences in DDSs' and ALJs' approach to their decisions. First, medical staff have different roles at the two levels. In addition, DDSs and ALJs respond differently to (1) the opinions of claimants' physicians and (2) claimants' statements about symptoms such as pain.

DDSs and ALJs Use Medical Expertise Differently

Medical experts play different roles in the DDS and ALJ decision-making approaches. At the DDS, medical or psychological consultants assess RFC of applicants. In contrast, ALJs may consult with medical experts but have sole authority to make the RFC finding. ALJs sought the advice of medical experts in only 8 percent of cases resulting in awards, according to our analysis.

Both the Bellmon and DHQRP studies compared RFC assessments made by SSA medical staff using the DDS approach with those made by awarding ALJs. According to both studies, medical staff tended to find that claimants had higher capacities to function in the workplace than the ALJs found.

DDSs and ALJs Seem to Differ in Their Reliance on Treating Physicians' Opinions

Under SSA regulations, adjudicators must consider the opinions of treating physicians who have an ongoing treatment relationship with the claimant. Such an opinion might include, for example, a statement that a claimant "cannot stand or walk for more than two hours total in a day." In the disability determination, adjudicators must give controlling weight to these treating source opinions provided they are (1) well supported by medically acceptable clinical and laboratory diagnostic techniques and (2) consistent with the other substantial evidence in the record. A treating physician's statement, however, that a claimant is "disabled" or "unable to work" does not bind adjudicators.

Treating physicians' opinions, however, seem to influence DDSs and ALJs differently. The DHQRP study found that the treating physician's report was one of the five most frequent reasons for ALJ awards. This implies that ALJs tended to give controlling weight to the treating physician's opinion, while DDS adjudicators were more likely to focus on assessing that opinion in conjunction with other medical evidence in the case file.

Reports of Symptoms and Claimants' Credibility Also Affect Differences, but Extent Is Unknown

A second factor contributing to differing DDS and ALJ decisions is the impact of symptoms (for example, pain, fatigue, or shortness of breath) reported by the claimant but not identifiable in laboratory tests or confirmable by medical observation. Like the opinions of the claimant's

own physician, assessment of symptoms is important in the disability decision. Adjudicators must assess symptoms by determining (1) whether the medically determinable impairments could reasonably be expected to produce such symptoms and (2) the intensity, persistence, and functionally limiting effects of the symptoms. According to SSA, adjudicators must assess the claimant's credibility on the basis of the entire case record to make a determination about these symptoms' effects. DDSs generally make such assessments on the basis of the case file (for example, statements made by applicants on the application or reports from medical sources that record applicants' comments). ALJs have additional evidence because they have the opportunity to consider the claimant's testimony in a hearing. Moreover, claimant credibility has a significant impact on ALJ decisions.

The DHQRP study identified the credibility of the claimant and claimants' allegations about pain as two of the top five reasons for an ALJ allowance decision. The impact of these reasons on DDS decisions is more difficult to assess. However, during the DHQRP study, reviewers using the DDS approach listened to tapes of claimant testimony in a small sample of 50 cases. The study concluded that claimant testimony had no or minimal impact on those adjudicators.

Effect of Differences in Policy Documents Difficult to Assess

SSA adjudicators use two different sets of documents as criteria for disability decisions, which some believe contributes to inconsistent decisions. DDS adjudicators must follow a detailed set of policy guidelines, called the Program Operations Manual System (POMS). The POMS for disability contains detailed interpretations of laws, regulations, and rulings as well as procedural instructions on deciding cases. ALJs, on the other hand, rely directly on the laws, regulations, and Social Security Rulings (SSR) for guidance in making disability decisions. The latter documents are generally shorter and much less prescriptive than the POMS.

This difference in policy documents, along with the difference in decisions between the DDSs and ALJs has led to the belief by some that there are two standards—or at least two different interpretations of policy. A 1994 Inspector General survey of DDS and ALJ opinion found that the DDSs' strict application of POMS—as opposed to the ALJs' direct application of disability law and regulations—was considered to have a strong effect on allowance rates by over half of those surveyed. Similarly, the Bellmon Report stated that, "SSA has long recognized that the standards and procedures governing decisions by DDSs and ALJs are not entirely consistent."

The type and extent of these differences have proven difficult to quantify, however. For example, the Bellmon Report identified significant differences in DDS and ALJ decisions based on impairments considered not severe. The study then identified differences in the regulations and POMS on this issue. The study concluded, however, that the two written standards, "while different, (were) not widely divergent." As such, it remains unclear whether the differences derive from the standards or from their differing application. Nevertheless, although their relative impact has not been quantified, policy differences cannot be discounted as a potential reason for inconsistent decisions.

SSA Is Taking Actions to Improve Consistency of Decisions

SSA has taken or planned several initiatives to make disability decisions more consistent. In July 1996, SSA issued nine SSRs to address several of the factors we identified as contributing to inconsistent decisions. For example, one of the new rulings reminds ALJs that they must obtain expert medical opinion in certain types of cases.[19] Another ruling clarifies when adjudicators must give the opinion of a treating physician special consideration. A third ruling states that an RFC of less than the full range of sedentary work is expected to be relatively rare. SSA also plans to issue a regulation to provide additional guidance on assessing RFC for both DDSs and ALJs, specifically clarifying when a less-than-sedentary classification is appropriate.[20]

In addition, partly on the basis of the nine rulings, SSA completed nationwide process unification training between July 10, 1996, and February 26, 1997. SSA officials pointed out that this training was the first time that the agency had brought together DDS and ALJ staff to share their views. The training represented a major effort—15,000 adjudicators and quality reviewers received 2 full days of training, coordinated by facilitators in SSA headquarters using a broadcast system. SSA has also started to evaluate the impact of the new rulings and training by collecting data before and after the new rulings and training.

Furthermore, SSA recently compared the policy language in the POMS with disability law, regulations, and SSRs and concluded that no substantive differences in policy existed. SSA did find some differences in wording and

[19]The ruling reinstates a previous SSA policy that an ALJ or Appeals Council member must obtain expert medical opinion before determining that an impairment or group of impairments that do not meet a specific listing are equivalent to the level of severity implied by the listings.

[20]In April 1997, SSA told us that the notice of proposed rulemaking on the less-than-sedentary regulation is ready for release but did not provide a date when it would be issued.

detail, however, that could lead to a perception of differences. To address this matter, SSA plans to develop a single policy presentation to be used by both DDSs and ALJs. To this end, the agency is using exactly the same words in any new regulation, ruling, and POMS publication. It has already done this, for example, for the SSRs on which the process unification training was based. SSA eventually plans to have all adjudication policy in the form of regulations or SSRs so that they are binding on ALJs as well as DDS adjudicators.

In the longer term, SSA also plans under redesign to develop new, more valid, and reliable functional assessment/evaluation instruments relevant to today's work environment. Because current differences in RFC assessments are the main reason for inconsistent decisions, however, SSA should proceed cautiously and test any new decision-making methods to determine their effect on consistency as well as on award rates before widespread implementation.

DDS Evaluations of Limited Use to ALJs

ALJs often cannot fully understand how DDS denial decisions have been made because DDS written evaluations provide neither clear explanations nor justifications for the findings and conclusions reached. Therefore, the evaluations often do not lay a solid foundation for subsequent appeals. For instance, the basis of the DDS' residual functional capacity (RFC) assessment is often unclear, leaving the ALJ without full understanding of the reasoning that led to the DDS denial. Furthermore, explanations of how the DDS considered evidence that ALJs might later rely on, such as the opinions of the claimants' own physicians, may often be missing from the case file or are not fully developed. As a result, ALJs often cannot rely on the evaluations as developed by the DDSS.

SSA has plans to change the process to improve the documentation of DDS evaluations so they can better serve as a foundation for ALJ decisions. These plans include requiring clear DDS explanations of the reasoning used to support reconsideration denials and improving development of evidence at the DDS. SSA also plans to return a selected number of cases involving new evidence from the ALJ level to DDSS for their reconsideration. Together, these changes in procedures will better serve as a foundation for appeals, improving the consistency of DDS and ALJ decisions.

DDS Medical Consultants Often Inadequately Explain RFC Assessments

As discussed in chapter 3, inconsistent decisions between DDSS and ALJS are due mainly to differences in RFC assessments. Studies show that DDS medical consultants often inadequately explain their conclusions, including those about an applicant's RFC. Such explanations, if improved, could be more useful in ALJ decision-making. In fact, SSA's policy is that an ALJ, when making an RFC assessment, must consider the opinion of the DDS medical consultant.

