Disability Evaluation Study Design: First Interim Report

Gooloo S. Wunderlich and William D. Kalsbeek, Editors; Division of Health Care Services, Institute of Medicine and Committee on National Statistics, National Research Council

ISBN: 0-309-59112-0, 43 pages, 8.5 x 11, (1997)

Disability Evaluation Study Design

First Interim Report

Gooloo S. Wunderlich and William D. Kalsbeek, Editors

Committee to Review the Social Security Administration's Disability Decision Process Research
Dorothy Rice, *Chair*
Division of Health Care Services
INSTITUTE OF MEDICINE
Committee on National Statistics
Commission on Behavioral and Social Sciences and Education
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
Washington, D.C. 1997

National Academy of Sciences 2101 Constitution Avenue, N.W. Washington, D.C. 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce Alberts and Dr. William A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

The project that is the subject of this report is supported by Contract No. 600-96-27893 between the National Academy of Sciences and the Social Security Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the agency that provided support for this project.

International Standard Book Number 0-309-05889-9.
Additional copies of this report are available for sale from the National Academy Press, 2101 Constitution Avenue, N.W., Box 285, Washington, D.C. 20055. Call (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area), or visit the NAP's home page at *http://www.nap.edu*. This report is also available via the Internet at *http://www.nap.edu*.

For more information about the Institute of Medicine, visit the IOM home page at *http://www2.nas.edu/iom*. For more information about the Committee on National Statistics, visit the CNSTAT home page at *http://www2.nas.edu/cnstat*.

Printed in the United States of America.

COMMITTEE TO REVIEW THE SOCIAL SECURITY ADMINISTRATION'S DISABILITY DECISION PROCESS RESEARCH

DOROTHY P. RICE* *Chair*, Professor Emeritus, Department of Social and Behavioral Sciences, Institute for Health and Aging, University of California at San Francisco

MONROE BERKOWITZ, Professor of Economics, Emeritus, and Director, Disability and Health Economics Research, Rutgers University, New Brunswick, New Jersey

RONALD S. BROOKMEYER, Professor of Biostatistics, The Johns Hopkins University School of Hygiene and Public Health

GERBEN DEJONG, Director, National Rehabilitation Hospital Research Center, and Professor of Family Medicine and Adjunct Professor, Georgetown University Institute of Public Policy, Washington, D.C.

MARSHAL F. FOLSTEIN, Chairman and Professor of Psychiatry, Tufts University School of Medicine; and Psychiatrist-in-Chief, New England Medical Center, Boston

DAVID B. GRAY, Professor and Associate Director for Research, Program in Occupational Therapy, Washington University School of Medicine

ROBERT M. GROVES, Professor and Program Director, Survey Research Center of the University of Michigan at Ann Arbor, and Director of the Joint Program on Survey Methodology at the University of Maryland at College Park

WILLIAM D. KALSBEEK, Professor of Biostatistics and Director, Survey Research Unit, University of North Carolina at Chapel Hill

JERRY L. MASHAW, Sterling Professor of Law and Management and Professor, Institute for Social ant Policy Studies, Yale University

CATHARINE C. MASLOW, Director, Initiative on Alzheimer's and Managed Care, Alzheimer's Association, Washington, D.C.

DONALD L. PATRICK,* Professor of Health Services, University of Washington School of Public Health and Community Medicine

HAROLD ALAN PINCUS, Deputy Medical Director and Director, Office of Research, American Psychiatric Association, Washington, D.C.

JOHN A. SWETS, † Chief Scientist, Information Sciences, BBN Corporation, Cambridge, Massachusetts

EDWARD H. YELIN, Professor of Medicine and Health Policy, Department of Medicine and Institute for Health Policy Studies, University of California at San Francisco

Study Staff

GOOLOO S. WUNDERLICH, Study Director
KRISTEN N. ROBINSON, Program Officer
MARGO L. CULLEN, Project/Research Assistant

*Member, Institute of Medicine.
†Member, National Academy of Sciences.

BOARD ON HEALTH CARE SERVICES INSTITUTE OF MEDICINE

COMMITTEE ON NATIONAL STATISTICS NATIONAL RESEARCH COUNCIL

Contents

1

Introduction

This report reviews the scope of work for a contract to be issued by the Social Security Administration (SSA) to conduct a Disability Evaluation Study (DES). This is the first in a series of short interim reports by the Committee to Review the SSA's Disability Decision Process Research.[1] This interim report is limited to a preliminary examination of the general features of the proposed survey design, data collection plans, coverage, and sampling as described in the draft scope of work dated July 30, 1996 (SSA, 1996b), submitted to the committee by SSA, as well as subsequent related discussions with, and clarifications by, SSA staff in response to inquiries. The committee has made no attempt to comment on the content of the questionnaires, specific measures of functional capability, or the content of the medical examinations and medical and diagnostic tests proposed for the DES.

This review is based on the draft scope of work, the protocol for the study developed by Westat, Inc., under contract with SSA, as well as on other relevant internal documents prepared by SSA in response to the committee's requests for information. SSA staff also made presentations on work done to date on the project, the information goals and objectives of the DES, and on SSA's plans to integrate a redesign of the disability decision process effort and the DES. The committee plans to review and comment further on the final survey—its design, approach, content, and plans for analysis as proposed by the survey contractor—and perhaps again later on once the survey is conducted and data become available. SSA, however, must make some important decisions now about the survey design and other basic features before it can issue a request for proposals (RFP) for the survey. The committee, therefore, believes that a timely, but preliminary, assessment of the draft scope of work for the survey contract and recommendations for changes in several basic aspects of it are appropriate and needed.

[1] The committee conducted much of the work on this report through a subcommittee composed of William Kalsbeek *Chair*, Ronald Brookmeyer, Gerben DeJong, Robert Groves, and Catharine Maslow. It was reviewed by the full committee, and subsequent revisions were made in response to comments from committee members. Thus, the report reflects the collective thinking of the committee on the issues addressed.

2

Background

STATEMENT OF THE PROBLEM

The Social Security Disability Insurance (SSDI) program (Title II of the Social Security Act (hereafter, "the act") and the Supplemental Security Income (SSI) program (Title XVI of the act) are the two largest federal programs providing cash benefits and medical assistance to persons with disabilities. SSDI is an insurance program that provides payments to persons with disabilities based on their having been covered previously under the Social Security program. SSI is a means-tested income assistance program for disabled, blind, and aged persons who have limited income and resources regardless of their prior participation in the labor force. The definition of disability and the decision process are the same for both programs. The act defines disability as an "inability to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment which can be expected to result in death or which has lasted or [is] expected to last for a continuous period of not less than 12 months..." (Section 223 (d)(1)).

During the last decade, SSDI and SSI have experienced unprecedented growth. Between 1989 and 1995, the number of working-age people receiving SSDI benefits rose from about 2.9 million to nearly 4.2 million, an increase of 45 percent. Likewise, the increase in the number of working-age people receiving SSI disability benefits grew from about 2.1 million to about 3.4 million, an increase of nearly 62 percent (SSA, 1996 a, b). This growth reflects increases in the number of people applying for and entering the programs and a decrease in the number leaving the programs, due at least in part to the younger age of persons with mental impairments. As a consequence of this growth, SSA has faced large workload increases in both the SSDI and SSI programs without a concomitant increase in administrative resources. This imbalance has resulted in significant delays in processing for disability claims and appeals determinations.

Many factors have contributed to fluctuations in the growth in the number of people receiving disability benefits under SSDI and SSI: economic changes in the United States; structural shifts in the labor market; changes in the size, composition, and characteristics of the working-age population (i.e., the baby boom generation, and increasing proportions of younger workers; women and persons with limited English proficiency in the workforce); changes in public policies and the types of disabling impairments that are recognized and diagnosed for disability benefits; a decrease in the average age of persons with disabilities with a resulting increase in duration of benefits; program outreach; and cost shifting by states associated with cuts in state and locally funded general assistance and other welfare programs as well as with deinstitutionalization of people with

mental disorders and mental retardation, and other disabilities previously cared for and financed by state hospital systems.

Despite all these factors and the resulting workload increases, the procedures for processing disability claims have not changed in any important way since the beginning of the SSDI program in the 1950s. Many people have complained that SSA's current disability decision process is lengthy and complicated, resulting in untimely and inconsistent decisions that are often based on subjective judgments (GAO, 1994; 1995; 1996).

SSA'S REDESIGN PLAN FOR DETERMINING PROGRAM ELIGIBILITY

Recognizing the need to improve its claims process, SSA has developed a long-term strategy for reengineering the disability decision process (SSA, 1994). The redesigned process is meant to be simpler than the one currently in place. SSA's basic objectives for the redesigned disability decision process are that it should be user-friendly, prompt and accurate, and that the work should be satisfying for its employees.

As one part of the reengineering effort, SSA has proposed a new decision process for determining whether individuals are "disabled" as defined by the act and SSA's implementing regulations. Briefly, under the present system, the disability decision process consists of four stages: (1) the initial claim, (2) a reconsideration, (3) a hearing before an administrative law judge, and (4) Appeals Council review. SSA's plans call for elimination of two of these stages, namely, reconsideration and Appeals Council review. At present, the decision process for initial claims involves a five-step sequential process (SSA, 1994).[2] Claimants whose applications are denied can have their claims reconsidered. If benefits are denied after the reconsideration, the claimant may request a heating before an administrative law judge at the SSA. Further appeals options include a request for the Appeals Council or the federal district courts to review the decision.

