DOCUMENT RESUME

ED 218 854 EC 143 116

AUTHOR Roistacher, Richard C.; And Others
TITLE Toward a Comprehensive Data System on the Demographic and Epidemiological Characteristics of the Handicapped Population. Final Report.
INSTITUTION Bureau of Social Science Research, Inc., Washington, D.C.
SPONS AGENCY National Inst. of Handicapped Research (ED), Washington, DC.
PUB DATE 21 Dec 81
GRANT 654AH00100
NOTE 128p.; Print is poor in parts.

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Disabilities; Elementary Secondary Education; *Incidence; *Information Systems; Management Information Systems; Program Development; Research Needs; *Statistical Data

ABSTRACT

The project was designed to investigate sources of available data on the incidence, prevalence, diagnostic conditions, functional limitations, and activity restrictions of the disabled population in the United States, and to design a comprehensive data system. Difficulties, both conceptual and practical in estimating incidence and prevalence are cited, including divergent data sources, varying definitions of disability and handicap, and inconsistent methodologies. Guidelines for extrapolating from existing data are given, including techniques of aggregation and merging, and linking files by entering distributions and summary figures into multidimensional tables. The report recommends steps to be taken by the National Institute of Handicapped Research to improve disability statistics, including advocacy for a uniform functional definition of disability and handicap and for inclusion of disability data in census and other national surveys. The development of a Disability Data Archiving and Analysis Center is detailed, and such a project's activities in nine categories are described: acquisition, cataloging, abstracting, reference, circulation, file processing, analysis, dissemination, and data improvement. Functional components as well as procedures, examples of user requests, and administration considerations are noted. Among five appendixes are papers on machine readable data files and a selected bibliography on the definition of measurement and disability. (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED218854

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it

Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy

Toward A Comprehensive Data System On the Demographic and Epidemiological Characteristics of the Handicapped Population

Final Report

Richard C. Roistacher
Engin Inel Holmstrom
Albert H. Cantril
John Terry Chase

A Research and Demonstration Project for the
National Institute of Handicapped Research
(Under Grant No. 654AH00100)

Bureau of Social Science Research, Inc.
1990 M Street, NW
Washington, DC 20036

December 21, 1981

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Bureau of Social Science Research Inc.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EC143116

EXECUTIVE SUMMARY

In November, 1980, the National Institute of Handicapped Research (NIHR) awarded the Bureau of Social Science Research (BSSR) a one-year grant to investigate the feasibility of establishing a comprehensive data system on the handicapped population of the United States. In the course of this work, BSSR staff members investigated the scope and content of over 300 machine-readable data files. We were able to provide NIHR with detailed descriptions of 42 major files and with summaries of 45 others. The initial purpose of our investigation was to determine which files could be used as the basis for a demographic information system. We concluded that in the absence of the Post-Census Disability Survey, no existing file could serve as the basis for an information system which would provide immediate answers to queries from a terminal. We further concluded that it is not possible to construct an omnibus "demographic master file" on the disabled population from existing files. We concluded that the goals implicit in the establishment of a comprehensive data system could best be pursued in four ways:

Data improvement. NIHR should develop a convincing argument to Congress and the Administration for implementation of a uniform functional definition of disability, and the inclusion of disability data in census and other national surveys. The basis of better information is better statistical data. At present, data on the disabled population are fragmented, incomplete, and incompatible. We recommend a set of statistical, data archiving, and technical policies to NIHR. The NIHR statistical policy should reconcile NIHR's statistical needs with the availability of data, thereby generating a set of data collection needs and priorities. The data archiving and access policy establishes machine-readable data files as bibliographic entities, much like publications and other information media. The policy is the basis for NIHR's participation in the federal "information economy" as a colleague of established statistical agencies.

A core multipurpose analysis agency. NIHR should become a core multipurpose analysis agency concerned with disability statistics.

DIDAAC. NIHR should establish a Disability Data Archive and Analysis Center (DIDAAC), a facility for answering statistical questions concerning the disabled population. DIDAAC would provide the best available answers through a comprehensive collection of files, advanced statistical and computing techniques, and serious scholarship by its staff. We provide a functional and organizational analysis of DIDAAC, some project cost and staffing levels, and a set of alternatives for its establishment in NIHR, in another federal agency, or as a contract with a non-government agency.

Demonstration project. NIHR should support a demonstration project to test DIDAAC's capability in making projections of national estimates from existing data sources.

ACKNOWLEDGMENTS

From the project's inception we realized that developing a plan for a comprehensive data system for the National Institute of Handicapped Research would require not only the best efforts of the project staff, but also the assistance of many others outside as well as inside the government.

First we would like to thank the members of the BSSR Advisory Committee:

Dr. Elizabeth M. Boggs, Hampton, New Jersey;

Mr. Paul Dziedzic, Chairperson, Governor's Committee on Employment of the Handicapped, Kirkland, Washington;

Dr. Byron Hamilton, Director of Research, Rehabilitation Institute of Chicago;

Dr. Joseph B. Heil, Jr., Assistant Manager, Services for Disabled People, American Telephone and Telegraph Co., Parsippany, New Jersey;

Ms. Judy Heumann, Center for Independent Living, Berkely, California;

Mr. Charles W. Hoehne, Attorney, Consultant Services, Inc., Austin, Texas;

Dr. Paul Hoffman, Assistant Dean, School of Education, University of Wisconsin-Stout;

Professor Saad Nagi, The Ohio State University, Columbus, Ohio;

Mr. Peter Rosenstein, Executive Director, American Academy of Physicians Assistants;

Mr. Tom Stripling, Associate Research Director, Paralyzed Veterans of America;

Dr. Martha Redden, American Association for the Advancement of Science, Washington, D.C.;

Dr. Don Thurman, Assistant Commissioner, Division of Rehabilitative Services, Little Rock, Arkansas.

We would also like to thank the members of the NIHR Interagency Demographics Subcommittee, and the Members of the National Council. We owe a particular debt of gratitude

to Dr. Elizabeth M. Boggs and Dr. Lawrence Haber, both of whom gave most generously of their wise counsel, and to Tom Stripling whose comments on the final draft were extremely helpful. It is our hope that the National Council on the Handicapped will find this document of particular service as they seek to shape the policies of NIHR.

There are a number of individuals whom we wish to thank for their counsel: Professors Saad Nagi of Ohio State, Professor Don Galvin at Michigan State, Professor Marcus Fuhrer at Baylor University; Dr. Jack Noble and Dr. Helga Roth at the Department of Education; Dr. Joan Van Nostrand and Dr. Mary Grace Kovar at the National Center for Health Statistics, Evan Kemp, Director of the Disability Rights Center; and Margaret Brown of the Rehabilitation Indicators Project at New York University Medical Center.

We would also like to thank Dr. Nathan Ed Acree, our first project monitor at NIHR, and his successor as project monitor, Dr. Betty Jo Berland, and Susan Thompson and Gail Perry for their assistance.

Finally, among BSSR staff, we would like to thank Dr. Robert T. Bower who was involved in the development of the basic concept of this project. Our thanks also go to Bruce B. Dunning who was in charge of the early phases of cataloging of data sets, Tim Muzzio who completed the first draft, and Ann J Gumbinner, Don Ricketts, and Scott Kuster, who did much of the work on indexing and classifying data files. We owe a great debt to our librarians, Mary K. Hartz and Lucy Duff, for their diligent work.

Toward A Comprehensive Data System

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF APPENDICES

A. Technical considerations in the evaluation of machine-readable data files. (Richard C. Roistacher)

B. Technical standards for machine-readable data files submitted to NIHR. (Richard C. Roistacher)

C. A data access and archiving policy for the National Institute of Handicapped Research. (Richard C. Roistacher)

D. Data files on the demographic and epidemiological characteristics of handicapped people.

E. The definition of measurement and disability: A selected bibliography. (Lucy W. Duff, Mary K. Hartz, Bruce B. Dunning, and John Terry Chase)

Toward A Comprehensive Data System
On the Demographic and Epidemiological Characteristics
Of the Handicapped Population

Final Report

I. INTRODUCTION

There is presently an urgent and unmet need for accurate statistical information on the handicapped population of the nation. Currently, there may be as many as ten million Americans so severely disabled that they need constant assistance. Another 20 to 30 million Americans may suffer from some form of disabling or handicapping condition. Such numbers rise and fall as Americans enter and leave the disability stream. Increased medical knowledge and technological innovations, coupled with better services and legislation responsive to the needs of the handicapped, have made it possible for many people with disabilities to live productive lives.

Disability is here defined as a reduction in a major life activity resulting from an acute or chronic condition.[1] The severity of disability is defined by the degree of functional limitation a condition imposes on a person, and the extent to which it affects a major life activity. Clearly, the definition must begin with the pathology which causes an impairment. However, clinical conditions acquire significance only if they lead to a functional limitation of a major life activity such as employment, schooling, or housework. The intervening variables linking impairment and disability are the unique characteristics of the person and/or the physical, psychological, and socioeconomic environment in which the person acts. Disability in this sense is a relative and dynamic phenomenon. What is a disabling condition for one person may not be so for another. Further, changing circumstances, attitudes and aspirations, as well as rehabilitation, may reduce or alter the impact of a condition that was previously disabling.

---

[1] For a discussion of definitional and measurement problems, see John Terry Chase, Engin Inel Holmstrom, and Albert H. Cantril, Research Memorandum No. 1: An Approach to the Review of Data on the Demographic and Epidemiological Characteristics of the Handicapped for Purposes of Developing a Comprehensive Data System. Washington, D.C., Bureau of Social Science Research, 1981.

For the purpose of this paper, we do not differentiate between disability and handicap. The current use of the term "handicap" is tied to eligibility criteria, reflecting legal definitions written into legislation. However, there is in this country a growing interest in the adoption of terminology developed by Philip H. N. Wood for the World Health Organization. Wood defined "disability" as "any restriction or lack (resulting from impairment) of ability to perform an activity in the manner or within the range considered normal for a human being." He defined "handicap" as "a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfillment of a role that is normal (depending on age, sex, and social and cultural factors) for that individual."[2] Handicap, in this sense, becomes a social phenomenon, based on the interaction of the disabled person with his environment. Wood suggests that the term "disablement" be used as a collective descriptor, referring to any experience identified by the terms "impairment," "disability," and "handicap." Woods' suggestion is currently under review, and the problem of nomenclature and classification is yet to be resolved.

During the 1970's, the public became increasingly concerned with the plight of handicapped people. Congress passed many legislative acts, resulting in many new and enlarged programs for handicapped people. By 1978, over 80 different Federal programs provided services to handicapped people either directly or indirectly.[3] Many of these programs collect information, but each has a different legislative mandate and has developed different definitions of their service populations. No single Federal agency or mechanism, such as the Census, collects uniform and consistent information on the handicapped people of the nation. As a result, only estimates are available on the number of people with different disabilities, and relatively little is known about their demographic and clinical characteristics, their attitudes and aspirations, needs, and service experiences. National data on persons with disabilities are needed to make informed policy decisions, and to guide research to identify and serve better the handicapped population of the nation.

Most policy decisions are allocation decisions. They implicitly or explicitly state a set of goals, their relative importance, and the relative resources to be expended in pursuit of each goal. Many policy decisions on

---

[2] World Health Organization. International Classification of Impairments, Disabilities, and Handicaps. Geneva: World Health Organization, 1980:27-28.

[3] Unpublished interagency communication.

handicapped people concern allocations of resources. The prudence and equity of such decisions depend heavily on an understanding of the demography of handicapped persons.

One set of data essential to the equitable allocation of resources among the disabled pertains to the incidence and prevalence of various handicaps. How many people suffer from a particular condition at present? How many will incur a condition over the course of a year? Which conditions are "disabling" for persons with what kinds of demographic characteristics? Accurate estimates of carefully determined diagnostic conditions are needed, along with information on functional limitations and activity restrictions. What is the degree of functional limitation a condition imposes? What is the best method by which limitations can best be overcome by taking a particular action or set of actions? What are the demographic and background characteristics of people with disabilities, their geographic distribution, racial and ethnic background, educational and economic status, family situation, etc.?

Statistical data on the incidence and prevalence of various handicaps and their demographic correlates constitute only one kind of information essential to policy making. A second level of more complex and relatively subjective information is required to understand the disability phenomenon and to establish the mix of support services and programs that will be most successful in helping people to cope with or overcome particular conditions. We need to know about diagnostic condition(s) from which the disabled person suffers, and the functional limitations and activity restrictions with which he or she has to cope. Further, we must examine the social situation in which he finds himself. The attitudes and resources of the family members and those of the person himself play a great role in defining the situation and determining the types and extent of his service needs.

Both quantitative and qualitative information are required to assist in the allocation of resources and in the conduct of research essential for the development of better services. Further, such information helps to inform and shape public opinion by establishing a common ground of understanding about the size and character of the nation's handicapped population. At present there is no single source of such data. The available data sources suffer from various definitional and measurement problems, and the information that they provide is of dubious quality, lacking the accuracy required of national statistics.

The legislation establishing the National Institute of Handicapped Research (NIHR)was designed to provide the means for developing an adequate statistical base for policy. It specifically gives NIHR the mandate for:

> producing, in conjunction with the Department of Labor, the National Center for Health Statistics, the Bureau of the Census, the Social Security Administration, and other federal departments and agencies, as may be appropriate, statistical reports and studies on the employment, health, income, and other demographic characteristics of handicapped persons and disseminating such reports and studies to rehabilitation professionals and others to assist in the planning and evaluation of vocational and other rehabilitation services for the handicapped.[4]

It is against this backdrop that NIHR supported a one-year planning project at the Bureau of Social Science Research (BSSR). As specified in the program announcement under which the grant was awarded, "the purpose of the project is to design a comprehensive data system which incorporates data from a number of sources (systems) on the demographic and clinical characteristics of the disabled population." The program announcement described the purpose of the comprehensive data system as:

1. To identify the epidemiological patterns of disability and related problems, including background and identifying information, type of disability and functional limitation, (e.g., vocational, independent living), as well as specific types of physical or mental impairment.
2. To identify patterns of need and problem identification as perceived by the disabled population, including need for services and benefits and problems of health and restoration care, income, transportation, etc.
3. To identify the residual capabilities and strengths of the disabled, particularly in relation to vocational and social potentials and aspirations;
4. To identify the experience of disabled clients with service organizations, problems and needs concerning independent living, housing, legal matters, work, transportation, recreation, etc.

During the course of the project it became evident that there were very few, if any, national sources of information on needs, residual capacities and strengths, and service experiences of disabled populations. The emphasis of the project was shifted to identifying data sources that could

---

[4] Section 202(b)(8) of The Rehabilitation, Comprehensive Services and Developmental Disabilities Amendments of 1978 (P.L. 95-602).

provide national estimates of the incidence and prevalence of disability, as well as other identifying information on the demographic characteristics of disabled Americans.

This report is the summary document of the project. It is in four parts: In Chapter II, we highlight the difficulties involved in providing reliable estimates of the incidence and prevalence of disabilities in the United States, discuss the shortcomings of available data sources, and present a brief description of information needs and gaps in this field. In Chapter III, we suggest techniques for making national estimates from existing data. In Chapter IV, we present our recommendations to NIHR, including a discussion of statistical policies and coordination that NIHR should develop in this area. In Chapter V, we present a detailed description of the proposed data archiving and analysis center, including alternative institutional arrangements and other considerations relevant to its implementation.

Appended to the report are several planning documents in support of various aspects of the operation of a central statistical facility.

II. ESTIMATING INCIDENCE AND PREVALENCE

At present, we have no single source of adequate data on the incidence and prevalence of disabling conditions for the population of the United States. On the one hand, there is both underenumeration and overenumeration of disabling conditions because of multiple impairments, and because of variation in definitional thresholds that govern inclusion in or exclusion from programs. On the other hand, the statistics that do exist cannot be effectively aggregated because of inconsistencies in data sources and in methods employed in collection and representation. As a result, only gross national estimates are available, and only in selected instances can a demographic profile of those with particular impairments be obtained.[5]

This section of the report identifies the conceptual and practical problems associated with currently available data sources. This section discusses obstacles to obtaining national estimates from such sources, and briefly outlines major information gaps. The next section includes a discussion of techniques that might be used in reducing the uncertainty of estimates.

Conceptual And Practical Problems

How many handicapped people are there in the United States? According to the Social Security Administration's 1972 Survey of Disabled and Nondisabled Adults, 7.7 million adults of working age had functional limitations that can be defined as severely disabling. The same survey identified another 3.4 million with less severe occupational disabilities, and 4.3 million with secondary work limitations, raising the total estimate of work disability among adults to just over 15 million.[6] Four years later, the Bureau of the Census' Income and Education Survey reported 16.6 million people with some degree of work disability. It also reported 28.2 million persons with some form of activity restriction.[7]

---

[5] Rehab Group, Incorporated. Digest of Data on Persons With Disabilities. Washington, D.C. U.S. Government Printing Office, 1979.

[6] Donald T. Ferron. Disability Survey 72. Washington, D.C.: U.S. Government Printing Office, 1981, p. 9.

[7] Rehab Group, Inc. Op Cit, p. 16.

Estimates for work-disabled people have ranged from 12 million at the beginning of the 1970's[8] to 21 million in 1978.[9] Estimates of the prevalence of disability usually vary from 13 percent to 31 percent depending on the reference source used.[10] The Rehab Group, incorporated, has noted that there are over 41 million impairments in the United States, but because multiple impairments are common, the number of persons with impairments is not known.[11]

Given the disparate nature of these national estimates of the incidence of "handicaps," it is not surprising that detailed breakdowns of kinds of impairment or the demographic attributes of persons with one or another impairment pose exceedingly complex questions regarding sources of data, measurement techniques, and estimation procedures.

The ability to make national estimates and provide an overview of the demographic and epidemiological characteristics of the disabled population would be greatly facilitated by the existence of an integrated data file. Such a file would contain records on individuals linking data on medical conditions; impairments; limitations of function; and work, school, and self-care disability to data on income, service experience, service costs, etc. It would carry identifying information on each record which would allow aggregation into geographic and other functional units, and linking to other files based on those units. Such a file could be the basis for a "query system" in which those requesting information could retrieve demographic, cost, income, and benefit statistics in response to commands entered at the keyboard.

---

[8] Saad Z. Nagi, An Epidemiology of Disability Among Adults in the United States, The Milbank Memorial Fund Quarterly, 1976, Vol.54, No.4, Fall, 439-467.

[9] U. S. Department of Health and Human Services, Social Security Administration. Work Disability in the United States: A Chartbook. Washington, DC: U. S. Government Printing Office, 1980. (SSA Publication No. 13-11978.)

[10] R. A. Dudek and W. M. Marcy. "Public policy and social responsibility with regard to rehabilitation and maintenance of disabled persons." Technological Forecasting and Social Change, 1980, Vol.17, 61-72.

[11] Rehab Group, Inc. Op Cit, p. 2.

In the course of the project, we identified more than 300 machine-readable data files,[12] providing information on incidence and prevalence rates as well as needs, aspirations, and service experiences of disabled individuals. We undertook a major effort to collect background information on all of these files in order to conduct a methodological screening. However, such a task proved to be too time-consuming for a one-year planning grant. We were able to provide NIHR with detailed descriptions of 42 major files and with summaries of 45 others. Appendix D presents a listing of 274 data sources on which some background information was obtained. These data sources are organized according to their sampling design and sponsorship as: National-Federal, National non-Federal, Multi-regional, State, Local, and International.

Our investigation of these data sources showed that no existing file could fulfill the functions of an integrated file as initially envisioned. In lieu of such a single file, we explored the feasibility of constructing a "demographic master file" from existing statistical data on the disabled. Our aspiration was to ascertain whether such a master file could be constructed from other existing data files by record merging[13] and concatenation.[14] The file would identify disability in the most general functional sense (i. e., functional limitations restricting a major or related life activity) and would minimally contain information on sex, race/ethnicity, income, education, family status, and geographic residence.

We examined 29 relatively recent national or multi-regional surveys conducted since 1975 that were concerned with disability either directly or indirectly. Table 1 summarizes information provided by these surveys. Our review of these 29 data files showed that they offered little prospect for the creation of a "demographic master file." The major difficulty arises from the limited purpose of various surveys and the resulting inconsistencies or

---

[12] Machine-readable data files (MRDF) are sets of information generally stored on punched cards or on magnetic media such as tape or disk.

[13] Record merging is the construction of a record from two or more records containing information on the same unit of observation. For example, a single file could be constructed by merging records containing medical information on individuals with records containing income information on the same persons. Both the two original files and the merged file would have one record per person.

[14] Concatenation is the combining of two files end-to-end. The files must be in the same format.

Table I

NIHR INFORMATION NEEDS AND NATIONAL OR MULTIGEOGRAPHIC DATA SOURCES (1975-1981)

BSSR ID#	Survey Title	Agency	Population Covered[a]	Year of Data Collection[b]	NIHR's Information Needs																						
					Prevalence & Incidence Rates, Epidemiological Patterns & Characteristics of Disabled People																Residual Capacities/ Coping Mechanisms	Needs				Aspirations	Service Experience
					Definitional Items				Other Infor.			Demographic Information															
					Diagnostic Categories	Impairments	Work Limitations	Other FL/Act Restrict.	Age on Onset	Etiology	Secondary	Age	Sex	Race	Ethnicity	Education	Income	Employment Status	Marital Status	Household Status		ADL 4	Use of Aids	Med. Care	Other		
1020	Decennial Census	DC	All	1980			■	■				■	■		■	■	■	■	■	■							
1021	Current Population Survey	DC	C,NI	C			■					■	■	■	■	■	■	■	■	■							
~~1023~~	~~Survey of Income & Education~~	~~DC~~	~~R~~	~~A~~		■	■	■				■	■	■		■	■	■	■	■							
1085	Continuous Longitudinal Manpower Survey	DC	R	C			■					■	■	■		■	■										
1101	State-Fed Programs of Voc. Rehab.	DE	R	A	■	■	■	■		■	■	■	■	■	■	■	■	■	■	■							
1109	Nat. Spinal Cord Injury Data Bank	DE	R	C	■	■			■	■	■	■	■	■	■	■	■		■				■	■			
1046	Comp. Service Needs Study - Rejected Applicants	DE	R	1975	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■		■	■		■	■	
1047	Comp. Sec. Needs Study - Participants	DE	R	1975		■	■	■				■	■	■	■	■	■	■	■	■		■					
1058	Survey of Agencies Serving the Develop. Disabled	DE	R	1979	■							Aggregate															
1304	Children and Youth Referral Survey	DE	R	1980	■	■							Aggregate														
1001	Nat. Health Interview Survey	DHHS	C,NI	W	■	■	■	■				■	■	■				■	■	■							
1003	Health & Nutrition Exam. Survey	DHHS	C,NI	W	■	■	■	■				■	■	■		■	■	■	■	■							
1006	Nat. Survey of Personal Health Practices & Consequences	DHHS	20-64 Yr.	1979-1980	■							■	■	■	■	■	■	■	■								
1013	National Nursing Home Survey	DHHS	R	1977	■	■	■	■				■	■		■				■			■	■	■			
1014	Survey of Institutionalized Persons	DHHS	R	1976	■		■	■				■	■	■		■			■								
1015	Nat. Ambulatory Med. Care Survey	DHHS	R	A	■							■	■	■													
1016	National Natality Survey	DHHS	R	P	■								■	■													
1073	Nat. Med. Care Expenditure Survey	DHHS	C,NI	1977-1978	■		■	■	■	■		■	■	■		■	■	■	■	■							
1031	Longitudinal Sample of Disability Insurance Applicants	SSA	18-64, R	C	■	■			■	■	■	■	■	■		■		■	■								
1032	Continuous Disability History	SSA	R	A	■	■	■	■				■	■	■			■	■	■	■							
1033	Longitudinal Retirement History Study	SSA	Elderly	1969-1979			■	■				■	■				■	■	■	■							
1042	Survey of Disability & Work	SSA	C,N 18-64	1978			■	■				■	■	■			■	■	■	■							
1043	Survey of Blind Disabled Children Receiving SSI Benefits	SSA	1-18, R	1978	■	■		■	■	■	■	■	■	■	■		■			■							
1174	AFDC Recipient Characteristics Study	SSA	R	1967-1979		■	■			■	■	■	■	■		■	■	■	■	■							
1103	Longitudinal Supplemental Security Income File	SSA	R	1976	■	■	■	■				■	■	■		■	■	■	■	■							
2098	Annual Survey of Hearing Impaired Children	NonF	0-21 Yr,R	C	■	■		■	■	■	■	■	■	■									■				
2323	Birth Defects Information System	NonF	R	C	■			■	■	■	■	■	■				■					■					
2118	Developmental Disabilities Project on Residential Services & Community Adjustment	DHHS	R	1977-1980	■			■				■	■							■							
3108	RAND/HEW Health Insurance Study	HEW	R	1974-1984	■	■	■	■				■	■	■		■	■										

a/ C=Civilian
I= Institutionalized
NI=Non-Institutionalized
R=Restricted (e.g.Program Beneficiaries)

b/ C=Continuous
W=Weekly
A=Annual
P=Periodic
AH=AdHoc

incomparability among items used in defining disability. There is also some inconsistency in the representation of demographic information. A related problem is that of estimating national population parameters, since most of the surveys are not based on samples of the U. S. population, but represent specially defined subgroups within the population. For instance, out of 29 major files, Decennial Census is the only survey based on a national sample; sampling designs of five other surveys are based on the total civilian, non-institutionalized population of the United States. Most of the Social Security Administration surveys are limited to program beneficiaries; other surveys may only cover a limited age group, or people with specific problems, or in specific settings.

Three classes of problem confound the ready linkages of these files: (1) Divergent sources of data, (2) varying definitions of disability, and (3) inconsistent methodologies. Next, we explore the nature of each of these problems.

Divergent Sources Of Data

Demographic information on the handicapped population is gathered from four sources: service eligibility statistics, service delivery statistics, population surveys, and ad hoc studies.

Service eligibility statistics. Service eligibility statistics are generated by applicants to agencies concerned with providing for handicapped and disabled persons. In order to have any data trace at all, the person must be brought into contact with the agency through his or her own efforts, the agency's efforts, or the efforts of some third party. During the initial period of contact, the person's characteristics that are of interest to the agency are recorded and a decision made as to whether the person is eligible to be a client. If the person is accepted as a client, then a complete set of personal data will be obtained. If the person is rejected, then the data record may be left incomplete.

There are many problems inherent in service eligibility statistics. First, a person applying to several agencies will generate several eligibility records and may thus appear more than once in a file constructed by concatenating eligibility records. Second, the quality of service eligibility statistics suffer from lack of standardization in both decision making and record keeping processes. For example the quality of the RSA-300 data collected by state agencies for state-federal programs of vocational rehabilitation varies greatly from state to state, and reduces the comparability of merged state data within the national RSA-300 system. While service eligibility statistics can provide much valuable information about some characteristics of persons with disabilities, it is unlikely that such records can provide national estimates of the incidence and prevalence of disability.

Service delivery data. Once a client has been enrolled, the agency must maintain records on what services have been provided. Service delivery data present many problems for making national projections. The problem of overcount may be less significant than is the case for eligibility statistics, but the quality of data suffers from lack of standardization, and the population covered includes service beneficiaries only. No information is collected on others who may be qualified but do not participate in the program for one reason or another, nor on others who are not eligible for such services. Examples of service delivery data are those collected by the Social Security Administration and the Veteran's Administration.

Sample surveys. A number of surveys sampling the U.S. population have collected data on disabilities, either directly or indirectly. The Decennial Census and the National Health Interview Survey are examples of such efforts. Most of these surveys present some problems. For example, the Decennial Census asks limited questions regarding disability. In the 1980 Decennial Census, only two very broad questions were asked concerning people with "any activity restrictions" and public transportation limitations. The 1980 Post-Census Survey that would have provided a rich and accurate source of disability information was cancelled after pretesting.

Short of such targeted census-type surveys, the cost of teasing disability data from other national surveys may be prohibitive. The weekly National Health Interview Survey (HIS) is rich with diagnostic information; however, unless cross-tabulated with various other questions concerning disability days, restricted-activity days, etc., such data might exaggerate the extent of disability within the U.S. population. The HIS was recently redesigned in an effort to alleviate such over reporting. The new format includes detailed information on conditions resulting in activity restrictions and functional limitations. However, impending budget cuts may delay the implementation of this new format.

In general, surveys that provide detailed diagnostic information on respondents (usually targeted health surveys) do not provide adequate information on functional limitations or activity restrictions. Since the existence of a chronic condition cannot always be equated with disability, the results of such surveys must be carefully interpreted. On the other hand, most of the national studies that ask for major life activity restrictions or limitations do not attempt to determine the etiology of disabling conditions (see Table 1). As a consequence, the currently available national data sources provide either inaccurate or incomplete information on disabled persons.

Ad hoc studies. In addition to these data sources, there are occasional or ad hoc surveys, primary or secondary in nature, that are undertaken by demographers working in universities or national organizations, particularly those representing various disability groups (such as the National Association for the Deaf). There are also a rich array of state and local studies which, although not sampling national populations, provide descriptive information on various aspects of disability, such as attitudes of others toward people with disabilities and problems of disabled people related to housing, transportation, recreation, etc. (See Table 2 for example of such studies.)

Since most of these ad hoc studies are not based on national samples, they cannot be used for enumerative purposes but do provide some useful descriptive information on the characteristics and experiences of people with disabilities.

Varying Definitions Of "Disability" And "Handicap"

A consequence of the programmatic origin of most data on impairments and disabling conditions is that definitional issues tend to be resolved at the stage in which a person's eligibility for participation in a program is determined. Agencies providing services to the disabled quite legitimately pursue the legislative intent of their special mandate. Consider three definitions written into law:

1. The Social Security Act defines disability as "the inability to do any gainful activity by reason of any medically determinable physical or mental impairment which can be expected to last for a continuous period of not less than 12 months. To meet this definition, you must have a severe impairment, which makes you unable to do previous work or any other substantial gainful activity which exists within the national economy." (Section 404.1505) Further, applicable regulations stipulate: "If you are working and the work you are doing is substantial gainful activity, we will find that you are not disabled regardless of your medical condition or your age, education, and work experience."

2, The "504" regulations of the Department of Education and of the Department of Health and Human Services pertaining to civil rights matters stipulate: "'Handicapped persons' means any person who (i) has a physical or mental impairment which substantially limits one or more major life activities, (ii) has a record of such impairment; or (iii) is regarded as having such an impairment." In the regulations operationalizing this definition, an impairment is defined as: "any physiological disorder or condition, cosmetic disfigurement, or anatomical loss affecting one or more ... body systems." (CFR, Vol. 45, No. 92, p. 30937) Further, "has a record of such an impairment" can refer either to having a history of substantial limitation or having been "misclassified as having a mental or physical impairment that substantially limits one or more life activities." And, "is regarded as having an impairment" can refer to an impairment "which does not in fact substantially limit major life activities, but that is treated by a (government) recipient as constituting such a limitation."

3. The definition of the Rehabilitation Services Administration views a "handicapped individual" as someone who "(i) has a physical or mental disability which for such an individual constitutes or results in a substantial handicap to employment and (ii) can reasonably be expected to benefit in terms of employability from vocational services."

   The RSA regulations add the statement "or for whom an extended evaluation of rehabilitation potential is necessary for the purpose of determining whether he might benefit in terms of employability from the provision of vocational rehabilitation services." The vocational rehabilitation definition emphasizes the rehabilitative and employment potential of the handicapped person, and excludes those who do not show such potential.

Individual disposition and attitude are obviously critical in this determination. Additionally, the final decision regarding eligibility resides with the State vocational rehabilitation counselors.

Thus according to the Social Security Administration, one is disabled only if gainful employment has been precluded for a year by an impairment. The Rehabilitation Services Administration's definition factors in a person's disposition and motivation in determining reasonable expectations of the person's employability potential as a result of participation in vocational rehabilitation program. The Department of Education and the Department of Health and Human Services have the most liberal definition for civil rights purposes: Any physical or mental impairment substantially limiting "major life activities" is considered as a handicap. Further, section 504 of the Rehabilitation Act includes not only those now restricted by an impairment, but those who may be restricted by having had or having been regarded as having such an impairment. Haber states that "Since the language of the legislation goes beyond work and

VR services to include a broad array of social services, the population covered against potential discrimination may include infants, students, and the elderly, as well as those whose handicaps preclude them from employment."[15] Indeed, the regulations are unequivocal in including these age groups.

The practical consequences are twofold: different kinds of people fall within the administrative record systems of these agencies, and the kinds of information recorded about each person differ according to the agency's mission. Thus, while the administrative files of the vocational rehabilitation agencies do record the nature of the client's specific impairment (e.g., hearing loss), they will not provide a true demographic record of the incidence and prevalence of disabling conditions, as both most and least severe instances of handicaps are excluded from consideration.

Inconsistent Methodologies

Definitional questions almost always have implications for how measurements are made. At least three facets of the problem should be noted: what variables are measured; what questions are asked or criteria invoked; what units of measurement are employed? Some surveys rely on existence of a condition (based on diagnostic categories and/or impairments) and evidence of limitation in life activities. Others concentrate on limitations without identifying the specific physical or mental condition or conditions leading to disability. Thus, it is difficult to obtain accurate national estimates on the numbers of disabled people who have specific handicapping conditions or impairments. It is easier to estimate the number of people who have work limitations or other limitations. However, limitations are not always identified in a consistent manner. Some surveys simply ask: "Does your condition limit the amount or type of activity in which you engage?" Others ask detailed questions on limitations affecting areas ranging from transportation to daily living needs, recreation, etc. Further, surveys differ in their coverage of reasons or conditions that result in work limitations or other activity restrictions. Such variations make any aggregation or comparison of results very difficult.

---

[15] Lawrence D. Haber. "Disability Concepts: Implications for Program and Policy Development" in Definitions of Pathology, Impairment, Functional Limitation, Disability. Mary E. Switzer Memorial Series No.1. Washington, D.C., National Rehabilitation Procedures, 1980, p. 24.

Other methodological differences between surveys, such as variations in demographic questions, also make aggregation difficult. Some surveys ask for family or household income, while others inquire about personal income. Educational levels are reported in a variety of ways, which are not always compatible across files. As a result, there are serious gaps in our knowledge of disability and in our ability to make projections for handicapping conditions.

Information Needs

Information needs of the NIHR include data on:

a. The prevalence and incidence of handicapping conditions, including background and identifying information,

b. Patterns of need and problem identification,

c. Residual capabilities and strengths of the disabled, particularly in relation to vocational and social potentials and aspirations,

d. Experience of disabled clients with service organizations, problems, and needs concerning independent living, housing, work etc.

Minimally, background data on disabled persons should include information on geographic location, age, race/ethnicity, and family income. Knowledge of these characteristics is essential to an understanding of handicapping conditions, to needs assessment, and to the development of appropriate programs. Knowledge about the location of affected persons indicates where and through what political channels resources will be allocated. Information on the age and income of affected persons gives an idea of the length of time resources will be consumed in their support and on the degree to which disabled persons can contribute financially to their own relief. Such information is also useful in assessing the overall economic consequences of disability and the potential costs and benefits of averting it. Information on race and ethnicity addresses questions concerning the social equity of the distribution of government aid to the disabled. Such information can also pinpoint specific vulnerability, such as the higher susceptibility of blacks to high blood pressure, and of fair-skinned persons to skin cancer.

One of the paradoxes of data for national decision-making is that the data of interest to those with relatively broad responsibilities are often collected and maintained by those with specific programmatic responsibilities. The situation results in three types of information gaps for policy makers and researchers: Gaps in population covered; gaps in information covered, and gaps in disability categories covered. The last is perhaps the most serious, arising from the lack of consistent and systematic use of definitional terms and assessment methods.

Gaps In Population Covered

There are serious gaps in the population of disabled people covered by many of the national surveys. A truly comprehensive data system on handicapped people requires information on the full range of disability for people of all ages. A major information gap concerns those people with disabilities who do not fall within the domain of various programs, either because they fail to meet program eligibility criteria, do not know about the programs, or despite their limitations and disadvantaged situation, do not consider themselves handicapped and in need of services.

Another information gap concerns the elderly and the very young. Much of the data on the elderly come from surveys of institutionalized persons. Since most elderly people are not institutionalized, but live by themselves or with family members, such surveys provide only limited and biased information about the number of disabled elderly and conditions they face. Similarly, surveys on the very young tend to focus on restricted populations, such as those with congenital defects detectable at birth or those with hearing impairments. In recent years, much additional information has been developed on handicapped children of school age but the classifications are usually stated in educational terms and are not well correlated with etiology. More information is needed on the children and youth of the nation and on what constitutes handicapping conditions at various ages.

Gaps In Information Covered

With minor exceptions, none of the national surveys provide information on residual capacities and coping mechanisms, needs, or service experiences of handicapped people--all priority areas as stated in the NIHR grant announcement. Some notable exceptions are the National Survey of Transportation of Handicapped People, sponsored by the Department of Transportation; the Sheltered Workshop Study, sponsored by the Department of Labor; and the Comprehensive Service Needs Study, sponsored by the Department of Education.

Various local studies also provide descriptive information: e.g., the Kentucky Assessment of Vocational Education Needs of the Disadvantaged and Handicapped, the Wisconsin Survey of Paratransit System Serving the Elderly and Handicapped, and the Assessment of the Unmet Needs of the Aged Blind Population in New Jersey. Table 2 presents example of such state and local studies.

It is also difficult to aggregate the results of these studies, as each varies in the ways in which essential demographic information such as age or income, is asked. Finally, very few national studies collect information on etiology, age on onset, and secondary conditions, although such information is essential to understanding and tracing the dynamic nature of handicapping conditions.

Gaps In Disability Categories Covered

The multiplicity of classificatory schemes and the lack of standardized criteria for the assessment of disability, coupled with the programmatic focus of many national data-collection agencies have resulted in data sources that are inconsistent, incomparable, and riddled with problems. While some surveys provide information on diagnostic categories, with insufficient reference to any accompanying activity restrictions, others concentrate on functional limitations with insufficient information on the etiology of such conditions. Currently, no national studies provide information on all conditions, and on the type and extent of functional limitation that such conditions impose. An excellent source for such information is a well-designed census or national sample survey. The population of interest in gathering basic data on the incidence and prevalence of handicapping conditions is that of all persons living in the United States. It is known that the kinds and degrees of disability vary with geographic location. It is apparent that reliable results require, at a minimum, a sophisticated sampling design using samples of adequate size and a cadre of well-trained interviewers. Such efforts are costly. If they are to be cost-effective, such surveys must be planned with a view to the multiple use of their results. Such planning requires a national perspective and the kind of consideration that only national authority can provide. Unfortunately, no agency concerned with the handicapped, including the NIHR, has any mandate to survey this population. It is important to realize that NIHR's mandate to produce reports in conjunction with other agencies that routinely collect statistical data is not the same as having the mandate to survey the disabled population of the United States.

The definition of "handicapped" is currently tied to eligibility criteria. It is defined in terms of one's entitlement to a program or service. Yet disability should be approached as a real personal characteristic of human performance and activity, and as such, should serve as an important indicator of the health status of the nation. The process of identifying handicapped persons and handicapping conditions should be separated from the process of

Table 2

EXAMPLES OF STATE & LOCAL STUDIES WITH INFORMATION ON THE HANDICAPPED POPULATION[a]

Prevalence & Incidence Rates & Characteristics	Needs	Aspirations	Service Experiences	Attitudes of Others
Registry of Deaf-Blind Children (AK) DES-DDDMRS Client Information System (AZ) Crippled Children's Services Program (AR) California Disability Survey (CA) Case Reporting of RH(D) Hemolytic Disease (CA) Reporting Systems for Handicapped Students (CO) Developmental Disabilities Information System (IL) Survey of Severely Disabled in Iowa (IA) Louisiana Survey of Hearing Impaired and Deaf-Blind Children and Youth (LA) Massachusetts Mental Retardation Service Coordination Battery (MA) Services for Children and Handicaps Management Information System (MN) Survey of Sensory Impaired Children and Youth (NE) Developmental Disabilities Information Survey (NY) 1977 Telephone Survey on Developmental Disabilities (OH) 1974-75 Follow-up of the Physically Handicapped (PA) Patient-Client Information System (PA) Brown University Health Surveillance Studies (RI) Handicapped (Children) Information Center (TX)	Services for Handicapped Children and Youth (IA) Adult Restorative Services (KS) Kentucky Assessment of Vocational Education Needs of the Disabled and Handicapped (KY) 1976 Needs Assessment Study (Mass. Commission for the Blind - MA) Minnesota Disability Assessment (MN) Survey of Handicapped Individual's Housing Preferences (MN) Survey of Paratransit Systems Serving the Elderly and Handicapped (WI) Assessment of Vocational Needs of Handicapped Students (WI)	Postsecondary Educational Status of Physically Disabled Adults (PA)	Hot Springs Rehabilitation Center Follow-up of Former Rehabilitation Clients (AR) UCLA Intervention Program for Developmentally Handicapped Infants & Young Children (CA) Household Survey on Effectiveness of Health Services (TN) RSA-300 Series	Pennhurst Telephone Survey on the Attitudes Toward the Mentally Retarded (PA)
San Diego County Teratogen Registry (CA) Medical and Allied Health Services Delivery System for Substantially Handicapped Developmentally Disabled Adults (Chicago - IL) Rochester Project (MN) 1979 Survey of Disability (Houston, TX) Statistics on Spinal Cord Injury Patients (TX)	Western Arkansas Study on Aging (AR) Aging Needs Assessment (CA) Transportation Needs of the Mobility Limited (ILL) Health, Income & Related Problems of the Elderly (KT) Howard County Handicapped Needs Assessment (MD) Transportation Needs of the Handicapped (Boston, MA) The Elderly in the Inner City (NYC)	Gerontology Retirement Study (CA) The Job Hunt of the Disabled (NY)		Survey of Children's Attitudes Toward Handicapped Peers (HI) Cook County Survey of Teacher Attitudes Toward Handicapped Children (IL) Survey of Landlord Attitudes Toward Housing for Mentally Retarded Persons (NYC)

[a] No data sets identified discribing residual capacities & coping mechanisms.

determining service eligibility. If disability measures are to be used as national indicators of health status, such measures must be uniform in definition and measurement. Data collection activities should be based on census-type national samples.

The 1980 Post-Census Disability Survey was designed with these considerations in mind. It has been pretested and offered great promise of producing the kind of demographic master file that could be employed in addressing general questions of incidence and prevalence. The Post-Census Disability Survey would also have served to answer many specific inquiries regarding the attributes of those with particular impairments. For budgetary reasons, however, the Survey was not conducted. Additionally, in early 1981, eight other surveys were either in the planning or developmental stage. At the time of this writing the surveys and their statuses are:

1. The Survey of Impaired Individuals in Households (DHHS), reduced in scope;
2. The Medical Examination Study of Extent of Impairment and Potential SSDI Eligibles (SSA), funded for 1982;
3. The Work Incentive Experiments (ORS/SSA), cut down and to be done in-house;
4. The Framingham Study Disability Analysis (ORS/SSA), to consist primarily of data archiving;
5. Alternative Practices in Disability Beneficiary Rehabilitation (SSA), cancelled;
6. The Survey of Income and Program Participation (ORS/SSA), cancelled;
7. The Survey of Individuals in Residential Institutions (DHHS), cancelled;
8. A Feasibility Assessment for a Longitudinal Survey of the Disabled (SSA), cancelled.
9. The Survey of Impaired Individuals in Households (DHHS), cancelled.

Further, the newly reformatted National Health Interview Survey (HIS) might also not be implemented. The new format is designed to provide much needed detailed information on activity restrictions and functional limitations of various impairments. It also places a greater emphasis than before on determining disability in those 60 or over and those between the ages of 5 and 17, both of which are relatively under-surveyed populations.

However, budget cuts threaten HIS as well as other well-established national data systems such as RSA-300.

This section discussed the problems associated with currently available data sources on disabled persons. Ideally, the most accurate source of national information on disabled persons and on the incidence and prevalence rates can be obtained from a census-type survey targeted to identifying and describing disability. Since there are no such surveys and the likelihood of undertaking one in the near future is almost nil, we now turn our attention to teasing estimates and demographic profiles of disability from existing data sources.

III. TECHNIQUES FOR THE USE OF EXISTING DATA

Throughout the course of this project, we have been driven by three objectives: identification and description of data files, consideration of how best to estimate incidence and prevalance, and design of a "comprehensive data system." The task of actually estimating incidence and prevalence was beyond the scope of our project. We have, however, outlined the procedures that could be employed in such an exercise.

General Guidelines

Two general rules should be followed while attempting to make national projections from available data sources on the disabled: First, the focus of inquiry must be well formulated; second, acceptable levels of confidence must be precisely stated.

Narrowing The Focus

To achieve an economy of effort in pursuit of estimates of incidence and prevalence, the focus of the inquiry must be precisely formulated. Not only is there a diversity of the populations subsumed in various data files, but data collection efforts have rested on varying definitions of what conditions should be included and how they are described.

A machine-readable data file's content can be described in terms similar to those describing any other information medium. In the case of numeric data files, however, the description of content is usually more specific than is the case for print materials. A machine-readable data file either does or does not contain income figures for residents of Nevada. It either does or does not contain people's age. The file may contain information which could be used to estimate such figures, or may contain keys which would allow it to be merged with a file containing such figures. A machine-readable data file, however, cannot allude to, imply, or metaphorically describe things. Thus, a machine-readable data file is usually either highly relevant or completely irrelevant to answering a particular question. Such specificity in a numeric file also implies that a description of the file's content is unequivocal. The variables, their unit of measure, and the population which they describe are as they are, regardless of the purposes for which the information is used.

In attempting to answer the question about the incidence and prevalence of disability, data files of interest are those of national scope and have information on disabling conditions as well as geographic location, age, income, and race of respondent. Given the definitional problems, it is likely that multiple definitions of disability will be required, and that the definitions of interest to the secondary users are not necessarily those of interest to the data's producer.

The initial question to be asked about each source of data is its relevance to the question we want to ask and the population about which we want to ask the question. We must know five things about each source of data:

1. What is its domain of discourse? e.g., does it contain a variable which indicates disability? Does it contain information on income?
2. How is information of interest expressed? In what units of measure?
3. What populations can be estimated from the available information? e.g., can we use a file based on a national sample to estimate disabling conditions in a certain region or disability in people of a certain age range?
4. What results are obtained from the file? What does it say about disabilities in the population of interest to us?
5. What confidence can we express in these results? To what degree are they uncertain? To what degree are they biased?

Statistical Goals

In addition to errors of coverage, definition, and data collection technique, we must attempt to cope with two distinctly statistical problems, uncertainty and bias. Given the present state of data on the disabled, it does not seem wise to make our primary goal that of maximizing the precision of estimates. At this stage, we should try to minimize the uncertainty of our estimates and to attempt to correct for bias.

Uncertainty is the more tractable problem. Suppose that it is estimated that 30,000 people have a particular disabling condition. In fact, this number is not meant to be exact, but is the best estimate of the actual number. If pressed more closely, we would say that although 30,000 is our best guess, we are 90 percent sure that the actual number of people with the condition is between 29,500 and 30,500. If the data were unbiased but had a lower degree of certainty, we would still specify that our best guess was 30,000. With the less certain data, we might state that we were 90 percent sure that there were between 29,000 and 31,000 affected people. Thus, minimizing uncertainty decreases the size of the "confidence interval" which surrounds a statistical statement.

If the data were biased, rather than uncertain, then our best guess would deviate from 30,000, even though the confidence interval around that estimate might be unchanged. Bias might occur because some data are omitted in an unknown but systematic fashion. Errors might be introduced as a result of interview bias or self-reporting bias. While, uncertainty produces fuzzy results, bias produces incorrect results. Our task is to state the degree of uncertainty surrounding our results and to indicate, if possible, the magnitude and direction of any bias.

Finally, we must give some idea of the sensitivity of our conclusions to our assumptions. Suppose that there were several common definitions of our hypothetical disabling condition. It is statistically prudent to do estimates of prevalence using more than one definition of the condition. If the estimates varied by only a few percentage points, then we could describe our estimate of prevalence as "robust." If, however, prevalence under one definition varied "greatly" from prevalence under the other definition, then the statistical presentation would require more qualification.

The most salient finding from our review of current estimates of the disabled population of the United States was the wide variation in such estimates. It is not surprising that the estimates vary widely, since they are not based on comparable assumptions. Our conclusion is that a satisfactory estimate of the number of disabled persons in the United States can be performed only if the following conditions are satisfied:

1. The data are from a survey which yields unbiased estimates of the U.S. population;
2. The survey records information on impairments, functional limitations, and vocational and activity limitations;
3. That estimates of the number of disabled persons be with respect to specific definitions of disability;

4. That analyses be based on microdata[14] rather than on tables or aggregates.

Techniques For Estimation

A number of techniques may prove useful for extending the utility of available data. The efficacy of a particular technique must be assessed in terms of the particular question at issue, the data available, the skill of the analyst in making the estimate, and the skill of the client in interpreting and using it. The following discussion of techniques is thus neither authoritative nor exhaustive.

Imputation On A National Sample

It is sometimes possible to infer information where it is missing from a record, or even from a file. For example, suppose we were estimating the number of people suffering from each of two disabling conditions. Suppose further that the best available data were in two files, one for each condition and that each survey asked about only one disabling condition. One problem is that persons suffering from both conditions would be counted in each file, but that the double counting would not be directly measurable. One way of correcting such double counting would be to locate studies of the co-occurrence of the two conditions by age, sex, occupation, etc. The results of such studies could be used to correct the overall estimates by adjusting the counts for persons by age, sex, occupation, etc. Similar "imputation" techniques used within a file can sometimes overcome problems of missing data.

Adjustment Of Sample Weights

It may sometimes be possible to use files based on one population to derive estimates for another population when the differences in the two populations' composition are known. By the use of such weighting, a file of microdata with one age distribution might be used to derive estimates for a population with a different age distribution. Records of people with under-represented ages would be multiplied by a weight factor to increase their representation. Records

---

[14] Microdata are records based on individual units of observation. A microdata record might be based on a person, case, service contact, etc. Conversely, an aggregate data record always contains a count of the number of units described by the record. Microdata records have an implicit count of one.

of people with an over-represented age would be weighted so as to decrease their contribution to the estimate. The efficacy of such weighting depends on our knowledge of the differences in the compositions of the populations and our confidence that other sampling conditions are "equal."

Aggregation And Merging

Very often files which cannot be merged on a record-by-record basis at one level of aggregation can be merged at some higher level of aggregation. For example, it may not be possible to match records between two files, each of which contain 10,000 records of individuals. However, if each person's record contained a county code, it would be possible to produce an aggregate file containing up to 3,042 records. Such records would show the number of persons in each county having given characteristics. In this case, 3,042, the number of U. S. counties is a "greatest common divisor" between the two 10,000-record files.

The trouble with such aggregation and merging is that the "greatest common divisor" can become arbitrarily small. If individual records in one of the two hypothetical files are identified by state rather than by county, then the greatest common divisor becomes the number of states rather than the number of counties. However, even very small merged files may be of interest in answering particular questions.

It is unlikely that the use of several files to make national estimates would involve any record matching. Such matching is appropriate only when two files contain data for the same set of persons. The most likely processing technique in such a case is to construct a file containing records derived from each of the original files. The concatenated file is then used as a pooled sample. Such a procedure requires comparability by variable and comparability by value, but not comparability by observation. (See Appendix A.) Thus, for the construction of national estimates, the issue of record linkages between files is moot.

Statistical Matching

Most record-by-record linkages between machine-readable data files are done on the basis of an exact match. Records are considered linked only if the values on key variables on one set of records matches that of key variables on the second file. For example, the RSA-SSA data link is matched on peoples' Social Security numbers.

Statistical matching is sometimes used where it is not possible to obtain a match key such as the Social Security number. In statistical matching, the record keys are attributes of the statistical population which will make records from one file "equivalent to those in the other file." Thus, microdata from two files might be matched on the basis of age, sex, race, and income in the hope that other information in the two files would be matched appropriately. The resulting file might be used to match, for example, a geographic location variable from one file with an occupation code from the other file. In general, statistical matching is difficult and is not nearly as powerful as exact-match procedures.[17]

Linkage Within Tables

At the highest level of aggregation, it may be necessary to link files by entering their distributions and summary figures into multi-dimensional tables. In its most extreme form, where there are no redundancies in the data, each of the cells in a table may be a summary figure from a separate file. Where several files contain data about the same phenomenon, it is possible to construct estimates through a number of table-fitting procedures. In this case, bivariate frequency tables are constructed from data in each of the two files and a set of fitting procedures used to make for the greatest degree of fit between the two sets of data. The project on Interagency Data Linkages has produced several working papers on the construction and fitting of such tables. This project represents the highest possible degree of organization and elegance in the matching of records from diverse agencies. It should serve as a standard and a model of efforts to merge data files.

The Interagency Data Linkage project in 1963 was established by the Bureau of the Census, the Social Security Administration, and the internal revenue service. The purpose of the project was to develop files of microdata containing information from files held by all three agencies, but devoid of personal identifiers. Such files, containing information on demographics, income, and earnings history, would be of immense value to researchers and policy-makers. The project begins by sampling the Census Bureau's Current Population Survey and linking CPS records with tax return records and Social Security earnings records for the sampled persons. In order to do the matching, selected personnel from each agency were sworn in as special agents of the other agency. The project produced a series of technical and substantive reports on the matching

---

[17] Report on Exact and Statistical Matching Techniques, Statistical Policy Working Paper 5, U.S. Department of Commerce, Office of Federal Statistical Policy and Standards, 1980.

process, techniques for performing and evaluating matching, and on the substantive results obtained from such matching.[18]

Characteristics Of Data For National Estimates

If any reasonable national estimate is to be derived, then we must begin with a sample estimating a national population. Many of the high quality and complete files, such as the RSA-SSA data link, are of no use in preparing national estimates because they do not have any intrinsic relation to the national population. At best, the RSA-SSA data link is a self-representing sample of people who are clients of both agencies. While it is possible to approximate a national population estimate by adjusting the weights of a non-national sample, such efforts suffer from three defects. The first is that such adjustments will probably add unacceptable bias to an already difficult enterprise. A second problem is one of sensitivity with respect to adjustments in sampling weight, which can only have a malignant affect on estimates of underrepresented segments of the national population. Finally, national estimates derived from non-national samples are not likely to be credible to a sophisticated audience, especially one with a political agenda of its own. A lack of robustness due to biased sampling is easy to discover, explain to others, and refute with statistics drawn from a more congenially weighted sample.

Measurement Problems

Nagi[19] has discussed sources of error in the measurement of disability. All of the usual problems of reliability, validity, comparability, and bias are present in survey data. Obviously, the designers of surveys do their best to minimize their measurement error. However,

---

[18] Fritz Scheuren, H. Lock Oh and Wendy Alvee, Matching Administrative and Survey Information: Procedures and Results of the 1963 Pilot Link Study, Report No. 3, Studies From Interagency Data Linkages, June 1978.

Joseph Steinberg, Some Observations on Linkage of Survey and Administrative Record Data, Studies from Interagency Data Linkages, U.S. Department of Health, Education, and Welfare, Social Security Administration, Office of Research and Statistics, August 1973.

[19] Saad Z. Nagi Disability in the United States: A Plan for an Information System prepared for the Office of Handicapped Individuals, DHEW, October 1977.

the definitional complexity of disability requires that the analyst be in a position to second guess the intentions of the survey, adjusting category bounds and relationships as needed. The only level on which such adjustments can be made is that of the individual data record. Thus any adjustment for measurement biases requires that the original microdata be used.[20]

Definition And Sensitivity

Disability estimates must be made for particular definitions of disability. Such definitions are constructed by manipulating and combining variables in microdata records. The robustness of a national estimate can be tested by noting the affect of "slight" perturbations in the definition. What is the effect on the prevalence estimate when a slightly more or less restrictive definition is used?

In summary, the problems of making national estimates of disability seem more inherent in the specification of definitions than in a lack of available data. However, the process of projecting national estimates from available data sources is not an easy one. It requires the assembling of data sources from public and private organizations and working out problems associated with confidentiality and proprietary. It will require on-hand experience and much experimentation. The cost-effectiveness of such an exercise will depend on such factors as the level of confidence to be tolerated for the estimates and the generality (or specificity) with which disability is defined.

The next section presents our recommendations to NIHR, including the pretesting and eventual establishment of an archiving and analytic system that will have the capability for making such national projections.

---

[20] Aggregate data can be used in the special circumstance where the aggregation results in no loss of information. For example, a microdata file containing only age and sex variables can be aggregated to a file of records giving the count for each combination of age and sex. The aggregate file contains exactly these same information as the microdata file.

IV. TOWARD A COMPREHENSIVE DATA SYSTEM

The objective of this planning project was to assess the feasibility of the establishment of a comprehensive data system on the handicapped people of the nation. Our investigation showed that a comprehensive data system, in the sense of one integrated file, was not feasible with currently available data sources. Instead, we recommend the testing and eventual establishment of a more flexible system with analytic and archival capabilities that could improve and build upon existing data sources.

This section includes our recommendations as well as a discussion of the statistical policies that NIHR should develop. We describe steps NIHR should take to achieve its congressional mandate of improving disability statistics through the power of coordination given to the Director of NIHR in the Interagency Committee on Handicapped Research.

Recommendations

The mission of NIHR is to carry out the Congressional mandate to "promote and coordinate research with respect to handicapped individuals" (P.L. 95-602). In interpreting this legislative mandate, the National Council on the Handicapped (NIHR's policy-making body), envisioned NIHR as spearheading joint sponsorship of research activities by "coordinating related federal programs of research" and by "producing, in concert with other federal agencies, demographic and other statistical information relative to handicapped individuals and disseminating such information.[21] In addition, NIHR is expected to focus on research areas unsupported or "not adequately covered in the missions of other federal agencies." [22]

The function of coordination is an important one, and gives NIHR the authority to address definitional and assessment issues essential to the improvement and establishment of a data collection system on the disabled of the nation. The major problem with existing data sources is that handicap is not uniformly defined, but is measured in terms of entitlement to programs or services. Disability or handicap is a phenomenon that affects a large proportion of

---

[21] National Council on the Handicapped, "Policies Governing the National Institute of Handicapped Research," Review Draft dated May 5, 1981.

[22] Ibid., Policy No.9, p. 6.

the population, some of whom may not be eligible for, know of, or choose to participate in, such programs. The definition and determination of disability should be freed from programmatic concerns of eligibility and handicap and tied to the functional assessment of ability to carry out a major life activity. NIHR is the only federal agency with a broad and non-programmatic responsibility in this area. NIHR should therefore promote the concept that disability is an indicator of the health status of the nation and as such, it should be part of any national survey, whether of general (such as the Census) or specific (such as the National Health Interview Survey) scope.

Recommendation 1. NIHR should develop a convincing argument to Congress and the Administration for implementation of a uniform functional definition of disability and handicap, and the inclusion of disability data in census and other national surveys.

In order to fulfill its mission as coordinator and initiator of research, that NIHR must develop the resources and staff capabilities to function as a core multipurpose analysis agency in this area. Examples of other core multipurpose analysis agencies in the federal government are the Economics Units in the Economics Statistics and Cooperative Service in the Department of Agriculture; the Bureau of Economic Analyses in the Department of Commerce; the Social Security Administration's Office of Research and Statistics; the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services; and the Federal Reserve Board.

Each agency has a responsibility for data analysis and dissemination in a specific area. NIHR should assume a similar function as the authoritative voice in the area of disability statistics.

Recommendation 2. NIHR should become a core multipurpose analysis agency concerned with disability statistics.

One of its most immediate tasks as a core multipurpose analysis agency should be the establishment, or support, of the Disability Data Archiving and Analysis Center (DIDAAC). We believe that NIHR's information needs--and indeed information needs of most others concerned with disability--would be better served through the establishment of this flexible system that will have broad capabilities to improve diverse sources of data. The system should be a continuing facility, driven by user demands and responsive to the needs of many types of users. Such a system will enhance NIHR's role as the coordinator of research and will eventually lead to standardization of data collection, analysis, and dissemination activities in the area.

Recommendation 3. NIHR should establish or support the establishment of a Disability Data Archiving and Analysis Center (DIDAAC).

Before fully implementing DIDAAC, however, it would be advisable to have a demonstration project to test the feasibility and cost-effectiveness of the facility. Experience gained from this project in the utilization of available data sources for making national projections would be most useful in determining the institutional and staffing characteristics of DIDAAC.

Recommendation 4. NIHR should support a demonstration project to test DIDAAC's capability in making projections of national estimates from existing data sources.

To move NIHR toward becoming a core multipurpose analysis agency it is essential to develop basic statistical policies and to improve its coordinator role vis-a-vis other statistical agencies, as discussed below.

Statistical Policies

In order to function as a core multipurpose analysis agency, NIHR must adopt some statistical policies. These statistical policies must address both good statistical practice and the coordinative mandate given to NIHR in P.L. 95-602. The most basic statistical policies concern the coordination of data collection efforts and a data archiving and access policy.

Data Coordination Policy

The Office of Federal Statistical Policy and Standards' Framework[23] gives a detailed set of goals for the coordination of statistical programs. The increasing interaction, and in some cases conflict, between agency programs and goals makes it advisable to coordinate the collection and compilation of statistical data. Tighter budgets and increasing respondent burdens make it necessary to provide for a maximum of power and comprehensiveness at the same time as expense, duplication, and respondent burden are minimized. An NIHR data policy must begin with a review of agency needs and a review of data presently being collected. The current project gives a comprehensive catalogue of available data relevant to a single set of

---

[23] Office of Federal Statistical Policy and Standards. A Framework for Planning U.S. Federal Statistics for the 1980's. Washington, D.C.: U.S. Government Printing Office, 1978.

demographic questions. We also outline the areas in which statistical data seem egregiously deficient.

An NIHR data collection and statistical policy thus serves to define a list of NIHR data needs and to reconcile that list of needs with a set of efforts to collect and compile needed data.

A Data Archiving And Access Policy

While a data collection and statistical policy is concerned with ends, a data archiving and access policy is concerned with the means to reaching those ends. The primary purpose of a data archiving and access policy is to transform machine-readable data files into bibliographic entities. At present, a large section of the economy is dedicated to the production, maintenance, and dissemination of printed materials. Most machine-readable data files are treated as auxiliary, private notes to printed reports. Throughout the government and the academic community, there has been a movement to treat machine-readable data files as recognized bearers of information, much like publications. Data files are being provided with a bibliographic identity: a title, author, producer, edition statement, etc. Once provided with a bibliographic identity, files are being accessioned by libraries using specially designed cataloguing rules much like those applicable to publications.

A data archiving and accessing policy states that machine-readable data files produced under the auspices of NIHR, its grantees, or its contractors will be assigned a bibliographic identity, will be preserved, and will be distributed to NIHR or its designated archive. Once so distributed, files will be available for analyses by others subject only to restrictions of confidentiality. The promulgation of a data archiving and accessing policy will bring NIHR into conformity with a current major trend in the community of data producers and consumers. (See Appendix C.)

Technical Standards For Machine-readable Data Files

Underlying a data archiving and access policy should be a set of technical standards for machine-readable data. Such standards provide guidance to data producers concerning the exact requirements of the agency. While the archiving and access policy treats files as publications to be archived and disseminated, the technical standard treats files as collections of records with numeric codes of various sorts, and states such things as the kinds of coding schemes to be used, the way in which records are to be formatted, and other information which, if implemented properly, keeps analysts and computerniks on reasonably friendly terms.

Technical standards should be consonant with the Federal Information Processing Standards (FIPS) but should also address needs specific to NIHR. For example, the FIPS give a set of standard codes to be used for states, counties, and other locations within the United States. NIHR technical standards should recommend the use of FIPS codes for such location variables. However, FIPS has no standard coding systems for variables relevant to disability. Therefore, it might be appropriate for NIHR technical standards to cite ICD-9, DSM-III, or the World Health Organization classification for disability. Thus, the technical standards for machine-readable data files are the lowest level of implementation of a statistical and data policy. (See Appendix A.)

Coordination With Other Federal Statistical Agencies

If NIHR is to become a general-purpose analytical agency, then it must be coordinated with other parts of the federal statistical system. How such coordination is to be established is an administrative matter for NIHR. However, some general areas of activity can be described.

Standards And Practices

One of the best ways to develop a good statistical and data collection policy is to adopt the best policies and practices of other agencies. The drastically low personnel level of the Office of Federal Statistical Policy and Standards has led to the establishment of large numbers of interagency coordinating committees. Such an interagency committee could be formed to help NIHR with the formulation of a statistical and data collection policy. Such a committee, composed of experienced statistical specialists, could also advise on standards and practices to be established within NIHR subsidiary to the statistical policy. The establishment of such practices would have the effect of bringing NIHR into compliance with the best practice, and of providing for commonality of standards and practices with other advanced statistical agencies. The current Subcommittee on Demographics and Statistics of the Interagency Committee on Handicapped Research is the logical basis for beginning such a liaison.

NIHR Primary Data

The NIHR statistical policy should specify a set of statistical series which can be obtained from other agencies or which can be generated in collaboration with other agencies. NIHR will then have to assume responsibility for the establishment and collection of statistical series primary to its needs which are not available from others.

Collegial Relations With Statistical Agencies

After the establishment of a statistical policy and the planning of its own collection efforts, NIHR will have changed its relationship to other statistical agencies from client to colleague. It is likely that advisory committees will now be replaced by coordinating bodies of various sorts, and that NIHR will participate in helping other agencies to plan data collections of mutual interest.

Joint Work On Data Collection

NIHR, as a collegial agency, will now be able to collaborate with other agencies on definition and collection of statistical information. In its new position, NIHR may find it easier to help service agencies to make their administrative data more useful to policy-makers and researchers.

V. INTRODUCTION TO DIDAAC

If NIHR is to become a multi-purpose analysis agency, then it must have a data archiving and analysis capability. Such capability would reside in a disability data archive and analysis center (DIDAAC). In particular, the establishment of such a facility is the best way to provide continuing answers to questions regarding the demographics, services consumed by, and needs of the disabled.

This section describes in detail functions and capabilities of the proposed center and presents alternative ways of implementing it, including cost estimates for such a center. In view of the financial constraints on NIHR, we do not propose that DIDAAC be implemented immediately. We recommend that the feasibility and cost-effectiveness of DIDAAC be tested in a demonstration project, designed to use existing data sources to provide national estimates of incidence and prevalence rates of various disabling conditions.

Characteristics And Functions Of DIDAAC

DIDAAC could perform the following functions for NIHR:

1. Perform analyses on primary and secondary data files for NIHR acquiring or using its own archive files as necessary;
2. Perform data analyses and respond to queries for other public and private clients;
3. Advise NIHR on technical matters relevant to data collection, including contracting requirements;
4. Promote the collection of disability data that provide information on etiology as well as functional limitations and activity restrictions;
5. Provide NIHR with technical advice concerning the monitoring of adherence to data and statistical standards, and eventually leading to less duplication of data collection efforts.

Internal Functions Of DIDAAC

DIDAAC's activities can be classified under nine headings: acquisition, cataloging, abstracting, reference, circulation, file processing, analysis, and dissemination, and recommendations for data improvement. The first five of these functions are similar to those in a library of print materials; the next three are specific to a data library and dissemination system, and the last to government-wide needs for improvement in the statistical system.

Acquisition. The acquisition function is concerned with locating machine-readable data files and their documentation. Some files, such as those maintained by RSA, can be acquired on a recurring and automatic basis. DIDAAC should be kept aware of NIHR grants and contracts which will result in the production of machine-readable data files. Contractors and grantees producing such files should be required to furnish copies to NIHR. DIDAAC can serve as an archive for such files and can supply contractors and grantees with technical assistance in meeting technical standards for machine-readable files submitted to NIHR. (See Appendix B.)

Those charged with the data acquisition function must maintain continuing contact with other government data producers and with data archives. There should be a continuous exchange of information with other government data sources such as the National Center for Health Statistics, the National Center for Educational Statistics, etc.

In addition to formal contacts with NIHR grantees and contractors and with other government agencies, DIDAAC should develop a network of informal contacts with data producers and users in government, private industry, and academia. It is a safe guess that there will be occasions when such informal expertise will provide materials which would be unavailable through "normal" channels.

A part of the initial list for the acquisition program is to be found in our interviews with potential users, who suggested many useful machine-readable data files. Another starting point is the catalog of machine-readable data files compiled over the last six months. Beyond the catalog, which has concentrated on files which appear to be directly relevant to problems of the handicapped, DIDAAC should subscribe to the catalogs of all potentially relevant data collection and archiving organizations. Such catalogs include those periodically produced by the National Technical Information Service and the core multipurpose collection and analytical agencies. In addition, DIDAAC should have access to general population statistics. Such statistics can be used in estimation equations; they also provide the necessary background information essential to understanding the significance of the disability phenomena.

Cataloging. DIDAAC must maintain a catalog of machine-readable data files, documentation, and relevant publications. The catalog should be maintained as a "union list" of materials in the collection, materials obtainable from other sources, and materials which could be obtained if necessary. The current edition of the Anglo-American Cataloging Rules (AACR-II)[24] gives extensive information on techniques for cataloging machine-readable data files. While not in itself a library catalog, the inventory of machine-readable data files compiled under this grant provides a set of information which can be used as the basis for a catalog.

Abstracting. A retrieval system for abstracts of data files and publications could be combined with the library catalog. In most cases, the abstracting task will require far more substantive knowledge of the data and the area of investigation then will be available to the usual library cataloger. If DIDAAC is to support an abstracting function, it will require that disability specialists be included on the library staff.

Circulation. The circulation function is concerned with inventory control and disclosure of information. The DIDAAC collection consists of an inventory of tapes and disk files. The circulation function is concerned with knowing the location and current status of such files, and with supplying file information to library staff and users. DIDAAC would not "lend" files in the usual sense. Tapes and disk files would be made available to users of DIDAAC's computing facility. Files for use at other computing installations would be copied to export tapes for shipment. Thus, at no time would a file ever be "on loan."

An important task for those concerned with circulation is to maintain conformity with various privacy and confidentiality standards. If DIDAAC is to be a "protected enclave," able to exchange confidential data with other government agencies, then someone must protect confidential data from unauthorized public disclosure.

Reference. The reference function mediates between the client and the DIDAAC collection. The reference function may be performed by a librarian, a substantive specialist,

---

[24] American Library Association. Anglo-American Cataloging Rules. Chicago, IL: R. R. Donnelly & Sons Co., AACR-II, 2nd Ed., 1978.

or a computing or statistical specialist. At the outset the reference function would probably be performed by those involved in data acquisition.

The reference function encompasses most of what are also called "user services" and "marketing." New clients will deal with Reference, who will introduce them to the other functional groups within DIDAAC as necessary. In many cases, brief requests will probably be handled by Reference without the need for referral to other DIDAAC groups.

File processing. File processing is almost exclusively the province of computing and statistical specialists. File processing is for two main purposes: to bring the file to a desired state of readiness; and to prepare it for a specific analysis. Initial processing includes the preparation of a machine-readable dictionary and user's guide, and checking the file's completeness, consistency, and accuracy. (See Appendix D.) In some cases, a set of classification variables may be created before the file is made available for use. Initial preparation may also include the deletion or encryption of personal identifiers.

Many users will require the creation of special variables or indices, the aggregation of microdata files, or the merging or linking of sets of files. In some cases, the files thus produced may be suitable for inclusion in the permanent collection.

Analysis. DIDAAC staff may include statistical specialists and people with substantive expertise in areas of disability. It would then have the capability of providing statistical consultation to clients and of writing research reports in response—to queries.

Dissemination. Regular reports could be produced to present results of DIDAAC efforts and findings and to suggest future directions of activity for data collection agencies. DIDAAC's purpose, and availability would be widely announced to facilitate use of the system. These functions would be geared toward enhancing NIHR's role as a coordinator in the area of disability leading to a more integrated data collection effort among all agencies.

Data improvement. It is reasonable to assume that DIDAAC will be instrumental in improving the quality of disability statistics. Its file processing activities should eventually lead to accumulation of highly comparable data files on disabilities. Through its statistical consulting and technical dissemination activities, DIDAAC should encourage changes in various collection efforts which would improve the comparability of data. DIDAAC could demonstrate and promote the necessity of collecting information on the etiology of a condition as well as its consequences in terms of functional limitations and activity restrictions. Such information would greatly improve our ability to more accurately relate disability to its medical and behavioral antecedents.

User Requests And DIDAAC Services

The following are some of the types of requests which might be made of DIDAAC and examples of the types of services needed to fulfill such requests. (Neither the list of requests nor of services is meant to be exhaustive.)

"Hot line" requests for a fact. Many requests to information providers are for a fact, i.e., something which can be expressed in a single sentence. The classic example of such a request is "How many disabled people are there in the U. S.?". In some instances, the person responding to the request may simply know the answer, or may be able to look it up on the spot. A rough definition for a "hot line" request might be that it can be answered in less than one hour and that the answer can be transmitted in a normal telephone conversation.

Requests for a bibliographic reference. DIDAAC will not ordinarily serve as a bibliographic reference service, since such functions are already well performed by a variety of services. A client with only a bibliographic request would probably be referred to the appropriate agency. A bibliographic request arising in the course of a client's continuing relation with DIDAAC would probably be handled by DIDAAC staffmembers as a courtesy to the client.

DIDAAC itself would probably be both a client and a provider of materials to bibliographic reference services. Much data documentation and most machine-readable data files have not been properly cataloged and abstracted. DIDAAC would make its own catalog records and abstracts available to bibliographic organizations.

Requests for publications. DIDAAC will serve as a primary source for its own analytical publications and for documentation for machine-readable data files in its collection. Recurring requests for the same information can be satisfied by the publication of a series of reports on the demographics and epidemiology of disability.

Requests for analysis. The reference facility will pass requests for analysis on to the analysis facility, which will be responsible for developing cost estimates and performing the work. Analysis can be orders of magnitude more costly than the provision of facts, publications, or bibliographic references. The latter products and services may be provided to a clientele without charge, but the former may entail sufficient expense as to make necessary a fiscal justification. In some cases, the client may pay the expenses of the analysis; in other cases, internal funds may be used. In either case, there must be some budgeting mechanism, so that an analysis task does not consume unwarranted resources.

Requests for a machine-readable data file. DIDAAC should provide files and documentation to clients who prefer to use their own computing facilities. In some cases, DIDAAC may acquire and prepare files for a client, who then has the files analyzed at another computing facility. In such a case, a copy of the file and documentation would be added to the DIDAAC collection.

Requests for special reports. DIDAAC personnel with substantive expertise in areas of disability can write reports for clients, but such activities must be budgeted and approved on an individual basis.

Functional Components Of DIDAAC

A full specification for a disability data archive and analysis center (DIDAAC) center requires a specification of hardware, software, data, procedures, and people. The specifications given here are only estimates, but can serve as a guideline for evaluating proposals.

Hardware. Primary hardware of DIDAAC is a medium or large scale computer housed in a computing center with extensive data storage facilities. An equally important piece of "hardware" is a library and cataloging facility for hard copy of documentation and publication, and microform. It is not necessary for the facility to have its own computer. The amount of computing that any conceivable facility would do is easily within the capability of any large research-oriented computing center. The economies of scale incurred by using a large computer especially one in a center with facilities for tape librarianship, far outweigh any possible advantages of possession of the machine.

Reasonable candidates for a DIDAAC computer center include the National Institutes of Health computing facility, the University of Michigan or Wayne State University Computing Centers, or any of the local Washington Computing Centers such as American Management Systems, COMSHARE, etc.

Software. A major factor in the choice of the computing center is the range of software available. Presently available software is fully adequate to do the computing, bibliographic, and publication needs of DIDAAC. There is no need for the writing of any appreciable amount of software, nor should DIDAAC be required to import and maintain its own programs. However, DIDAAC can be expected to make heavy use of software for file manipulation, statistical analysis, graphics, information retrieval, document processing, and communications.

File processing in statistical software are often combined in these same package. File processing programs are used for aggregation, merging, and the creation of new variables and indices. Statistical programs are used for the analysis of files created by the file processing programs. The most popular statistical and file processing package is SPSS, a well documented package but one with serious shortcomings as a file management package. More powerful packages for file management include the Statistical Analysis System (SAS) and OSIRIS IV. SAS has a powerful set of analytic routines and is growing in popularity. OSIRIS IV has as its major advantage the ability to handle multilevel, structured, and hierarchical files.

Some DIDAAC users may prefer to do their own statistical work using an interactive statistical system. Statistical systems with interactive capabilities include P-Stat, the SPSS Conversational Statistical System (SCSS), and the Michigan Interactive Data Analysis System (MIDAS). At least one interactive statistical system should be provided for the use of DIDAAC clients.

An increasing number of statistical results are being disseminated as graphics rather than numeric tables. Some statistical packages such as SAS and SPSS have integral graphics routines, and offer a choice of graphic and tabular output formats. Such facilities should be offered to DIDAAC clients either as hands-on option or as outputs generated by the technical staff.

Graphic output requires a suitable output device. DIDAAC users should have access to a variety of high and low resolution plotters. Inexpensive low resolution devices can be purchased by DIDAAC for remote connection to the computer. Expensive high resolution devices should be provided by the computing facility on a chargeable basis.

If DIDAAC is to have a statistical publications mission, then it should employ the computer as a publishing mechanism. Computer based text processing facilities allow the incorporation of numeric and graphic output into publication without the need for transcription. The computing facility should have available multi-font devices like the Xerox 9700 Page Printer and a phototypesetter.

DIDAAC differs from a contract research computing service in maintaining a continuing collection of data and documents. Such a collection becomes useful only with a proper catalog, a catalog which is best maintained as a data base. DIDAAC should be able to make its catalog directly available on-line to its users as well as to its staff. The information retrieval software for doing so should be provided by the computing center, while the file of catalog information will be provided by DIDAAC.

An electronic message system and an on-line conferencing system have proved valuable for facilities such as DIDAAC and should be made available to the facility and its clients. An electronic message system serves as a valuable way for clients to transmit work orders, and to seek and receive aid. On-line conference programs provide a medium for on-going technical and administrative seminars.

Data. The core of DIDAAC is its data archive. A continuing program of data acquisition should result in a definitive archive of data on all aspects of disability. The growth of a linkage between NIHR and other statistical agencies should result in regular procedures for the acquisition of government data. In addition, DIDAAC should serve as NIHR's official repository for data gathered under NIHR grants and contracts.

Data analysis projects conducted within DIDAAC should differ from those conducted in the usual course of research contract by being oriented toward "improving the collection." Since the DIDAAC archive will be a continuing facility, it will be worth a continuing investment in the upgrading of the files in the collection. Thus, a client request for a specific file creation job may in fact result in the creation of a more general file. The investment in the more expensive process of creating the general file is justified because the file contributes both to the client's needs and to the quality of the collection.

While machine-readable data files are the raison d'etre, few if any files will be maintained in the same location as the DIDAAC staff. Data tapes will be maintained at the computing facility under the custodianship of the computing facilities tape librarian. The responsibility for back-ups, tape maintenance, and vault storage of irreplaceable data will remain with DIDAAC. Arrangements for the care of files containing personal identifiers or other confidential information must be worked out between DIDAAC and the computing facility. It is not likely that such files will amount to any significant proportion of the DIDAAC archive nor that any extraordinary procedures will be required to handle such files.

Procedures

The operating procedures of DIDAAC will resemble those of libraries, computing centers, and research services. Procedures will, of course, change with experience. It is possible at the outset to specify some broad classes of procedures which must be established.

User services. DIDAAC is primarily driven by the needs of its users. Accordingly, procedures must be established for determining what services are available, to whom they are available, and which DIDAAC staff will be made available to handle particular kinds of requests. It may be necessary to establish a priority system for classifying customers and one for classifying the sorts of work to be done. Some sorts of analyses or data may be classified as beyond DIDAAC's scope of work. (The problems of high priority clients with low priority tasks or of low priority clients with high priority tasks are too baroque even to contemplate.) At the outset, the technical staff should be responsible for customer relations, once a user has been accepted. If there is significant growth in the use of DIDAAC, then it may be necessary to have one or more persons who serve as a gatekeeper, allocating technical and staff resources to users.

Costs. There has been some controversy over the degree to which DIDAAC should function on a cost recovery basis. At one extreme is the model of the contract research organization which recovers all operating expenses from users. At the other extreme is the model of the public or university library (or the Congressional Research Service) which offers "unlimited" services to its user community. DIDAAC has the potential for incurring an extraordinarily wide range of costs. Some requests may be answered in the course of a one minute telephone call, while others may require several person months of effort and large amounts of computing.

The most satisfactory way to allocate resources among clients is by the use of an internal dollar-cost system. All expenditures, whether in personnel time or other resources, should be expressed in dollars and charged to a budget. The degree to which the internal cost structure is presented to users is a decision for DIDAAC management and NIHR. In some cases, the user may be entirely ignorant of the internal cost function. For example, what is advertised externally as a "free" public inquiry service could be but accounted for internally as a set of $20 budgets available to public users. Any person with a casual inquiry could be serviced up to the $20 limit, could be asked to provide justification for a higher expenditure, or could be asked to pay costs beyond public use budget. Similar budgets could be set up for other classes of users and may be made manifest to them as required.

Experience has shown that such an internal dollar cost system is the only rational way of allocating resources among a multitude of users.

Marketing. It is in the interest of both DIDAAC and NIHR that it serve as wide a constituency as possible. Broad use of the facility will tend to encourage a flow of statistical data and ideas about disability, to improve the quality of data, and to improve the quality of DIDAAC's collection. The marketing function should concentrate on identifying potential users of statistical data and analyses. Accordingly, cooperative arrangements should be made with other data producers and with coordinate agencies such as the National Rehabilitation Information Center (NARIC). Such cooperative marketing arrangements will allow DIDAAC to maximize its potential users base while maintaining friendly and cooperative relations with other suppliers of statistical and bibliographic data.

One of DIDAAC's most important publications will be it's list of new acquisitions of data and documentation. Such a list not only informs potential users of new DIDAAC capabilities, but also serves to inform them of research and statistical work in the field of disability.

Publication. One of the main purposes of DIDAAC will be as a center for reliable statistical information on the disabled. While most DIDAAC reports will be generated in response to specific user queries, it is expected that a core set of disability statistics will be of general interest. Requests for such "best selling" information can best be filled by a regular program of publication. Although such publications will be used to fill large numbers of user requests, the analyses supporting the publication must be scheduled by the DIDAAC management. Thus, a schedule of recurring publications must be matched by a schedule of recurring analyses of data to support the publication.

Data acquisition. DIDAAC differs from a contract research computing center in that it has a continuing responsibility to maintain and improve its collection. Accordingly, significant attention must be given to maintaining a liaison with other producers of statistical data. NIHR should require all contractors and grantee's to supply DIDAAC with copies of machine-readable data files produced in the course of research.

Data quality control and maintenance. DIDAAC must have internal quality control function so as to inform users of the quality of data and analyses. If DIDAAC is to serve as a repository of data produced under NIHR grants and contracts, then it can serve to monitor the quality of such data. NIHR might included as a special condition of research agreements that machine-readable data files must be submitted and must pass certain quality control checks. (Such data checking will require more than trivial amounts of resources, but will result in better research results.) NIHR efforts to improve the quality of data can only serve to ameliorate the quality of statistical data throughout government.

In addition to data quality control DIDAAC must have a data maintenance program. It has become apparent that machine-readable storage media such as magnetic tape are susceptible to deterioration over time, especially if they remain unused. DIDAAC must have a policy of tape inspections and back-ups, including secure vault storage of irreplaceable statistical data.

Accounting. DIDAAC must have facilities for the accounting of cost, effort, and results. The internal cost accounting system is a necessity for balancing the competing needs of users. A system for accounting of productivity is necessary to determine the extent to which DIDAAC is a successful investment.

Administrative Organization Of DIDAAC

DIDAAC's administration divides into two major areas, the technical administration of data and the use of those data to produce knowledge about disabilities. These two major functions could be further divided into data archiving, data operation, statistical consulting, and client services. These four groups are discussed in more detail below.

Management. Initially the management functions of DIDAAC can be performed as an additional duty by one of the technical staff. As the organization grows, a specialized management staff can be developed.

Staff. The DIDAAC staff must have a knowledge of the disability area and disability statistics, as well as general knowledge of computers and programming. The number of people to be employed and the division of labor is a function of availability of funds and DIDAAC's potential market. It could usefully employ people with background in computation, librarianship, statistics, and disability. In some cases, DIDAAC could function with the help of specialists called in on a consulting basis. The next section gives a more detailed description of an administrative organization for DIDAAC and potential levels of staffing.

Data Archive. The actual maintenance of data tapes will be done by the computing center. DIDAAC's own data archiving operation will consist of accessioning tapes, constructing catalog records, and pursuing a schedule of back-ups and file saving operations. Data archiving functions are best handled jointly by a computing specialist and an librarian. A person of either specialty can be designated as DIDAAC's official data archivist.

Data operations. Data operations are the core of DIDAAC. The actual operation of the computer is left to the computing center. DIDAAC's data work should be performed by technical people with Bachelor's or Master's level training in computer science and in the social sciences. As technical people gain proficiency, they will be able to perform both as computer people and as substantive specialists in disability statistics. Since clients with complex jobs will usually have their own expertise in disability, it can be expected that growth in DIDAAC staff will occur primarily among its data operations people.

Research and client services. The research and client services group is the repository of DIDAAC's substantive expertise in disability and disability statistics. The group serves as a liaison between clients and the computer people in data operations and as a client of data operations for the production of a DIDAAC research reports. Necessary skills for personnel in client services and research include the social sciences, public health, disability, rehabilitation, statistics, and librarianship.

Implementation Of DIDAAC

A data archive and analysis center could be established in several ways. These include:

1. A facility established within NIHR, using government personnel and government or contract computing facilities;
2. A facility or activity within another federal agency supported by an interagency agreement with NIHR;
3. A new center within the Federal government, but outside of NIHR;
4. A center within a private facility supported by a federal grant or contract.

DIDAAC's responsibilities will involve it with both the supply of and the demand for statistical data. DIDAAC's staff members will be involved with other government statistical agencies and with researchers and contractors while locating and accessioning data files. They will also be concerned with the longer range matters of setting requirements for data and setting and enforcing data standards. On the demand side, DIDAAC's staff members will be involved with NIHR and other executive branch clients, as well as congressional, state and local government, and private-sector clients.

One of DIDAAC's most important characteristics is its continuity. The facility must be given time to establish a base of expertise and a data collection. Without a continuing basis of support, it will offer little that an expert researcher or contractor could not provide for themselves.

Criteria

The alternative implementations for DIDAAC can be evaluated using criteria of staff expertise, excellence of collection, continuity of existence, independence, timeliness of response, and cost.

Staff expertise. DIDAAC must be able to field an expert senior staff and to maintain a reputation for expertise. This criterion is more than the fatuous remark that people should be good at their jobs. DIDAAC must have access to recognized experts in disability and disability statistics, either as staff members or consultants. Lack of such of demonstrable expertise may not adversely affect the quality of DIDAAC's work, but it will decrease its credibility.

Excellence of the collection. DIDAAC must maintain a definitive and comprehensive collection of files on disabilities. The requirement for a collection of demonstrated excellence is exactly analogous to the requirement for staff expertise.

Continuity of existence. It may be possible to begin operations with a staff of wide reputation and undisputed excellence. However, the development of a full staff will probably take some time. The development of a definitive data collection will require considerable time under any circumstances. Once DIDAAC begins operations, it will require some time to develop a reputation for institutional excellence and to become generally known. All of these processes will be facilitated by a continuity of support and of institutional identity. Changes of name, institutional affiliation, or an excessively high turnover rate of senior staff will tend to work against an impression of competence. Interruptions in funding will work against the reality as well as the impression of competence.

Independence. DIDAAC must be independent from the excessive influence of program specific statistical agencies. It must also be perceived as independent from such influence. Such independence can be facilitated in several ways. Establishment of DIDAAC in an independent institution can insulate it from undue influence by its clients. The development of a wide client base can provide some stability when some costs are recovered from non-NIHR clients. However, it is unlikely that the majority of DIDAAC's costs will be reimbursed by non-NIHR clients.

Probably the best way to safeguard DIDAAC's independence is through the reputations of its senior staff members. People with established reputations for professional integrity are unlikely to exhibit bias.

Timeliness of response. While DIDAAC may perform a number of recurring or long-range studies, most of its work will probably be directed to satisfying one-time requests of short to moderate duration. DIDAAC should have sufficient staff on hand to respond to high priority requests without undue turbulence in scheduling and staffing. It should also have adequate communication facilities, including sufficient telephone lines, terminals, etc.

Cost. Although a truism, it should be said that DIDAAC should operate as efficiently and inexpensively as possible.

Costs For Several Levels Of Staffing

It is beyond our expertise to prepare a budget for a government facility. Table 3 shows some possible levels of staffing, ranging from a minimal staff to a very large staff. Figure 1 shows estimated costs for each level of staffing. The salary, expense, and indirect cost multipliers used are BSSR's, but are representative of what most contractors will charge.

Table 3: Examples of DIDAAC Staffing Levels
(Full Time Equivalents)

	Minimum						Maximum
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
Facility manager	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Senior Researcher	-	0.5	0.5	1.0	1.0	1.5	2.0
Junior Researcher	1.0	-	1.0	1.0	2.0	2.0	2.0
Senior programmer	-	1.0	-	1.0	1.0	1.5	2.0
Junior Programmer	0.5	0.5	1.0	0.5	1.0	1.5	2.0
Librarian	-	-	-	0.5	1.0	1.0	1.0
Secretary	0.5	1.0	1.0	1.0	1.0	2.0	2.0
Total	3.0	4.0	4.5	6.0	8.0	10.5	12.0

Table 4 shows the relation between the alternative institutional arrangements and the criteria.

Conclusion

There is presently a serious disparity between the chaotic state of disability statistics and the order and serviceability implied in the call for a comprehensive data system on the disabled population. Even if the Post-Census Disability Survey were completed, and we recommend that it be completed, there would still remain the problem of coordinating the collection, analysis, and dissemination of disability statistical data. Congress has given the Director of NIHR the mandate to produce better statistical data through cooperation with major data collection agencies. We view the step-by-step development of the Disability Data Archive and Analysis Center (DIDAAC) as the best means of addressing the need for better data through coordination.

Initially, DIDAAC would concentrate on the improvement of disability data through the archiving, analysis, and dissemination activities described above. As time goes on and the expertise of the DIDAAC staff and its overseers (e.g., the Subcommittee on Demographics and Statistics of the Interagency Committee on Handicapped Research, and the National Council on the Handicapped) strengthened, DIDAAC's function would be increasingly to help the Director of NIHR

Figure 1: Annual Cost of DIDAAC at Seven Levels of Staffing

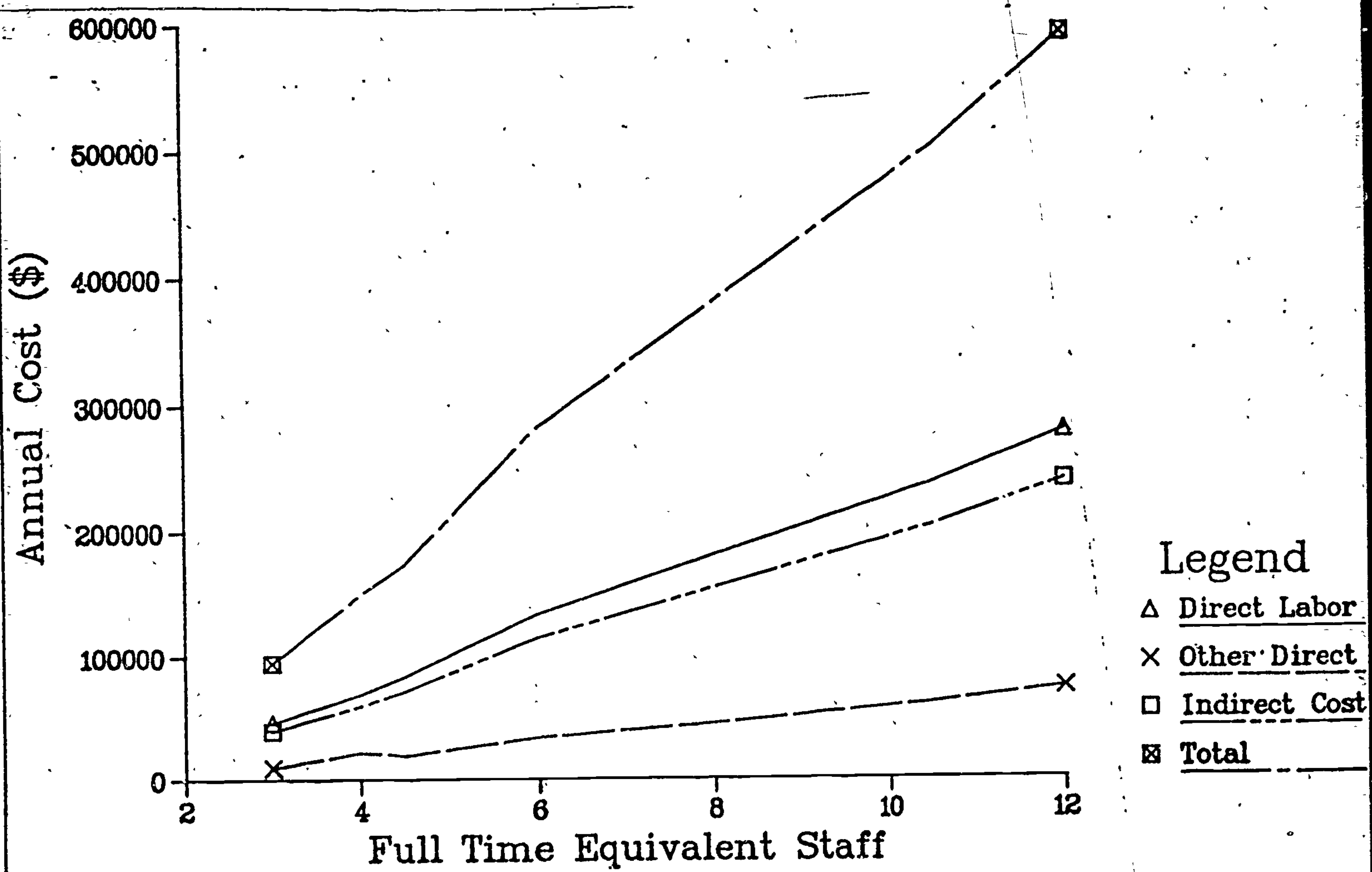

Table 4: Alternative implementations of DIDAAC

Criterion	Within NIHR	In Other Federal Agency	In New Federal Agency	Contractor
Staff expertise	No NIHR staff available at present Staff would be entirely under NIHR's control. New staffing not likely	Other agencies have good people with strong reputations. Other agencies are not likely to have all necessary skills.	Same staffing problem as at NIHR.	Skills are available and can be obtained by subcontracting or consortia.
Collection	NIHR the owner and permanent home of the collection. No available staff with acquisition and maintenance skills.	Agency would be a custodian, not an owner. Agency would have necessary maintenance skills (or else no agreement.)	Agency would be an owner, not a custodian. Would have the requisite skills for maintenance.	Contractor would be a custodian, not an owner. NIHR would need repossession arrangement if contract terminated. Would have maintenance skills (or no contract.)
Continuity	By definition, would exist as long as NIHR.	Subject to NIHR funding and willingness of other agency.	?	Contingent on funding of contract.
Independence	Dependent on staff reputation and performance.	Dependent on staff reputation and performance.	Dependent on staff reputation and performance.	Dependent on staff reputation and performance.
Responsiveness	High incentive to be responsive. Subject to staffing limitations.	Subject to conflicts with internal agency requirements and staff limitations.	Subject to competing requests of other clients. Moderate intention to be responsive	High intention to be responsive. Possibility of competing requests from other clients.

to evolve statistical policies that will improve the quality and comparability of disability statistics throughout the federal government.