To this end, SSA requires DDS medical consultants to record explanations of their reasoning. In particular, the agency asks medical consultants to fully describe how they used the medical evidence to draw their conclusions about an applicant's RFC. RFC forms and procedures require that medical consultants discuss in writing how the medical evidence in the case file supports or refutes an applicant's allegations of pain or other symptoms. Finally, the RFC forms also require medical consultants to explain how conflicts among treating physician opinion and other medical evidence in the case file were resolved.

Disability Hearings Quality Review Process (DHQRP) data, however, indicate that existing SSA procedures do not ensure that DDS decisions are

well documented. Specifically, procedures require the disability examiner to prepare supplementary explanations when the resolution of key issues is not well documented elsewhere in the case file. The DHQRP study of appealed reconsideration denials found that in about half the cases that hinged on complex issues—such as conflicts with the treating physician's opinion, assessment of RFC, and weighing of allegations regarding pain or other symptoms—DDS documentation failed to explain how these issues were resolved. The insufficient documentation of the underlying medical analyses limited their usefulness during the appeal process.

ALJ Awards Are Often Based on Information Not Available to DDSs

Although ALJs use the medical evidence assembled by DDSs, they often base their decisions on additional documentary or testimonial evidence. This both contributes to inconsistent decisions and makes it difficult to reconcile those differences. Procedures at the hearings level, such as longer time frames for evidentiary development and permitting the introduction of new information, result in the availability of new documentary evidence for appeal cases. In addition, testimony during the face-to-face hearing and the opportunity it provides for further assessing the claimant's credibility provide new information not in DDS case files.

Additional Medical Evidence Results in ALJ Awards

SSA studies show that in many instances introducing additional documentary evidence at the hearing level results in an ALJ's awarding benefits. DHQRP data show that about three-quarters of the appealed cases sampled contained new evidence. The study estimated that 27 percent of the hearing awards hinged on additional evidence, resulting in an assessment of a more severe impairment or a more restrictive RFC. In addition, the Bellmon Report found that when new evidence was removed from the case file, the ALJ award rate decreased from 46 to 31 percent. This study also found that approximately three-quarters of new documentary evidence was medical in nature rather than, for example, statements of friends and associates.

One reason that appeals cases have additional evidence is that ALJ procedures allow for more time to be spent on evidence development. Although SSA regulations stipulate that "every reasonable effort" be made to obtain necessary evidence, DDS guidelines state that evidence should generally be gathered within 30 calendar days. ALJ guidelines, however, provide a time frame for evidence gathering that is almost twice as long and can be extended if necessary.

In addition, ALJs responding to an Inspector General (IG) survey believed that DDSs often fail to adequately develop evidence to show the true nature and extent of an applicant's disability. The ALJs attributed some of this to a lack of adequate resources at the DDSs and pressures to dispose of cases.[21] Also, surveyed ALJs said that DDS problems with developing evidence, particularly medical evidence, contribute to their reversals of DDS denials. In an earlier survey we conducted of DDS administrators, almost two-thirds responded that workload and staffing pressures had affected the accuracy of denial decisions.[22] Seven DDS administrators (14 percent) said the harmful effect on the accuracy of denial decisions was great or very great.

Finally, the presence of attorneys or others who represent the claimant's interests may also result in the presentation of new evidence during an appeal. Because attorneys are generally paid only when decisions favor their clients, they are motivated to find and present additional evidence. Although few claimants hire attorneys or other representatives at the DDS level, DHQRP data showed that representatives attended 81 percent of ALJ hearings.

Claimant Testimony Appears to Result in ALJ Awards

With few exceptions, ALJ hearings present a claimant's first opportunity for face-to-face contact with a disability adjudicator. Studies show that face-to-face encounters with claimants appear to account for a significant number of ALJ reversals. Specifically, in the DHQRP study, reviewing ALJs believed that a favorable assessment of the claimant's credibility is a factor in 34 percent of sampled hearing allowances. Although DDSs and ALJs also assess credibility from case file information, testimony received at a hearing appears to especially influence ALJs when assessing the credibility of a claimant's subjective allegations such as the effect of pain on functioning.

The IG's 1994 report showed that nearly 60 percent of ALJs surveyed believed that the claimant's appearance before an ALJ strongly affects awards; 90 percent believed it has a moderate to strong effect. Furthermore, the Bellmon Report found that the ALJ award rate decreased by about 17 percentage points when evidence from the claimant's record of testimony was removed from the case file.

[21]The Disability Appeals Process: Administrative Law Judge Perspectives, Department of Health and Human Services Office of the IG (Washington, D.C.: May 1994).

[22]Social Security: Increasing Number of Disability Claims and Deteriorating Service (GAO/HRD-94-11, Nov. 10, 1993).

New Impairment Claims Also Result in ALJ Awards

Because claimants may offer new documentary and testimonial evidence at an ALJ hearing, they can also change their impairment type or add a new, secondary impairment, which also affects consistency of DDS and ALJ decisions. Moreover, in about 10 percent of cases appealed to the ALJ level, claimants switch the basis of their primary impairment from a physical claim to a mental claim. Under current procedures, the DDS lacks the opportunity to routinely consider these switched claims, then incorporate this consideration in their analysis, thus providing the ALJ with a basis for confirming or rejecting the new impairment claim.

Effect of Other Factors Does Not Appear Major or Is More Difficult to Substantiate

In addition to inadequately explained RFC assessments and new evidence submitted on appeal, we examined other factors that could affect inconsistent decisions. We could not attribute any significant effect, however, to other factors, such as worsening condition of claimants and the lack of government representation at hearings.

Claimant's Worsening Condition Does Not Appear to Be a Major Contributor

Because claimants must often wait several months—on average almost a year—for an ALJ hearing, it seems reasonable to conclude that some ALJ awards could be explained by the claimants' condition deteriorating during that time. Worsening conditions, however, are not a major contributor to ALJ awards, according to our examination of program data. About 93 percent of ALJ awards had onset dates—dates on which the ALJ had determined the individual had become disabled—that preceded the DDS decision, suggesting that the ALJ had decided the individual had been disabled when the DDS denied the case. If worsening conditions were a major factor contributing to ALJs awarding benefits, we might expect to see ALJ-determined onset dates coming after the date of the final DDS denial. Because such onset dates are relatively rare, however, little basis seems to exist for concluding that worsening conditions influence many ALJ awards. Moreover, neither the Bellmon Report nor the DHQRP study discussed worsening conditions as a key factor influencing ALJ awards.

An ALJ award based on a worsening condition may have also followed a DDS denial based on the assumption that a claimant's impairment would improve within 12 months (individuals are not disabled if their impairment is expected to last less than 1 year), SSA officials noted. If expected improvement did not, in fact, occur, then the ALJ award would have correctly been based on the original alleged date of onset. About 10 percent of ALJ awards are made to individuals whose claim the DDS had

denied on the basis of the duration requirement, according to our analysis of program data. This 10 percent, however, represents a maximum amount because available program data did not allow us to isolate the impact of other factors—such as new information introduced at the ALJ level—which could have been the main reason for the ALJ award.

Lack of Government Representation at Hearings Not Fully Evaluated

Although the ALJ is expected to consider SSA's interests during the hearing, the agency is not formally represented. The presence of a government attorney or other advocate to represent SSA at hearings has been discussed over the years as a way of improving the ALJ hearing process. Although claimants have the right to representation, SSA relies on the ALJ to fully document the case, considering the claimant's as well as the government's best interests.

In the early 1980s, SSA initiated a pilot project at selected hearing offices to test the effect of SSA representation at hearings. At a 1985 congressional hearing, SSA released preliminary information from the pilot that suggested that ALJ awards made in error could be cut by 50 percent if SSA were represented at appeal hearings.[23] Acting under a July 1986 court injunction, however, SSA halted the pilot project. The court concluded that the entire notion of SSA representation, as implemented, violated procedural due process. In May 1987, SSA decided to end the project, stating that the administrative resources committed to it could be better used elsewhere. As a result, the preliminary results were never verified, and a final report was never issued.

SSA's Planned Improvements in Procedures

SSA plans to take several actions so that DDS and ALJ procedures better ensure decision-making consistency, including requiring more detailed DDS rationales, returning selected appealed cases to the DDS for consideration of new evidence introduced at ALJ hearings, and using a "predecision interview" by a disability examiner.