The concept of disability in recent years has generally shifted from a focus on diseases, conditions, and impairments to one on functional limitations caused by these factors (Adler, 1996). "The goal of the new decision process is to focus decision making on the functional consequences of an individual's medically determinable impairment(s)" (SSA, 1994, p. 21). Although the presence of a medically determinable impairment will remain the central requirement for eligibility as required by law, the redesigned process will focus directly on the applicant's functional ability to work and will rely on standardized instruments for measuring functioning to reach decisions.

Because of its complexity and far-reaching impact, SSA has concluded that the re-engineering effort, including the decision process, requires extensive research, testing, validation, and further development of some of its components prior to implementation (SSA, 1996c). Consideration of the specific functional and vocational criteria to be used in the new decision

[2] The decision process involves (1) determining whether the claimant is engaging in substantial gainful activity (GSA), (2) determining the severity of the claimant's limitation, (3) determining whether the claimant's impairment meets or equals the medical listings and how long the impairment has lasted or is expected to last, (4) determining the claimant's residual functional capacity (RFC) to perform past relevant work, and (5) determining the claimant's RFC to perform another type work in the nation's economy.

process has begun, but much remains to be done before satisfactory measures are available. Moreover, the effect of reengineering on the number and characteristics of future beneficiaries also needs further study. The agency, therefore, is now engaged in a multiyear, complex research effort to develop and validate the redesigned decision process. Considerable thought, study, and effort is being invested in developing a research agenda.

SSA has requested the National Academy of Sciences to conduct an independent, objective review of, and make recommendations about, the reliability, validity, adequacy, and appropriateness of its current and proposed research activities as they relate to the proposed redesigned disability decision process. Included in the review is the approach, survey design, and content of the complex multiyear DES. The results of that review are the subject of this interim report and of the committee's ongoing work. (See the Appendix for the study mandate.)

3

SSA's Proposed Design for the Disability Evaluation Study

In the previous sections we described briefly the background leading SSA to redesign the process by which eligibility for disability payments is determined, as well as the agency's approach toward the redesign plan. In this section we briefly discuss the general features of the Disability Evaluation Study (DES) as planned by SSA; identify the key survey design and sampling plan, data collection, and operational decisions made to date; and discuss some of their limitations as they relate to the efficiency of the sampling plan and accepted statistical principles and practices. We also lay out alternatives for SSA to consider and rule on prior to making final decisions about the design and operation of the DES in the scope of work for the survey RFP.

INFORMATION GOALS OF THE DES

In response to a recommendation made by the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds in 1992 (DHHS, 1992), SSA has initiated a significant research effort, with both a short- and long-term focus, aimed at understanding the growth of the disability programs. DES is the cornerstone of SSA's long-term disability research plan. It is a complex multiyear national survey of the population 18-69 years of age.

The adequacy of any survey design depends on what questions the data collected are intended to answer. SSA has identified in the draft scope of work (SSA, 1996a, pp. 11-12) four major information goals for the survey:

1. Estimate the total number and characteristics of persons in the United States with medical, vocational, and functional limitations or impairments, that is, persons with impairments severe enough to meet SSA's definition of disability; this group would represent the universe of potentially eligible nonbeneficiaries who could apply and meet the current criteria, but who are not now receiving benefits.
2. Identify the number and characteristics of people who are not eligible under the current SSA definition of disability, but who could be included as a result of any changes in the disability decision process.

3. Identify the factors (i.e., accommodations, social support, and other factors) that permit persons with similar impairments who could qualify for benefits to continue working.
4. Identify the variables needed to monitor and assess in a cost-effective manner future changes in the prevalence of disability.

The committee has reviewed these goals in the light of alternatives to conducting a survey on a nationally representative sample. It judged that the informational needs of SSA regarding the size and characteristics of the population eligible for benefits, factors permitting them to work, and assessments of future changes in the prevalence of disability could not be fulfilled easily using other existing information sources.

RECOMMENDATION 3-1. The committee strongly endorses the conduct by the Social Security Administration of a well-designed, carefully pretested and statistically sound Disability Evaluation Study.

Large scale surveys like the DES are complex undertakings. They require careful planning prior to the data collection phase and refining during and after data collection. A primary principle of survey design is to determine in considerable detail what data are to be collected and how the data will be used so that decisions can be made on sample design, reliability, and collection procedures. Many federal agencies, when mounting such surveys, establish a committee of technical advisors to offer in the short-term technical input to ensure that management decisions include a careful consideration of possible alternatives. Reliance on the contractor alone to provide such input might not serve SSA well. The committee urges SSA to avail itself of such technical input in a structured manner.

In making this recommendation the committee also urges SSA to place the DES in the context of the ongoing statistical needs of the agency. For example, SSA has stated that it hopes the DES will permit better forecasting of the changes in size of the beneficiary population. This implies ongoing measurement of the size of the pool, with updated instrumentation to reflect any changes in the SSA eligibility protocol. The DES as currently designed does not provide such a capability. Thus, SSA needs to articulate, preferably before all aspects of the DES have been fixed, the methods of providing ongoing monitoring of the size of the eligible pool of beneficiaries.

The committee also believes that a detailed plan for the statistical analysis of the DES is required. This plan should include the development and validation of models that forecast the size of the disabled population and that synthesize the DES data with other data sources.

GENERAL FEATURES OF THE DES DESIGN

As originally conceived, the principal information goal of the DES was (a) to estimate the size of the population potentially eligible for disability benefits in order to assess the upper bounds for the growth of SSDI and SSI programs, and (b) to identify the factors that enable some persons with severe impairments to remain in the workforce.

A national probability sample survey to obtain this information is planned. In order to efficiently identify a relatively large group of potentially eligible nonbeneficiaries a decision was made to use a screening mechanism. These nonbeneficiaries are persons whose impairments are severe enough that they would likely be eligible for disability benefits if they applied. Other subgroups—current beneficiaries, people with lesser impairments (the "borderline" group), and nondisabled persons—are included in the survey to ensure full coverage as well as to provide the data needed to meet the DES objectives. Westat, Inc., under contract with SSA, developed a protocol for the conduct of the DES, setting out and weighing the relative merit of different options to address key aspects of the survey design. The protocol included an appraisal of the basic parameters of the overall size and scope of the survey (Westat, Inc., 1995b).

DES in Three "Stages"

While the DES was being developed, efforts to reengineer the process were underway and the disability decision process was being redesigned on a parallel but separate track. DES assumed an additional role of evaluating the proposed redesigned process and of serving as a source for testing functional assessment instruments and the decision process itself. The protocol proposed by Westat was modified to accommodate these additional roles of the DES. A major alteration, as conceived by SSA, is to conduct the survey in three distinct consecutive stages: a pilot test followed by two separate national data collection stages. Table 3-1 summarizes SSA's survey plan in three "stages" as described in their draft scope of work.

As indicated in Table 3-1, SSA plans to conduct a comprehensive pilot test about 6 to 9 months after awarding the contract and completing the development work for the survey, and prior to launching stage 1. The pilot is meant to test operational methods and data collection instruments, response rates, respondent burden, and to determine what changes are required. The draft scope of work and subsequently reported timelines indicate a 3-month period for the pilot study, but the committee understands that final decisions on the scope and duration of the pilot study have not been made. A 3-month period would be sufficient for little more than a dress rehearsal of a close-to-final-form survey design. Given the present state of the DES protocol, this use of a pilot test stage seems premature. The committee believes it would be more realistic for SSA to use the entire time period set aside for the current stage 1 to develop and field test a DES protocol that would be fully implemented on a national basis in what SSA calls stage 2.

SSA also indicates in the draft scope of work that the pilot test cases and data collection will be used as part of the stage 1 sample, if possible. Considerable changes in the final questionnaires and in operations are usually required as a result of pretests and pilot tests. The committee therefore thinks that including pretest cases in a major national survey like the. DES is not wise.

As currently planned, both stage 1 and stage 2 of the survey are to be designed as complementary national samples so that nationally representative estimates can be generated from each individual sample as well as from both samples combined. SSA's design specifications for the two samples require that they be approximately equal in size and drawn from separate, nonoverlapping primary sampling units (PSU). SSA expects to combine the data from the two stages for most of the survey analysis.

The stated purpose of the two separate data collection stages is to accommodate the reality that some of the components of the revised disability decision process that need to be tested in

stage 1 and will not be completed until after the fieldwork for stage 1 is well under way. SSA, therefore, plans to collect data in stage 1 to facilitate early testing and refining some of the proposed functional assessment criteria for the redesigned disability decision process.