We have outlined here the technical and policy steps needed to establish DIDAAC. The initial phase of a demonstration grant should be designed to address the immediate need for good statistics, and the longer-term need of coordinating federal disability statistical efforts.

APPENDIX A

Technical Considerations in the Evaluation of Machine-Readable Data Files

Richard C. Roistacher
Bureau of Social Science Research
Washington, DC 20036

Technical Consideration in the Evaluation of Machine-Readable Data Files

Richard C. Roistacher
Bureau of Social Science Research
Washington, DC 20036

December 31, 1981

This research memorandum describes general technical considerations that are essential for evaluating the internal consistency and comparability of any data files to be included in a comprehensive data system. The specifications in this document fit within the broader, more content-bound and project-specific set of requirements.

The acquisition of data begins with a search for available data files or with a client's request for a specific file. In the latter case, there is prima facie reason for considering the data file relevant to the purposes of the comprehensive data system. Files obtained as a consequence of a search by the staff of the data system must be evaluated on substantive grounds for their relevance and adequacy. Once a substantive evaluation has resulted in a decision to acquire, a data file of unknown quality or of doubtful provenance must be subjected to an assay to determine its completeness and consistency. This report describes the rationale and methods of a data assay.

The Data Assay

The technical evaluation of a set of data is called a data assay. A data assay serves four major purposes:

1. To assess our confidence in the quality of the data;
2. To investigate the costs and benefits of correcting errors in the data;
3. To provide guidelines as to which corrections are feasible and which are not;
4. To provide error estimates for data which cannot be corrected.

The assay has the following institutional uses:

1. It serves as a guide to data editors and those who will be correcting the data;

2. It provides secondary analysts with a statement of the data's quality;

3. It provides valuable feedback to data producers on the quality and stylistics of their product;

4. It provides an overall evaluation of the quality of a research or data production project.

The assay of machine-readable data files(MRDF) is primarily formal; and is affected very little by the file's substantive content. Part of the evaluation of a data file concerns its content and relevance. Such evaluation, while most important, is not within the scope of this discussion. The assay of machine-readable data can be classified under the linguistic headings of syntactics, semantics, and pragmatics.

The syntax of data is its formal structure without reference to its message. An example of a syntactic characteristic of the data set would be that the variable SEX is in position 91 of each record and is represented as an IBM-format numeric character.

The semantics of a data set is the message or information it carries. Some semantic characteristics of a data set might be that it has a record for each of 35,000 hospitalized patients, and that the sex of patient 10954 is male.

The pragmatics of a data file is the mapping of syntactic elements onto semantic ones. An example of a pragmatic characteristic of the data set would be that sex is represented by a "1" for male, "2" for female, and "9" for missing data.

File-level Analysis

A data set in machine-readable form is usually referred to as a "file." A file is distinguished with a name and other information which allows it to be located, and processed with the computer. Even though pragmatics are at the lowest level of the hierarchy, files cannot be evaluated until some pragmatic criteria have been met. Primarily, tapes must be readable. While the IBM 360/370 hardware has provided a _de facto_ standard for magnetic tape formats, some old tapes may be written in formats which require explicit conversion. A conversion may be complicated by the fact that it is sometimes not clear what the conversion's target should look like. When there are only summary statistics for an unreadable file of microdata, it may be difficult (or impossible) to tell whether the file has been properly deciphered.

Some tapes, while written in readable format, are so old or so poor in quality of recording medium as to make reading difficult or impossible. In some cases, it may be desirable to attempt to reconstruct the data, either from the tape as provided, or with the help of the file's producer.

Structure error. The highest level of error is what can be called a file structure error, in which the file is not as described. In some cases, a structure error is the result of having been given the wrong file. At the beginning of an assay we note the record length, the number of records in the file, and the content of fields listed as record identification in the file description. A file with too few or too many records is suspicious; a file with records that are too long or too short is unsuitable for further analysis until an explanation or new documentation has been obtained.

Complex Files. A "complex" file is one that contains more than one type of record. A complex file, for example, might contain one record for each patient in a hospital and one record for each prescription written for a patient. Complex files may incur structure errors at the level of the file or at the level of the individual record type. Most complex files are structured, having a logical relation between records of different types.[1]

One common error in hierarchical files is that the number of records at a lower level does not match the count that is contained as data in a higher level. While such errors may not be as severe as having a file with missing or extra records, such disparities cast doubts on the file's quality.

---

[1] It should be noted that not all complex files need be structured and not all structured files need be hierarchical. While somewhat bizarre, it would be possible to combine the records of two unrelated files into a single structured file. A file is structured when it includes a logical description of the relation between record types. A file of data from a longitudinal study could consist of parallel records from different observation periods, but with no hierarchical relation.

Record- And Variable-level Analysis

The first detailed analysis of a set of data is syntactic. Before any analysis of semantics is feasible, it is necessary to determine whether the data file conforms to its description. The possible errors in a data set fall into several levels.

Syntactic Errors

The evaluation of a set of data includes a complete analysis of the syntactic characteristics of every variable. A variable's "syntactic range" is the set of representations a variable may take. For example, a variable called NAME might be restricted to upper case printing characters, with a maximum length of 20. A variable called AGE may have its syntactic range given as 3 numeric characters with leading zeros. Accordingly, "FRED JONES" is within the syntactic range of NAME, but "Fred Jones" is not. Similarly, "045" is within the syntactic range of AGE, but "TEN" and " 45" are not, (Blanks are not included in the syntactic range of AGE.) Violations of the syntactic range are serious errors, indicating potentially grievous flaws in the quality of the data. Syntactic errors are almost always at the level of the individual variable or data item. Syntactic errors are usually fairly easy to discover. Some syntactic errors are primarily stylistic, as when a file contains leading blanks where leading zeros have been prescribed. In such cases, the repair or accommodation to such errors is straightforward.

Semantic Errors

A variable's "semantic range" is the set of messages it may validly convey. While the syntactic range of SEX is from "0"to"9", its semantic range may be restricted to 0, 1 and 9 indicating "male," "female," and "missing data." Semantic errors are often described as "wild" or "out of range" values.

Joint Semantic Ranges. Semantic ranges apply to sets of variables as well as to individual variables. The investigation of semantic ranges of sets of variables is usually called "consistency checking." For example, "Male" is in the semantic range of SEX and "Caesarian section" is in the semantic range of PROCEDURE. However, the joint range of SEX and PROCEDURE does not include these two values. The concept may be extended to any set of variables and is exactly parallel to the construction of joint probability spaces.

Statistical Semantic Ranges. A statistic is a variable which describes a set of values for another variable. What is commonly called the range of a statistic is here called the syntactic range. Where there is knowledge of a set of data, it is possible for a statistic to have a semantic as well as a syntactic range. For example, a proportion has a syntactic range from 0 to 1. However, if we are examining a set of data from obstetrical patients, then any proportion of males above 0 is outside of the semantic range for this statistic. The investigation of semantic ranges of statistics requires a considerable increment in commitment to the assay of the file, since substantive expertise is necessary to define valid ranges. A data specialist can determine the syntactic and semantic ranges of variables and sets of variables from the file documentation. Data specialists can also compare obtained distributions and statistics with those given in the documentation. However, it is beyond the scope of a data specialist's expertise to say whether or not a particular statistical value or distribution is plausible given a data set's domain of inquiry.

Comparability And Mapping

Most assays of data do not include a specific assessment of comparability with other data sets. (One strong exception to this rule is in the case of statistical time series.)

Since a major task in developing a comprehensive data system on the handicapped is the incorporation of data from diverse sources, a special effort must be made to investigate the comparability of newly acquired data to data already in the system. We can adopt a unitary or a pluralistic standard of comparability. The unitary standard is the minimum data set, which will serve as the comparability target for all accessioned data. The assay of every file must include an evaluation of its potential for comparison with or inclusion in the minimum data set. At the beginning, there may be some modification of specifications for the minimum data set so as to be able to accommodate existing files. However, once the master file has been defined, it will serve as the comparability target. Some sets of data will be comparable with each other, even though none is comparable to the minimum data set. Comparability may be in terms of observation, variable, or value.

Comparability by observation. Two data sets are comparable by observation if their records are derived from the same unit of analysis. Thus, income records on a group of individuals are comparable by observation to health records on those same people. (It is assumed that the personal identifiers on the records can be matched.)

Comparability by variable. Data are comparable by variable when records referring to different units of observation contain identical variables. Mortality data from different cities are comparable by variable where they contain information on date, cause of death, and age at death.

Comparability by value. Comparability by variable may not be of any great help unless we can establish comparability by value. The trivial case of establishing comparability by value is where the pragmatics of one variable must be matched to another. If two sets of data record race as "white," "black," and "other," then it is a relatively simple matter to map the numeric values used in one data set on to those used in the other.

What is more common is that one set of values is a finer partition of the data space than another. If one data set gives five classes of ethnicity while another records ethnicity as "white" and "non-white" then the only way to merge the two files is to aggregate the five ethnic categories into "white" and "non-white."

Often the problem is not the fineness of a partition but the degree to which two partitions of a data space map on to each other. A geographic example may help. A variable called ZIP CODE may be mapped on to a variable called POSTAL REGION by extracting the high-order zip code digit. A subset of zip codes could be matched to cities by extracting the first three digits. However, many three digit combinations do not refer to cities but to areas that include several different post offices.

It would be extremely difficult to map zip codes onto counties, because postal planners do not in general use counties as components in the construction of postal routes and regions. In fact, the only way to provide such a mapping would be to construct a much finer variable indexing towns and unincorporated areas, which could then be mapped onto both zip codes and counties.

Similar considerations hedge the mapping of data sets indexed by medical diagnostic category with those indexed by impairment and with those indexed by functional limitation. In some cases, it may be possible to construct mappings from one system of indexing to another, but such a mapping will require considerable effort and the collaboration between those with expertise in handling data and those with expertise in research on the handicapped.

REFERENCES

Roistacher, R. C. A general consistency check for machine readable data. Sociological methods and research. 1976, 4:301-320.

_____ Technical standards for machine readable data submitted to the Law Enforcement Assistance Administration. Washington, DC: U. S. Department of Justice, Law Enforcement Assistance Administration, 1979.

_____ A style manual for machine readable data and their documentation. Washington, DC: U. S. Department of Justice, Bureau of Justice Statistics, 1980.

_____ Data archiving techniques at the LEAA Research Support Activity, Center for Advanced Computation, (CAC Document No. 211), 1976. (With B. B. Noble)

APPENDIX B

Technical Standards for Machine-Readable Data Supplied to the National Institute of Handicapped Research

Richard C. Roistacher
Bureau of Social Science Research
Washington, DC 20036

Technical Standards for Machine-Readable Data Supplied to the National Institute of Handicapped Research

Research Memorandum Number 5

Richard C. Roistacher
Bureau of Social Science Research
Washington, DC 20036

December 31, 1981

This standard sets forth technical requirements for data that may be included in the National Institute of Handicapped Research's Comprehensive Data System (CDS).

Any data file submitted for inclusion in the library of the Comprehensive Data system should meet these technical standards. All tapes maintained in the CDS library will conform to this standard. Although most of these standards are simply good data processing practice, some of the requirements are specific to the needs and facilities of the CDS. Additional information about these standards and the Comprehensive Data System, or assistance in meeting these standards, may be obtained from NIHR.

Another aid to data producers is the Government Printing Office publication, "A Style Manual for Machine-Readable Data Files and Their Documentation," which describes optimal practices and techniques for formatting and documenting machine-readable data files. This document is available as publication NCJ-62766 from the National Criminal Justice Reference Service, Box 6000, Rockville, Maryland, 20850.

Tape Recording Standards

Data supplied for possible inclusion in the CDS should be on 9-track magnetic tape written at a density of 800, 1600, or 6250 bytes (characters) per inch. Since more reliable encoding methods are used at the higher densities, 6250 BPI tapes are preferred to 1600 BPI tapes, which are preferred to 800 BPI tapes. The tape should be labeled with IBM or ANSI volume and file labels. If 9-track equipment is not available, then files may be submitted on 7-track magnetic tape. Either fixed- or variable-length records are acceptable, and files should be blocked to a length that will provide for efficient handling of the data. Physical record lengths of from 80 up to a maximum of 32,767 characters are acceptable. Experience has shown that physical record lengths of approximately 16,000 characters provide for the most efficient processing.[1] In general, data producers and users should consult with the technical staff of their computing facility to determine the facility's preferred size for physical records. The American National Standards Institute recommends a maximum physical record size of 2400 characters, which is considerably less efficient than the maximum permitted by this standard. Volume and file-naming conventions should be agreed on by NIHR and the data supplier before tapes are written. While it is impractical to devise a set of file-naming conventions that are universally applicable, it will often be in the interest of NIHR and its data suppliers to establish naming conventions for a particular set of data files.

Data Types

All machine-readable data records supplied for possible inclusion in the CDS will consist entirely of alphanumeric, EBCDIC, or ASCII characters. (Length fields in binary format that are part of the IBM variable-length record are not included in this restriction.)

For the purpose of this standard, variable will be considered either computational or noncomputational. Noncomputational items contain information such as names and labels that are never used in any arithmetic or numerical operation. Noncomputational variables may contain any EBCDIC or ASCII printing characters. Noncomputational items will be left justified and padded to the right with alphabetic blanks.

---

[1] Data on magnetic tape are divided into physical and logical records. A logical record contains data from a single "unit" (e.g., patient, gas bill, library catalog card, etc.). Data processing is often more efficient if several logical records are written end-to-end onto the tape to form a single "physical" record. This process is called "blocking." and a physical record is sometimes called a "block."

Computational items contain numeric information that is designed to be used in computations. Computational variables may contain only the characters 0-9, ".," "+," and "-." The only exception to this is where D- or E-format floating-point data are represented.[2] Wherever possible, data should be represented in integer format with implicit decimal places noted in the documentation. Computational variables should be right justified and padded to the left with zeros or blanks. If a computational field is signed, the sign must immediately precede the left-most numeric character. Computational fields must contain at least one numeric character. In particular, computational fields consisting only of a sign or of blanks are not permissible. Fields containing only a signed zero are not acceptable, since some computers cannot represent a signed zero.

Missing data. All variables for which there may be missing data must have an explicit missing data value or values. In particular, it will never be assumed that a value of zero or a field consisting entirely of blanks indicates missing data. Missing data values must be indicated in the documentation as a list or a range of values, and may include zero, but may not include values of -0, a blank field, or a field consisting only of a sign.

Missing data values must occur in the same field as the variable to which they refer. If an alternate value is to be used in place of a missing data value, the original variable must carry an appropriate missing data code, while the alternate value will be shown in a separate variable that has been declared for that purpose. In no case will the alternate value be carried in the original variable, with an explanation code in another variable. The rationale for this standard is that the meaning of a variable should be determinable without reference to a second variable. If the true value of a variable has been suppressed or modified, then the value of the variable should indicate such suppression or modification. If an alternate value is to be offered in such cases, the appropriate variable in the record can then be read if the analyst so wishes.

---

[2] Floating-point data are used to represent very large or very small numbers. The number 2,625,000,000 is written in floating-point notation as "2.625 E+09." The "E+09" is called the exponent, and indicates that the decimal point is to be shifted nine places to the right. D-format floating-point notation is similar, but uses a "D" rather than an "E" to mark the exponent. The number "0.00000000000465" is written in floating-point notation as "4.65 D-12." the "D" indicates that the number was computed with double precision arithmetic, which allows more significant digits to be represented.

A brief example may be helpful. Suppose that the variable "Number of Occupants" is sometimes estimated rather than counted directly, and that estimated values are to be flagged. "Number of Occupants" might be assigned values of 999 for "missing data" and "998" for "estimated". A second variable, "Estimated number of occupants" should be used to supply the estimated value for cases in which "Number of occupants" is 998. Where "Number of occupants," the original variable is not an estimate, "Estimated number of occupants" should have a value which denotes "inappropriate." In no case should these variables be organized so that "Number of occupants" carries both counted and estimated values, while "Estimated number of occupants" carries, e.g., a 0 for observations in which the number of occupants was counted, and a 1 for observations in which the number of occupants was estimated.

File Organization

Variables. Records transmitted for possible inclusion in the CDS should contain no undocumented or irrelevant fields. The width of a variable should be sufficient to accommodate the entire range of variation that may be expected of the item, but should not be excessive. Conversely, fields need be no wider than is required to accommodate the maximum expected value of a variable. Thus, the variable "number of diagnoses" for a clinical record probably need not be larger than two characters, and certainly no larger than three.

In any case, the maximum allowable field width for an integer variable on tapes submitted to NIHR is 15 characters. Larger numbers should be represented by supplying a scaling factor[3] in the documentation, or by the use of floating-point format. A noncomputational variable may be up to 32,767 characters long.

Dates. The Federal Information Processing Standard (FIPS) format should be used for the recording of dates in machine-readable form. A date in FIPS format is expressed as six numeric characters, YYMMDD, where YY is the last two digits of the year, MM is the ordinal number of the month, and DD is the day of the month. In this format, February 3, 1983, is written as 830203. Where conciseness is paramount, dates may be expressed in the alternate FIPS format of YYDDD, where YY is the last two digits of the year, and DDD

---

[3] A scaling factor allows fewer characters to be used in representing a variable. For instance, $50,000,000 can be represented as "50" if the variable is defined as "Dollars in millions."

is the ordinal number of the day in the year. In this format, February 3, 1983, is written as 83034.

Record type identification. Each record in a file containing more than one record type, e.g., a hierarchical file, will carry a variable that identifies the type of record. The record identifier must be the first variable in each record, and must have the same format for all records in the file. Even where each record in a file has a different length, a variable identifying the type of record will be included. A new value of this identifier will be used whenever there is a significant change of the procedures used to generate that type of record. Such changes include changes not only in record layout, but also in instrument design, data collection, and coding. The need for a new record type is obvious when coding instructions or code values are changed. However, even when such changes consist only of the addition of new coding categories to existing variables, a new record type should be produced. Otherwise, analysts may not properly interpret the absence of a particular response.

Record identification items. Each record in a file submitted to NIHR must carry an identification number unique to the data file. If no existing variable will suffice as a unique identifier, then a sequence number will be assigned by the CDS library. Where a file contains more than one type of record, each record shall carry a variable identifying the type of record, as well as a unique sequence number. Where the records in a file represent a hierarchy or tree, a record will carry unique identification sufficient to identify it and its position in the hierarchy. In particular, it should not be necessary to infer the location of a record in a hierarchy solely from its position in the file.

For example, consider a file consisting of household, person, and incident records. Each record in the file must begin with the same four data items: a record type indicator, a household identifier, a person identifier, and an incident identifier. An incident record will carry a type and an incident identifier, as well as the identifiers of the person and household to which it belongs. A person record will carry type identifier, the identifier of the household to which it belongs, its own identifier, and a dummy incident identifier. Record identifiers should be positive integers. Dummy identifiers should be fields of zeros. Blank identification variables are not permitted.

In general, identifiers of lower-level records need be unique only within level, since concatenating identification variables generates a unique identifier.[1] Questions as to whether sequence numbers should be unique within record type or within levels should be resolved by agreement between the data supplier and NIHR before the file is generated.

The rationale underlying the assignment of a unique identifier to each record is that users of a data file should be able to perform arbitrary sorts and to subset the file without requiring the use of any facilities other than sort program and a file-copying utility.

Standardization Of Data Codes

Where possible, standard data codes should be used. There are no universal standards, but it is often possible to find an existing and well-constructed set of codes that is appropriate. There are two reasons for emphasizing the adoption of existing coding schemes. The first reason is that the adoption of standard coding schemes facilitates the linking of files. The second, and more important, reason is that the development of coding schemes for variables with large numbers of categories is a considerable project in itself. It is risky and uneconomical for a project to undertake the development of its own coding scheme for such things, as diagnoses, geographical locations, occupations, mental diseases, etc. The choice of which "standard" coding scheme is used is not so important as that *ad hoc* coding schemes not be used if equally good coding schemes are already in use.

Data suppliers subject to this standard are encouraged to consult the following sources of information on standard coding schemes.

Federal Information Processing Standards (FIPS). The Federal Information Processing Standards published by the National Bureau of Standards are the official Federal standards for several coding schemes. FIPS coding schemes should be considered for geographic area variables such as states, counties, Standard Metropolitan Statistical Areas (SMSAs), populated places, etc. FIPS incorporate several American National Standards Institutes (ANSI) standards which have been adopted for official Federal use. Data suppliers can obtain the current list of FIPS publications,

---

[1] To concatenate identifiers is to string them together into a single identifier. Consider a file of households with people in them. Suppose that each household has a three-digit identifier which is unique within the file, and each person has a two-digit identifier which is unique within his or her household. Then person number 4 in household number 207 is uniquely identified as person number 20704.

NBS Publications List 58, from the Office of Standards Administration, Institute for Computer Sciences and Technology, National Bureau of Standards, Washington, D. C. 20234, telephone (301) 921-3157.

Bureau of the Census. The Bureau of the Census has its own set of geographical area codes as well as a continually revised set of occupation and industrial codes. Data suppliers whose files are related to Census products should communicate with the Data User Services Division, Bureau of the Census, Washington, D. C. 20233 to receive copies of the latest Census coding schemes.

Office of Federal Statistical Policy and Standards. The Office of Federal Statistical Policy and Standards, U. S. Department of Commerce, serves as a central clearinghouse for data representations and coding schemes. Data suppliers who are unable to locate a satisfactory data coding scheme are encouraged to communicate with the Office of Federal Statistical Policy and Standards, U. S. Department of Commerce, Washington, D. C. 20230, telephone (202) 673-7956.

Abstract

When data files are submitted to the CDS, the producer must provide both NIHR and the CDS with an abstract of the file's form and contents. The abstract must include bibliographic and technical information sufficient to meet the requirements of the Office of Federal Statistical Policy and Standards concerning public use data files released by Federal Agencies. The writing of the abstract is entirely the responsibility of the file's producer. The abstract must contain all elements listed below which are relevant to the form and content of the file.

1. Date of the abstract.

2. Unique file number (producers number).

3. Bibliographic citation.

    3.1. Title. The title should be descriptive of the contents of the data file and should be followed by the date of the data program, if applicable. The title may be changed by NIHR if it does not clearly delineate the file from other files.

    3.2. Statement of responsibility. The agency or agencies responsible for the file's substantive content. If a contractor produces the file, both the contractor and the contracting agency should be included.

3.3. Place of production. The formal mailing address of the organization which produced the file.

3.4. Producer. The organization which actually produced the file. This entry is usually redundant to the statement of responsibility.

3.5. Date of production. The date the file became operational in a computerized form and available for possible inclusion in the CDS.

4. General description.

4.1. Type of file. Types of files include text files, aggregate data files , microdata files, etc.

4.2. Universe description. A description of the population covered by the data in the file, including what, if any, sampling method was used.

4.3. Subject matter description. A description of the types of variables included on the file and any other comments about the purpose of the data collection, the scope of the study, or any special characteristics of the study.

5. Time coverage. The time period to which the data in the file refer, including dates of collection.

6. Geographic coverage. Information concerning the total geographic scope of the file and additional levels of geographic coding provided within the file.

7. Technical description. A brief description of the file structure (hierarchical,rectangular, etc.), the file size (the approximate number of logical records in the file), and any information about the format of the file (SPSS, SAS) or dependency on specific computers or software.

8. Reference or related materials. The bibliographic citation of all materials that relate to the machine-readable file or its technical aspects such as the technical documentation or that include analysis, results, or tables derived from the data.

9. Related machine-readable files. The bibliographic citation of other related data or software files that are available to the public or qualified users.

10. Office of Management and Budget clearance number. The number given by OMB to any form(s) or plan(s) used in the data collection. Please indicate if OMB clearance was not required.

11. Terms of availability. Any restrictions on access to the file such as privacy certification and a contact person(s) in the producing organization who can discuss the substantive and computing aspects of the file. The name, full address, and telephone number of the contact person(s) must be included.

Documentation

Each tape submitted for possible inclusion in the CDS must be accompanied by documentation giving the physical characteristics of the volume and files as well as the logical composition of each type of record. Wherever possible, the documentation should be in machine-readable form and supplied both in hard copy and as a file on the tape. Machine-readable documentation is preferred because it prevents the separation of documentation from data and because the physical quality of the documentation will not be degraded by repeated copying.

Machine-readable documentation should be in printer image. If a document processing program is used to format the source text, it is requested that the source text of the documentation and a reference to the text processing software be included as well. (This standard makes no specification of any particular document processing program, nor does it require that such a program be used at all.) The minimal documentation of a data file consists of a tape volume table of contents, a character and octal or hexadecimal dump of a sample of records of each file, information sufficient to construct a bibliographic citation to the machine-readable data file, an abstract of the file's form and content, and a minimal codebook.

Where appropriate, data files should be accompanied by copies of the original collection instruments, including survey questionnaires and interview schedules. Copies of editing and coding instructions used in the creating the data file should also be included.

Tape Table Of Contents

The tape volume table of contents listing should include all information from the volume label and from the file labels. Information should be in an easy-to-read form, rather than a dump of the text of the labels. The following figure is an example of a table volume table of contents listing.

If possible, the tape table of contents should be produced by a program which also verifies the readability of the tape. If no tape listing program is available, then the tape table of contents should be produced manually, using information from the job which produced the tape or a dump of the tape. (Since tape listing programs are becoming increasingly available, it is suggested that data centers without such programs attempt to acquire them.)

Minimal Codebook

The codebook included with the file must contain at least the following information for each variable:

1. A reference number.

2. An unambiguous name for the item.

3. A textual description of the item, or the text of the question, if from a questionnaire.

4. The starting location, width, location of implicit decimal point, or scale factor.

5. Missing data codes and their meanings.

6. The mode in which the variable is represented, i.e., numeric character, alphanumeric string, floating-point binary, etc.

The codebook must also contain a list of the valid values for categorical items and valid ranges for continuous items. Missing data codes must be documented in the same fashion as other values and not left implicit.

Frequency Tables

A frequency distribution for each categorical variable must accompany each file submitted for possible inclusion in the CDS. The mean, standard deviation, range and number of cases of continuous variables should also accompany the file. Values that fall outside of those defined in the codebook should be annotated if they cannot be corrected.

```
+-----------------------------------------------------------------------------------------------------------
|IBM-LABELED 6250-BPI   9TP VOLUME=RSW009 OWNER=CAC.U-ILL  RACK#=C4524
|LP=ON  BLK=ON  RING=OUT DTCHK=ON  RETRY=10
|FILE                     BLOCK RECORD TAPELTH  RECORD            BLOCK LTH    CREATED    EXPIRES   USER      BATCH
|  # DATA SET NAME        COUNT  COUNT  (FEET)  FORMAT             AV.  MAX. DD MMM YY DD MMM YY I.D.      RECEIPT#
|
|  1 NCS73.NAT.CQ1         1233 125858  283.07 VB(15250.305)      15141 15250 23 MAR'77           SGDA
|  2 NCS73.NAT.CQ2         1234 126442  283.17 VB(15250.305)      15135 15250 23 MAR'77           SGDA
|  3 NCS73.NAT.CQ3         1187 123279  272.60 VB(15250.305       15146 15250 23 MAR'77           SGDA      691127
|  4 NCS73.NAT.CQ4         1212 125436  278.21 VB(15250.305)      15139 15250 23 MAR'77           SGDA      691127
|  5 NCS74.NAT.CQ1         1086 112135  249.39 VB(15250.305)      15139 15250 23 MAR'77           SGDA      691127
|  6 NCS74.NAT.CQ2         1083 112086  248.84 VB(15250.305)      15148 15250 23 MAR'77           SGDA      691127
|  7 NCS74.NAT.CQ3         1072 111090  246.18 VB(15250.305)      15138 15250 23 MAR'77           SGDA      691127
|  8 NCS74.NAT.CQ4         1100 113455  252.62 VB(15250.305)      15141 15250 23 MAR 77           SGDA      691127
|   TOTAL TAPE LENGTH = 2114.07 FEET
|<*><*><*> END OF TAPE  <*><*><*>
|
+-----------------------------------------------------------------------------------------------------------
```

Figure 1: A tape volume table of contents.

APPENDIX C

A Data Access and Archiving Policy for the National Institute of Handicapped Research

Richard C. Roistacher
Bureau of Social Science Research
Washington, DC 20036

A Data Access and Archiving Policy For the National Institute of Handicapped Research

Research Memorandum Number 4

Richard C. Roistacher
Bureau of Social Science Research
Washington, DC 20036

December 31, 1981

The National Institute of Handicapped Research has the mission of coordinating research and information on the nation's handicapped. Policy #13 of the National Council on the Handicapped's Policies Governing the National Institute of Handicapped Research states:

> Strategic support goal. In order to advance its goals for research and training, NIHR will initiate the development of a core capacity for in-house data management, knowledge transfer, technology evaluation and assessment, and public education designed to assure more effective planning and to accelerate translation of research results into practice by consumers, providers, and the public. This policy should be implemented in conjunction with Policy #32.

In order to manage data and transfer knowledge, it is necessary that NIHR have a policy concerning data access and archiving. Without such a policy, data collected at NIHR's expense may be lost to further use after the data's original purpose has been fulfilled. Much of the data on the handicapped is what can be called operational data. Operational data are subject to continual change as they are used to represent a current state of affairs. However, large amounts of static data are also collected. Static data are collected during a finite period of analysis and are then analyzed in a completed form.

Some static data are collected on a recurring basis as part of legislatively mandated statistical series. Others are collected as part of large one-time studies, while still others are collected by individual investigators under grant or contract.

In spite of the enormous amounts of money and effort invested in their collection, most data are analyzed only once and then forgotten. Data are left in whatever condition they were in at the time the initial analysis was completed: data and documentation are often destroyed or mislaid.

An increasing amount of emphasis is now being placed on "secondary analysis," research on data collected by others and for other purposes. Secondary analysis not only provides for economy, but also allows the testing and verification of earlier conclusions and the conduct of comparative research, even where such research was not contemplated by the original collectors of the data.

The policy and scientific potential of secondary analysis, as well as its great economy, make it important that all government data collection and analysis policies include a consideration of data archiving, secondary access, and secondary analysis. Indeed, present federal law provides that copies of all federal data sets be deposited with the National Archives. However, unless policies, standards, and procedures for implementing archiving and secondary access are produced, the depositing of tapes will be of little use to future researchers and analysts. This paper suggests a data archiving and accessing policy for the National Institute of Handicapped Research.

The Life Cycle Of Research Data

Research data begin as the responsibility of their collector. Even though the collector may be working under the strictest of mandates, it is impossible for the supervising agency to maintain any supervision of the actual data. Generally, a primary collection of data is followed by an assay of the file resulting from the primary collection. Where the assay shows systematic deficiencies, the collection methods are altered and collection proceeds in one manner or another to a conclusion. Most of the time, the sponsor's only contact with the data is through a report. In some cases, a copy of the data file may go back to the sponsor for in-house analysis. In a few cases, the sponsor and the data collector collaborate in a joint analysis of the data.

In most cases, once the report has been written, the data are forgotten. However, if proper policies have been implemented, data and documentation can be passed at relatively modest cost from the original analyst to a data preparation facility and thence to a final repository. In some cases, the sponsor may wish to use data after the conclusion of the original analysis project. In other cases, the sponsor has no immediate use for the data but wishes to make them available to others. The original analysis project should leave data and documentation in reasonable order and ready for at least minimal processing by an archival facility. At a minimum, the archive will assign bibliographic control information to the data and documentation, and copy them to an archival storage medium. In some cases, the archive facility may perform extensive assays and recodes of data, or may write extensive additional documentation. However, data collected under the aegis of a good secondary access policy should require relatively little work from the archiving facility. The data's permanent repository can be either an archive or a library. An archive provides prospective users with access to the data, but does not provide analytical services. A data library provides both data and services. In some cases, services are limited to reformatting and subsetting of data and documentation, while in other cases a full range of analytical services is offered.

Primary Data Use By NIHR

The National Council on the Handicapped's Policy #32 states:

> Data system. NIHR will develop a national data system designed to be 'comprehensive' from the point of view of NIHR's mission. NIHR will develop capacity to maintain access to, evaluate, utilize, store, and retrieve, on a cost-effective basis, selected primary data and data on other banks relevant to the mission of the NIHR. Such data will include, but may not be limited to demographic and related data on handicapped persons and data on the results of research. NIHR will take steps to make such data available, usable, and accessible to investigators supported by the NIHR and other Federal agencies, and, as appropriate, to consumers and providers and their organizations. NIHR may charge reasonable fees for such services.

NIHR thus is mandated to be the primary user and an archival facility for its own data. However, it is neither necessary nor optimal for NIHR to be responsible for all secondary use of its data.

There are five major categories of secondary users for NIHR data: NIHR itself; other government agencies; colleges and universities; industrial and public users; and posterity. For each of these categories of users, there is a primary depository for data. NIHR itself can maintain temporary control of data which are of current interest. Other government and public users would probably use National Technical Information Service as their supplier for government data. The Inter-University Consortium for Political and Social Research is the country's largest data and archive and distributor for academic users. Finally, the National Archives is the legal repository for government data and can provide for their redistribution to all categories of users.

Responsibility For Static Data

Dynamic data are those which are undergoing continuous revision either because they are part of an on-going collection effort, or because they are designed to reflect a changing situation. While the policies outlined here are applicable to dynamic data, such data are so strongly affected by local technical factors that a general data policy must be applied with extreme care. However, at some point a dynamic data base will become static, either because the collection effort has ceased or because the data has been abstracted from the dynamic data base for research or statistical purposes.

The Processing Cycle

A research data set passes through primary, archival, and secondary phases of processing. Primary processing is done by the collection or analysis organization. Primary processing should produce two classes of products: research reports and a file of data suitable for archival processing or secondary use. A NIHR policy on data archiving and access should insure that investigators produce a machine readable data file and accompanying documentation which conform to basic standards of syntax and style.

Archival processing includes the assay and editing of data, preparation of documentation in a form suitable for distribution, the addition of bibliographic identity and control. In some cases archival processing may include only a quality control check and the addition of bibliographic information. Large data collection projects may do their own archival processing; small projects may do some archival processing, but will seldom be able to write their own bibliographic control and cataloging information. Archival processing ends with the transfer of the machine readable data file and documentation to a permanent depository.

The NIHR data archiving and access policy should insure that data, documentation, and bibliographic control information conform to appropriate standards and that data files are deposited in the appropriate archives.

Secondary processing is done as part of analytical efforts independent of the initial analysis. A NIHR data archiving and access policy should be directed to insuring that machine readable data files, documentation, and bibliographic information are disseminated to the widest possible audience.

The policy should insure that to the greatest extent possible, secondary analyses can be carried out without assistance from the archive, from the primary analysis organization, or from NIHR.

Tasks And Responsibilities

A data archiving policy must assign responsibility in five major areas of data processing and use: setting of standards, quality control, processing and analysis, access, and archival deposit.

Setting of standards. NIHR must be responsible for setting of data and documentation standards for both primary and secondary analysis. Independent investigators should be provided with a set of minimal data and documentation standards included as part of their research grant or contract conditions. At the end of the grant or contract, they should provide NIHR with a tape of their research data accompanied by minimal documentation. Large contractual projects may be required to meet data and documentation standards up to and including preparation of complete archival files. The standards set by NIHR during the primary and archival processing provide the basis for standards during secondary processing, since the initial standard can only be maintained or degraded.

Quality control. During primary processing, responsibility for quality control of data rests entirely with the collector or analyst. Maintenance of data and documentation quality is intrinsic to the conduct of good science and cannot be delegated.

During archival processing, the archive facility may conduct additional assays of the data, both to serve as a reject on the original collection effort and to correct errors which may have evaded the original analyst.

Processing. Processing for analysis is intrinsic to the primary analysis and has no analog in the archival phase. A data archiving and access policy should be directed to seeing that methods of analysis are disclosed in order to facilitate replication and comparison. Analytical methods should be documented in research reports and publications.

Access. During primary processing, data are usually the intellectual property of the collector and analyst. In the case of large contractual efforts, NIHR may consider the contractor to be an agent, with ownership of the data residing directly with NIHR. During the primary analysis, access should be at the discretion of the data's owner.

After analytical processing, either an active or passive access to data may be provided. Passive access to data is much like that provided by a library. Data files and documentation are deposited in public facilities, and bibliographic information added to catalogs, information retrieval systems, and publication of data availability. However, the depository does nothing more than provide copies of tapes and documentation on demand. Active access to the data is provided by a facility which augments passive access by performing data processing and analysis tasks at the direction of clients. Active access to data requires not only additional skills on the part of the access facility, but also a continuing contractual relationship with clients. Passive access can be provided on the basis of one time fees for the copying of tapes and documentation.

Archive deposit. A data access and use policy should be directed at ensuring that copies of machine readable data files and documentation are deposited in institutions which will provide them to the widest possible appropriate audience. At the end of the primary processing, data should be deposited with an archive processing facility under contract to NIHR. The facility should have the mandate to assay data and to demand information necessary to improve files which do not meet the prescribed standard. Primary collectors and analysts should be required to respond to requests of primary archiving facility.

After archival processing, machine readable data files should be deposited in appropriate archives. At a minimum these archives should consist of a machine readable archive division of the National Archives and Records Service, the National Technical Information Service, and the Inter-University Consortium for Political and Social Research. In addition, bibliographic information concerning data and documentation should be disseminated to the library of congress. Such a dissemination policy will make NIHR data available to the widest possible audience, while at the same time relieving NIHR of the task of providing continuing data support to secondary analysts.

APPENDIX D

Data Files on the Demographic and Epidemiological Characteristics of Handicapped People

NATIONAL/FEDERAL

DEPARTMENT OF AGRICULTURE

- Nationwide Food Consumption Survey (1125)

DEPARTMENT OF COMMERCE

- Bureau of the Census
  - 1980 Decennial Census of Population and Housing (1020)**
  - 1970 Decennial Census of Population and Housing (1025)
  - Current Population Survey (1021)**
  - Annual Housing Survey (1022)
  - 1976 Survey of Income and Education (1023)**
  - 1967 Survey of Economic Opportunity (1066)
  - Continuous Longitudinal Manpower Study (CETA) (1085)**
  - Research Panel, Income Survey Development Program, 1978 and 1979 (1127)
  - Post-Census Disability Survey: Richmond Pretest (1024)

DEPARTMENT OF DEFENSE

- Office of Civilian Health
  - Medical Program of the Uniformed Services (CHAMPUS) Data Base (1175)

DEPARTMENT OF EDUCATION

- Bureau of Education for the Handicapped
  - National Speech and Hearing Survey (1083)**
  - Handicapped Children's Early Education Program (1301)
  - Bureau of Education for the Handicapped Counts (1166)
- Office of Special Education

  Child Count Data Base (1302)
- Rehabilitation Services Administration
  - State-Federal Program of Vocational Rehabilitation (1101)**
  - National Spinal Cord Injury Data Base (1109)**
  - 1974 Sheltered Workshop Survey, Greenleigh Asso. (1182)
  - Survey of Hearing Impaired and Mentally Retarded Persons in State Institutions (1170)
  - 1969 National Survey of Medicaid Impact (1160)
  - 1973-74 Comprehensive Service Needs Study (Urban Institute for the Rehabilitation Services Administration) (1046 to 1048, see below)
  - Vocational Rehabilitation Rejection Survey (1046)**
  - CMRC Service Recipient Survey (1047)**
  - Survey of Providers of Rehabilitation (1048)
  - RSA-SSA Data Link (1027)

---

*Cover page available.

**Detailed discription available.

- Office of Planning, Budget, and Evaluation/Office of Education
  - Assessment of Selected Resources for Severely Handicapped Children and Youth (1086)
  - Project VO-ED: A National Survey (1164)
- Administration on Developmental Disabilities
  - State plans for Developmental Disabilities (1348)

DEPARTMENT OF HEALTH AND HUMAN SERVICES (DHHS)

- Administration of Aging
  - National Study to Assess the Service Needs of the Hispanic Elderly (1123)
  - Identifying Opportunities for Improving the Quality of Life of Older Americans (1121)
  - National Survey of the Aging (1124)
  - Program Research for Serving the Indian Elderly (1126)
  - With a Little Help From My Friends (Two National Telephone Surveys of the Elderly) (1130)
- Health Care Financing Administration
  - Medicare Bill File (1100)
  - Early and Periodic Screening, Diagnosis, and Treatment Data (1171)
- Office of Human Development Service--Administration of Children, Youth, and Families
  - Annual Survey of Headstart Handicapped Efforts (1973 to date) (1044)*
  - National Survey of Agencies Serving the Developmentally Disabled (1058)**
  - Social Services Reporting Requirements (1094)
  - Follow-up Study of Former Vocational Rehabilitation Clients (1120)
  - Developmental Disabilities Project on Residential Services and Community Adjustment (1118)**
- Office for Civil Rights
  - Survey of Individuals in Residential Institutions (in planning) (1061)
  - Children and Youth Referral Survey (1304)**
  - Elementary and Secondary School Civil Rights Survey (1165)*
  - Special Purpose Facilities (1306)
  - Survey of Impaired Individuals in Households (1062) (In Planning)

DHHS--PUBLIC HEALTH SERVICE

- National Center for Health Statistics
  - National Survey of Personal Health Practices and Consequences (1006)**
  - Health Interview Survey (HIS) (1001)**
  - Health and Nutrition Examination Survey (HANES) (1003)**
  - The National Hospital Discharge Survey (1007)
  - Resident Places Survey I - Nat'l Survey of Nursing Homes - 1963 (1008)
  - Resident Places Survey II - " " " - 1964 (1009)
  - Resident Places Survey III - " " " - 1969 (1010)
  - 1969 National Nursing Home Survey (1011)

- 1973-74 National Nursing Home Survey (1012)
- 1977 National Nursing Home Survey (1013)**
- 1976 Survey of Institutionalized Persons (1014)**
- 1970 Health Services Utilization and Expenditure Survey (1112)**
- Master Facilities Survey (1019)
- Health Services Utilization and Expenditures (1112)
- 1975 National Ambulatory Medical Survey (1015)**
- National Natality Surveys (1016)*
- National Infant Mortality Surveys (1017)*
- National Mortality Surveys (1018)*
- Congenital Anomalies and Birth Injuries Among Live Births: United States, 1973-74 (1184)

- Health Resources Administration
  - BCHS Reporting Requirements (1028)
  - BCHS Medically Underserved Area Data Base (1029)
- Health Services Administration
  - Indian Health Service Health Care Statistics System (1308)
- Center for Disease Control
  - Birth Defects Monitoring Program (BDMP) (1309)*
- National Center for Health Services Research
  - 1977 National Medical Care Expenditure Survey (1073)**
  - 1970 Survey of Health Services Utilization and Expenditure (1112)**
  - 1978-79 National Medical Care Utilization and Expenditure Study (1056)
  - Analysis of National Survey Data on Access to Medical Care (1026)
- ADAMHA
  - Survey of Patient Characteristics--State and County Mental Hospital (1200)
  - Sample Survey of Additions to Inpatient Services--State and County Mental Hospitals (1201)
  - Sample Survey of Inpatient Additions--Private Psychiatric Hospitals (1202)
  - Sample Survey of General Hospital Inpatient Psychiatric Unit Discharges (1203)
  - Inventory of Comprehensive Community Mental Health Centers (1204)*
  - Inventory of Transitional Mental Health Facilities (1205)
  - Sample Survey of Additions to Outpatient Psychiatric Services or Clinics (1206)
  - Inventory of General Hospital Psychiatric Services (1209)
  - Sample Survey of Additions to Partial Care Programs (1210)
  - Sample Survey of Clients Under Care in Outpatient Psychiatric Services or Clinics (1211)
  - Sample Survey of Clients Under Care in Partial Care Programs (1212)
  - Sample Survey of Admissions to Veterans Administration Medical Centers (1213)
  - Sample Survey of Discharges from Residential Treatment Centers for Emotionally Disturbed Children (1214)
  - Biomedical and Behavioral Study of Aging (1167)
  - The Epidemiologic Catchment Area (ECA) Program (1183)

- Inventory of Mental Health Facilities (1208)
- Mental Health Demographic Profile System (1207).

DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

- Longitudinal (ten-year) Survey of Housing Adjustments (in planning) (1059)
- Public Housing Reporting System (1104)

DEPARTMENT OF LABOR

- Rearranged Work Schedules for Handicapped Employees in the Private Sector (1045)
- 1973-76 Sheltered Workshop Study (1051)
- National Longitudinal Surveys (1105)
- Survey of Disabled Veterans of the Vietnam Era (1049)
- BLS Occupational Injury and Illness Records (1169)**

DEPARTMENT OF TRANSPORTATION

- Survey of Transportation Problems of Transportation Handicapped Persons in Urban Areas (1129)
- 1976-77 National Survey of Transportation Problems of Handicapped Persons (1097)

ACTION

- Senior Companion Program (SCP) Evaluation (1128)
- Foster Grandparent Program Evaluation (1151)

LIBRARY OF CONGRESS

- 1974 Braille Reader Survey (1050)

OFFICE OF PERSONNEL MANAGEMENT

- OPM Central Personnel Data File (CPDF) (1093)

SOCIAL SECURITY ADMINISTRATION

- Continuous Work History Sample (1030)
- Longitudinal Sample of Disability Insurance Applicants (1031)**
- Continuous Disability History Sample (1032)**
- 1969-79 Retirement History Study (1033)**
- 1966 Survey of Disabled Adults (1034)**
- 1967 Survey of Institutionalized Adults (1035)**
- 1968-70 Survey of Newly Entitled Beneficiaries (1036)**
- 1969 Followup of Disabled Adults (1037)
- 1971 Survey of Recently Disabled Adult (1038)
- 1972 Survey of Disabled and Nondisabled Adults (1039)**
- 1974 Followup of Disabled and Nondisabled Adults (1040)**
- 1974 Survey of Low Income, Aged, and Disabled (1041)**
- 1978 Survey of Disability and Work (1042)**

- 1978 Survey of Blind and Disabled Children Receiving Supplemental Security Income Benefits (1043)**
- Supplemental Security Income (SSI) Program (1103)
- Retirement, Survivors, and Disability Insurance Claims Study (1102)
- Survey of Income and Program Participation (in planning) (1060)
- AFDC Recipient Surveys (1174)**
- Feasibility Assessment for a Longitudinal Survey of the Disabled (in planning) (1155)
- Alternative Practices in Disability Beneficiary Rehabilitation (in planning) (1154)
- Work Incentive Experiments (in planning) (1153)
- Medical Examination Study of Impairments and SSDI Eligibles (in Planning). (1152)

VETERANS ADMINISTRATION

- Veterans Administration Annual Patient Census (1089)
- Housing Assistance Program (1106)
- Condensation and Pension Reporting System (1107)
- 1971 VA/Harris Survey of Vietnam Era Veterans (1172)
- 1979 VA/Harris Survey of Vietnam Era Veterans (1173)

NATIONAL/NON-FEDERAL

- American Association of University Affiliated Programs
  - American Association of University Affiliated Programs Data Base (2315)
- American Foundation for the Blind (AFB)
  - Print-Limited Population Survey (2312)
  - Consumer Needs Survey of the Visually Impaired (2311)
- Association of State and Territorial Health Officials
  - National Public Health Program Reporting System (2317 and 2318, see below)
  - Inventory of Programs and Expenditures of State and Territorial Health Agencies (2317)
  - Maternal and Child Health Supplement (2318)
- Epilepsy Foundation of America
  - Basic Statistics on the Epilepsies (2322)
- Gallaudet College
  - Annual Survey of Hearing Impaired Children and Youth (2098)**
  - National Hearing Impaired Achievement Testing Projects (2099)
  - Survey of Families of Hearing Impaired Children (2305)*
- National Association for the Deaf
  - American Annals of the Deaf (2319)
  - National Census of the Deaf Population (2310)**
- National Association for Retarded Citizens
  - 1971 Survey of Mentally Retarded Students (2180)
  - Followup Survey of Mentally Retarded Citizens (2181)
- National Association of State Directors of Special Education
  - State Profiles in Special Education (2332)
- National Committee for Citizens in Education
  - Survey on Parent Involvement in the IEP Process (2334)
- National Easter Seal Society
  - National Easter Seal Society Program Services Reporting System (2329)*
- National Tuberous Sclerosis Association, Inc.
  - Tuberous Sclerosis Medical registry (2335)*

- United Cerebral Palsy Association, Inc.
  - Programming for Atypical Infants and Their Families (2331)*
  - Data Retrieval System (2303)

- Paralyzed Veterans of America
  - Survey of SCI Veterans '79 (2185)

- Other
  - Birth-Defects Information System (BDIS) - Tufts-New England Medical Center (2323)*
  - Institute for Child Behavior Research (ICBR) Data Bank--Institute for Child Behavior Research (2328)
  - 1972 Ohio State University National Survey--Ohio State University (2065)**
  - Idiot Savant Survey--N.Y. State Institute for Basic Research (2071)
  - National Survey of Handicapping Conditions Found in Juvenile Correctional Institutions - S.C. Department of Correctional Institutions (2074)

MULTIGEOGRAPHIC AREA

- Multistate Information (Mental Health) System--DHEW (3064)
- 1968-71 Baseline Surveys of Urban Health Center Service Areas (3096)
- RAND Health Insurance Study (for DHEW) (3108)**
- Multistate (California, Colorado, Nevada) Survey of Residential Status of Developmentally Disabled Persons (3084)
- Experimental Health Services Delivery Systems Family Health Survey (for DHEW) (3111)
- Collaborative Perinatal Project, National Institute of Neurological and Communicative Diseases and Stroke (3339)*
- Computerized Patient Profile System, National Foundation of Dentistry for the Handicapped (3340)*
- Individual Data Base (IDB) Mental Retardation Research Center (3341)*
- Juvenile Rheumatoid Arthritis Data Bank, Special Treatment Center for Juvenile Arthritis (3342)*
- Kent Infant Development (KID) Scale Protocols, Kent Development Metrics, Kent State University (3343)*
- New York Hospital--Cornell Medical Center Regional Comprehensive Hemophilia Diagnostic and Treatment Center Data Base (3344)*
- Patient Registry for Actual or Suspected Epilepsy, Comprehensive Epilepsy Program (3345)*
- Protection and Adcocacy (P & A) Data Retrieval System, Institute for Comprehensive Planning (3346)*
- Southeast Registry of Deaf-Blind Children and Youth, The Southeast Regional Center for Deaf-Blind Children (3347)*
- 1975 Followup of 1968-71 Baseline Survey of Urban Health Center Service Areas--NCHSR (3162)
- Long-term Care Reimbursement Experiments-Evaluation of Homemaker and Day Care Services--NCHSR (3147)
- Cost Effects and Benefits Associated with Domiciliary Care and Intermediate Care--Administration on Aging (3117)
- Informal Social Networks in Support of Elderly Blacks in the Black Belt--Administration on Aging (3122)
- Independent Living Rehabilitation (ILR) Projects--Urban Institute (3177)

STATE

ALASKA

- Registry of Deaf-Blind Children (4401)*

ARIZONA

- DES-DDDMRS Client Information System (4402)*

ARKANSAS

- Crippled Children's Services Programs (4403)*
- Service Management Project (4133)

CALIFORNIA

- Assessment of California State University, Northland Higher Education Mainstreamed Deaf Program (4068)
- California Survey of School Administrators: Mainstreaming Orthopedically Handicapped Children (4081)
- Children Services Caseload and Expenditure Information (4404)*
- California Disability Survey (4178)
- Case Reporting of RH(D) Hemolytic Disease (4417)*

COLORADO

- Client Oriented Record Entry (CORE) (4416)*
- Reporting System for Handicapped Students (4405)*

FLORIDA

- Client Information System (4406)*

ILLINOIS

- Developmental Disabilities Information System (4407)*

IOWA

- Iowa Survey of the Severely Disabled (4078)**
- Services for Handicapped Children and Youth (4408)*

KANSAS

- Adult Restorative Services--Administration on Aging (4131)

KENTUCKY

- Kentucky Assessment of Vocational Education Needs of the Disadvantaged and Handicapped (4075)

LOUISIANA

- Louisiana State Survey of Hearing Impaired Children and Youth (Gallaudet College) (4054)
- Louisiana Survey of Hearing Imparied and Deaf-Blind Children and Youth (4409)*

MASSACHUSETTS

- 1976 Needs Assessment Study--Massachusetts Commission for the Blind (4052)
- Division of Family Health Services Client Data Base (4410)*
- Massachusetts Service Coordination Battery (4411)*

MINNESOTA

- Minnesota Disability Assessment (4070)
- Survey of Handicapped Individuals' Housing Preferences (4090)
- Services for Children with Handicaps Management Information System (4412)*

NEBRASKA

- Nebraska State Survey of Hearing Impaired Children and Youth (Gallaudet College) (4057)**

NEW YORK

- New York Developmental Disabilities Information Survey (DDIS) (4413)*
- Automated Goal-Oriented Medical Records for the Mentally Retarded (4063)

OHIO

- 1977 Ohio Telephone Survey on Developmental Disabilities (4053)

PENNSYLVANIA

- Pennsylvania Survey of Postsecondary Educational Status of Physically Disabled Adults (4082)
- Pennsylvania Followup of the Physically Handicapped (4087)
- Pennhurst Telephone Survey on Attitudes Toward the Mentally Retarded (4088)
- Patient Client Information System (PCIS) (4414)*

RHODE ISLAND

- Brown University Health Surveillance Studies--Population Research Laboratory (4161)

TEXAS

- Texas State Survey of Hearing Impaired Children and Youth (4077) Gallaudet College
- Handicaped Information Center (4415)*

WISCONSIN

- Wisconsin Survey of Paratransit Systems Serving the Elderly and Handicapped (4055)
- Wisconsin Assessment of Vocational Needs of Handicapped Students (4076)

LOCAL

ARIZONA

- Community Services Program (5132)

ARKANSAS

- Western Arkansas Study on Aging; USDA (5150)
- Hot Springs Rehabilitation Center Followup of Former Rehabilitation Clients (5156)

CALIFORNIA

- UCLA Intervention Program for Developmentally Handicapped Infants and Young Children (5427)*
- San Diego County Teratogen Registry (5425)*
- Family and Long Term Care--NCHSR (5144)
- Aging Needs Assessment (5135)
- USC Gerontology Retirement Study (5157)
- Stanford Research Institute Interview Data on Chronic Conditions Compared with Information Derived from Medical Records, NCHSR (5429)
- Follow-Up Study of Spinal Cord Injury Patients (5436)

CONNECTICUT

- Epidemiologic Catchment Area Program, NIMH (5143)
- New Haven City Government Employee Survey (5176)

DISTRICT OF COLUMBIA

- The Mental and Nervous Disorder Utilization and Cost Survey (5216)

GEORGIA

- Metropolitan Atlanta Congenial Defects Program (5421)*

HAWAII

- Survey of Children's Attitudes Toward Handicapped Peers (5092)

ILLINOIS

- A Study of the Medical and Allied Health Services Delivery System for Substantially Handicapped Developmentally Disabled Adults, Chicago (5115)**
- Community Ecology and Adaptation of Older Persons, Chicago, National Institute on Aging (5140)
- Pre-Start Data Collection/Retrieval System for Perinatal Project (5423)*
- Metropolitan Chicago Perinatal Data Project (5422)*
- Cook County Survey of Teacher Attitudes Toward Handicapped Children (5067)
- Transportation Needs of the Mobility Limited, Illinois University (5217)

KENTUCKY

- Health Income and Related Problems of the Elderly in Powell County (5146)

MARYLAND

- Baltimore Health Maintenance Program Evaluation Study (5136)
- Howard County Handicapped Needs Assessment (5080)
- Caregiver Training for the Elderly Impaired in Rural and Urban Settings (5137)
- Baltimore Longitudinal Study--The Normality of Aging (5434)
- Ambulatory Care for Chronic Conditions in an InnerOCity Elderly Population (5435)

MASSACHUSETTS

- Decision-Making for Home Care (5142)
- Geriatric Assessment and Research Center Model Project (5145)
- Framinghan Heart Study: Functional Disability Study (5116)
- Transportation Needs of the Handicapped, Boston (5114)

MICHIGAN

- Community Care: The Chronic Disease Service Module, NCHSR (5139)
- Roles for Homes for the Aged in Meeting Community Needs (5148)

MINNESOTA

- Rochester Project, Mayo Clinic (5424)*

NEW JERSEY

- An Assessment of the Unmet Needs of the Aged Blind Population in New Jersey (5186)

NEW YORK

- Community Based Project for the Mentally Frail Elderly (5138)
- Fetal Life Study - Perinatal Data Book (5420)*
- Automated Developmental Disabilities Out-Patient Treatment Review System (5418)*
- Survey of Landlord Attitudes Toward Housing for Mentally Retarded Persons (5072)
- Aging as a Rural Phenomenon, Administration on Aging (5134)
- Day Hospital Service in Rehabilitation Medicine (5141)
- The Elderly in the Inner City, NYC, Administration on Aging (5119)
- The Health Insurance Plan of Greater New York (HIP) Study of Health Interview Responses Compared with Medical Records (5428)
- The Job Hunt of the Disabled; An Exploratory Study (5179)
- Experimental Transportation for Elderly and Disabled, NCHSR (5431)

NORTH CAROLINA

- Epidemiology of Dysphoria and Depression in an Elderly Population (5215)

2

OHIO

- Miami (Ohio) Valley Health Service Area Survey (5158)
- Survey of Older Americans Living in Cleveland, Ohio - 1976 (5430)

PENNSYLVANIA

- Hershey Household Health Survey (5095)
- Pennsylvania Special Education Followup Employment Survey (5079)
- Longitudinal Chronic Care Study in an Acute Care Hospital (5163)
- Team Demonstration Project for Noninstitutional Care for the Elderly, Philadelphia (5149)

TENNESSEE

- Tennessee Household Survey on Effectiveness of Health Services NCHSR (5159)
- Health Status, Household, and Community Integration, Nashville (5432)

TEXAS

- 1979 Survey of Disability in Houston, Texas (5110)
- Statistics on Spinal Cord Injury Patients (5091)
- Longitudinal Follow-Through Study of 88 Deaf Children (5437)

VIRGINIA

- Community Services - Northern Virginia (5419)

WASHINGTON

- Uniform Performance Assessment System, Child Development and Mental Health Center, Seattle (5426)*

INTERNATIONAL

- Multinational (U.S. & Canada) Survey of Disruptive Behavior Group Care Facilities (7069)
- International Collaborative Study of Medical Care Utilization (WHO) (6113)
- International Pilot Study of Schizophrenia (IPSS) (6168)**

APPENDIX E

The Definition of Measurement and Disability: A Selected Bibliography

Lucy W. Duff, Mary K. Hartz, Bruce B. Dunning
and John Terry Chase

The Definition and Measurement of Disability
A Selected Bibliography

This bibliography was prepared as background material for Research Memorandum One and is not intended as exclusive coverage of all the materials available in this area. Materials chosen simply reflect the purpose of the first phase of the comprehensive data system project: to review the thinking of researchers, policy-makers, and practitioners on how to conceptualize, classify, and measure disability. Emphasis is on disability in the generic sense, across traditional diagnostic and impairment categories. The following and somewhat arbitrary classification was used to organize the bibliography:

I. The Concept Of Disability--General

A. Conceptualization And Classification

Badley, Elizabeth M. "Implications of classification schemes for information flow in caring systems." Paper presented at symposium of Society for Social Medicine, September 1976.

Berkowitz, Edward D., ed. Disability Policies and Government Programs. New York: Praeger, 1979.
Several chapters focus on definitions and classifications. See especially those by Saad Z. Nagi, Edward D. Berkowitz, John J. Dempsey and Jerry L. Mashaw.

Coet, L. T. "Defining term handicap--function of sex, race, religion, and geographic location." Psychological Reports 41:783-87, 1977.

deJong, G. Movement for Independent Living: Ideology, and Implications for Disability Research (Occasional paper no. 2). Boston: Tufts University Medical Rehabilitation Research and Training Center and East Lansing: Michigan State University, Center for International Rehabilitation, 1979.

Granger, Carl V. "Fellowship report [to World Rehabilitation Fund]." 1979.
Fellowship project explored British approaches to disability terminology development, measurement of severity of disability and uses of epidemiologic methods for investigating disability.

Haber, Lawrence D. "Disabling effects of chronic disease and impairment." Journal of Chronic Diseases 24:469-87, 1971.

Another edition appears as: "Epidemiological factors in disability: I. Major disabling conditions." In Compilation of Social Security Survey of the Disabled Population: 1966 (Report no. 6). Washington: U.S. Social Security Administration, Office of Research and Statistics, 1977.

Haber Lawrence D. "Haber Session." In transcript of Workshop Conference on Disability (San Juan, 1969). Sponsored by American Sociological Association, Medical Sociology, Section, Committee on Disability, and Instituto Psicologico de Puerto Rico, 1969.

Discussion centers on the role of social agencies in labelling disabled clients.

Haber, Lawrence D. "Identifying the disabled: concepts and methods in the measurement of disability." Social Security Bulletin 30:17-34 (Dec.), 1967.

Haber, Lawrence D. "Social planning for disability." Journal of Human Resources 8:33-55 (Supp., Feb.), 1973.

Hershenson, David B. "A three-dimensional model for client evaluation." Rehabilitation Counseling Bulletin 20:308-10 (June), 1977.

Koshel, Jeffrey J. and Carl V. Granger. "Rehabilitation terminology: who is severely disabled?" Rehabilitation Literature 39:102-06 (April), 1978.

Nagi, Saad Z. Natural History of Disability. Columbus: Ohio State University, 1971.

Nagi, Saad Z. Disability and Rehabilitation--Legal, Clinical, and Self-Concept and Measurement. Columbus, Ohio State University Press, 1969.

Nagi, Saad Z. Disability in the United States: A Plan for an Informational System. Washington: Office for Handicapped Individuals, 1977.

Nagi, Saad Z. "Epidemiology of disability among adults in the United States." Milbank Memorial Fund Quarterly 54:439-67 (Fall), 1976.

Nagi, Saad Z. "Gate-keeping decisions in service organizations: when validity fails." Human Organization 33:47-58, 1974.

Nagi, Saad Z. "A study in the evaluation of disability and rehabilitation potential: concepts, methods, and procedures." American Journal of Public Health 54:1568-79, 1964.

Neisworth, John T. and Robert T. Smith. "An analysis and redefinition of 'developmental disabilities.'" Exceptional Children 40:345-47 (Feb.), 1974.

Noble, John H. "Social inequity in he prevalence of disability: projections for the year 2000." In Population and Development Problems Relating to Disability Prevention and Rehabilitation, pp. 23-32. Geneva: World Health Organization, 1980

Rehab, Inc. An Analytical Appraisal of the Definition of Severe Handicap. 2v. Arlington, VA, 1976.

Salbreux, R., J. M. Deniaud, S. Tomkiewicz and M. Manciaux. "Classification and prevalence of severe and multiple handicaps." Neuropsychiatrie de l'Enfance et de l'Adolescence 27,1-2:5-28, 1979.

Seethamma, Hampapur N. and Ranjit K. Majumder. "New rehabilitation law requires a clearer definition of handicap." Rehabilitation Literature 40:340-41f (Nov.-Dec.), 1979.

Sunshine, Jonathan. Disability (OMB Staff Technical Paper). Washington: U.S. Office of Management and Budget, 1979.
Describes definitions used by, and scope of, U.S. disability programs.

U.S. Working Group on Uniform Definition of Disability. "Proposals to reduce the inconsistencies in concepts, criteria, and definitions of disability and handicap--decision memorandum [and attachments]." Washington: U.S. Dept. of Health, Education, and Welfare. Assistant Secretary for Planning and Evaluation, 1978.
Decision memorandum, drafted by Henry Aaron, presents findings and recommendations of the Working Group. Attachments include John DeWilde's "Discussion paper: disability definitions," Geoffrey E. Bock's "A compilation of definitions pertaining to the handicapped and the disabled in federal statutes and regulations," and drafts and summary of several agency commentaries on the proposals of the Working Group.

U.S. National Institute of Handicapped Research. Definitions Task Group. "Definitions for NIHR, general and specific: government usage." (Draft). Washington, n.d.

Whitten, E. B. ed. *Pathology, Impairment, Functional Limitation and Disability--Implications for Practice, Research, Program and Policy Development and Service Delivery*. Report of the First Mary E. Switzer Memorial Seminar (Cleveland, 1975). Washington: National Rehabilitation Association, n.d.

Wood, Philip H. N. "Appreciating the consequences of disease: the international classification of impairments, disabilities and handicaps." *WHO Chronicle* 34, 10:376-80 (Oct.), 1980.

Wood, Philip H. N. and Elizabeth M. Badley. "Setting disablement in perspective." *International Rehabilitation Medicine* 1, 1:32-37, 1978.

Wood, Philip H. N. and Elizabeth M. Badley. *People with Disabilities--Toward Acquiring Information Which Reflects More Sensitively their Problems and Needs* (Monograph No. 12). New York: World Rehabilitation Fund, 1981.

World Health Organization. *International Classification of Diseases*. 2v. Geneva, 1967.

World Health Organization. *International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease.* Geneva. 1980.

Manual "published for trial purposes in accordance with resolution WHA 29.35 of the Twenty-Ninth World Health Assembly, May 1976."

B. Measurement Methodology

Ableson, Janet. "Measurement of disability in the Canada Health Survey." Ottawa: Health and Welfare Canada, 1978.

Alexander, J., Marcus J. Fuhrer and C. D. Rossi. 1979 Survey of Disability in Houston, Texas: research project report (project duration: Sept. 1979 to Sept. 1981). Houston: Baylor College of Medicine Medical Rehabilitation Research and Training Center, n.d.

Attkisson, C. C., W. A. Hargreaves, and W. A. Horowitz eds. *Evaluation of Human Service Programs*. New York: Academic Press, 1978.

Badley, Elizabeth M., Roger P. Thompson and Philip H. N. Wood. "The prevalence and severity of major disabling conditions--a reappraisal of the Government Social Survey on the Handicapped and Impaired in Great

Britain." International Journal of Epidemiology 7, 2:145-51, 1978.

Bolton, Brian, ed. Handbook of Measurement and Evaluation in Rehabilitation. Baltimore: University Park Press, 1976.

Bolton, Brian. Introduction to Rehabilitation Research. Springfield: C. C. Thomas, 1974. Bunten-Mines, Elaine. Measurement Standards for the Handicapped (Planning report 4). Washington: U.S. National Bureau of Standards, 1980.

Burns, J. R. and W. M. Marcy. Human Rehabilitation Techniques: Modeling Rehabilitation (Volume VI). Lubbock: Texas Tech University, 1977.

Ellis, David. "Methods of assessment for use with the visually and mentally handicapped: a selective review." Child Care, Health and Development 4:397-410 (Nov.-Dec.), 1978.

Haber, Lawrence D. "Disabling effects of chronic disease and impairment II: Functional capacity limitations." Journal of Chronic Diseases 26:127-51 (April), 1973.

Another edition appears as: "The epidemiology of disability: II. The measurement of functional capacity limitations." In Compilation of Social Security Survey of the Disabled Population: 1966 (Report no. 10). Washington: U.S. Social Security Administration, Office of Research and Statistics, 1977.

Koss, Margo, et al. Social Services: What Happens to the Client? Washington: Urban Institute, n.d.

Presents procedures for monitoring changes in client functioning and satisfaction.

Luft, Harold S. "Household surveys and causes and consequences of disability." Journal of Occupational Medicine 17:173-76 (March), 1975.

Luft, Harold S. "Probability of disability--the influence of age, race, sex, education and income." Paper presented at the 1975 Annual Meeting of the American Public Health Association (Chicago), 1975.

Nicholls, William L. II. Technical Report for the California Disability Survey. Berkeley: University of California, Survey Research Center, 1979.

Pelzer, Inge and Mayeda Tadashi. Item Contents of Performance Measurement Instruments for the Developmentally Disabled. Los Angeles: University of California, Neuropsychiatric Institute Research Group, 1980.

Reagles, Kenneth W. and Alfred S. Butler. "The human service scale: a new measure for evaluation." *Journal of Rehabilitation* 42:34-8 (May-June), 1976.

Solomonidis, Dorothee. "Baseline data for the psychotechnological examination of the handicapped." *Bulletin du C.E.R.P.* (Centre d'Etudes et Recherches Psychotechniques) 23:199-215 (July-Sept.), 1975.

Sullivan, Daniel F. "Disability components for an index of health." Hyattsville, MD: National Center for Health Statistics, 1971.

U.S. National Committee on Vital and Health Statistics. *Comparison of Data Elements from Ten National Center for Health Statistics Data Sets*. Hyattsville, MD, 1977.

U.S. National Committee on Vital and Health Statistics. *Long-Term Health Care Minimum Data Set: Report*. Hyattsville, MD, 1980.

Wolinsky, Fredric D. and Marty E. Zusman. "Toward comprehensive health status measures." *Sociological Quarterly* 21:607-21 (August), 1980.

Wright, George N. et al. "Who are the tough rehabilitation cases?" Madison: University of Wisconsin, Regional Rehabilitation Research Institute, 1970.

II. Disability As Functional Limitation

A. Conceptualization And Classification

Brown, Margaret et al. *Rehabilitation Indicators: A Method for Enhancing Accountability and the Provision of Rehabilitation Services.* (A series of project methodological reports.) 1979.

Brown, Margaret et al. "Rehabilitation indicators: their nature and uses for assessment." In *Rehabilitation Client Assessment*, pp. 102-17. Edited by Brian Bolton and D. W. Cook. Baltimore: University Park Press, 1980.

Bury, Michael R. and Philip H. N. Wood. "Sociological perspectives in research on disablement." *International Rehabilitation Medicine* 1, 1:25-32, 1978.

Haber, Lawrence D. "Some parameters for social policy in disability: a cross-national comparison." *Milbank Memorial Fund Quarterly* 51:319-40 (Summer), 1973.

Haber, Lawrence D. and Richard T Smith. "Disability and deviance: normative adaptations of role behavior." *American Sociological Review* 36:87-97, 1971.

Indices, Inc. *Functional Limitations: A State of the Art Review*. Falls Church, VA, n.d.

Raynes, N. V. "Issues in the assessment of the health status, needs, and treatment of the mentally and physically handicapped." *Medical Care* 14:53-9 (5 Supp., May), 1976.

Slater, Samuel B. et al. "The definition and measurement of disability." *Social Science and Medicine* 8:305-08, 1974.

Sussman, Marvin B., ed. *Sociology and Rehabilitation*. American Sociological Association, 1965.

Includes several chapters on concepts and models of disability, among them ones by Saad Nagi, Eliot Freidson, Robert Straus and Edward A. Suchman.

Twaddle, Andrew C. "The concept of health status." *Social Science and Medicine* 8:29-38 (Jan.), 1974.

Twaddle, Andrew C. "Illness and deviance." *Social Science and Medicine* 7:751-62 (Oct.), 1973.

Twaddle, Andrew C. *Sickness Behavior and the Sick Role*. Boston: G. K. Hall, 1979.

Urban Institute. *Report of the Comprehensive Service Needs Study*. Washington: U.S. Rehabilitation Services Administration, 1975.

Discusses issues in defining classes of the handicapped and methods of measuring their incidence.

B. Measurement Methodology

Bassett, P. T. "Measurement of outcomes: a report." Paper presented at the 1st Institute on Rehabilitation Issues, Denver. Morgantown: West Virginia University, Regional Rehabilitation Research and Training Center, 1975.

Fuhrer, Marcus J., David Cardus and Robert M. Thrall. "Proposed rehabilitation research and demonstration projects: estimating target population size." *Archives of Physical Medicine and Rehabilitation*62, 4:156-60 (April), 1981.

Ruesch, Jurgen, et al. "Measurement of social disability." *Comprehensive Psychiatry* 13:507-18, 1972.

Stewart, Anita L., John E. Ware, Jr., and Robert H. Brook. *A Study of the Reliability, Validity and Precision of Scales to Measure Chronic Functional Limitations Due to Poor Health*. Santa Monica: Rand Corp., 1977.

III. Disability As Impairment

A. Physical Disability--Mobility, Sensory, And General Physical Impairment

Dawis, R. V. et al. "Methodological problems in rehabilitation research: Minnesota studies in vocational rehabilitation: V." Minneapolis: University of Minnesota, 1958.

Gonnella, C. et al. *Normative Criteria for Cancer Rehabilitation* (Rehabilitation Research Monograph Series no.1). Atlanta: Emory University Medical Rehabilitation Research and Training Center, 1978.

Granger, Carl V. and David S. Greer. "Functional status measurement and medical rehabilitation outcomes." *Archives of Physical Medicine and Rehabilitation* 57:103-09 (March), 1976.

Granger, Carl V., Gary L. Albrecht and Byron B Hamilton. "Outcome of comprehensive medical rehabilitation: measurement by PULSES profile and the Barthel Index." *Archives of Physical Medicine and Rehabilitation* 60:145-54 (April), 1979.

Hanman, Bert. "Evaluation of physical ability." *New England Journal of Medicine* no. 258:986-93, 1958.

Kass, Irving. "Disability benefits for chronic lung disease patients--the system is uneven and often unfair." *American Lung Association Bulletin* 64, 1:7-10 (Feb.), 1978.

Reports efforts of the American Lung Association to develop disability standards that adequately reflect the existence and extent of pulmonary disability and criteria for workers disabled by chronic lung disease.

Jones, Robert and Susan Jay Spungin. "A survey of national recordkeeping systems for visually handicapped children." *Journal of Visual Impairment and Blindness* 71:353-4 (Oct.), 1977

McHale, John. "Mental retardation and the future: a conceptual approach." In The Year 2000 and Mental Retardation, pp. 19-69. Edited by Stanley C. Plog and Miles B. Santamour. New York: Plenum, 1980.

Nagi, Saad Z. and John L. Melvin. "Factors in behavioral responses to impairments." Archives of Physical Medicine and Rehabilitation 51:552-57, 1970.

Nagi, Saad Z. "Rehabilitation and visual impairments: some issues in research and practice." In Blindness and Services to the Blind in the United States, pp. 149-63. Cambridge, MA: Organization for Social and Technical Innovations, Inc., 1971.

Sabatino, David A. "The definition and assessment of visual and auditory perception." Journal of Clinical Child Psychology 8:188-94 (Fall), 1979.

Smith, R. T. "Correlates of rehabilitation and recovery among the physically disabled." Baltimore: Johns Hopkins University, 1979.

Vail, M. and J. Townsend. "Services to the blind: a community concern." Paper prepared for the 11th Institute on Rehabilitation Services (New Orleans), 1973.

Provides methodologies used in estimating prevalence and incidence of blindness.

B. Mental Disability--Intellectual And Emotional Impairment

Bolton, Brian. Handbook of Measurement and Evaluation in Rehabilitation. Baltimore: University Park Press, 1976.

Readings summarize the current status of psychological measurement principles and practices in evaluation of disabled clients.

Botterbusch, K. F. Use of Psychological Tests with Individuals Who Are Severely Disabled. Stout: University of Wisconsin, 1976.

Clausen, John. "Quo Vadis, AAMD?" Journal of Special Education 6:51-60 (Spring), 1972.

Clausen's review of the American Association of Mental Deficiency definition of "mental deficiency" focuses on its inclusion of adaptive behavior. It is followed in the same journal issue by several commentary articles.

Davitz, Joel R. et al. Terminology and Concepts in Mental Retardation. New York: Columbia University, Teachers College, 1964.

Grossman, Herbert J., ed. *Manual on Terminology and Classification in Mental Retardation.* rev. Washington: American Association on Mental Deficiency, 1977.

Huberty, Thomas J. et al. "Adaptive behavior in the definition of mental retardation." *Exceptional Children* 46:256-61 (Jan.), 1980.

Jackson, Robin. "The ecology of educable mental handicap." *British Journal of Mental Subnormality* 20:18-22 (June), 1974.

Jaslow, Robert T. and Sanford V. Smith. "A proposal for a new conceptual use of the term mental retardation." *Mental Retardation* 10:36-7 (Feb.), 1972.

Kauppi, Dwight R. *The Application of General Semantics to the Classification of Mentally Retarded.* (Work Adjustment Project research report no. 17). Washington: U.S. Social and Rehabilitation Service, 1968.

Keogh, Barbara K. et al. "Proposed markers in learning disabilities research." *Journal of Abnormal Child Psychology* 8, 1:21-31, 1980.

Lakin, K. C. *Demographic Studies of Residential Facilities for the Mentally Retarded: An Historical Review of Methodologies and Findings.* Minneapolis: University of Minnesota, n.d.

Luckey, Robert E. and Ronald Neman. "Practices in estimating mental retardation prevalence." *Mental Retardation* 14:16-18 (Feb.), 1976.

Mercer, Jane R. *Labeling the Mentally Retarded: Clinical and Social System Perspectives on Mental Retardation.* Berkeley: University of California Press, 1973.

Milgram, Norman A. "MR and mental illness--a proposal for conceptual unity." *Mental Retardation* 10:29-31 (Dec.), 1972.

Orvaschel, Helen, Diane Sholomskas and Myrna M. Weissman. *The Assessment of Psychopathology and Behavioral Problems in Children: A Review of Scales Suitable for Epidemiological and Clinical Research (1967-1979).* (Series AN: Epidemiology, no. 1). Rockville, MD: U.S. National Institute of Mental Health, 1980.

Parker, J. L. "Need for new strategies for identification and assessment of mental retardation." *Slow Learning Child* 18:131-41 (Nov.), 1971.

Roman, Paul M. and Harrison M. Trice. "Defining and creating disability: the case of 'alcoholism.'" Paper presented at Annual Meeting of Southern Sociological Society, 1970.

Schainblatt, Alfred H. and Harry P. Hatry. Mental Health Services: What Happens to the Clients? Washington: Urban Institute, 1979.
Presents procedures for monitoring mental health clients' distress, social functioning and satisfaction with services of agencies.

Stein, Zena and Mervyn Susser. "Mutability of intelligence and epidemiology of mild mental retardation." Review of Educational Research 40:29-67, 1970.

U.S. Mental Health Statistics Improvement Program. System Design and Implementation Task Force. "The design and content of a national mental health statistics system." (Draft submitted to the ad hoc Advisory Group of the Mental Health Statistics Improvement Program.) Rockville: U.S. National Institute of Mental Health, Division of Biometry and Epidemiology, 1980.

U.S. National Institute of Mental Health. Definitions for Use in Mental Health Information Systems. Rockville, MD, 1980.

Ware, John E. Jr. et al. Conceptualization and Measurement of Health for Adults in the Health Insurance Study: Vol. III, Mental Health. Santa Monica: Rand Corp., 1979.

Wolman, B. B., ed. Clinical Diagnosis of Mental Disorders: A Handbook. New York: Plenum Press, 1978.

IV. The Life-Stage Dimension Of Disability

A. Infancy And Early Childhood

Hatcher, Roger P. "The predictability of infant intelligence scales: a critical review and evaluation." Mental Retardation 14:16-20 (Aug.), 1976.

Henney, R. Lee. The Impact of the Amendment of the Definition of "Developmentally Disabled" on the DD Program in FY '79 and FY '80. Washington: Institute for Comprehensive Planning, 1980.

Katoff, Lewis and Jeanette Reuter. "Review of developmental screening tests for infants." *Journal of Clinical Child Psychology* 9:30-4 (Spring), 1980.

Keogh, Barbara K. "Early ID: selective perception or perceptive selection?" *Academic Therapy* 12:267-74 (Spring), 1977.

Lessen, Elliott I. and Terry L. Rose. "State definitions of preschool handicapped populations." *Exceptional Children* 46:467-69 (March), 1980.

MacTurk, Robert H. and John T. Neisworth. "Norm referenced and criterion-based measures with preschoolers." *Exceptional Children* 45:34-9 (Sept.), 1978.

Sower, Ruth and Robert Covert. "Identifying preschoolers with special needs: A county-wide project to help plan future special services in the public schools." *Education and Training of the Mentally Retarded* 10:84-90 (April), 1975.

B. The School-Age Population

Abt Associates. *Assessment of Selected Resources for Severely Handicapped Children and Youth.* 3v. Cambridge, MA, 1976.
See especially for treatment of definitions and measurement issues chapter 2 of volume 1.

Barnard, Kathryn E. and Helen Bee Douglas, eds. *Child Health Assessment, Part I: A Literature Review.* Hyattsville, MD: U.S. Health Resources Administration, Division of Nursing, 1974.

Buck, Mildred R. "The multi-dimensional model for the assessment of children referred for classes for mental retardation." *Journal of Afro-American Issues* 3:91-102 (Winter), 1975.

Button, James E. and Edward W. Sontag. "Child find: Part I--methodology." *Education and Training of the Mentally Retarded* 14:222-27 (Oct.), 1979.

Button, James E. and Edward W. Sontag. "Child Rind: Pt. II--process." *Education and Training of the Mentally Retarded* 14:299-303 (Dec.), 1979.

Cordoni, Barbara Keene. *A Century of Research Related to Specific Learning Disabilities.* (Ed.D. dissertation, Duke University). Ann Arbor: University Microfilms, no. 77-1062, 1976.

Eisen, Marvin et al. Conceptualization and Measurement of Health for Children in the Health Insurance Study. Santa Monica: Rand Corp., 1980.

Hobbs, Nicholas. "An ecologically oriented, service-based system for the classification of handicapped children." In Ecosystem of the "Sick" Child, pp. 271-90. New York: Academic Press, 1980.

Hobbs, Nicholas. The Futures of Children: Categories, Labels, and their Consequences. San Francisco: Jossey-Bass, 1975.

Holt, Kenneth S. "Assessment of handicap in childhood." Child Care Health and Development, 5:151-62 (March-April), 1979.

Nagi, Saad Z. and P. C. Luken. "Childhood disability: a social epidemiology." Columbus: Ohio State University, 1975.

Ozer, Mark N. "A cybernetic approach to assessment: a problem-solving planning system." In A Cybernetic Approach to the Assessment of Children. Edited by Mark N. Ozer. Boulder: Westview Press, 1979.

Ozer, Mark N. Solving Learning and Behavior Problems of Children: A Planning System Integrating Assessment and Treatment. San Francisco: Jossey-Bass, 1980.

Richmond, Bert O. "Non-cognitive development of mentally retarded children." In Proceedings of the Eleventh Invitational Conference on Measurement in Education (1972), Section I. Compiled by Thomas M. Goolsby, Jr. New York: Harcourt Brace Jovanovich, 1972.

Rittenhouse, Robert K. "The usability of incidence figures as indices of successful identification of handicapped children." Distributed by Educational Documents Reproduction Service, ERIC. 1978.

Salzinger, Suzanne et al., eds. The Ecosystem of the "Sick" Child: Implications for Classification and Intervention for Disturbed and Mentally Retarded Children. New York: Academic Press, 1980.

Shotick, Andrew L. "Evaluating achievement of the mentally retarded: a comprehensive process." In Proceedings of the Eleventh Southeastern Invitational Conference on Measurement in Education (1972), Section I. New York: Harcourt Brace Jovanovich, 1972.

U.S. Office of Education, Bureau of Occupational and Adult Education. Resurge '79: Manual for Identifying, Classifying and Serving the Disadvantaged and Handicapped under the Vocational Education Amendments of 1976 (P.L. Law 94-482). Washington, 1980.

Weber, Douglas L. and Margaret B. Weber. "A suggested evaluation strategy for P.L. 94-142." Paper presented at the Annual Meeting of the Eastern Educational Research Association (Kiawah Island, SC), 1979.

C. Aging And The Aged

Andres, Reubin. The Normality of Aging: The Baltimore Longitudinal Study. Bethesda, MD: U.S. National Institutes of Health, 1979.

Gellman, W. Use of a Modified Scale of Employability for Handicapped Persons with Older, Disabled Workers. Chicago: Jewish Vocational Service, 1971.

Granick, Samuel and Robert D. Patterson. Human Aging II: An Eleven-Year Follow-up Biomedical and Behavioral Study. Rockville, MD: U.S. Alcohol, Drug Abuse and Mental Health Administration, 1971.

Haber, Lawrence D. "Age and capacity devaluation." Journal of Health and Social Behavior 11:167-82, 1970.

Nagi, Saad Z. and B. King. Aging and the Organization of Services (rev.) Columbus: Ohio State University, 1976.

Pfeiffer, E. Multidimensional Functional Assessment: The OARS Methodology. 2d ed. Durham: Duke University Medical Center, Center on Aging, 1978.

Smith, J. M. and A. Sheiham. "How dental conditions handicap the elderly." Community Dentistry and Oral Epidemiology 7, 6:305-10, 1979.

U.S. Interagency Statistical Committee on Long-Term Care for the Elderly. Data Coverage of the Functionally Limited Elderly. U.S. Office of Management and Budget, 1980.

V. The Environmental Dimension Of Disability

A. Physical Facilities--Residential Environment

Falcocchio, John C. *A Methodology for Evaluating the Effectiveness of Transportation Improvements for the Elderly and Handicapped*. Washington: U.S. Dept. of Transportation, Research and Special Programs Administration, 1979.

Lifchez, Raymond and Barbara Winslow. *Design for Independent Living: The Environment and Physically Disabled People*. New York: Whitney Library of Design, 1979.

McLain, Richard E., A. B. Silverstein, Linda Brownless and Mini Hubbell. "Measuring differences in residential environments among institutions for retarded persons." *Psychological Reports* 41:264-66 (Aug.), 1977.

Stoudt, Dennis W. *1976 Survey of Institutionalized Persons: Methods and Procedures*. Washington: U.S. Bureau of the Census, 1978.

B. Education

Baca, Leonard and Hermes Cervantes. "The assessment of minority students: are adaptive behavior scales the answer?" *Psychology in the Schools* 15:366-70 (July), 1978.

Forness, Steven R. "Behavioristic orientation to categorical labels." *Journal of School Psychology* 14:90-6 (Summer), 1976.

Discusses noncategorical programs in two special education settings for children with a variety of disability types.

Franken, Marion E. *Identifying Handicapped Students and Their Vocational Needs for 1977-1982*. Madison: Wisconsin Vocational Studies Center, 1977.

Garrett, John E. and Nettye Brazil. "Categories used for identification and education of exceptional children." *Exceptional Children* 45:291-92 (Jan.), 1979.

Goodman, Joan F. "The diagnostic fallacy: a critique of Jane Mercer's concept of mental retardation." *Journal of School Psychology* 15:197-205 (Feb.), 1977.

Haggerty, D. E. and E. S. Sacks. "Education of handicapped--towards a definition of an appropriate education." *Temple Law Quarterly* 50:961-94, 1977.

Justen, Joseph E., III and Gregory E. Brown. "Definitions of severely handicapped: a survey of state departments of education." American Association for the Education of the Severely/Profoundly Handicapped: *AAESPH Review* 2:8-14 (March), 1977.

McGrady, Harold J. and Douglas E. Wiseman. "Problems with definitions related to learning disabilities." Paper presented at the 56th Annual International Convention of the Council for Exceptional Children (Kansas City), 1978.

McNutt, Gaye. "Procedures for identifying specific learning disabilities." Distributed by ERIC Documents Reproduction Service, 1978.

Page, Joyce C. *A Survey of the Criteria Used by Elementary Level Regular Classroom Teachers to Identify Children with Special Needs.* (Ed.D. dissertation, University of Massachusetts). Ann Arbor: University Microfilms, no. 77-6396, 1976.

Schenck, Susan J. and William K. Levy. "IEPs: The state of the art in Connecticut." Hightstown, NJ: Northeast Regional Resource Center, 1979.

Sontag, Ed, Judy Smith and Wayne Sailor. "The severely/profoundly handicapped: Who are they? Where are they?" *Journal of Special Education* 11:5-11 (Spring), 1977.

Ysseldyke, James E. and Richard R. Regan. "Nondiscriminatroy assessment and decision making: Embedding assessment in the intervention process." Minneapolis: Institute for Research on Learning Disabilities, 1979.

C. Employment

Backer, Thomas E. "New directions in rehabilitation outcome measurement: emerging issues in rehabilitation." In *Annual Review of Rehabilitation*, vol. 1, pp. 193-230. Edited by Elizabeth L. Pan, Thomas E. Backer and Carolyn L. Vash. New York: Springer, 1980.

Bolton, Brian et al. Twelve Years Later: The Vocational and Psychosocial Adjustment of Former Rehabilitation Clients. Fayetteville: University of Arkansas, Vocational Rehabilitation Research and Training Center, 1979.

Bonge, D. "Pattern-probability approach to predicting client outcomes." Fayetteville: University of Arkansas, Vocational Rehabilitation Research and Training Center, 1975.

Ehrle, Raymond A. "The criterion of success problem in providing employability services." Vocational Guidance Quarterly 18:306-10 (June), 1970.

England, G. W. Measurement of Vocational Needs: Minnesota Studies in Vocational Rehabilitation, XVI. Minneapolis: University of Minnesota, 1964.

Esser, Thomas J., ed. Client Rating Instruments for Use in Vocational Rehabilitation Agencies. Stout: University of Wisconsin, Dept. of Rehabilitation and Manpower Services, 1975.

Halpern, A. S. "Principles and practices of measurement in career education for the handicapped." Eugene: University of Oregon, Mental Retardation Rehabilitation Research and Training Center, 1977.

Hasselkus, Betty Risteen and Margaret J. Safrit. "Measurement in occupational therapy." American Journal of Occupational Therapy 30:429-36, 1976.

Hawryluk, Alex. "Rehabilitation gain: a new criterion for an old concept." Rehabilitation Literature 35:322-28 (Nov.), 1974.

Hills, W. G. and K. W. Hills. "Consumer's measurement of vocational rehabilitation in Alaska." Norman: University of Oklahoma Research Institute, 1975.

Lester, J. T. "Relationships of certain personality and demographic variables to success in vocational rehabilitation." Los Angeles: Los Angeles Orthopaedic Foundation, 1970.

McCray, P. "Learning assessment in vocational evaluation." Stout: University of Wisconsin, Materials Development Center, 1979.

Medvene, Louis J. and Sheila H. Akabas. The Job Hunt of the Disabled: An Exploratory Study. New York: Columbia University School of Social Work, Industrial Social Welfare Center, Regional Rehabilitation Research Institute, 1979.
Includes methodology for measuring job outcomes.

Noble, John H. Jr. "Rehabilitating the severely disabled: the foreign experience." Journal of Health Politics Policy and Law, 4:221-49 (Summer), 1979.
Sets forth a model involving fourteen policy and environmental variables affecting rehabilitation outcomes.

Polka, J. A. Household Survey of Non-Institutionalized Handicapped Population--City of New Haven: Final Report. New Haven: Southern Connecticut State College, Dept. of Sociology, 1977.

Schecter, E. S. Follow-Up of Disabled and Nondisabled Adults, Number Two: Work Experience of the Disabled: 1972 and 1974. Washington: U.S. Social Security Administration, 1979.

Vest, L. and P. Cooper. "Data analysis techniques for case weighting systems." Fayetteville: University of Arkansas, Vocational Rehabilitation Research and Training Center, 1978.

D. Attitudes Of/about The Disabled

Bolton, Brian and Marceline E. Jaques, eds. The Rehabilitation Client. Baltimore: University Park Press, 1979.
Several essays examine methods for measuring attitudes of or about the disabled.

Donaldson, Joy. "Changing attitudes toward handicapped persons: a review and analysis of research." Exceptional Children 46:504-14, 1980.

English, R William and Judson B. Oberle. "Toward the development of new methodology for examining attitudes toward disabled persons." Rehabilitation Counseling Bulletin 15:88-95 (December), 1971.

Finkelstein, Victor. Attitudes and Disabled People: Issues for Discussion. New York: World Rehabilitation Fund, 1980.
Explores disability as a "special kind of discrimination, or social oppression."

Haber, Lawrence D. "Some problems of research in disability." In Proceedings of the Conference on Medical Sociology and Disease Control (University of Chicago, 1964), pp. 19-25. Edited by Gerald Gordon. Chicago: University of Chicago, Graduate School of Business, n.d.

Focus is on attitudes of family members toward the disabled member.

Nagi, Saad Z. "Congruency in medical and self-assessment of disability." Industrial Medicine 38:27-36, 1969.

Nooe, Roger M. "Measuring self-concepts of mentally retarded adults." Social Work 22:320-22, (July), 1977.

Viscardi, H. Measurement of Attitudes toward Disabled Persons (Special Report). Albertson, NY: Human Resource Center, 1967.