To improve explanations of DDS decisions, SSA plans to require more detailed DDS rationales. New guidelines for all reconsideration denials are to require DDS adjudicators to write rationales explaining how they made their decisions, especially how the medical consultants assessed RFC, treating physician opinion, pain, and other factors. On the basis of feedback from the process unification training, SSA plans further

[23]Hearing before the Select Committee on Aging, House of Representatives, 99th Congress, 1st session, Mar. 18, 1985.

instructions and training for the DDSs on the bases for their decisions and where in the case files this information should go. SSA issued a ruling in July 1996 clarifying that ALJs consider the findings of fact made by DDS medical and psychological consultants as expert opinion evidence of nonexamining sources and plans to issue a regulation to further clarify the weight given by ALJs to the DDS medical consultants' opinions.[24]

To ensure that DDSs have an opportunity to review all relevant evidence before an ALJ hearing, SSA plans to return selected appealed cases to the DDS for consideration of new documentary evidence introduced at ALJ hearings. This would avoid the need for a more costly and time-consuming ALJ decision in cases where the DDS would award benefits. If the DDS cannot allow the returned claim, however, the DDS medical consultant must provide a revised assessment of the case's medical facts. SSA plans to implement this project in May 1997, at which time it would begin selecting about 100,000 of the roughly 500,000 appealed cases per year for such claims.[25] Moreover, SSA's decision to limit such claims to about 100,000 cases may need to be reassessed in light of the possible benefits that could accrue from this initiative.

SSA also plans to test the use of a "predecision interview" by a disability examiner with the claimant before denying a claim. This interview would provide an opportunity for the DDS to routinely obtain and consider testimonial evidence. It would also allow the DDS the chance to better ensure that claimants understand how decisions about their cases are made and what evidence might be relevant. This could improve the claimants' ability to provide complete and relevant information and make all relevant disability claims earlier in the disability determination process.

[24]In April 1997, SSA told us, the notice of proposed rulemaking on the DDS medical consultants' opinions was in final clearance within SSA.

[25]DHQRP data show that 76 percent of appealed cases contain new evidence.

Quality Reviews Do Not Focus on Inconsistency Between DDS and ALJ Decisions

SSA could use its ongoing quality reviews to better focus on differences in DDSs' and ALJs' assessments of functional capacity and of procedures to improve its management of the decision-making process and reduce inconsistent decisions between DDSs and ALJs. Current quality reviews, however, focus on the DDS and ALJ decision-making processes in isolation from one another and do not reconcile differences between them. In addition, to better manage the process and reduce inconsistencies, SSA also needs a quality review system that focuses on the overall process and provides feedback to all adjudicators on factors that cause differences in decisions. SSA has data and mechanisms in place that it could use to begin integrating its quality reviews and to provide feedback to DDSs and ALJs. In the longer term, SSA plans to systematically review decision-making at all levels through a new quality review system.

Quality Reviews Not Designed to Address Differences

SSA has several quality review systems that review disability DDS and ALJ decisions. As shown in table 4.1, each of the reviews has a different purpose. None was developed to identify and remedy the factors that contribute to differences in DDS and ALJ decisions.

Table 4.1: SSA Reviews Differ by Organization Reviewed and Purpose

		Fiscal year 1996 cases reviewed (approximate)	
Type of review	**Purpose**	**Award**	**Denial**
DDS level			
Quality assurance	Determine whether DDS decisions comply with written standards and criteria, including performance standards	27,000	33,000
Pre-effectuation[a] review (PER)	Protect the solvency of the DI trust fund by intercepting DDS award errors before payment	235,000	Not applicable
ALJ level			
Own-motion[a]	Review ALJ award decisions before payment	4,000[b]	[b]
Appealed ALJ denials	Ensure supportability of ALJ denial before possible court appeal by applicant	Not applicable	57,000

[a]The PERs cover DI and DI/SSI concurrent cases; own-motion reviews cover cases involving DI only.

[b]Excludes reviews of "bureau protests," which are generally cases with technical problems related to insured status, but includes a small number of denials, which are being phased out.

At the DDS level, staff who report to SSA's Office of Program and Integrity Reviews (OPIR) perform a quality assurance review to promote the accuracy and consistency of DDS determinations. The review uses continuous random samples of completed award and denial actions. On the basis of errors found during this review, SSA computes accuracy rates for each DDS, which it compares with performance standards. DDSs that fall below standards for two consecutive quarters are subject to increased SSA oversight and may be removed from making disability decisions. In addition, DDS staff also perform a pre-effectuation review or PER (that is, a review before benefits payments are paid) of awards to protect the solvency of the DI trust fund. Under this review, staff review 50 percent of DI awards (not SSI-only cases) to prevent payment of erroneous awards.

At the ALJ level, quality review heavily focuses on the review of claims denied by ALJs and appealed to SSA's Appeals Council. Claimants whose claims are denied by an ALJ and want to appeal the denial must apply to the Appeals Council before bringing their claim to a federal court. The purpose of this final agency review is to ensure that the case file fully supports the ALJ denial decision before possible court appeal by the claimant. On the basis of this review, the Appeals Council may, among other things, reverse the denial decision or remand the case to the ALJ for further action. In addition, like the PER at the DDS level, the Appeals Council performs a PER of ALJ awards. Unlike the 50-percent sample used for the PER, however, this Appeals Council review samples only a portion of DI-only awards totaling about 3 percent of all ALJ DI awards to people under age 59.

Review of Awards and Denials Is Imbalanced, but Effect on Decision Bias Not Evident

As shown in table 4.1, DDS reviews emphasize awards; the ALJ reviews, however, emphasize denials. This may inappropriately give DDSs an incentive to deny claims and ALJs an incentive to award claims in both instances to avoid scrutiny by quality reviewers. Available evidence, however, does not support this conclusion.

Before SSA instituted the PER of DDS award determinations in fiscal year 1981, national accuracy rates were generally higher for initial denials than for awards. After the PER was instituted, this situation reversed. By 1983, award rates were more likely to be accurate than denial rates. This trend may suggest that instituting the review caused a decline in the accuracy of denials, while increasing the accuracy of awards. Other factors could have influenced these accuracy trends, however, including workload pressures and program changes. In addition, the difference between the denial and

award accuracy rates is slight. In fiscal year 1996, the denial accuracy rate was only 2.9 percentage points lower than the award accuracy rate.

Moreover, data from the DHQRP study suggest that the evidence supports ALJ awards and denials equally. As part of that study, reviewing ALJs assessed 3,000 ALJ awards and 3,000 denials and found virtually the same support rates for both types of cases: 81 percent of awards and 82 percent of denials were supported by substantial evidence.

Current Quality Reviews Mirror Differences in Approach and Procedures

How DDS and ALJ quality reviews operate reflects the differences in how decisions are made at the two levels. First, quality reviewers use the same decision-making approach as those they are reviewing. Therefore, they sustain the differences in approach discussed earlier rather than reconcile them. For example, the Appeals Council, mirroring the approach of the ALJs, infrequently consults with medical experts. Second, DDS reviews do not examine the possible impact at the ALJ level of weaknesses in evidence or the explanation of the decision. As a result, SSA misses the opportunity to use quality reviews to strengthen procedures so that DDS decisions better serve as a basis for ALJ consideration.

Differences in Approach Not Identified and Reconciled

The staff and approach used in SSA's quality reviews of DDS decisions mirror those used in the DDS process. SSA review teams, composed of disability examiners and physician consultants, assess the quality of DDS decisions using the same policies and procedures that DDSs use in making their decisions. For example, when review staff examine a DDS decision, a physician consultant on the team has final authority regarding the correctness of the residual functional capacity (RFC) assessment made by the DDS medical consultant.

Likewise, SSA's Office of Hearings and Appeals (OHA) staff perform ALJ reviews in a manner that mirrors the ALJ process. Staff at OHA screen decisions for conformance with the same standards and procedures used by ALJs, then refer cases that merit further review to the Appeals Council, which consists of attorneys. Similar to ALJs, Appeals Council reviewers have sole authority for assessing a claimant's RFC, and they seek medical input infrequently. The Appeals Council's medical staff and contract physicians consulted in about 17 percent of the cases reviewed by the Appeals Council, according to our analysis of available SSA data.