TABLE 3-1 Current Survey Design Proposed by the Social Security Administration

STAGES	ACTIVITY
Pilot Test	A comprehensive pilot test of all operational methods and data collection instruments to assess what if any, protocol alterations are required.
Stage 1	**Half of Data Collection**—First half of DES data collection (including telephone screening, interviews, clinical examinations, functional assessments, other tests, etc.) applied to a national half sample.
	Final Changes in Redesign Method—made in conjunction with the pilot, stage 1 data collection, as well as with separate laboratory testing, this research would facilitate final resolution of issues tied to the redesign decision process, e.g., developing an acceptable functional assessment screening instrument, to experiment with the use of Department of Labor's O*NET for job requirements.
	Supplemental Research—As needed and separate from DES data collection, further research would be undertaken to address other process and design issues.
Stage 2	**Half of Data Collection**—Second half of DES data collection (including telephone screening, interviews, clinical examinations, functional assessments, other tests, etc.) applied to a complementary national half sample; part of the stage 2 survey instrumentation would be determined by the outcome of stage 1.
	Analysis—of national and subnational profiles of disability using stage 1 and 2 data combined, assessment of the redesigned decision process.

SOURCE: SSA, 1996b.

In stage 2, a second nationally representative sample will employ the functional assessment instruments presumably improved from stage 1. SSA expects stage 2 to provide a direct field test of the redesigned disability decision process under development. The interview, medical examination, and functional assessment instruments used in stage 1 will also be administered to the stage 2 sample, as will any new or modified functional assessment instruments required for testing the redesigned disability decision process that were not available for inclusion in stage 1.

In addition, stage 2 will complete the total sample size required to meet the goals of estimating the potential pool of eligible nonbeneficiaries, those of that group who continue to work, and the accommodations they receive. SSA states that most analytical uses of the DES will

combine the data from both stage 1 and stage 2 samples (SSA, 1996a, p. 16). The committee seriously doubts that pooling the data from the two samples is scientifically defensible. The dynamic state of the DES survey instruments is bound to make the stage 1 and stage 2 questionnaires substantially different, thus leading to differential context effects in a large number of survey questions (Schuman and Presser, 1981; Turner and Martin, 1984; Groves, 1989; Smith, 1991; Tanur, 1991). Because stage 1 and stage 2 will involve somewhat different methods, operations, and instruments, it will not be possible to combine the data as planned without introducing "stage effects" in DES estimates resulting from data inconsistencies. After changes are made in stage 2, the data that are collected from the two stages will not be comparable. Content and administration of the questionnaires and functional assessment instruments will differ in the two stages and some questions and tests (e.g., the final functional assessment instruments) will be administered in stage 2 but not in stage 1.

The committee understands that stage 1 is meant to identify and test the measures of functional ability, and to inform the development of the final instruments for use in stage 2. However, a nationally representative sample is not needed for these purposes. Much of the work on definitions and procedures for data collection could be tested with small samples using different approaches including the use of laboratory research. Increasingly survey operations, especially for large scale surveys, are being preceded by small scale, carefully structured field testing coupled with laboratory research, so that definitions and concepts critical to the results can be better understood. Geographically, socially, and educationally dispersed nonprobability samples, laboratory subjects, or both are adequate to develop and pretest the functional assessment instruments. Indeed, several existing sources can be used for such testing for the DES.[3]

RECOMMENDATION 3-2. The committee recommends that the current stage 1 and pilot study be merged, expanded, and extended into a research, development, and testing phase of the survey with application to samples of the type that are more traditionally used in methods testing. Only When the development and refinement of the functional assessment instruments, survey operations, and other issues are tested and resolved should a national sample survey be launched using a single protocol.

Recommendations 3-2 and 3-3 are part of a broad revision of the survey design for the DES. Table 3-2 presents a summary of the committee's recommended survey design.

A significant advantage of these recommended changes in survey design for the DES is that they will ensure a more statistically sound design and accompanying estimates, with no significant delay in the award of the contract. The research, development, and testing phase will be part of the survey contract. In addition, they are likely to result ultimately in appreciable cost savings (since one rather than two national surveys would be conducted (Table 3-3).

[3] Some examples of these sources are state vocational rehabilitation agencies, labor unions, large companies, federal government agencies, and national associations that advocate for those with impairments such as Independent Living Centers, the Association of Retarded Citizens, the Learning Disability Association, the Manic-Depressive Society, the National Amputee Coalition, People First, and the Paralyzed Veterans of America.

TABLE 3-2 The Committee's Recommended Survey Design

RESEARCH, DEVELOPMENT, AND TESTING	NATIONAL PROBABILITY SURVEY
Pretests—Field testing of instruments, procedures, modes of examinations, alternative approaches to screener and response rates, and functional assessment instruments on small subnational, nonprobability samples.	**Data Collection**—Data collection (including screening, interview, examinations, functional testing and other tests, etc.) applied to a national probability sample of size equaling or exceeding the size of the samples for stages 1 and 2 combined.
Comprehensive Pilot Test—Testing of the survey instruments, examination, functional assessment criteria, and survey procedures, possibly followed by a dress rehearsal of the survey.	**Analysis**—National and subnational profile of disability using DES data; assessment of redesigned decision process.
Final Changes in Redesign Method—Done as a result of the pretesting and pilot testing, as well as relevant laboratory testing. This research would facilitate, prior to the conduct of the main survey, final resolution of issues tied to the redesigned decision process, e.g., developing an acceptable functional assessment instrument and experimenting with the use of Department of Labor's O*NET for job requirements.	
Analysis of Existing Data—Using existing national data sources (e.g., NHANES, NHIS, SIPP, etc.), complete analyses to inform SSA and the contractor about sample issues and that at least partially address some of the DES information goals.	

TABLE 3-3 Advantages of the Recommended Survey Design

- **Retain current timeline**—SSA can still award the contract following the recently revised timeline.
- **More time for preparation and resolution of issues for DES data collection**—More needed time would be available to resolve remaining decision process and survey design issues, including development of functional assessment instruments, in advance of beginning the national DES data collection.
- **Larger sample size for national estimates**—National estimates would be derived from the full national sample drawn from a larger number of primary sampling units, as opposed to the half samples containing data based on different questions and testing instruments.
- **Uniform data sets**—One uniformly gathered set of data would be available, rather than two half samples with less than fully comparable data.
- **Stronger design/redesigned** comparisons—Data would be available for assessing the impact of design/redesigned decision processes applied to the same set of respondents.
- **Lower data collection cost**—Although the presurvey period will be longer then currently planned and the cost for this activity may be higher, ultimately cost savings would likely arise from having one rather than two data collection startups and tests done on non-national samples.

SAMPLE DESIGN

The sample design for the DES (both stage 1 and stage 2 of the survey) as planned by SSA is driven by the following four objectives (SSA, 1996a, p. 20):

1. produce precise estimates of the various subgroups of the working-age population with disabilities and who among them are disabled enough to be likely to be eligible for disability benefits if they applied;
2. yield a sample of nonbeneficiaries with disabilities sufficient to permit analyses by a number of subclasses;
3. yield a sample of the "borderline" group of persons with disabilities sufficient to permit estimates of the number and characteristics of persons who might become eligible, or cease to be eligible, if the current SSA disability decision criteria are altered, and
4. yield a sample of nondisabled persons sufficient to permit comparisons with the population with disabilities on measures of physical and functional performance and medical conditions in the population.

TABLE 3-4 Summary of Sample Design Proposed by the Social Security Administration

SAMPLING STEP*	SAMPLING UNITS	STRATIFICATION	TYPE OF SELECTION METHOD	APPROXIMATE SAMPLE SIZE
1	Contractor-specified primary sampling units (PSUs)	**Unspecified by SSA** (typically by region, racial/ethnic mix, and degree of urbanization)	**Unspecified by SSA** (typically with probability proportional to population size or number of households)	100-110 PSUs mentioned by Westat—half of which will be used for each data collection stage
2 (Phase 2)	**Households** (from a sample of residential phone numbers; beneficiary sample from SSA records)	**Unspecified by SSA** (beneficiary status, if SSA records are used as a frame source)	**RDD Sampling** (SSA's choice, although face-to-face screening was considered)	171,000 persons 18-69 years of age (in 100,000 screened households)
3 (Phase 2)	**Person** (every person in household up to 3 persons)	**Yes** By disability groupings: 1. Potentially eligible nonbeneficiaries 2. Borderline nonbeneficiaries (persons with lesser impairments) 3. Nondisabled persons 4. Current beneficiaries	**Unspecified** (Stratified random sample, with variable sampling rates set to achieve targeted sample sizes for the four disability groups)	**5,500** By disability groupings (in Col. 2): 1. Potentially eligible nonbeneficiaries (3,000): • who are working persons (700) • who are younger (450-600) • who are 62-69 years old unspecified) • with mental, emotional or behavioral conditions (750-900) • from minority groups (500 blacks) 2. Borderline nonbeneficiaries (persons with lesser impairments) (1,500) 3. Nondisabled persons (500) 4. Current beneficiaries (500)

* These levels of sampling, called "stages" in sampling statistics, are referred to here as "steps" to minimize confusion with SSA's reference to "stages" when referring to "rounds" of DES data collection.
SOURCE: SSA, 1996b.

The sample sizes proposed by SSA appear to be driven primarily by the first objective and by cost considerations. With those two factors in mind, SSA has set a goal of identifying a sample of 3,000 nonbeneficiaries with severe disabilities (the likely eligible group), out of a total sample of 5,500 persons. Table 3-4 summarizes the sample design as planned by SSA and the size of the subgroups within the total sample.