In addition, although SSA's Office of Disability is responsible for promulgating a uniform decision-making policy, management control of reviews is split between OPIR, which reports to the Deputy Commissioner for Finance, Assessment, and Management, and the Appeals Council, which reports through OHA to another Deputy Commissioner. The two review groups have not routinely met to identify and resolve issues related to inconsistent decisions.[26]

Quality Reviews Do Not Ensure That DDS Decision Builds a Solid Foundation for ALJ Decision

SSA's quality reviewers examine the evidence gathered by the DDS to determine if the end result complies with SSA regulations and guidelines. Although SSA's reviewers assess the adequacy of the DDS's explanation of the initial decision, the reviewers consider the DDS to have made an accurate decision whether it is well explained or not. If a DDS medical consultant fails to adequately explain the basis for the RFC assessment—but nonetheless the decision appears correct and based on adequate evidence—the reviewers do not charge DDS with an error affecting its performance accuracy.

This approach focuses on performance accuracy; it does not provide DDSs with routine, systematic feedback on inadequate RFC explanations because SSA does not return cases to DDSs for correction solely because RFC explanations are inadequate. Instead, if reviewers return a case to a DDS because of other types of errors, such as inadequate evidence to support the decision, the returned case would include comments on inadequate RFC explanations by DDS medical consultants, according to SSA officials. Otherwise, the only way that reviewers might provide feedback on inadequate RFC explanations is during periodic visits to DDSs. Consequently, SSA lacks a routine, systematic mechanism for giving DDSs timely information on the adequacy of their RFC explanations.

Likewise, Appeals Council reviews have not emphasized ALJs' consideration of DDS medical consultants' opinions. First, the Appeals Council samples few ALJ awards for review. Such reviews could identify differences between the DDS medical consultant's opinion and the ALJ view. Second, even if the Appeals Council might want to consider the views of DDS medical consultants, the lack of explanation gives the Council little to review.

In addition, SSA's quality reviews of DDSs' performance accuracy do not focus on weaknesses in DDS evidence gathering from the standpoint of

[26]Recently, SSA has started such meetings under its process unification effort.

whether the evidence could later contribute to ALJ reversals. Instead, reviewers of DDS decisions focus on whether the evidence in the file supports the DDS's own decision. They do not consider whether gaps in evidence may become significant in a later appeal. For example, if the file indicates that the claimant has a treating physician, but the treating physician's report is missing from the file, quality reviewers do not automatically cite this as a performance accuracy error. Instead, they determine whether the totality of evidence in the file supports the DDS's decision. If the decision is supported adequately—despite the missing evidence—the reviewers do not charge the DDS with a performance accuracy error, though this lack of evidence could become significant at the ALJ level. Although the DDS decision may be technically accurate, it may also be vulnerable to reversal on appeal, a factor that the current quality assurance system does not consider in assessing the overall quality of DDS decisions.

In keeping with procedures, DDS reviewers also determine whether the DDS has made a reasonable effort to obtain the evidence. In assessing the reasonableness of the effort, however, the DDS reviewers again do not focus on the potential impact of the missing information if the case were to be appealed. Such a focus would be necessary for both identifying and reconciling differences in decisions.

SSA Has Plans to Improve Quality Reviews

SSA has taken or planned several actions to reduce decisional inconsistency, including addressing factors that we identified as important contributors to the inconsistency. First, the agency has started to systematically gather information on this subject. In 1992, SSA established the Disability Hearings and Quality Review Process (DHQRP), which collects data on ALJ decisions and on the DDS reconsideration denial decisions that preceded them. DHQRP provides a data-driven foundation to identify inconsistency issues and focus on strategies for resolution. According to quality reviewers, SSA has continued this process and anticipates issuing more reports in the future.

In addition, SSA is completing work on a notice of proposed rulemaking, with a target issue date of August 1997 for a final regulation to establish the basis for reviewing ALJ awards, which would require ALJs to take corrective action on remand orders from the Appeals Council before benefits are paid. As envisioned, disability examiners and physician consultants as well as reviewing judges will review ALJ awards. In November 1996, SSA began an initial start-up period for this effort and after

the regulation is issued plans to target about 10,000 cases for review during the first year.

Unlike existing quality reviews, the new process aims to identify and reconcile factors that contribute to differences between DDS and ALJ decisions. When the reviewers find ALJ awards they believe are unsupported, they send these cases to the Appeals Council. If the Appeals Council disagrees with the conclusions of the quality reviewers, the case is referred to a panel of SSA disability adjudicators from various SSA units. This review process can reveal significant policy issues because the panel will receive cases in which the reviewing Appeals Council judge disagrees with the reviewing examiner and medical consultant. On the basis of issues identified, SSA could issue new or clarified policies or provide adjudicators with additional training. In addition, SSA's process unification effort calls for returning certain cases to the DDS when new evidence is provided at the hearing level.

In the longer term, SSA envisions instituting a new quality review system that will systematically review decision-making at all levels. One focus of the new system is making the right decision the first time. SSA estimates this new system will help reduce the percentage of awards made by ALJs, while increasing the percentage made by DDSs. Under SSA's model, when this redesign is fully implemented, the percentage of all awards made by ALJs would decline from around 29 to 17 percent, and the percentage made by DDSs would increase from 71 to 83 percent. The agency has not explicitly established this as a goal, however.

Conclusions, Recommendations, Agency Comments, and Our Evaluation

Conclusions

Inconsistent decisions between DDSs and ALJs are a long-standing problem for SSA management with implications for the fairness, integrity, and costs of both the decision-making process and the program overall. The award rate of appeals raises questions about the fairness of the process because many claimants are awarded benefits only after a lengthy appeal. Moreover, persistent inconsistencies between the two levels can undermine confidence in the integrity of the decision-making process. Furthermore, the later the case is finally decided in the appeals process, the more expensive it is to adjudicate.

SSA can make more progress than it has in the past by unifying the decision-making process at both the DDS and ALJ levels. Meanwhile, reducing inconsistent decisions will be limited to some extent by factors inherent in the program. Disability decisions are inherently complex and require adjudicators to exercise judgment on a range of issues. As a result, expectations about the level of agreement possible in such a program should acknowledge this reality. Moreover, the process involves large numbers of decisionmakers with more than 15,000 adjudicators, quality reviewers, and others, including over 1,000 ALJs, making these complex decisions nationwide.

SSA has developed process unification initiatives that, if implemented, could significantly improve the consistency of decisions. Competing workloads at all levels of adjudication, however, could jeopardize progress in this important area. SSA should capitalize on the momentum it has recently gained and give consistency of decisions the sustained attention it requires as an essential part of redesign. For example, the agency has ongoing data gathering and review mechanisms in place that could produce real progress in this area. SSA has not established explicit outcome-oriented goals or measures, however, to assess its progress in achieving consistent decisions. We believe the strategic planning process required under the Government Performance and Results Act can be a useful vehicle to help focus management attention on the results SSA hopes to achieve through process unification and to monitor its progress toward reaching these results. In this context, SSA needs to establish performance goals to measure its progress in shifting the proportion of cases awarded from the ALJ to the DDS level. SSA could then monitor its progress and make corrections if its actions do not achieve the desired results. Using quantifiable performance goals to measure results would place a high priority on this issue and bolster public confidence in SSA's commitment to achieve more consistency in DDS and ALJ decision-making.

Under process unification, SSA plans to ensure that the DDS decisions are better explained and thus more useful to ALJs. Workload pressures at the DDSs, however, may make full and thoughtful explanations of their decisions difficult. SSA will need to consider ways to reduce these pressures if the agency's plans are to be effective. At the ALJ level, SSA's plans to return cases to the DDSs are important, given the significance of new evidence as a possible reason for awards. SSA's decision to limit such returns to about 20 percent of cases, however, could reduce the effectiveness of this initiative.

In addition, SSA plans to improve its quality reviews but could move more quickly to implement these plans. Historically, SSA has never had a unified system of quality reviews, despite studies documenting inconsistent decisions. Specifically, in 1982, the Bellmon Report identified problems in the consistency of less-than-sedentary residual functional capacity (RFC) assessments, and the Disability Hearings Quality Review Process (DHQRP) reinforced this finding in 1994. However, SSA has not effectively used its quality reviews to focus on this problem or taken action to resolve it. Similarly, DHQRP identified problems with DDS rationales, but no systematic feedback has been provided on this issue. The DHQRP results give SSA an adequate foundation and an ongoing review mechanism to begin unifying quality reviews between the DDSs and ALJs without further delay. SSA could, for example, use the DHQRP findings on less-than-sedentary awards to sharpen and focus current Appeals Council reviews. The agency could also focus on the adequacy of DDS decision explanations in its unified quality review program.

We are also concerned that, without adequate planning and evaluation, some redesign initiatives could have unintended consequences. For example, under redesign, SSA intends to develop new, more valid, and reliable functional assessment/evaluation instruments that are relevant to today's work environment. The agency intends to rely heavily on these instruments in decision-making. But, because differences in RFC assessments are the main reason for ALJ awards, SSA should proceed cautiously. As such, it should test any new decision methods to determine their effects on consistency as well as on award rates before widespread implementation.

Recommendations

SSA is beginning to implement initiatives to reduce inconsistent decisions between DDSs and ALJs, realizing that the lengthy and complicated decision-making process and inconsistent decisions between adjudicative

levels compromise the integrity of disability determinations. We support these initiatives and recommend that SSA take immediate steps and be accountable for ensuring that they are implemented as quickly as feasible. For example, using available quality assurance systems, SSA should move quickly ahead to improve feedback to adjudicators at all levels. In addition, to better ensure that adjudicators review the same record, the agency should increase the number of cases it plans to return to DDSs when new evidence is submitted on appeal.

In addition, we recommend that, given the magnitude and seriousness of the problem, the Commissioner should, under the Results Act, articulate the process unification results that the agency hopes to achieve and establish a performance goal by which it could measure and report its progress in shifting the proportion of cases awarded from the ALJ to the DDS level.

SSA's Comments and Our Evaluation

SSA officials generally agreed with the conclusions and recommendations in this report and stated that the report would be useful to SSA in its efforts to reduce inconsistent decisions between DDSs and ALJs. SSA agreed with our recommendation that the agency take immediate steps and be accountable for ensuring that its process unification initiatives are implemented as quickly as feasible.

Regarding our other recommendation, SSA said that the goal of making a greater proportion of awards at the DDS level and fewer on appeal was laudable and would promote good customer service. But SSA disagreed about taking steps to be accountable for attaining this goal. Agency officials believed that the natural outcome of SSA's process unification initiatives would effect an increase in DDS awards and a decrease in ALJ awards. Because process unification is the linchpin of the disability determination process, however, not just disability redesign, we continue to believe that SSA needs to establish a performance goal for achieving process unification and that the Results Act is the appropriate mechanism to do this.

SSA took exception to our remarks suggesting that its proposal for a new decision methodology could exacerbate inconsistent decisions. We do not agree. Under redesign, SSA plans to reduce medical determinations to a relatively small number of claims, while expanding the functional component of the decision-making process. Because it is unlikely that the new decision methodology will eliminate all adjudicator judgment needed

in making functional determinations, we continue to believe that SSA should proceed cautiously and test any new decision methods to determine their effects on consistency as well as award rates. In its comments, SSA stated that it is committed to using research results to dictate which, if any, changes will be made in the decision methodology. We support this commitment.

The full text of SSA's comments and our response appear in appendix III. In addition, SSA provided technical comments, which we incorporated in the report as appropriate.

DDS and ALJ Disability Decisions and Operations

Table I.1: DDS Decisions—September 1992 Through April 1995

Impairment type	Number of decisions	Percent of total	Number of awards	Award rate (percent)	Listing awards	Functional awards	Percent of awards—functional	Awards, unknown basis	Denied
All impairments	4,592,595	100	1,396,717	30	850,633	541,135	39	4,949	3,195,878
Total physical	3,148,186	69	919,668	29	570,089	347,470	38	3,424	2,228,518
Musculoskeletal	1,083,437	24	168,958	16	26,277	142,146	84	535	914,479
Back	660,747	14	69,663	11	5,330	64,080	92	253	591,084
Under age 50	387,252	8	8,719	2	2,970	5,685	65	64	378,533
50 and above	272,486	6	60,823	22	2,352	58,282	96	189	211,663
Unknown age	1,009	0	121	12	8	113	93	0	888
Other musculoskeletal	422,690	9	99,295	23	20,947	78,066	79	282	323,395
Other physical	2,064,749	45	750,710	36	543,812	205,324	27	1,524	1,314,039
Total mental	1,066,217	23	450,176	42	272,795	175,856	39	1,525	616,041
Illness	860,482	19	338,447	39	186,085	151,504	45	858	522,035
Retardation	205,735	4	111,729	54	86,710	24,352	22	667	94,006
Unknown impairment	378,192	8	26,873	7	7,749	17,809	66	1,315	351,319

Source: GAO analysis of SSA administrative data.

Table I.2: ALJ Decisions—September 1992 Through April 1995

Impairment type	Number of decisions	Percent of total	Number of awards	Award rate (percent)	Listing awards	Functional awards	Percent of awards—functional	Awards, unknown basis	Denied
All impairments	759,999	100	586,821	77	142,267	439,663	75	4,890	173,179
Total physical	589,955	78	439,404	74	57,793	377,986	86	3,625	150,551
Musculoskeletal	288,795	38	217,153	75	21,464	194,061	89	1,628	71,642
Back	202,752	27	151,758	75	13,949	136,919	90	890	50,994
Under age 50	111,544	15	76,105	68	8,534	66,989	88	582	35,439
50 and above	69,864	9	57,679	83	3,838	53,556	93	285	12,184
Unknown age	21,344	3	17,974	84	1,577	16,374	91	23	3,371
Other musculoskeletal	86,043	11	65,395	76	7,515	57,142	87	738	20,647
Other physical	301,160	40	222,250	74	36,330	183,924	83	1,996	78,910
Total mental	169,353	22	146,899	87	84,467	61,402	42	1,030	22,454
Illness	137,566	18	120,191	87	63,876	55,482	46	833	17,375
Retardation	31,788	4	26,709	84	20,591	5,920	22	198	5,079
Unknown impairment	691	0	518	75	7	276	53	235	173

Source: GAO analysis of Disability Hearings Quality Review Process study data.

Table I.3: DDS and ALJ Operations, FY 1986 Through First Quarter, FY 1997

Fiscal year	1986	1987	1988
DDS			
Applications received	2,248,432	2,107,221	1,594,833
Initial decisions	2,007,130	2,010,996	1,516,873
Awards	757,943	698,324	540,135
Denials	1,249,187	1,312,672	976,738
Initial award rate (percent)	38	35	36
Reconsideration receipts	533,776	594,698	457,402
Reconsideration appeal rate	43	45	47
Reconsideration decisions	501,631	589,810	438,251
Awards	82,914	83,846	60,600
Denials	418,718	505,964	377,651
Reconsideration award rate (percent)	17	14	14
ALJ			
Appeals received	225,273	270,241	274,779
Appeal rate (percent)	54	53	73
Appeal decisions	170,661	216,916	238,815
Awards	104,371	130,832	150,744
Denials	66,290	86,084	88,071
ALJ award rate (percent)	61	60	63
Percent of all awards	11	14	20

1989	1990	1991	1992	1993	1994	1995	1996	1997 first quarter
1,589,652	1,737,533	2,014,194	2,392,644	2,564,163	2,609,498	2,488,878	2,438,498	516,483
1,489,534	1,589,311	1,802,896	2,258,980	2,513,709	2,551,210	2,551,953	2,298,801	542,368
547,397	621,223	759,120	981,504	974,868	860,578	787,455	707,204	182,159
942,137	968,088	1,043,776	1,277,476	1,538,841	1,690,632	1,764,498	1,591,597	360,209
37	39	42	43	39	34	31	31	34
472,551	525,689	546,294	627,892	769,948	823,641	864,415	798,668	194,958
50	54	52	49	50	49	49	50	54
442,218	484,499	502,561	603,681	746,241	793,689	858,999	766,775	185,269
67,636	80,988	86,998	102,829	106,787	100,173	112,094	100,107	30,151
374,582	403,511	415,563	500,852	639,454	693,516	746,905	666,668	155,118
15	17	17	17	14	13	13	13	16
281,478	297,326	312,892	372,073	488,173	515,148	557,350	497,933	128,690
75	74	75	74	76	74	75	75	83
251,991	248,237	266,818	302,660	319,789	354,173	444,350	485,737	100,241
167,786	177,571	197,758	226,959	238,094	265,776	324,611	323,266	65,424
84,205	70,666	69,060	75,701	81,695	88,397	119,739	162,471	34,818
67	72	74	75	74	75	73	67	65
21	20	19	17	18	22	27	29	24

Sources: DDS data from State Operations Reports; ALJ data from the Office of Hearings and Appeals Key Workload Indicators.