Severe impairments are relatively rare in the general population. In fact, the severity and prevalence of a disabling condition are inversely related; the higher the prevalence of a condition, the lower the severity, and vice versa (LaPlante, 1991). Because SSA's eligibility criteria tend to filter out people with less severe disabilities, SSA is faced with many low-prevalence disabling conditions all of which cannot be screened adequately into the sample. The exceptions may be mental conditions and low back conditions. SSA is cognizant of this situation, and therefore it has built into its sampling plan provision for oversampling persons with severe disabilities.

Accordingly, the sample will contain:

- a "core" group of nonbeneficiaries with severe disabilities (3,000),
- persons with significant but lesser impairments, the "borderline" cases (1,500),
- nondisabled persons (500), and
- current SSDI and/or SSI disability beneficiaries, who will be included primarily for the purpose of benchmarking the distinctive characteristics of the core group (500).[4]

SSA assumes that the core group sample of 3,000 will be sufficient to analyze several subgroups of particular policy interest. These subgroups will include potentially eligible nonbeneficiaries who are working; younger nonbeneficiaries with disabilities; nonbeneficiaries aged 62-69 years; nonbeneficiaries with mental, emotional, or behavioral conditions; and nonbeneficiaries with disabilities from minority groups.

The committee has several questions and concerns about the adequacy of the size of the total combined sample, of stage 1 and stage 2 samples, and of the allocations among the four subgroups. How did SSA arrive at this particular disproportionate sample design? The allocation clearly favors the 3,000 nonbeneficiaries with severe disabilities. What is the basis for choosing the sample sizes for the four groups? What precision targets were used to arrive at the sample sizes? What response rates are to be reached for each component of the DES and for the overall survey? It appears that the sample sizes discussed in stages 1 and 2, both separately and combined, will lack the condition specificity that SSA would require for estimation and analytical purposes. The cells will be much too small, especially if SSA stratifies on more than one disabling condition and/or demographic or socioeconomic characteristics such as age, gender, minority status, or working nonbeneficiaries with specific disabling conditions.

Similarly, the proposed sample size for the borderline group of persons with less severe disabilities may not be sufficient in its analytical strength for assessing how alternative decision processes or definitions of disability would affect outcomes. The differences in outcomes resulting from changing the decision process is likely to be minimal if any for persons with severe disabilities, but some real differences could show up among borderline cases under alternate

[4] According to the draft scope of work, this sample of current beneficiaries may be drawn from screened cases, from SSA administrative records, or both.

decision processes. Clearly, SSA needs to give more thought to issues of sample selection, size, and allocation.

RECOMMENDATION 3-3. The committee recommends that the national survey should be conducted with one sample large enough to estimate the sizes of the populations at risk with acceptable levels of statistical precision.

This recommendation, along with Recommendation 3-2, reflects the committee's concern about the plausibility of merging data from stages 1 and 2 of the DES. Also, they reflect the inadequacy of a 1-year period after award of the contract for planning, developing, and testing the survey methods that will be used for the national DES. The impact of these two recommendations on the timeline of the DES and on SSA's research on the proposed changes in the disability decision process are shown in Figure 3-1. There one can see that the committee's recommendation to gather DES data as part of a single national survey, rather than as part of two potentially incompatible half samples, implies a period of at least 24 months instead of 12 months for the research, development, and testing activities. These activities will be a key to the success of the DES.

ANALYSIS OF DATA FROM EXISTING NATIONAL SURVEYS

Over the past many years, several national surveys have included a significant number of questions through special modules or supplements about chronic disease, disability, employment status, and receipt of disability benefits. Findings from the surveys provide some information about the relationships among these factors, but the usefulness of the information for purposes of SSA's disability decision process redesign is limited by several problems, including the small number of working-age subjects with severe disabilities, the lack of sufficiently detailed questions about their impairments, and the absence of a medical examination and tests of functional capacity. However, data from these surveys could provide some idea of the prevalence of disability in this country and the proportion and characteristics of those people who are working despite their disabilities. Such data could be useful to guide SSA in questionnaire development and in designing the sampling plan for the DES. Some examples of major surveys are the National Health Interview Survey (NHIS) and the National Health and Nutrition Examination Survey (NHANES) both conducted by the National Center for Health Statistics (NCHS), and the Survey of Income and Program Participation (SIPP) conducted by the Bureau of the Census.

The 1994-1996 Disability Supplement of the NHIS was conducted by the NCHS in two phases with a nationally representative sample of people 18-64 years of age. It addresses many of the issues of interest to DES. The Phase I and II questionnaires include detailed questions about chronic disease, disability, employment status, and receipt of disability benefits. The sample for Phase II of the survey includes more than 10,000 working-age persons with disabilities based on an extensive interview in Phase I.[5]

[5] The 1994 Phase I data collection was completed in December 1995, and data from the 1994 Phase I interviews are currently available for analysis. The 1994 Phase II data collection is almost complete. The 1994 Phase II data on adults will be available by July 1997. The 1995 Phase I data collection will be completed by June 1997, and Phase II will be done by the end of summer.

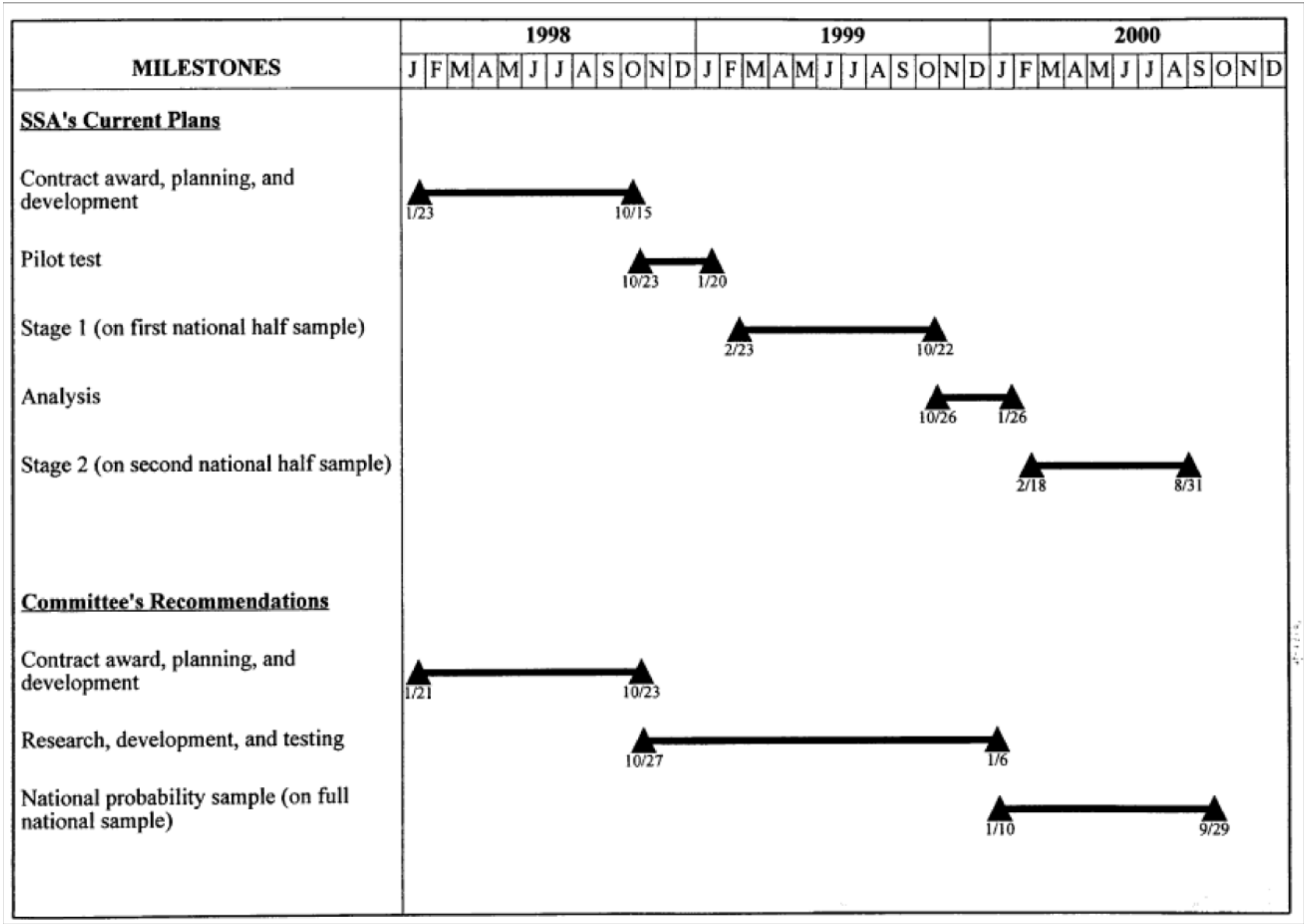

Figure 3-1
Alternative timeline for the Disability Evaluation Study. (The timeline for the decision process research will not be affected.)

The committee's preliminary review of the questionnaires for Phases I and II of the NHIS Disability Supplement yielded many questions that seem relevant for the development of the DES. Table 3-5 shows the topics covered. Within each topic, the questionnaires include many detailed questions.