SSA Studies Addressing Differences Between DDS and ALJ Decision-making

Two major SSA studies have found that DDS and ALJ adjudicators systematically reach different results, even when considering the same evidence. In a report known as the Bellmon Report, SSA issued the results of the first study in January 1982. The second study, known as the Disability Hearings Quality Review Process (DHQRP), is an ongoing quality review of ALJ decisions for which SSA periodically issues reports on review results. The most recent DHQRP report was issued in June 1995.

The Bellmon Report

The Bellmon Report's major finding was that even when reviewing the same evidence from the same cases, DDSS and ALJS often reach different conclusions on whether claimants are disabled. SSA issued this 1982 report to comply with a provision known as the Bellmon amendment in the Social Security Disability Amendments of 1980 (P.L. 96-265). This provision required SSA to conduct ongoing reviews of ALJ decisions to ensure that the decisions conform to statute, regulations, and binding policy. The requirement grew out of congressional concerns about (1) the increasing number of DDS denials being appealed to the ALJ hearing level, (2) the high percentage of DDS denials being reversed by ALJS, and (3) the accuracy and consistency of ALJ decisions.

Initiated in October 1981, SSA's study was designed to examine whether two separate sets of reviewers—one using the DDS decision-making approach and the other using the ALJ approach—would reach different conclusions when considering the same evidence for the same case (see table II.1). Under the study, each set of reviewers reached its own conclusions on each case without knowledge of the decision by the original ALJ or the other reviewers and without personal contact with the claimants.

Table II.1: Design of SSA's Bellmon Study

Two sets of reviewers	Review criteria	Decision(s) reviewed	Quality reviewers' responsibilities
SSA medical consultants/ disability examiner teams	SSA's Program Operations Manual System (POMS)	Original ALJ decision	Evaluate medical evidence Assess adequacy of evidence Determine impairment severity Assess residual functional capacity (RFC) Conclude whether claimants are disabled on the basis of the same evidence
Reviewers from SSA's Appeals Council	The act, regulations, Social Security Rulings (SSR), and guidance handbooks	Original ALJ decision	Conclude whether claimants are disabled on the basis of the same evidence

As shown in table II.1, the first set of reviewers were teams of SSA medical consultants and disability examiners who applied the standards and procedures found in SSA's POMS, which governs DDS decision-making. According to SSA, POMS contains SSA's official program policy and program operations guidance, which is binding on DDSs and all SSA components except ALJs and the Appeals Council. POMS is based on, and consistent with, the Social Security Act, SSA's regulations, and SSRS. POMS is also consistent with circuit court case law. Thus, the conclusions of the SSA medical consultant/disability examiner teams represented the correct application of DDS standards.

The second set of reviewers were from SSA's Appeals Council, which is SSA's final administrative review authority on all appealed disability decisions. These reviewers applied the standards and procedures governing ALJ decisions. These governing standards and procedures consisted of the Social Security Act and SSA's regulations and rulings, along with guidance provided in various handbooks. Because they applied these decision-making criteria, the reviewers' conclusions from the Appeals Council represented the correct application of the standards and

procedures that apply to ALJs. The review was conducted in three phases addressing three different questions.

Phase I: Do the Standards, Procedures, and Practices of DDSs, ALJs, and the Appeals Council Result in Different Decisions for the Same Cases?

To address this question, the medical consultant/disability examiner teams reviewed a representative random sample of 3,600 recent ALJ decisions, of which 64 percent had awarded benefits. The Appeals Council reviewers then reviewed all cases in which the medical consultant/disability examiner teams disagreed with the original ALJ decision, plus an additional 300 cases in which no disagreement existed, for a total of 2,183 cases. The Appeals Council—using the ALJ decision-making approach—awarded benefits in 48 percent of the cases; the medical consultant/disability examiner teams—using the DDS approach—awarded benefits in only 13 percent of the cases. The report identified three possible causes for the different results.

First, standards and procedures differed. The ALJ approach, unlike the DDS approach, often resulted in a finding that the claimant's RFC was "less than the full range of sedentary work," which is the most restrictive RFC possible and usually results in benefits being awarded. The Appeals Council reviewers concluded that claimants in 9 percent of the cases had an RFC that restricted them to less than the full range of sedentary work, while the original ALJs had found that 18 percent—twice as much as the Appeals Council—had that RFC. In contrast, the medical consultant/disability examiner teams concluded that none of the claimants had an RFC that restricted them to less than the full range of sedentary work. ALJs and the Appeals Council also awarded benefits more often than did medical consultant/disability examiner teams because of severe pain combined with significant impairments or nonsevere mental disorders combined with significant physical impairments. The report also noted that ALJs apparently gave considerable evidentiary weight to treating physician conclusions that claimants are medically disabled.

Second, standards were inconsistently applied. Although reviewers from the Appeals Council applied the same standards and procedures that ALJs used in making decisions, the Appeals Council denied benefits in 37 percent of the cases in which the original ALJs had awarded benefits and awarded benefits in 21 percent of the cases in which the original ALJs had denied benefits. These inconsistencies were even more pronounced than these percentage differences indicate because even in those cases in which the Appeals Council agreed with the ALJs on whether benefits should be awarded or denied, they disagreed on the basis for the decision.

For example, when the original ALJs based their decisions to award benefits on SSA's medical listings, the Appeals Council agreed with that basis in only 41 percent of the cases; when the original ALJs based their decisions to award benefits on vocational criteria, the Appeals Council agreed with that basis in only 38 percent of the cases. Agreement on other criteria for awarding benefits was significantly lower.

Third, other contributing factors included subjectivity, organizational trends, and management emphases. The report noted historical trends in DDS and ALJ rates of awarding benefits, SSA program and management focus on certain aspects of the process, and the subjective judgment inherent in determining whether an individual can engage in substantial gainful activity. The report stated that one manifestation of the subjectivity in the process may be different decisions produced by different organizational levels.

Phase II: Do Claimants' In-Person Appearances Before ALJs Have an Effect on ALJ Decisions?

This phase involved selecting a special subsample of 1,000 cases from the 3,600 cases used in the first phase. Written transcripts of the hearings were prepared that retained the testimony of expert witnesses but excluded any testimony by the claimant or observations about the claimant's personal appearance. A representative sample of 48 other ALJs then reviewed the files.

The reversal rate of the original ALJs had been 63 percent, but the ALJs who reviewed the sample of cases for this study reversed only 46 percent. Because information on claimants' in-person appearances was the only information that had been removed from the file, the report concluded that claimants' in-person appearances do affect ALJ decisions. The study also observed differences in the reversal rates for claimants with legal representation (61 percent) and those without representation (48 percent).

Phase III: Does the Submission of Additional Evidence After the DDS Reconsideration Decision Affect ALJ Decision-making?

This phase used the same sample of 1,000 cases used in phase II. All medical and vocational evidence added to the file after the DDS reconsideration decision was removed from the files. Another group of 48 ALJs reviewed the files.

When the additional evidence submitted after the DDS decision was removed from the files, the ALJ reviewers' overall reversal rate of 46 percent in phase II declined to 31 percent in phase III. The difference was solely attributed, on the basis of a statistical test, to additional

medical evidence that had been submitted in 74 percent of the cases. Additional vocational evidence did not affect reversal rates. The medical consultant/disability examiner teams also reviewed these cases with and without the additional medical evidence. Their reversal rate was 15 percent when the additional evidence was present and 12 percent when it was not.

DHQRP

SSA's DHQRP study found that differences between DDS and ALJ decisions noted in the Bellmon Report continued into the 1990s. SSA instituted DHQRP partly as a result of our 1992 report on disparities in the ALJ award rates between black and other claimants.[27] Although we did not conclude that racial bias was the factor responsible for these disparities, we could not rule it out, which raised questions in many sectors, including the Congress, about the extent to which SSA has fulfilled its mandate to have a fair, unbiased ALJ hearing process.

SSA responded that our report did not draw a meaningful conclusion about the impartiality of hearing decisions. SSA stated that it could not immediately address all of our specific findings, however, because it did not maintain an ongoing quality review database of ALJ hearing decisions. As a result, SSA implemented a series of initiatives to address questions we raised. Among the initiatives was the creation of DHQRP.

The SSA Commissioner directed DHQRP to examine not only any racial differences but also program issues through an ongoing quality review of ALJ hearing decisions. Implemented in March 1993, DHQRP's objectives are to promote fair and accurate hearing decisions and to collect sufficient data to permit analysis of other adjudicative issues. From DHQRP results, SSA intends to identify areas of the ALJ decision-making process that may require some fine tuning through continuing legal education or program-specific training for ALJs and other adjudicators. Thus far, SSA has issued two reports—in October 1994 and June 1995—covering the results of reviews conducted through March 25, 1994.

DHQRP is a three-tier review process involving (1) medical consultants from SSA's Office of Disability, Office of Medical Evaluation; (2) disability examiners from SSA's Division of Disability Hearings Quality; and (3) ALJs from SSA's Office of Hearings and Appeals (OHA) who serve as reviewing judges (see table II.2).

[27]See Social Security: Racial Difference in Disability Decisions Warrants Further Investigation (GAO/HRD-92-56, Apr. 21, 1992).

Table II.2: Three-Tier DHQRP

Quality reviewers at each tier	Review criteria	Decision(s) reviewed	Quality reviewers' responsibilities
SSA medical consultants	SSA's POMS	DDS reconsideration denial ALJ decision	Evaluate medical evidence Assess adequacy of evidence Determine impairment severity Assess RFC
SSA disability examiners	SSA's POMS	DDS reconsideration denial ALJ decision	Assess whether reconsideration denial is supported Assess whether ALJ decision is supported
Peer reviewer ALJs from OHA	The act, regulations, SSRs, and Hearings, Appeals, and Law Litigation Manual (HALLEX)	ALJ decision only	Assess whether ALJ decision is supported by substantial evidence Assess whether ALJ decision meets review criteria

In the DHQRP study, similar to the Bellmon Report, SSA's medical consultants and disability examiners' reviews represent the DDS approach to decision-making. Their reviews are based solely on criteria found in SSA's POMS, which contains the SSA decision-making policies and procedures for DDS decision-making.

The medical consultants evaluate the written evidence available to (1) the DDS examining team issuing the reconsideration denial that preceded the sampled hearing decision and (2) the original ALJ who rendered the sampled hearing decision. For each review, the medical consultant evaluates and assesses the adequacy of the file's medical evidence, determines the level of severity, and if necessary, assesses the claimant's RFC. The medical consultants do not listen to the audiotape of the testimony offered at the hearing.

After completion of the medical review phase, SSA's disability examiners use the medical consultants' assessments to review the DDS's

reconsideration denial and the ALJ's hearing decision. Using the medical review findings and other evidence in the file, the disability examiner assesses whether the reconsideration denial and the hearing decision are supported adequately. Like the medical consultants, the disability examiners do not listen to the audiotape of the testimony offered at the hearing.

In the last phase of DHQRP, working ALJS serve as peer reviewers who evaluate only the hearing decision as it was issued by the original ALJ. These peer reviewer ALJS measure the degree to which the original ALJ's hearing decision conforms to the Social Security Act, SSA's regulations and rulings, and SSA's HALLEX. In doing so, the peer reviewer ALJS apply the substantial evidence criterion from HALLEX, which defines substantial evidence as "that evidence which, although less than a preponderance, is sufficient to convince a reasonable mind of the credibility of a position taken on an issue when no evidence on the opposing side clearly compels another finding or conclusion."

The DHQRP results reported by SSA demonstrate how DDS and ALJ decisions differ or reveal reasons why they differ even when adjudicators base their decisions on the same documentary case file evidence.

First, although peer reviewer ALJS have concluded that about 81 percent of the ALJ reversals were supported by substantial evidence, the SSA medical consultant/disability examiner teams, who used the DDS decision-making approach, have concluded that only 41 percent of the ALJ reversals were supported adequately by written evidence in the case file. The medical consultant/disability examiner teams have concluded that (1) the evidence in the file actually supported an opposite decision in 48 percent of the ALJ reversals and (2) another 11 percent had insufficient evidence to make any decision.

Second, DHQRP has provided evidence that ALJS generally find claimants' RFCS to be significantly more restricted than do DDS medical consultants using POMS criteria. When peer reviewer ALJS have reviewed ALJ reversals, they have concluded that 56 percent of the claimants had RFCS that limited them to less than the full range of sedentary work, while the medical consultants, who have reviewed the same written evidence as the peer reviewer ALJS, have concluded that only 6 percent had such restricted RFCS.

Third, the peer reviewer ALJS have identified the top factors that influenced the original ALJS to award benefits. The reviewing ALJS have identified these

factors as (1) a more restrictive RFC assessment, (2) symptoms other than pain, (3) claimant credibility, (4) medical source statements, and (5) the impact of pain.

Fourth, DHQRP has shown that DDS medical consultants, when assessing RFC, sometimes have overstated the claimant's capacity to function in the workplace. In reviewing DDS reconsideration denial case files, the SSA medical consultant/disability examiner teams have found that in 18 percent of the cases the medical evidence supported awarding benefits. In many of those cases, SSA's review teams using POMS criteria have arrived at different RFC assessments than have the DDSs on the basis of the same written evidence.

These findings are based on cases reviewed through March 25, 1994. As of that date, the medical consultants and disability examiners had reviewed 9,600 ALJ decisions (5,500 that awarded benefits and 4,100 that denied benefits). The peer reviewer ALJs had reviewed a subsample of 6,000 ALJ decisions (3,000 that awarded benefits and 3,000 that denied benefits).

The sample for DHQRP is selected monthly from OHA's case control system. Although hearing allowances are reviewed after SSA starts sending benefit payments, all cases selected for review are administratively final and are not changed by the study results. Hearing denials selected for the study either already have been reviewed by the Appeals Council, or their appeal period has expired.

The sample design allows analyses and conclusions to be drawn at the national and regional levels but not at the state or DDS level. In addition, the DHQRP database does not identify the ALJs who issued the decisions being reviewed.

Comments From the Social Security Administration and Our Evaluation

Note: GAO comments supplementing those in the report text appear at the end of this appendix.

SOCIAL SECURITY

Office of the Commissioner

July 8, 1997

Ms. Jane L. Ross, Director
Income Security Issues
U.S. General Accounting Office
Washington, D.C. 20548

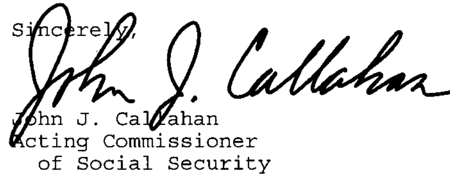
Dear Ms. Ross:

Thank you for the opportunity to comment on the draft report, Social Security Disability: SSA Must Hold Itself Accountable for Continued Improvement in Decision-Making, GAO/HEHS-97-102). We appreciate the General Accounting Office's acknowledgment of the complexity of the issues in determining disability entitlement. The report's findings are thoroughly researched and we highly regard your staff's input on the subject.

We wholeheartedly support your conclusion that the Social Security Administration (SSA) needs to take immediate steps to reduce decisional inconsistency between State Disability Determination Services' and SSA's Administrative Law Judges. I want to assure you that I am convinced that the goal of process unification is the linchpin to our disability redesign program. Your findings and suggestions can only serve to help us achieve this goal.

If you have any questions, please call me or have your staff contact Carolyn Colvin at (410)965-0100.

Sincerely,



John J. Callahan
Acting Commissioner
of Social Security

Enclosures

SOCIAL SECURITY ADMINISTRATION BALTIMORE MD 21235-0001

COMMENTS OF THE SOCIAL SECURITY ADMINISTRATION (SSA) ON THE GENERAL ACCOUNTING OFFICE (GAO) DRAFT REPORT, "SOCIAL SECURITY DISABILITY: SSA MUST HOLD ITSELF ACCOUNTABLE FOR CONTINUED IMPROVEMENT IN DECISION-MAKING" (GAO/HEHS-97-102)

Thank you for your report that examines the SSA disability program's multi-level decisionmaking processes. As you know, the Agency is in the midst of the arduous task of unifying the initial and appeal-level processes. We appreciate the thorough analysis given to the subject and are pleased that your report acknowledges the complexity of the subject, and the difficult task we face in trying to lessen decisional inconsistencies. We especially appreciate GAO's support of SSA's completed or ongoing actions.