The committee recognizes that the NHIS Disability Supplement does not include performance tests, medical examinations, or subjects' medical records. Thus, the survey findings cannot be used to simulate SSA's current disability decision process—one of SSA's objectives for the DES. On the other hand, analysis of the survey findings would provide valuable information from a nationally representative sample on (a) the proportion of working-age people with certain diseases, conditions, and functional limitations who do and do not work, (b) the proportion of those who do not work who receive disability benefits, and (c) respondents' perceptions of factors that affect their ability to work. This information would be useful in designing the screening instrument for DES, particularly in developing questions that would help to identify individuals in two of SSA's target groups: nonbeneficiaries who might be expected to meet the SSDI/SSI criteria if they were not working, and people in the borderline group, who might be expected to nearly meet the criteria.[6]

A topical supplement containing an extensive set of questions about disability was used as part of the sixth wave of the 1990 panel and the third wave of the 1991 panel of the SIPP. The total sample size for this study was approximately 30,000 interviewed households (McNeil, 1993).

Similarly, NHANES-III, conducted from 1988 to 1994 on a sample of 40,000 persons aged 2 months and older, included a relatively small number of working-age subjects with severe disabilities; it did, however, include medical examinations and some performance tests (NCHS, 1994). The committee believes that experience gained from this survey could answer important questions about response rates, refusals, and procedures for conducting medical examinations and performance tests on subjects with severe disabilities in a survey context, and about topics such as self-versus proxy reports.

Moreover, SSA could learn from the experience of the NHANES survey, in particular, about conducting medical examinations and performance tests on people with severe disabilities. For example, people with severe disabilities may not be willing to leave home for these procedures, and may not want to undergo a lengthy medical examination. Others may not be able to provide accurate information about their medical records. Proxy respondents and home examinations may be a necessity for persons with severe disabilities.

RECOMMENDATION 3-4. The committee recommends that the Social Security Administration use relevant data from the National Health Interview Survey Disability Supplement, National Health and Nutrition Examination Survey, Survey of Income and Program Participation, and other relevant surveys to assist in developing the sample design, survey operation, and questionnaire content for the Disability Evaluation Study.

[6] The 1994-1996 Disability Survey also includes a nationally representative sample of elderly persons. The Phase II interview for the elderly differs in some ways from the Phase II interview for the nonelderly. It is likely, however, that information about subjects age 65-69 could be used to answer some of SSA's questions about people in this age group.

TABLE 3-5 Topics and Variables from Phases I and II of the 1994-1996 Disability Survey, Population 18-64 Years of Age

PHASE I: TOPICS	PHASE II: TOPICS AND VARIABLES
Sensory impairments (problems with seeing, hearing, smelling, tasting)	Repeated
Communication problems	
Learning problems	
Specified diseases and conditions	Specified diseases and conditions
ADL impairments (need for hands-on help or supervision, extent of difficulty	Repeated with many variables, including difficulty, pain, and speed of performance
IADL impairments (need for hands-on help or supervision, extent of difficulty	Repeated with many variables, including difficulty, pain, and speed of performance
Other functions and activities (lifting, walking, standing, bending, reaching, using fingers to grasp or handle objects)	Repeated with many variables, including the difficulty with the function or activity
Mental and emotional problems (impact on ability to find or keep a job and kind or amount of work the person can do)	
Participation in sheltered workshop, work training, and supported employment	Repeated with many variables
Use of physical or occupational therapy	Repeated with many additional services
Use of vocational rehabilitation services	Repeated with many variables
Use of assistive devices	Repeated with many variables
Income and assets	
Receipt of disability payments	
Employment status	Employment status or work history (ever worked, type of work, volunteer work)
	Need for accommodations in order to work (e.g., ramps, parking, special workstation)
	Need for other assistance to work (e.g., interpreter, job coach, personal attendant)
	Reasons for not working or looking for work (e.g., fear of loss of disability benefits, or health insurance, lack of transportation)
	Type of living situation (e.g., apartment, group home, center for independent living); adaptations or modifications to home, services provided by living situation
	Transportation (ability to drive, use of regular public transportation, special transportation, mobility training)
	Social activities (e.g., visit friends or relatives, go out, talk on the telephone)

SOURCE: NCHS, 1995.

4

Data Collection Plan

Data collection for the DES involves:

- a screening interview;
- a personal interview and physical performance tests;
- an extensive medical, and if needed, psychological examination; and
- a series of core and special medical tests.

In addition, SSA plans to obtain all medical evidence of record identified by the respondent and by third party reports on all persons in the sample to supplement information from the interviews and medical examinations in order to determine if the person meets SSA's current definition of disability. SSA indicates that the medical evidence of record can provide more information about impairments and other health problems than can be obtained from a medical examination at a single point in time. Survey designs that seek the medical records of persons in the sample face the high cost of conducting record searches as well as the risks of missing data. SSA and the contractor should carefully develop and evaluate protocols for record acquisition during the early months of the DES contract and evaluate subsampling designs, possibly disproportionately sampling important target groups. Any subsampling scheme could reduce the cost of DES relative to a design that obtains medical records for all persons in the sample.

SCREENING INTERVIEW

As stated in the previous section, the survey design includes, as a first phase, a large-scale screening survey to identify a sample of working age persons to permit accurate identification of (a) current SSA disability beneficiaries; (b) nonbeneficiaries with severe disabilities who appear likely to meet SSA's medical/functional disability criteria; (c) "borderline" nonbeneficiaries who have some degree of impairment, and are likely to be candidates for disability benefits at some later time; (d) people who, on the basis of the screening, do not appear to be disabled. At the time of the screening interview, the interviewer is supposed to schedule an in-person main survey interview and medical examination.

The first issue in choosing sampling frames concerns the target population. In general, most national household population surveys routinely exclude people who are institutionalized and those homeless people who cannot be found in households or other living quarters visited during household surveys. SSA has decided to exclude from the DES the institutionalized population and the segment of the homeless population who cannot be found in households or other quarters at the time of the interview.

However, the question of including or excluding homeless people from the DES is not as straightforward as for the other household surveys. The committee recognizes the likelihood of relatively high rates of disability among homeless and institutionalized populations, and the resulting negative bias resulting from their exclusion. At the same time it has serious questions about the operational and methods issues. Can reliable information be obtained, feasibly and economically, from homeless and institutionalized populations? Techniques have been developed to locate, sample, and obtain data about each of these populations. Yet locating and screening respondents for eligibility require special efforts involving careful, and long-term planning, large amount of staff resources, considerable time, and high levels of funding. Homeless people present problems in scheduling, interviewing, and administering performance tests and medical examinations. Maintaining contact with them and getting them to participate in adequate numbers in the medical examination is also problematic. Likewise, obtaining permission from family members for the participation of people in long-term care institutions who are not able to grant permission themselves may be difficult.

The committee discussed these various issues and also reviewed a draft discussion paper prepared by Westat for the SSA (Westat, Inc., 1995a). The committee concurs with SSA that adding homeless and institutionalized populations to the sampling frame at this time would not be cost effective. Much research and testing are required to develop the necessary protocols and procedures for conducting the DES among homeless people and those living in different types of institutions. The costs of sampling and interviewing in the various types of institutions would be prohibitive. Thus, limiting the target population to the household population seems appropriate. The committee, however, urges SSA to undertake research as part of its long-term research plan leading to the inclusion of these populations in subsequent studies or a separate supplement to the DES in the near future.

The current proposal specifies the use of telephone number frames for DES. This decision by SSA appears to be primarily driven by cost considerations. The choice of sampling frame determines the nature of noncoverage error in any survey. Common choices in surveys in the United States are area frames, offering theoretically complete coverage of households and institutions; dual frame designs, combining telephone and area frames; dual frame designs combining area and institutional list frames; and telephone number frames.

If SSA uses a telephone frame as the screening tool, the DES will miss households without telephones. Approximately 5 percent of households in the United States are without telephones. Persons in households without telephones have a higher rate of disability than those in households with telephones. This rate is 17 percent for those without telephones compared with 15 percent among persons with telephones (Thornberry and Massey, 1988; LaPlante and Carlson, 1996). The availability of telephones also is negatively correlated with income. This noncoverage of persons in households with no telephones should be of particular concern with regard to the representativeness of the sample of persons with disabilities.

The nontelephone population can be covered through a dual frame approach, whereby nontelephone households are obtained through an area frame and telephone households through both an area frame and a telephone number frame. (Such a design is currently being used in the New Federalism Survey conducted by the Urban Institute and Westat.) Alternatively, a full area probability design can be implemented to provide such coverage. SSA's current decision omits the nontelephone population from coverage in the DES.[7] Because DES will be the principal instrument by which the United States will obtain prevalence data on disability as defined by law, the committee concludes that a national survey of disability that does not offer statistical estimates of the nontelephone population is statistically inadequate.

Further, telephone sampling and screening is likely to offer lower response rates than face-to-face screening (Groves, 89; Lessler and Kalsbeek, 1992). As a consequence the screening sample will need to be increased to compensate for the losses from the sample because of nonresponse; the higher nonresponse rates are likely to increase the risk of bias in the estimates. Thus, although telephone screening may be less expensive, some aspects of the quality of the data collected are more suspect. Careful study of mechanisms to increase the screener response rate is required. These mechanisms might include incentives, refusal conversion efforts, switches to alternative modes of data collection, and so on.

In addition, there is no indication about how SSA will deal with people with hearing loss, communication disorders, mental and cognitive impairments, and emotional disturbances. SSA also has the problem of response burden for the total household if more than one person in the household has a disability and proxy reporting is not encouraged. Similar problems will have to be faced in the main interview and in administering medical examinations and performance tests to persons with severe disabilities. The effect on response rates and bias could be significant. SSA should test several options dealing with these problems in pretests prior to the start of the national survey.

It is clear that in terms of coverage of the adult working-age population, survey response rates, and some features of the screening measurement, the preferred design is an area probability, face-to-face survey. It is also clear that the cost of such a design is higher than the alternative proposed by SSA. The additional costs for a survey of this importance and complexity should be considered in the context of the size of the program itself (SSDI and SSI)—which, if current trends continue, could cost $100 billion annually—and the implications of poor or imprecise information.

The committee, therefore, urges a careful review of the costs of a full area probability survey, in light of the cost savings proposed in later recommendations.

[7] SSA is considering the possibility of supplementing the telephone screening with face-to-face interviews on a small scale for the purpose of supporting some estimates of the bias introduced. No decision about that has been made.

RECOMMENDATION 4-5. The committee recommends that the Disability Evaluation Study be based on a design offering full coverage of the U.S. household population of adults. If resources are lacking to mount an area probability sample using face-to-face interviews, the Social Security Administration should use a multiple-frame design of a statistically optimum mix of the general population followed by face-to-face interviews of the eligible population.

Moreover, the cost effectiveness of various mixes of telephone and area probability frames for the screening phase should be investigated as part of the research, development, and testing phase of the study. For example, one possible dual frame design would use a frame of listed telephone numbers, and cluster the listed telephone sample geographically (to reduce the cost of face-to-face interviews). This design would supplement the telephone frame sample geographically with a clustered area probability sample to cover the unlisted telephone and the nontelephone households, as well as the listed telephone households. Thus, listed telephone households enter the sample in two ways; others enter only through the area frame sample. Appropriate statistical adjustments exist for combining the samples to obtain national estimates (Hartley, 1962; Lund, 1968; Lepkowski and Groves, 1986).

MEDICAL EXAMINATION

A general medical examination has been designated as a key component of the DES. According to the draft scope of work, the results of a general medical examination, combined with relevant "medical evidence of record" from health care providers, will be obtained for DES respondents. These two data sources will constitute the body of "medical, functional, and other pertinent data" from which "to make the most accurate possible prediction of a disability eligibility decision based on current program criteria" (SSA, 1996b, p.13). The committee assumes that the main purpose of the examination is to provide clinical input to establish a respondent's level of disability when determining program eligibility. The examination, however, may also serve as a validation standard for the data in the medical evidence of record collected from health care providers.

SSA has considered two options for such an examination: (1) employ local physicians, and (2) use a traveling team consisting of a physician, nurse, and mental health social worker with mobile examination centers of the type used in NHANES. SSA has determined that the mobile examination centers approach is the best way to obtain acceptable response rates and acceptable levels of quality and standardization. SSA further states that locating the mobile units in relatively close proximity to those people with severe impairments who will be examined, may yield an adequate response rate.

The amount of detail associated with the planned medical examination phase of the DES is not fully apparent from the draft scope of work, and the examination cannot be fully developed until procedures for the disability assessment in the DES are set. If the DES medical examination is to include an in-depth medical interview and performance test of all body systems, as well as a set of relevant laboratory and other diagnostic tests, the proposed examination would need to be administered by trained medical staff. This model for data collection would be similar to that of

NHANES. That survey uses traveling medical teams consisting of a group of physicians, nurses, and medical technicians who gather examination data in "stands" at mobile centers that move among sampled areas of the country (NCHS, 1994). If, on the other hand, the extent of medical information needed from the examination phase is far less (e.g., limited to administering simple clinical and performance tests), the NHANES model may not be cost effective. Current SSA thinking, therefore, leaves unresolved some of the key issues tied to this important and potentially very costly portion (in terms of dollars, time, and statistical precision) of DES.

Three fundamental questions arise in reaching closure on design aspects of the examination phase of DES: (1) what tests and diagnostic procedures will be included in the examination, with particular attention to the relationship between a person's claims of functional limitation and functional performance testing?, (2) what types of health care professionals will conduct the examinations?, and (3) in what setting will the examination take place? Because the answers to these questions depend on examination content (which is yet to be resolved and is also a later agenda item for the committee), the committee defers comment on the first question altogether, sets out some options for addressing the other two questions, and outlines a general strategy for reaching a final resolution for all three questions.

One aspect of the examiner issue is the mixture of professional skills required by the team as a whole. The mix of health care professionals needed to staff the examination teams would depend on the content of the examination. For instance, the more medically specialized the tests and procedures, the higher the level of professional training and experience required for the examination team. If, on the other hand, the examinations were to include a variety of procedures with varying ranges of professional competencies needed for their successful administration, the team might consist of persons with varying specialties (e.g., medical doctor, nurse, laboratory technician, and others).

An aspect related to the examiner issue is whether the examinations will be conducted by a few traveling teams or by separate teams of local health professionals for each sample area. The current DES plan adopts the NHANES model of traveling teams with mobile examination centers, the main argument being the potential for achieving improved data quality and examination response rates by having a group of professionals specifically trained for the DES examination. The SSA prefers the traveling team option over having local health professionals administer examinations at their own facilities. If there is a need for less comprehensive and specialized examination data than is required for NHANES, however, the plausibility of using local professionals for DES may be greater, though clearly the content of the examination must be determined before such decisions can be justified.

Several options have been raised regarding the setting for the examination. However, before making a final decision, each option needs to be carefully considered. True to the NHANES model currently adopted by SSA for the DES, the examination would be conducted in a mobile examination center which, like NHANES, would be a modified semitrailers or a recreational vehicle equipped to provide a suitable venue and facilities for all procedures and tests. At each sample location, examinations of area residents participating in the DES would be completed before the mobile center is moved to the next location.

An alternative to this option (tied to the use of local professionals) would have the examinations conducted at the locations where the professionals practice, thus decentralizing the examination process around the sample area. This alternative, with the accompanying need to

schedule visits by respondents at several locations to complete the examination, was rejected by SSA because of its presumed greater response and measurement problems.

A third option to consider is to have either a traveling or locally trained team set up an examination center in space that is rented or leased specifically for the DES. The advantage of this option is that like the mobile centers, the examination becomes a locally centralized process, thus affording greater quality monitoring and simplifying examination scheduling for study participants.

Because of the current state of uncertainty concerning the approach to determining the disability status of study participants, the committee strongly urges SSA to undertake a series of steps before final resolution of relevant design issues tied to the medical examination.

1. Specific components of the examination must be identified and details of their content developed based on the information needs tied to the medical and functional criteria needed to establish eligibility for disability benefits under (a) the existing decision process and (b) a redesigned process. Examination components may be limited and their contents general if medical information needs for determining disability status are limited, conversely, the number of components may be larger and their content broadly detailed if the medical information needs are great.
2. Once the examination content is set, the next step is to identify the appropriate knowledge, background, and skills needed by examiners to collect the required information. If performance tests and other medical aspects of the examination are relatively simple, nurses, physical therapists, or both (rather than physicians) could constitute the examination team. Where there is some uncertainty as to team composition, alternative team compositions might be set for later testing. Teams with appropriate knowledge, background and multidisciplinary skills including but not limited to occupational therapists, psychologists, social workers, and rehabilitation workers could be considered.
3. Lastly, field procedures should be developed that are sufficient for formal field testing, laboratory testing, or both of the various data collection options.

Illustrated below is an example of one possible menu of options for the DES examination:

Option	Team Composition	Team Origin	Examination setting
A	Physician + nurse + technician + psychiatric social worker	Traveling	Mobile center
B	Physician + nurse + technician + psychiatric social worker	Local	Team member's office
C	Physician + nurse + technician + psychiatric social worker	Traveling	Leased/rented space
D	Nurse + physical therapist + psychiatric social worker	Traveling	Leased/rented space

RECOMMENDATION 4-6. The committee recommends that once the options for using different combinations of team composition and origin, examination setting, and other dimensions are sufficiently set for assessments, a formal field experiment should be performed during the research, development, and testing phase of the survey to determine the validity and reproducibility of these options as well as the most cost- effective approach to meeting the objectives of the survey.

This methods-testing activity should include a comparison of both the relative operational efficiency by measuring productivity and unit costs (including the costs and time involved in building and equipping these units, moving them, renting space for them, bringing in electricity and sewer at each location, dismantling and putting them back together, and calibrating the equipment at each site), as well as the statistical merit of all options (e.g., measures like refusal rates, inter-rarer reliability, response validity, the cost impact on the number of sample PSUs, the impact on response rates resulting from the short period of time the centers can remain in each PSU). Also, studies are needed to address measurement issues, such as assessing functional status, quality of medical evidence of record, quality and validity of examinations by the various teams, and other similar issues. These tests call for repeated measurements and will need to be iterative to be able to test the validity and reproducibility of measurements. Taken as a whole, these validity studies as well as operational and statistical comparison measures would provide a more sound basis for setting the final approach to completing this important part of DES.

MISCELLANEOUS DATA QUALITY AND PROCEDURAL ISSUES

The committee notes the following important procedural issues that have been considered by SSA for the DES but not fully resolved: (a) the use of incentives to encourage study participation, (b) offering a home examination to those who are unable to travel to an examination site, and (c) assessing the quality of medical evidence of record data to support a person's claim of functional limitation. Various potentially significant sources of error in DES estimates are affected by these procedures. For example, nonresponse rates and bias are likely to be reduced by using monetary incentives (Lessler and Kalsbeek, 1992, p. 168) and offering a home examination. However, measurement error may be adversely affected by offering the home examination or relying on provider data for the medical evidence of record (Cohen and Carlson, 1994).

Although SSA appears to have given considerable thought to each of these procedural matters, none was sufficiently resolved in the draft scope of work to provide a useful response to the RFP. SSA asks the bidders to the RFP for a discussion of options for dealing with these issues and a statement of the rationale for the choice that is put forth, but it provides no solid basis for the contractor to come to a final resolution. Without explicit directives to the bidders concerning the means and method, the best possible proposals will not be forthcoming. To ensure a degree of rigor to the resolution of these three issues,

RECOMMENDATION 4-7. The committee recommends that the Social Security Administration require in the scope of work a rigorously designed experiment in the field testing and development phase of the survey to identify mechanisms for enhancing participation in the Disability Evaluation Study, to guide decisions on the use of a home examination for those unable to travel to an examination site, to establish the validity of the measures obtained, and to assess the quality of the medical evidence of record.

SSA should explicitly spell out the methods or procedures it wishes to have considered. Decisions on these design questions then can be applied to the plan for the full national survey to be conducted after the development and testing phase is completed. The committee urges SSA to incorporate testing for the three sample and data quality issues in the above recommendation and the examination issues in Recommendation 4-6 into the same field test.

Recommendation 4-7 implies the need for the field tests to be of sufficient size to allow testing along several dimensions of study. An illustrative example of options that might be considered for each of the incentive and home examination quality issues listed above follows:

Option	Home Examination Offered?	Incentive Offered?
1	No	No
2	No	Yes
3	Yes	No
4	Yes	Yes

Furthermore, suppose that an assessment of the quality of the medical evidence of record data were to be done through an adjudication process involving medical information from the medical examination, the respondent's medical history, and all health care providers visited. The DES examination options (A to D on page 23) combined with the options for incentives and home examination would imply the need for 16 (4 x 4) cells in a full factorial design. Each cell requiring say, 50 households would imply a sample of 800 (16 x 50) for the field test.

The committee recognizes that the rigorous field testing it proposes will significantly add to the cost of the research, testing, and development phase of DES, however, it is convinced that the benefit to the planning of the critical national survey will be substantial and ultimately may even reduce total costs by using the most cost-efficient and valid methods in the large national survey. Without this testing, the unique context of the DES leaves SSA with few solid sources in the methods literature for sound decisions on these and other important survey design questions.

The committee assumes that SSA has plans for advance arrangements and public awareness campaigns in the local areas to promote support and increase response rates prior to interviewing and administering medical examinations. This type of activity is commonly done for large surveys and is often vital in increasing response rates and local goodwill.

Because of the complexity and sensitive nature of this survey, SSA should pay special attention to advance arrangements. The plan of work should clearly delineate responsibilities of SSA and the contractor. Advance activities over a span of several months could include notifying

local officials, selecting field offices, sites for medical examinations, and gaining support from media and other contacts. Introductory letters and fact sheets are usually sent to local officials including local health directors and key staff, county executives, police, representatives of local advocacy organizations, and provider organizations. Meetings can be held with local officials and advocate groups to explain the purpose and procedures. Written endorsements are usually helpful in allaying fears and increasing participation. In fact, the presence of an advocate for persons with disabilities with the examination team may aid in increasing participation in medical examinations. Some examples of such campaigns prior to a survey are NHANES III conducted by NCHS and the decennial census conducted by the Bureau of the Census.

DATA LINKAGE WITH SSA FILES AND CONFIDENTIALITY ISSUES

Matching data files on individuals is of increasing value to analysts of survey data because it permits inexpensive expansion of the data record. In recent years concerns about the ethics of matching data on human subjects have become prominent in research circles (Duncan, et al., 1993). Several principles appear to be emerging from these concerns: (a) persons whose identifiable data records are matched should be informed about the matching as well as the content of the medical examination and several other items in the survey instruments; (b) the purposes of the matching must be described to the respondents; (c) permission from the respondents must be obtained for the matching; (d) nonstatistical uses of the data must not be threatened by the match; (e) and participation must be voluntary. Although such procedures are common for health surveys, in a survey such as DES there are added complications of obtaining informed consent from persons with mental and cognitive impairments who may not be competent to give their informed consent.

In addition, beneficiaries and potential beneficiaries may not want to respond to certain questions or participate in certain tests unless confidentiality is guaranteed because of a perceived threat to their receipt of benefits. Assurance of confidentiality would require the contractors to strip identifiable information from the files as soon as it is no longer needed for survey purposes. However, such a procedure does not allow SSA to link the data with other SSA files.

Laws that control access to administrative records, such as reports of earnings covered by Social Security, restrict their use for statistical purposes. In seeking to match individual cases in DES through a Social Security Number to the longitudinal earnings and other administrative files and to third parties (e.g., health insurance providers, physicians, employers, or pension plans), investigators can track longitudinally the work earnings of the sample. This would be a great benefit to the analytical value of the data. SSA views such a linkage as important to the success of the current DES and possible future extensions to the DES sample for measuring changes over time. The draft scope of work states that the contractor will turn over identifiable individual records to SSA, and that SSA will carry out the matching. It further requires the contractor to develop appropriate language for a waiver of confidentiality, reflecting the specific commitments to be made at the time of the interview. Although SSA recognizes these confidentiality issues, the committee believes that additional steps must be taken by SSA to obtain adequate informed consent from the participants, to ensure sufficient protection of confidentiality for study subjects, and to deal with these issues in a way that will not compromise the scientific integrity of the DES design. Current beneficiaries and the group potentially eligible for benefits particularly may have

concerns about participating. The impact of such waivers on the response rates and the resultant bias introduced by refusals also should be taken into consideration and evaluated during the research and testing phase of the survey.

The presence in a regulatory environment of statistical databases linked to earnings and other administrative files is precarious. It is not clear whether individuals in the sample would have legal protections from regulatory uses of their survey data when they are matched to the longitudinal earnings file. But they deserve such protection.

There are several accepted models of matching that remove from the regulatory agency direct access to the information that would allow actions to be taken against individuals. SSA is urged to review the procedures used in the Health and Retirement Survey sponsored by the National Institute on Aging and other relevant surveys to protect the respondents' confidentiality and yet permit linkage with other data files. For example, the matching could be performed by the contractor, and a file could be returned to SSA stripped of all identifiers. The contractor also could update the match over the years. Other arrangements to guarantee that matched files cannot be abused deserve attention.

RECOMMENDATION 4-8. The committee recommends that the Social Security Administration enhance its safeguards for matched data according to accepted practices by employing procedures used in recent federal surveys, and that it take into consideration the effect of such procedures on response rates.

5

Concluding Comments

The committee and SSA agree on the very important role that DES will play in SSA's long-term disability research in understanding the growth of the disability programs. The survey, if well designed, could also help SSA to develop and evaluate its disability decision redesign and address future policy questions. The committee, however, has strong reservations about the current plan for the DES. It generally believes that a more focused effort should be devoted to preparing a single, larger national survey with one large sample and one protocol, instead of two half samples threatened by mixed measurement methods and other nonequivalencies. The committee understands that SSA has started to address some of the issues that were discussed previously during committee meetings and that are addressed in this report. The committee, however, continues to be concerned about the adequacy of the technical advisory structure for the survey, and SSA's ability to get the necessary definitional and feasibility research done in time to mount the survey. It therefore urges SSA to establish a formal technical advisory body for the purpose of oversight of the planning and conduct of the survey.

As previously stated, the committee intends to comment further on the DES in subsequent reports. In its final report, the committee also intends to address issues surrounding the long-term analysis of and future directions in needed disability decision research. This report addresses only the general features of the survey design and some data collection and related issues described in SSA's draft scope of work. These decisions are critical to the conduct of a cost-effective, statistically sound DES. The committee is issuing this brief report at this time because it recognizes the agency's need to make immediate decisions about these topics in order to issue an RFP and award a contract for the DES in a timely manner.

The text of the committee's recommendations follows, keyed to the chapter in which they appear in the body of the report.

RECOMMENDATIONS

RECOMMENDATION 3-1. The committee strongly endorses the conduct by the Social Security Administration of a well-designed, carefully pretested, and statistically sound Disability Evaluation Study.

RECOMMENDATION 3-2. The committee recommends that the current stage 1 and pilot study be merged, expanded, and extended into a research, development, and testing phase of the survey with application to samples of the type that are more traditionally used in methods testing. Only when the development and refinement of the functional assessment instruments, survey operations, and other issues are tested and resolved should a national sample survey be launched using a single protocol.

RECOMMENDATION 3-3. The committee recommends that the national survey should be conducted with one sample large enough to estimate the sizes of the population at risk with acceptable levels of statistical precision.

RECOMMENDATION 3-4. The committee recommends that the Social Security Administration use relevant data from the national Health Interview Survey Disability Supplement, National Health and Nutrition Examination Survey, Survey of Income and Program Participation, and other relevant surveys to assist in developing the sample design, survey operation, and questionnaire content for the Disability Evaluation Study.

RECOMMENDATION 4-5. The committee recommends that the Disability Evaluation Study be based on a design offering full coverage of the U.S. household population of adults. If resources are lacking to mount an area probability sample using face-to-face interview, the Social Security Administration should use a multiple frame design of a statistically optimum mix of general population followed by face-to-face interviews of the eligible population.

RECOMMENDATION 4-6. The committee recommends that once the options for using different combinations of team composition and origin, examination setting, and other dimensions are sufficiently set for assessments, a formal field experiment should be performed during the research, development, and testing phase of the survey to determine the validity and reproducibility of these options as well as the most cost-effective approach to meeting the objectives of the survey.

RECOMMENDATION 4-7. The committee recommends that the Social Security Administration require in the scope of work a rigorously designed experiment in the field testing and development phase of the survey to identify mechanisms for enhancing participation in the Disability Evaluation Study, to guide decisions on the use of home examination for those unable to travel to an examination site, to establish the validity of the measures obtained, and to assess the quality of the medical evidence of record.

RECOMMENDATION 4-8. The committee recommends that the Social Security Administration enhance the safeguards of matched data according to accepted practices by employing procedures used in recent federal surveys and that it take into consideration the effect of such procedures on response rates.

References

Adler, M. Federal Disability Programs. *Encyclopedia of Financial Gerontology*, 1996.

Cohen, S.B., and Carlson, B.L. Comparison of Household and Medical Provider Reporting Expenditures in the 1987 Medical Expenditures Survey. *Journal of Official Statistics, Statistics Sweden* 10(1):3-29, 1994.

DHHS (U.S. Department of Health and Human Services). The Social Security Disability Insurance Program: An Analysis. Report of the Department of Health and Human Services, Washington, D.C., December 1992. (Unpublished.)

Duncan, G.T., Jabine, T.B., and deWolf, V.A. *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics.* Washington, D.C.: National Academy Press, 1993.

GAO (U.S. General Accounting Office). *Social Security Disability: SSA Quality Assurance Improvements Can Produce More Accurate Payments.* Report to the Chairman, Committee on Finance, U.S. Senate. Pub. No. GAO/HEHS-94-107. Washington, D.C.: GAO, June 1994.

GAO. *Social Security Disability: Management Action and Program Redesign Needed to Address Long-Standing Problems*. Testimony by Jane L. Ross to the Subcommittee on Social Security, Committee of the Ways and Means, House of Representatives. Pub. No. GAO/ HEHS-95-233. Washington, D.C.: GAO, August 3, 1995.

GAO. *SSA Disability Engineering: Project Magnitude and Complexity Impede Implementation*. Testimony by Diana S. Eisenstat to the Subcommittee on Social Security, Committee of the Ways and Means, House of Representatives. Pub. No. GAO/HEHS-96-211. Washington, D.C.: GAO, September 12, 1996.

Groves, R.M. *Survey Errors and Survey Costs.* New York: John Wiley, 1989.

Hartley, H.O. Multiple Frame Surveys. Washington, D.C.: American Statistical Association, *Proceedings of the Social Statistics Section,* 1962, pp. 203-206.

LaPlante, M.P. Disability Risks of Chronic Illnesses and Impairments. *Disability Statistics Report* (2). Washington, D.C.: National Institute on Disability and Rehabilitation Research, U.S. Department of Education, June 1991.

LaPlante, M.P., and Carlson, D. Disability in the United States: Prevalence and Causes, 1992. *Disability Statistics Report* (7). Washington, D.C.: National Institute on Disability and Rehabilitation Research, U.S. Department of Education, July 1996.

Lepkowski, J.M., and Groves, R.M. A Mean Squared Error Model for Dual Frame, Mixed Mode Survey Design. *Journal of the American Statistical Association* 81(396):930-937, 1986.

Lessler, J.T., and Kalsbeek, W.D. *Nonsampling Error in Surveys*, New York: John Wiley, 1992.

Lund, R.E. Estimates in Multiple Frame Surveys. Washington, D.C.: American Statistical Association, *Proceedings of the Social Statistics Section*, 1968, pp. 282-288.

McNeil, J.M. *Americans With Disabilities: 1991-1992. Data from the Survey of Income and Program Participation*. U.S. Bureau of the Census, Current Population Reports, P70-33. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, 1993.

NCHS (National Center for Health Statistics). Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94. *Vital and Health Statistics* 1(32). Hyattsville, MD.: NCHS, U.S. Department of Health and Human Services, 1994.

NCHS. Current Estimates from the National Health Interview Survey, 1994. *Vital and Health Statistics* 10 (193). Hyattsville, MD.: NCHS, U.S. Department of Health and Human Services, 1995.

Schuman, H., and Presser, S. *Questions and Answers in Attitude Surveys.* New York: Academic Press, 1981.

Smith, T.W. Context Effects in the General Survey. In Biemer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S., eds., *Measurement Errors in Surveys*. New York: John Wiley, 1991, pp. 57-71.

SSA (Social Security Administration). *Plan for a New Disability Claim Process* . Washington, D.C.: SSA, September 1994.

SSA. *Annual Statistical Supplement, 1996 to the Social Security Bulletin.* Washington, D.C.: SSA, September 1996a.

SSA. Draft Scope of Work: The Disability Evaluation Study. Washington, D.C.: SSA, July 30, 1996b. (Unpublished.)

SSA. Research Plan for the Development of a Redesigned Method of Evaluating Disability in Social Security Claims. *Federal Register* 61 (175):47542-47544, September 9, 1996c.

Tanur, J.M., ed. *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation, 1991.

Thornberry, O.T., and Massey, J.T. Trends in United States Telephone Coverage Across Time and Subgroups. In Groves R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., and Waksberg, J., eds., *Telephone Survey Methodology*. New York: John Wiley, 1988, pp. 25-49.

Turner, C.F., and Martin, E. *Surveying Subjective Phenomena*. New York: Russell Sage Foundation, 1984.

Westat, Inc. Including the Homeless and the Institutionalized Populations in the Medical Evaluation Study: A Discussion of Issues. Draft report submitted to the Social Security Administration, Office of Research and Statistics, Washington, D.C., January 1995a. (Unpublished.)

Westat, Inc. Planning and Developing a Protocol for the Medical Evaluation Study. Final Report to the Social Security Administration, Office of Research and Statistics, Washington, D.C., April 1995b. (Unpublished.)

Appendix

REVIEW OF THE SOCIAL SECURITY ADMINISTRATION'S DISABILITY DECISION PROCESS RESEARCH

Study Mandate

The study will review and provide advice on the scope of work, design, content of the survey, and the approach and scientific methods of completed and planned research as the Social Security Administration (SSA) develops the new disability decision process. The study will focus on the population 18-69 years of age. Although the committee is given latitude in setting its own agenda and designing its plan of work, the topics it explores will include:

- Review of the research plan and timeline for developing a new decision process for disability;
- Review of the preliminary design of the Disability Evaluation Study (DES) research efforts, the scope of work for the DES, and the design and content of the survey, as proposed by the survey contractor, as well as SSA's plans to integrate the decision method and DES research effort, identifying statistical design, methodological and content concerns, and other outstanding issues;
- Examine the results of completed research including research into existing functional assessment instruments and subsequently identified research for SSA's redesign efforts, and provide advice for adopting or developing functional assessment instruments or protocols for the redesigned disability process and the DES in particular; and
- Assess the results and findings of the research undertaken by SSA, comment on future research proposals, and offer advice on the analysis of the consequences of alternative disability determination processes. Some of the topic areas that might be considered include functional assessment of work-related limitations of physical and mental impairments, disability decision processes (including screening mechanisms); testing and validating decision processes for determining disability; and age, education, and work experience.

Acronyms and Abbreviations

ADL	activities of daily living
AIR	American Institutes for Research
The act	The Social Security Act
DES	Disability Evaluation Study
DHHS	U.S. Department of Health and Human Services
DOL	U.S. Department of Labor
GAO	U.S. General Accounting Office
IADL	instrumental activities of daily living
NCHS	National Center for Health Statistics
NHANES	National Health and Nutrition Examination Survey (NCHS)
NHIS	National Health Interview Survey (NCHS)
O*NET	Occupational Information Network (DOL)
PSU	primary sampling unit
RDD	random digit dialing
RFC	residual functional capacity
RFP	request for proposal
SGA	substantial gainful activity
SIPP	Survey of Income and Program Participation (U.S. Bureau of the Census)
SSA	Social Security Administration
SSDI	Social Security Disability Insurance
SSI	Social Security Insurance
VCU	Virginia Commonwealth University