We believe that several instances of the report's wording could be clarified. Therefore, in addition to our comments below, attached is SSA's mark-up of your draft report.

General Comments

Now on pp. 38, 52.
See comment 1.

We note the report's modification of the terminology "residual functional capacity" (RFC) to "functional assessments." The distinction is important because the term "functional assessment" has a specific meaning, especially in the medical community, that is quite different from an administrative determination of RFC. Using the term "functional assessment" to describe an adjudicator's findings can create confusion about the decisionmaking process. Generally, a functional assessment is an evaluation—performed in a clinical setting—of a patient's ability to perform a specific activity or set of activities. In SSA's disability process, a functional assessment may be included in the medical evidence and, therefore, must be evaluated along with all other evidence. An RFC assessment is an administrative finding of fact by the adjudicator based on his/her evaluation of the evidence. The consequences of obfuscating the distinction can be seen in the last paragraph of Chapter 3 in the draft report (and again duplicated just before the "Recommendations" section in Chapter 6). The narrative reflects a misunderstanding of the role of functional assessment (as opposed to RFC assessment) in the decisionmaking process, and of SSA's proposal for developing a new way to use functional assessments. SSA's proposal is not to expand use of adjudicator RFC assessments (as implied by the discussion at the end of Chapter 3) but, rather, to change the way we assess functioning in our decision process.

Functional assessments to be utilized in the new methodology would consist of direct measurement of function, and this is currently the subject of research. Such assessments are designed to be more objective and may actually ameliorate some of the variations cited by GAO.

Given this distinction, we recommend that GAO reconsider its remarks suggesting that the SSA proposal for a new decision process has the potential to exacerbate decisional inconsistency. We are committed to using our research results to dictate what, if any, changes are made in decision methodology. Therefore, we recommend that the report state that utility and suitability would be determined by the results of SSA's research and testing efforts.

GAO RECOMMENDATION

SSA is beginning to implement initiatives to reduce decisional inconsistency between DDSs and ALJs, stemming from a realization that the lengthy and complicated decision-making process and inconsistent results between adjudicative levels compromise the integrity of disability determinations. We support these initiatives and recommend that SSA take immediate steps, and hold itself accountable, to ensure that they are implemented as quickly as feasible. For example, using available systems, SSA should move ahead expeditiously to implement its specific quality initiatives for improving the consistency of feedback to adjudicators at all levels. In addition, to better ensure that adjudicators review the same record, the agency should step up the number of cases it plans to return to DDSs when new evidence is submitted on appeal.

SSA COMMENTS

We concur.

Process unification is perhaps the single most critical element in the disability process redesign. We intend to devote significant effort to achieve results in this area beginning immediately. In fact, we have already taken several actions that should ultimately have significant impact on achieving greater consistency. Among them, we have:

- Issued eight new Agency rulings to clarify policy on complex disability adjudication issues;

- Conducted nationwide training for over 14,000 State and Federal staff (including ALJs and DDS decisionmakers) involved in disability decisionmaking, and will conduct post-training follow-up with training facilitators;

- Developed procedures for returning some cases to the DDSs for medical consultant review. Actual case returns will begin later this month and, over the next 12 months, we plan to return approximately 100,000 cases in which new medical evidence is submitted prior to the hearing;

- Issued instructions to DDSs for improved documentation of decisional rationale; and
- Begun pre-implementation testing of a quality review of hearing decisions.

We also intend, as you suggest, to implement specific quality initiatives designed to provide feedback to adjudicators at all levels. Finally, in April 1997, we started testing the full-process model in eight States, which includes the elimination of the reconsideration step and inserts the pre-decisional interview.

RECOMMENDATION

The Commissioner should articulate in the agency's annual performance plan a goal for making a greater proportion of awards at the DDS level and fewer on appeal, consistent with the Government Performance and Results Act.

SSA COMMENTS

See comment 2.
Now on p. 51.

While the goal of increasing the percentage of awards made at the DDS level and decreasing the percentage at the appeals level is laudable and certainly promotes good customer service, we do not believe it advisable, nor necessarily legally defensible, to set quantitative goals for such percentages or for allowance rates at each level of our process. Claimants, under the Social Security Act, are entitled to an individual determination or decision on their claim at each adjudicative level. Uniformly understood and applied substantive disability policy should promote the making of correct decisions as early in the process as possible. However, numerical goals for the distribution of decisions cannot be set because a wide variety of factors influence the outcome of claims, including passage of time, worsening of condition, new evidence, etc. Setting numerical goals could be construed as dictating decisions of allowance or denial, while each claim must actually stand on its own merits.

However, we believe that the natural outcome of our process unification initiatives will bring about an increase in awards at the DDS level and decrease at the ALJ level. Moreover, our initiatives include monitoring the percentage spread between DDS and ALJ allowance rates to measure whether the Agency is making more allowance determinations at the DDS level and fewer allowance decisions at the ALJ level. In fact, we have already begun to see early indicators of this kind of change. Our latest data indicate that the DDS allowance rate has increased, from

4

30.9% in fiscal year (FY) 1995 and 30.8% in FY 1996 to 32.4% in FY 1997. At OHA, allowance rates have declined, from 65% in FY 1995 to 58.4% in FY 1996 and 55.7% through May in FY 1997.

Attachment

The following are GAO's comments on the Social Security Administration's letter dated July 8, 1997.

GAO Comments

1. We revised the report, where appropriate, to use the term "RFC" assessment rather than "functional" assessment. In addition, we clarified our discussion of SSA's plans for a new decision methodology to distinguish between (1) the development new instruments for clinical assessments of function and (2) the expansion of functional considerations in the administrative decision-making process envisioned under disability redesign. We agree that more objective assessments of function would ameliorate some differences in results if these assessments provided better evidence for decision-making.

We continue to believe, however, that the new decision-making process may exacerbate inconsistent decisions. Under disability redesign, SSA plans to reduce medical determinations to a relatively small number of claims, while expanding the functional component of the decision-making process. Because it is unlikely that the new decision-making method will eliminate all adjudicator judgment needed to make functional determinations, we continue to believe that SSA should proceed cautiously and test any new decision-making methods to determine their effects on consistency as well as on award rates. In its comments, SSA stated that it is committed to using research results to dictate which, if any, changes will be made to the decision-making methods. We support this commitment.

2. We revised our recommendation to emphasize the importance of SSA's committing itself under the Results Act to foster consistency in results and to monitor and report on its progress in shifting the proportion of cases awarded from the ALJ to the DDS level. Such a shift is the measure of SSA's achieving its qualitative goal to "make the right decision the first time." In it comments, SSA stated that it wholeheartedly supports our conclusion that it needs to take immediate steps to reduce inconsistent decisions and considers process unification the linchpin of disability redesign. But the agency has not taken steps to be accountable for the success of this effort. Because process unification is the linchpin of the determination process, not just disability redesign, we continue to believe that SSA needs to establish a performance goal for this initiative and that the Results Act is the appropriate mechanism to achieve the desired results.

SSA believes that it would not be legally defensible or advisable to set quantitative goals for increasing DDS award percentages and decreasing the

award percentages for hearings. It stated that setting numerical goals could be construed as dictating decisions to award or deny claims, while each claim should stand on its own merits. We believe that if properly designed and implemented, however, the use of a performance goal would direct the agency's overall management of the process and would not dictate individual decisions.

Furthermore, the agency states that a wide variety of factors influences claim outcomes, including worsening conditions, new evidence, and the like. As the report discusses, however, virtually all factors influencing differences in results—differences in approach, inadequate DDS written evaluations, and problems in the focus of quality reviews—are under SSA's management control. Therefore, we believe it is advisable for SSA management, consistent with the Results Act, to hold itself accountable for continued progress toward process unification.

GAO Contacts and Staff Acknowledgments

GAO Contacts

Cynthia A. Bascetta, Assistant Director, (202) 512-7207

Staff Acknowledgments

David F. Fiske
William E. Hutchinson
Ira B. Spears
Kenneth F. Daniell
Carolina M. Morgan
Ellen S. Habenicht
Carol Dawn Petersen
Mary Ellen Fleishman
John G. Smale, Jr.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000
or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

or visit GAO's World Wide Web Home Page at:

http://www.gao.gov

PRINTED ON RECYCLED PAPER

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested