

THE NATIONAL ARCHIVES OF THE UNITED STATES

LITTERA
SCRIPTA
MANET

★ 1934 ★

3-443-85-001

COLLABORATIVE PERINATAL PROJECT

1958 - 1974

DOCUMENTATION

Record Group 443

Records of the National Institutes of Health



CONTENTS

Abstract

Introduction

RCPP User's Guide



August 1983
3-443-85-1

United States National Institutes of Health. Collaborative Perinatal Project (NCP). 1958-1974 [Machine-Readable data file]. Created by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS). Washington: National Institutes of Health [producer] 1958-84; Washington: National Archives and Records Administration [distributor].

PROVENANCE: The National Institute of Neurological Diseases and Blindness (NINDS) of the National Institutes of Health instituted the Collaborative Study of Cerebral Palsy, Mental Retardation, and other Neurological and Sensory Disorders of Infancy and Childhood. From this study evolved the need for and the initiation of the NCP and the reorganization of the NINDS into the NINCDS. Data from medical histories, examinations, and observations on the participating individuals was supplied by twelve institutions throughout the United States between 1959 and 1974.

TYPE OF FILE: Microdot. The unit of observation is the individual. There is one record for each medical examination or clinical observation.

UNIVERSE DESCRIPTION: The universe consists of samples of 100 to 2000 pregnancies per year (1959-1965) from each of twelve participating institutions: Boston Lying-In Hospital; Children's Hospital, Buffalo; Charity Hospital, New Orleans; Columbia-Presbyterian Medical Center, New York; Johns Hopkins Hospital, Baltimore; Medical College of Virginia, Richmond; University of Minnesota Hospital, Minneapolis; New York Medical Hospital; University of Oregon Medical Center, Portland; Pennsylvania Hospital/Children's Hospital, Philadelphia; Providence Lying-In Hospital; University of Tennessee College of Medicine, Memphis; The sampling frame and selection process varied from institution to institution and is fully explained in the documentation. The NCP lasted sixteen years and included approximately 58,000 study pregnancies. Data on the women and their pregnancies and study children are included in the NCP data base.

SUBJECT MATTER DESCRIPTION: The Collaborative Perinatal Project was a prospective study. The data constitute an important resource for biomedical and behavioral research in many areas of obstetrics, perinatology, pediatrics and developmental psychology. The data also provide a prospective base for examining neurological and neurosensory defects and the relationship of pregnancy and perinatal factors on the health of individual children. The major categories of data collected include obstetrical, pediatric, pathological, serological, socio-economic and family, genetic history, psychological, speech, language and hearing. The mother was examined

during pregnancy, labor and delivery. The children were given neonatal examinations and follow-up examinations at four, eight, and twelve months, and three, four, seven, and eight years. Supplemental information was gathered throughout the study, including family linkages between related women participating in the NCFP. In addition to the master file there are variable files and work files arranged to facilitate researcher access of the 6700 data items.

GEOGRAPHIC COVERAGE: Geographic coding includes the city and state of the twelve participating institutions and the street address, city, and state of each patient.

TECHNICAL DESCRIPTION:

STRUCTURE: Fixed length, non-hierarchical

SIZE: Master File - 58,760 Administrative Records

SORT SEQUENCE: Records within the master file are sorted by the master file card number and a unique case number identifying the institution, mother and child. Records within the Variable file are sorted by variable and thereunder by unique case number. Records within the work files are sorted by unique case number.

REFERENCE MATERIAL:

Documentation Relating to the Collaborative Perinatal Project

Available from the Machine-Readable Branch (MSRB)

This is a guide to the machine-readable data files. It contains an introduction to the data; detailed history of data design, collection, and processing; the organization and content of the files; and codebooks.

RELATED PRINTED REPORTS:

The documentation includes an extensive bibliography of research reports based on the project data.

RELATED MACHINE-READABLE DATA FILES:

None

FILE AVAILABILITY:

The data in the NCFP contain personal identifiers and medical information protected under the Privacy Act and KARA General Restriction B4. Access to the data for statistical research may be permitted within the appropriate guidelines and restrictions.

These files are available from:

Machine-Readable Branch (MRRB)
National Archives and Records Administration
Washington, DC 20408
(202) 523-3267

DESCRIPTORS: TUBERIA, PREGNANCY, NEUROLOGY, CEREBRAL PALSY, OBSTETRICS,
PERINATOLOGY, PEDIATRICS, DEVELOPMENTAL PSYCHOLOGY, EMB.
DEVS, SEROLOGY, CONGENITAL MALFUNCTIONS

Work Files:	Description	Records	Notes
W1	Socioeconomic Index at Registration	55,910	records
W2	" " " " Seven Years	40,363	"
W3	Drugs Taken During Pregnancy, Trade Names	199,088	" (49,214 cases)
W4	" " " " Active Compounds	252,154	" (")
W5	Congenital Malformations, One & Seven Years	59,191	"
W6	Cerebral Palsy Diagnosis	53,251	"
W7	Abnormalities at Seven Years	41,869	"
W8	Speech, Language and Hearing at Eight Years	20,137	"
W9	Tennessia Classification	58,806	"
W10	Rupture of Membranes	49,188	"
W11	Survey of Viral, Bacterial, Parasitic and Fungal Infections During Pregnancy	10,224	"
W12	Serological Testing - Complement Fixation Tests	19,323	"
W13	Serological Testing - Toxoplasmosis & Rubella	39,645	"
W14	" " - Cord Blood	32,098	"
W15	" " - Abnormalities & Controls	2,255	"
W16	Serum Specimen Inventory	227,215	"
W17A	Family Linkage - Mother's Relationship	8,150	"
W17B	" " - Children's Relationship	12,094	"
W17C-	" " - Cousin, 1st	8829	"
D-	" " - Niece/Nephew	301	"
E-	" " - Aunt/Uncle	219	"
F-	" " - Cousin, maternal	4626	"
G-	" " - Cousin, maternal in-law	1467	"
H-	" " - Cousin, maternal in-law	4250	"
I-	" " - Niece/Nephew half	343	"
J-	" " - Aunt/Uncle, half	266	"
K-	" " - Cousin, half	775	"
L-	" " - Cousin-once-removed	1405	"
M-	" " - Cousin-2nd	1841	"
W18-	Visit Summary	267,128	(28,455 cases)



PROCEEDINGS TO REGISTRATION

The study of the etiology of cerebral palsy (CP) was a longitudinal, multi-center, multi-institutional effort sponsored by the Developmental Neurology Branch, Neurology and Neurosurgery Program, National Institute of Neurological and Communicative Disorders and Stroke (NIDCD), the National Institutes of Health. The primary goal of the study was to determine the etiology of the developmental disabilities of cerebral palsy and mental retardation. Through a series of case studies and studies which relate the events and circumstances of pregnancy, labor and delivery to the neurological and mental disabilities of these children, as well as the genetic causes and development of these disabilities, the study sought to determine the etiology of cerebral palsy, mental retardation, and other developmental disabilities, such as epilepsy, ataxia, brain malformations, and other developmental disabilities, or communicative disorders, which are associated with the perinatal period in order to develop effective methods of prevention and intervention.

The study included a total of approximately 1,000 women and the children born to them during the study. Additional demographic information was obtained on the fathers, and the siblings of the children and other members of the family.

Personal and other information was collected on the women within the study during pregnancy and delivery between 1953 and 1965 and on the children during their first three years of life from birth until eight years of age. The study included women who were hospitalized in the United States. The women in the study were patients of obstetric and gynecological medical centers across the United States. The study included women from 14 percent to 100 percent for the most part the women were of the following categories:

The study was designed to not duplicate the national population in two significant respects. They were nearly all from lower income groups and the racial composition was approximately half-white, half-black.

The project is unique in the magnitude of the study, the comprehensiveness of the information collected, the comparability of the information collected between cooperating institutions and across the two decades of the study, the provision for future data analysis, the prospective, longitudinal nature of the project, and the enormity of the cost (\$200 million). The medical, psychological and demographic information is not time restricted and can be linked to future medical studies. Further, since basic data collection did not involve any changes or additions in the practices or procedures at the participating institutions, the data also represent an overview of medical practices during the two decades of the project.

Medical information on the women includes history of previous pregnancies and infertility, daily cigarette consumption, drug use in pregnancy, perinatal observations, labor and delivery events, obstetrical reports, and anesthesia reports. Socioeconomic, family history and family health information also was collected.

Medical information on the children includes comprehensive examinations conducted at birth, newborn, four-months, one year and seven years; psychological testing at eight-months, four years and seven years; speech, language and hearing examinations at three years and eight years; and serological and pathological studies. Some socioeconomic and family history also is available on all siblings.

The seven volumes of documentation prepared by NINCDS to support future research in the NICHD data provide extensive detail on the background of the study, the history of the project, the sample, data collection and data processing, access to the data while in NINCDS custody and record layouts and coding for each variable within the master file, the variable file, and the work files. Appendices contain sample forms and data entry cards. The bibliography of published research based on NICHD will be updated periodically and refilled.

ACCESS TO DATA:

ICPP data was subject to the provisions of the Privacy Act of 1974 and the Freedom of Information Act, Exemption B6, while in HHD custody. Records in the physical and legal custody of the National Archives are subject to the Privacy Act. However, National Archives General Restriction 4 (105-61.5302-4 in accordance with 5 USC 552 and 44 USC 2104) restricts:

"records containing information about a living individual which reveal details of a highly personal nature that the individual could reasonably assert a claim to withhold from the public to avoid a clearly unwarranted invasion of privacy, including but not limited to information about the physical or mental health or the medical or psychiatric care or treatment of the individual, and that - (1) contain personal information not known to have been previously made public, and (2) relate to events less than 75 years old."

Such information may be released for statistical or quantitative research when the researchers have provided adequate written assurance about the protection and use of the information. Researchers seeking access to the ICPP data must complete an "Application for Statistical Research in Records Containing Personal Information" and agree in writing to the "Conditions of Use."

In keeping with the applicable statutes, the ICPP data provided for statistical or quantitative research use will not contain the name, address, or social security number of any individuals in the project. The HHD (HINDS) case number will be available on each data record to permit identification and linkage of all records related to a particular patient and her children. The case number includes the institution, type of patient selection, gravida identification number, order of pregnancy and child or gravida number.

LEVEL OF VALIDATION:

The elaborate data collection, input and verification procedures, as described in the User's Guide, and the extensive past use of the data both by NINCDS and by other researchers demonstrate the validity of the NICPP data. Routine preservation activities associated with accessioning the data into the National Archives also demonstrated the validity of the data. Further, due to the personal nature of the information in the data complete sample computer printouts could not be included as part of the documentation.

The documentation for the Collaborative Perinatal Project consists of the abstract, this introduction, the User's Guide, and the bibliography of published research based on NICPP data.

The documentation was prepared by Bruce Ambacher, Archivist, Machine-readable Branch.
September, 1985

**NINCDS COLLABORATIVE
PERINATAL PROJECT
A User's Guide to the Project and Data**

**Volume I
An Introduction to the
History, Scope and
Methodology of the Project**

**L. E. Sever
A. R. Olsen
N. R. Hinds**

**C. R. Watson
E. B. Perrin
J. S. Littlefield**

December 1983

**Prepared for
the National Institute of Neurological
and Communicative Disorders and Stroke
under Contract 2311105150**

 **Battelle**
Pacific Northwest Laboratories

LEGAL NOTICE

This report was prepared by Battelle as an account of sponsored research activities. Neither Sponsor nor Battelle nor any person acting on behalf of either:

MAKES ANY WARRANTY OR REPRESENTATION, EXPRESS OR IMPLIED, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, process, or composition disclosed in this report may not infringe privately owned rights; or

Assumes any liabilities with respect to the use of, or for damages resulting from the use of, any information, apparatus, process, or composition disclosed in this report.

NINCDS COLLABORATIVE PERINATAL PROJECT:
A USER'S GUIDE TO THE PROJECT AND DATA

Volume I.
An Introduction to the History,
Scope and Methodology of the Project

L.E. Sever	C.R. Watson
A.R. Olsen	E.B. Perrin
N.R. Hinds	J.S. Littlefield

December 1983

Prepared for
The National Institute of Neurological and
Communicative Disorders and Stroke
under Contract 2311105150

Battelle
Pacific Northwest Laboratories
Richland, Washington 99352

ABSTRACT

The purpose of this user's guide is to provide researchers with complete documentation of data gathered during the course of the NINCDS Collaborative Perinatal Project (NCPP). The NCPP lasted sixteen years and included approximately 58,000 study pregnancies. Data on the women and their pregnancies and study children are included in the NCPP data base, which is made up of three separate computerized files. The user's guide consists of seven separate volumes. Volume I provides background on the study and detailed procedures for requesting and obtaining data. Volumes II, III and IV provide complete documentation for data items contained in the master, variable and work files, respectively. Volume V, the master index to data items, is a computerized compilation of all data items included in the project. Volumes VI and VII are computerized indexes that allow a researcher to scan for data items in an alphabetical glossary (Volume VI) or to find data items arranged according to person, time of collection or measurement and general subject categories (mother, delivery, medication, etc.; see Volume VII).

ACKNOWLEDGMENTS

The authors would like to thank all those who contributed to the development of this user's guide. Dr. Joseph S. Drage, Dr. Halsey M. Marsden and LaVerne C. Edmondson of the Developmental Neurology Branch, NINCDS, provided valuable guidance on all aspects of the NINCDS Collaborative Perinatal Project. They, along with Drs. Sarah H. Broman, N. C. Myrianthopoulos, Karin B. Nelson, Paul L. Nichols and Ms. Charlotte Jackson, reviewed the initial draft of the guide. We received advice and guidance for making the guide useful to researchers from the following University of Washington faculty members: Dr. Irvin Emanuel, Professor, Department of Epidemiology, School of Public Health and Community Medicine; Dr. Alan G. Fantel, Research Associate Professor, Department of Pediatrics, School of Medicine; Dr. James L. Gale, Professor, Department of Epidemiology, School of Public Health and Community Medicine; Dr. Durlin E. Hickok, Clinical Assistant Professor, Department of Obstetrics and Gynecology, School of Medicine; Dr. Thomas D. Koepsell, Associate Professor, Department of Epidemiology, School of Public Health and Community Medicine; Dr. Clifford J. Sells, Associate Professor, Department of Pediatrics, School of Medicine; Dr. Philip S. Spiers, Department of Epidemiology, School of Public Health and Community Medicine; and Dr. Gerald Van Belle, Professor, Department of Biostatistics, School of Public Health and Community Medicine. Scott Smith, Research Assistant, provided expertise in constructing names for data items and in defining the categories used in Volume VII. Amy Brix, Paula Alley and Lori DesCamp entered data and assisted in organizing the volumes. Sharon Popp and Mary Lou Lomon typed the documents.

PREFACE

INTRODUCTION

The data from the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) Collaborative Perinatal Project are an important resource for biomedical and behavioral research in many areas of obstetrics, perinatology, pediatrics and developmental psychology. The data were collected as part of a prospective study, unique in its design and magnitude. The data constitute a repository of information of great value. Books and monographs based on analyses of these data and other publications number in the hundreds. Even so, the possibilities for the development of further knowledge based on this study are immense. It is unlikely that such a study will be undertaken again and it is thus of particular importance that the data be utilized as fully as possible.

OBJECTIVES AND READER ASSUMPTIONS

This User's Guide, NINCDS COLLABORATIVE PERINATAL PROJECT: A USER'S GUIDE TO THE PROJECT AND DATA, describes the NINCDS Collaborative Perinatal Project (NCPD) and the data base resulting from that project. The User's Guide objectives are (1) to provide comprehensive documentation of the NCPD, with information regarding the design and conduct of the study, and (2) to allow effective and efficient independent use of the NCPD data by researchers previously unfamiliar with the data. Documentation of the study and methods employed in data collection and processing will help researchers clearly understand the strengths and limitations of the data set.

The reader is assumed to be either a researcher (probably a Ph.D., M.D. or advanced graduate student) or a computer programmer associated with such a researcher. The NCPD data are most appropriate for epidemiological studies and hence the reader is assumed to have general interest in such studies.

DOCUMENT STRUCTURE

This document, NINCDS COLLABORATIVE PERINATAL PROJECT: A USER'S GUIDE TO THE PROJECT AND DATA, is divided into seven volumes. The volumes are:

- Volume I, AN INTRODUCTION TO THE HISTORY, SCOPE AND METHODOLOGY OF THE PROJECT
provides an overview and should be read prior to acquisition of any other volumes.
- Volume II, PROJECT STUDY FORMS AND DOCUMENTATION OF TRANSFER TO COMPUTERIZED DATA ITEMS IN MASTER FILE
is an exhaustive compilation of forms, instructions for completing the forms, definition of codes, punched card descriptions, and tabulation of data items related to each form. Over 5000 data items are defined in the 2000 pages of this volume.

- Volume III, VARIABLE FILE,
describes the major summary file, which contains 1200 data items.
- Volume IV, SELECTED NCPP WORK FILES,
describes 18 computer tapes related to specific research areas within the study.
- Volume V, MASTER INDEX TO THE NCPP COMPUTERIZED DATA ITEMS,
tabulates all data items in the order they appear on the various computer data files and assigns unique identification numbers to each item.
- Volume VI, ALPHABETICAL PERMUTED GLOSSARY OF NCPP COMPUTERIZED DATA ITEMS,
tabulates the data items in alphabetic order with multiple entries for selected words within the data item names.
- Volume VII, CATEGORIZATION OF DATA ITEMS, BY PERSON, TIME OF COLLECTION OR MEASUREMENT, AND GENERAL SUBJECT AREA,
presents the data item names and identification numbers in three separate orderings based on person, time, and general subject area.

The structure of Volume I is described in this preface; that of Volumes II through VII in Chapter 6 of Volume I.

Chapter 1 begins with a review of the history of the Collaborative Perinatal Project and its goals and objectives. Key to the study was its collaborative nature and the selection of the participating centers. The study sample selection process is discussed and resulting characteristics of each sample presented. This provides the researcher with an understanding of the composition of the NCPP subject population on which the data were obtained.

Chapter 2 includes a discussion of the data collection methodology for the study. Development of forms and manuals used to collect data in a standard format is described, as are the forms themselves and the data collection process. Because of the multicenter nature of the project, standardization of data collection and consistency and accuracy of the data across centers are important considerations.

Chapter 3 describes procedures used for processing the data. This includes consideration of the activities carried out at both the collaborating institutions and at the Perinatal Research Branch of NINCDS. The data are organized into various computer data files and these files are described.

Chapter 4 of the guide presents an overview of the data collected classified by categories. Because of the breadth and diversity of the data, the data items included in the study have been organized according to bio-behavioral category. A hierarchical system of classification that allows a researcher to determine the substantive areas included in the study is also included.

Chapter 5 presents the information needed by a researcher to generate a request for access to data from the NCPP. Included here is a discussion of the policies and procedures to be followed in requesting data. General information on the structure of the individual computer files is included to guide the researcher in assessing the utility of each computer file for specific research requirements.

Chapter 6 describes the contents and use of Volumes II to VII. In these volumes, data collection forms and manuals are described and reproduced; documentation of the transfer of data from the forms to computerized data items is provided as well (Volume II). The three types of computer files are documented and the individual data items contained in the files identified (Volumes III and IV). Also included are a master index (Volume V), an alphabetical permuted glossary of the computerized data items (Volume VI) and categorization of data items by person, time of collection or measurement, and general subject area (Volume VII).

A bibliography that lists all publications based on data from the NCPP is available from the Developmental Neurology Branch of NINCDS. We recommend that a prospective researcher review the bibliography to identify pertinent research that may have been conducted using this data set.

REQUIREMENTS FOR RESEARCHERS

To use data from the NCPP, the researcher must first satisfy the requirements for data access, as established by the NINCDS. These requirements are outlined in Chapter 5 of Volume I of the User's Guide.

Prior to initiating a data request, however, it is recommended that the researcher determine if his computing resources are adequate to process any NCPP data tapes requested. Depending on the data request, substantial resources may be required. We recommend that the discussion of data files in Chapter 3 be reviewed carefully by an individual with computer programming proficiency.

SUGGESTED RESEARCH PLAN

Based on our experience in developing this document, we suggest an approach to determining if the NCPP data files are of potential use in a research study. While alternative approaches or modifications to the following are possible, we feel that the proposed approach will be the most economical one, both in terms of time and other resources, particularly for the individual who has not used NCPP data before.

We suggest that the researcher first study this volume (Volume 1) to develop an understanding of the design of the study and the methods by which the data were collected, processed and stored. This will provide an indication of the potential usefulness of the data. At the same time, the categories of data available should be reviewed to determine if general areas of interest to the researcher are included in the NCPP.

If this review indicates that NCPP data are relevant to the researcher's needs, a copy of the current NINCDS Collaborative Perinatal Project Bibliography and the remaining volumes of the User's Guide appropriate to the researcher's project should be obtained. The bibliography will enable the researcher to determine the previously published work that is relevant. The other volumes of the User's Guide, described in Chapter 6, provide the researcher with specific information on the data collection forms, data items available, coding of data items, and location of data items on the various tape files. Using these volumes, the researcher will be able to generate a request to obtain specific data of interest.

When specific data items of interest are determined and known to be available, a formal request for the data should be submitted, following the procedures described in Chapter 5. Copies of computer data files will be provided to the investigator after approval of the request. Special tapes will not be created for researchers. If the researcher is interested in variables or data that are not computerized, access to the microfilmed copy of the original study forms may be requested. Microfilmed study forms are available for viewing at NINCDS only.

The researcher is expected to conduct analyses of the data requested independent of the Developmental Neurology Branch of NINCDS. The information provided in this guide regarding tape characteristics, field locations, and variable coding is designed to provide the researcher with the knowledge needed for independent use of these data.

In summary, by developing a thorough familiarity with this guide and the descriptions of the NCPP data, the investigator can address questions of research interest. We hope this guide meets its goals of allowing effective and efficient independent use of these data by researchers previously unfamiliar with the NINCDS Collaborative Perinatal Project.

CONTENTS

ABSTRACT	iii
ACKNOWLEDGMENTS	v
PREFACE	vii
INTRODUCTION	vii
OBJECTIVES AND READER ASSUMPTIONS	vii
DOCUMENT STRUCTURE	vii
REQUIREMENTS FOR RESEARCHERS	ix
SUGGESTED RESEARCH PLAN	ix
CHAPTER 1. THE NINCDS COLLABORATIVE PERINATAL PROJECT: AN INTRODUCTION	I.1.1
BACKGROUND OF THE STUDY	I.1.1
HISTORY OF THE COLLABORATIVE PERINATAL PROJECT	I.1.2
Collaborating Institutions	I.1.4
Withdrawal of Collaborating Institutions	I.1.5
COLLABORATIVE PERINATAL PROJECT ADMINISTRATION	I.1.6
STUDY SAMPLE SELECTION AND COMPOSITION	I.1.7
Women Lost to the Study	I.1.11
Children Lost to the Study	I.1.13
Sample Size Reduction at Later Ages: Attrition of Collaborating Institutions	I.1.14
THE EXPANDED RESEARCH PROGRAM AND RESEARCH AREA	I.1.15
THE NINCDS COLLABORATIVE PERINATAL PROJECT BIBLIOGRAPHY	I.1.16
SUMMARY	I.1.16
CHAPTER 2. DATA COLLECTION METHODOLOGY	I.2.1
THE PRETEST: FORM DESIGN AND DEVELOPMENT	I.2.1
CATEGORIES OF DATA COLLECTED	I.2.2

FORM MODIFICATION AND REPLACEMENT	I.2.3
PHASES IN DATA COLLECTION AND ASSOCIATED FORMS	I.2.11
Prenatal Phase	I.2.11
Labor and Delivery	I.2.12
Newborn Phase	I.2.12
Four Months Phase	I.2.13
Eight Months Phase	I.2.13
Twelve Months Phase	I.2.13
Three Year Phase	I.2.13
Four Year Phase	I.2.14
Seven Year Phase	I.2.14
Eight Year Phase	I.2.14
Supplementary Information	I.2.15
PROCEDURE AND INSTRUCTION MANUALS	I.2.15
PROCEDURES INSTITUTED TO ENSURE CONSISTENCY AND ACCURACY OF PRIMARY DATA	I.2.15
CHAPTER 3. DATA PROCESSING PROCEDURES	I.3.1
DATA PROCESSING AT THE COLLABORATING INSTITUTIONS	I.3.1
DATA PROCESSING AT THE PERINATAL RESEARCH BRANCH	I.3.1
ACCURACY OF THE DATA PROCESSING	I.3.2
POOLING OF DATA FROM THE COLLABORATING INSTITUTIONS	I.3.3
DATA FILES CONSTRUCTED	I.3.4
Master File	I.3.4
Master File Card Number and NINDB Case Number Rationale	I.3.5
Variable File	I.3.7
Work Files	I.3.9
NON-COMPUTERIZED DATA	I.3.9

CHAPTER 4. NCPP DATA: HIERARCHICAL CLASSIFICATION AND PERSON, TIME AND SUBJECT CATEGORIZATION	I.4.1
HIERARCHICAL CLASSIFICATION	I.4.1
PERSON, TIME AND SUBJECT CATEGORIZATION	I.4.19
SUMMARY	I.4.20
CHAPTER 5. PROCEDURES FOR OBTAINING NCPP DATA	I.5.1
CHAPTER 6. A RESEARCHER'S GUIDE TO THE NCPP DATA	I.6.1
DATA ITEM IDENTIFICATION AND NAMING	I.6.1
VOLUMES OF THE GUIDE	I.6.4
Volume II, The Master File	I.6.4
Volume III, Variable File	I.6.4
Volume IV, Work Files	I.6.6
Volume V, Master Index	I.6.7
Volume VI, Alphabetical Permuted Index	I.6.8
Volume VII, Categorization of Data Items	I.6.8
STEPS TO FOLLOW IN USING THE GUIDE	I.6.8
EXAMPLE USES OF THIS USER'S GUIDE	I.6.9
Example 1	I.6.9
Example 2	I.6.11
Example 3	I.6.13
SUMMARY	I.6.16
CHAPTER 7. REFERENCES	I.7.1
APPENDIX A. SUBJECT SELECTION	I.A.1
APPENDIX B. STANDARD COHORT DEFINITIONS	I.B.1
APPENDIX C. QUALITY CONTROL PROGRAMS	I.C.1
PSYCHOLOGY	I.C.1
SPEECH, LANGUAGE AND HEARING TESTS	I.C.2
SEVEN YEAR PEDIATRIC-NEUROLOGICAL EXAMINATION	I.C.4

FIGURES

2.1	Data Collection for the NCPP, 1959-1974	I.2.3
3.1	Cards on the Master Data File	I.3.6
6.1	Drugs in Pregnancy OB-15 Example Form	I.6.12

TABLES

1.1	Institutions, Location and Approximate Number of Core Registrants	I.1.5
1.2	Categorization of NCPP Cases	I.1.8
1.3	Sample Size by Institution and Ethnic Group in the NCPP Population (Cohort IID)	I.1.9
1.4	Comparison of Selected Characteristics of Study and Lost-to-Study Gravidas	I.1.12
2.1	Forms Used in the NCPP by Subject Area	I.2.4
2.2	Final Forms Used in the NCPP by Time of Administration	I.2.7
2.3	Summary of Form Replacement	I.2.10
3.1	Derivation of Master File Card Number and NINDB Case Number	I.3.5
3.2	Collaborating Institutions and Their Code Numbers	I.3.8
3.3	NCPP Work Files Documented	I.3.9
4.1	Hierarchical Classification: Primary Data Classes for the NCPP	I.4.3
4.2	Hierarchical Classification: Secondary Data Classes for the NCPP	I.4.4
4.3	Hierarchical Classification: Tertiary Data Classes for the NCPP	I.4.6
4.4	Person, Time and Subject Categories for the NCPP Data Items	I.4.21
4.5	Definition of Person, Time and Subject Categories	I.4.22
6.1	Abbreviations for Person, Time and Subject Categories	I.6.3
6.2	Structure of Volume II Parts	I.6.5

CHAPTER 1. THE NINCDS COLLABORATIVE PERINATAL
PROJECT: AN INTRODUCTION*

BACKGROUND OF THE STUDY

Plans for the NINCDS Collaborative Perinatal Project or NCPP, originally named the Collaborative Study of Cerebral Palsy, Mental Retardation, and other Neurological and Sensory Disorders of Infancy and Childhood, were initiated at the National Institute of Neurological Diseases and Blindness (NINDB) shortly after the Institute's founding in 1950. The general goal of the study was to investigate the relationship between perinatal complications and abnormal outcomes of pregnancy, since by the 1950s many such relationships had been reported in the literature. The factors which gave rise to the support for the investigation are summarized in the following quotation from the introduction to The Women and Their Pregnancies (Niswander and Gordon, 1972):

"There is an increasing emphasis today in the United States on improving the health of American citizens. One aspect of this endeavor is to improve reproductive efficiency in order to increase the likelihood of the birth of healthy babies free from disease and impairment, and capable of optimal physical and intellectual development. The achievement of this goal depends upon the enlightened and widespread application of measures to prevent perinatal mortality and the continuum of reproductive wastage, which includes mental retardation, congenital malformation, cerebral palsy, and handicapping neurosensory defect.

During the past half century, in many countries including the United States, maternal mortality has declined dramatically; the risk of death associated with pregnancy has, to a large extent, been eliminated. A sharp reduction in infant deaths from 28 days to the end of the first year of life has also occurred during the same period. However, the number of deaths occurring during the perinatal period has declined more slowly. In this Study, perinatal deaths include those fetal deaths occurring between the 20th week of gestation and the time of delivery, and deaths of liveborn infants during the neonatal period (to 28 days). It is the custom to report fetal and neonatal deaths separately, but frequently it is useful in some circumstances to combine them and to consider perinatal deaths as a unit.

The magnitude of the problem of perinatal death comes sharply into focus with the realization that until old age

*Much of the material for this introduction is taken from Broman et al. (1975) and Niswander and Gordon (1972).

the risk of dying is highest during the perinatal period. While the general mortality rate for the country approximates 9 per 1000 individuals in the population during the period 1955-65, the perinatal death rate is almost four times as great, approximately 35 per 1000 livebirths. The age specific risk of dying does not again approach a rate of 35 per 1000 until the 64th year is reached. Moreover, even at this rate the risks are not comparable because the age specific risk extends over a one-year time span while the risk of dying during the perinatal period is limited to about one-half, from 20 weeks gestation until 28 days after birth."

The high perinatal mortality rates and the associated human suffering were not the only causes for concern. Of far more importance to the individual, the family, and the community was the "continuum of fetal insult" which includes congenital malformations, cerebral palsy, mental retardation, deafness, blindness, and other neurosensory defects. Estimates of the numbers of people with such conditions are very large; when the NCPP was established it was estimated that approximately 20 million individuals in the United States had handicaps or defects which fell within this general category. Their special care, rehabilitation, and education paid for by family, community, and government cost billions of dollars each year. Effective means of preventing these defects were, and are, urgently needed.

HISTORY OF THE COLLABORATIVE PERINATAL PROJECT

It was against this background that plans for the NCPP were developed (Broman et al. 1975). After the establishment of the National Institute of Neurological Diseases and Blindness (NINDB) in 1950, concern about the etiology of cerebral palsy and other forms of neurological, sensory, and intellectual deficits led to the planning of a comprehensive study of pregnancy and its outcome, the outgrowth of which was the Collaborative Study of Cerebral Palsy, Mental Retardation, and Other Neurological and Sensory Disorders of Infancy and Childhood. NINDB, under the directorship of Dr. Pearce Bailey, became the focal point for the planning of research into the etiology of brain damage in childhood. Numerous organizations participated, including voluntary health agencies such as United Cerebral Palsy, the National Association for Retarded Children, and the Association for the Aid of Crippled Children, as well as professional organizations such as the Academy of Neurology and the Academy of Cerebral Palsy.

The Appropriations Subcommittee of the House of Representatives heard testimony from Dr. Bailey and other experts in 1953. These professionals emphasized that maternal infections, toxins, nutritional deficiencies, anoxia, and blood incompatibilities between mother and infant may account for certain forms of cerebral palsy and malformations in the offspring. In 1954, the NINDB developed a coordinated system of brain registries for cerebral palsy and other disorders resulting from brain injury. By making postmortem material from all areas of the United States available, the registries aided research on the brain and the development of improved methods of diagnosis and treatment of neurological diseases.

In 1955, at the hearings of the Appropriations Committee of the House of Representatives, Dr. Bailey announced plans to set up a collaborative study involving various institutions throughout the United States. The purpose of the proposed study was to provide an opportunity to correlate the clinical manifestations of different types of cerebral palsy with the underlying neuroanatomical damage in the brain. In March of 1955, a panel of experts was convened by Bailey to draw up a protocol for a collaborative study of cerebral palsy. The objectives were: (1) to make a more precise determination of fetal, environmental, and medical factors leading to the various forms of human cerebral palsy, and (2) to link the symptoms of this group of disorders to the causative brain damage.

In 1957, Bailey proposed that in addition to the clinical-pathological study of cerebral palsy, a longitudinal investigation of pregnant women and their children be conducted in various medical centers throughout the country by obstetricians, pediatricians, neurologists, neuropathologists, and other specialists. Early in the planning of the collaborative clinical-pathological study of pregnancy and its outcome, the inadequacy of hospital records was recognized. Standardization of records between hospitals and expansion of the information normally recorded was required for a study of this scope. Furthermore, in previous studies, information about pregnancy and perinatal events had been collected retrospectively from parents of children with defects. To offset this problem, the prospective approach was chosen for the NCPP. Data would be collected on pregnancy and perinatal events as they occurred, eliminating biases due to knowledge of pregnancy outcomes.

The need for detailed prospective data, systematically recorded, coupled with the rarity of neurological deficits in childhood, made the availability of a large group of pregnant women imperative. A population of 40,000 pregnancies was projected. From a population this size, approximately 80 cases of Down's syndrome, 200 cases of cerebral palsy, 640 stillbirths, 680 neonatal deaths, 2,800 premature infants, and 3,000 congenital malformations could be expected, the minimum numbers necessary to answer the projected questions. The crux of the research effort was to study a large number of cases in great detail in order to evaluate the effects of perinatal factors on the health of the individual child.

In the offspring, disorders of the nervous system and abnormalities of any other body systems appearing at the time of delivery, during infancy, or in early childhood were to be evaluated. Included were cerebral palsy, mental subnormality, behavioral disorders, and other specific neurological or sensory defects. Related areas of investigation included identification of: (1) factors operative in early and late fetal loss, prematurity, and infant and early childhood mortality; (2) the relationship of physical and mental development in early childhood to genetic, biological, and environmental factors; (3) events in the postnatal environment related to the later development of disturbances of structure and function in the nervous system; and (4) more precise definitions of the events of the reproductive process and the nature of its outcome.

The investigation of relationships between factors and conditions affecting parents and the occurrence and course of abnormalities in offspring was to be accomplished by analysis of pooled information. Data were to be

collected uniformly in the collaborating medical centers on women studied during pregnancy and from their offspring followed from infancy through early childhood. The research effort was directed toward the reevaluation of the effects of factors already suspected in the etiology of abnormal outcomes of pregnancy. In addition, the research sought to clarify mechanisms through which these factors were operative, and the discovery of factors not presently known or suspected.

Areas to be investigated in the parents included: (1) the conditions and complications of pregnancy, the normal and abnormal physiology of pregnancy, labor, and delivery; (2) environmental factors influencing the mother, including social and economic conditions, emotional stress, and medical care; (3) biological factors, such as age, parity, medical and reproductive history, and immunological characteristics; and (4) the genetic background of the parents.

Many managerial details had to be worked out in an undertaking that would cover a wide geographic area and involve teams of professional men and women working under many different systems of hospital and clinical management. In planning the NCPP, these details had to be considered carefully, and were often revised and then re-revised as efforts were made to gather uniform data across the country.

Collaborating Institutions

In 1956, NINDB personnel visited hospitals across the country to acquaint the hospitals' staffs with the general outline of a prospective study of pregnancy and the children thereof. Interest was generated and numerous applications from a variety of institutions were received.

In 1957, after staff members and consultants from the NINDB visited the major medical schools and centers of the United States and acquainted them with the research plans, the following fourteen institutions were approved to participate in the NCPP: Boston Lying-In Hospital and Children's Medical Center; Buffalo Children's Hospital; Charity Hospital, New Orleans; Children's Hospital, San Francisco; Columbia-Presbyterian Medical Center; Johns Hopkins Hospital; Medical College of Virginia; University of Minnesota; New York Medical College; University of Oregon; Pennsylvania Lying-In Hospital and Children's Hospital (Philadelphia); Providence Lying-In Hospital and Brown University; University of Tennessee; and Yale University.

Data collection for the NCPP was initiated in 1959. During the first three months of that year, nine collaborating institutions began registering obstetrical patients; all institutions had begun patient registration by October of 1960, when Buffalo Children's Hospital entered the study. Table 1.1 lists the collaborating institutions that enrolled subjects, the dates between which the registration of subjects occurred, and the approximate number of subjects enrolled.

Withdrawal of Collaborating Institutions

During the course of the study, four collaborating institutions dropped out for various administrative reasons. In 1959, Children's Hospital in San Francisco withdrew after participating in the pretest phase. In 1961, after registering about 900 pregnancies, Yale withdrew from active participation. In 1963, Columbia Presbyterian Hospital terminated obstetrical intake for the NCPP, but continued to follow the children through eight years of age. On July 1, 1970, New York Medical College, having completed the examination of children at age four, terminated because of insurmountable difficulties in achieving an adequate retention rate for children at age seven. While other institutions had a return rate of over 70%, the rate at New York Medical College was only 45% to 55%, and researchers held little hope of improvement because of the mobility of the Puerto Rican population in the sample.

In summary, twelve medical centers contributed obstetrical patients to the NCPP. All of the centers were in urban areas; six were in the Northeast, four in the South, one in the West, and one in the north-central region of the U.S.

TABLE 1.1. Institutions, Location, and Approximate Number of Core Registrants

NINDS INSTITUTION NUMBER	ABBREVIATION	LOCATION	INSTITUTION	REGISTRATION DATES		THOUSANDS OF SUBJECTS
				FROM	TO	
05	BO	Boston, MA	Lying-In Hospital Children's Medical Center	1/59	12/65	13.2
10	BU	Buffalo, NY	Children's Hospital, SUNY	10/60	12/65	3.0
15	CH	New Orleans, LA	Charity Hospital	3/60	12/65	2.6
31	CO	New York, NY	Columbia-Presbyterian	1/59	4/63	2.2
37	JH	Baltimore, MD	Johns Hopkins Hospital	1/59	12/64	3.8
45	VA	Richmond, VA	Medical College of Virginia	1/59	12/65	3.2
50	MN	Minneapolis, MN	University of Minnesota	1/59	12/65	3.3
55	NY	New York, NY	New York Medical College	2/59	12/65	4.7
60	OR	Portland, OR	University of Oregon Medical School	3/59	12/65	3.3
66	PA	Philadelphia, PA	Pennsylvania Hospital Children's Hospital of Philadelphia	1/59	12/65	10.3
71	PR	Providence, RI	Providence Lying-In	3/60	12/65	2.9
82	TN	Memphis, TN	University of Tennessee College of Medicine	10/59	12/65	3.6
						56.1

COLLABORATIVE PERINATAL PROJECT ADMINISTRATION

A central coordinating staff was established at NINDB and in 1959, when data collection began, its size increased. The Project Services Branch of NINDB provided professional, administrative, and logistic support, and the NINDB Biometrics Branch (later renamed Office of Biometry) was responsible for data management and retrieval, and statistical services. The central staff developed close liaison with the project directors at the participating hospitals.

In the fall of 1960, the Perinatal Research Branch was established to combine the capabilities required to direct and coordinate the NCPP and, at the same time, to explore leads emanating from the study. An infectious diseases laboratory and a pathology laboratory also participated in the NCPP.

During the conduct of the NCPP, a number of committees and task forces were established to provide guidance for the study. A Perinatal Research Committee was established to provide overall outside guidance. This committee was empowered with the authority to advise NINDB on fiscal matters relevant to the NCPP, as well as on research policy and scope. The Perinatal Research Committee also reviewed all NCPP data collection research grant applications and assisted NINDB in evaluating the performance of the NCPP. A Principal Investigators' Committee, consisting of senior professionals from all institutions involved in the project, was responsible for the conduct of the study in the respective institutions and in the project as a whole. The collaborating institutions were also represented by the Project Directors' Committee. This committee was composed of the project directors (managers) from the individual institutions. It decided on the administrative desirability and feasibility of various facets of study design, data collection and data analysis. The project directors were responsible for the managerial aspects of implementing policy decisions made by the Principal Investigators' Committee.

The Perinatal Research Branch and the Project Directors' Committee both set up advisory committees in the disciplines relevant to the project. Advisory subcommittees to the Perinatal Research Branch included obstetrics, pediatrics, psychology, pathology, speech-language-hearing, socioeconomic, sample maintenance, statistics, administration, interviewing, and editing. The Project Directors' Advisory Committees were made up of the senior professional representatives from every institution at the operational level in the disciplines of obstetrics, pediatrics, pathology, psychology, speech-language-hearing, sample maintenance, interviewing, and editing. The committees' functions included advising the Project Directors' Committee on professional issues recommended by the Perinatal Research Branch.

The size and complexity of the NCPP required a highly developed and integrated staff to conduct the research developed and directed by the above referenced committees. The participants at each collaborating institution are included in an appendix of Niswander and Gordon (1972).

A change in organizational structure within NINDB took place in 1967, when the Office of Biometry was established. This office was given responsibility for support in data retrieval and statistical consultation for

the NCPP. At the same time, the support of the collaborating institutions was changed from grants to contracts in response to the criticism that the study lacked strong central direction. This move coincided with the beginning of a formal inter-institutional quality control program to obtain test-retest reliability measures for the examinations given to the children at ages four, seven, and eight.

STUDY SAMPLE SELECTION AND COMPOSITION

The Collaborative Perinatal Project was not intended to develop incidence or prevalence rates of events of pregnancy, or of conditions in the offspring, and was not concerned with the selection of a study sample representative of the population of the United States or of the community in which the study center was located. Instead, the objectives required a broad spectrum of pregnancy conditions, free from biases based on special interests. The samples were chosen so as not to interfere with the routine of the maternity clinic, and so as to promote continuing follow-up of the children. A restriction was the case load a given institution could handle; this ranged from around 300 to 2,000 patients per year, with most institutions providing around 500 to 800 cases annually.

Some participating institutions selected all eligible women; others, only a random sample. Eligibility was defined by the sampling frame (a defined group of patients from which registrants for the study were chosen) for each institution. A common exclusion was "walk-ins" or patients with no prenatal visits prior to the day of delivery. All but one of the 12 sampling frames consisted of clinic patients. The socioeconomic and ethnic composition of the NCPP population was representative of the populations qualifying for medical care at the participating institutions. Detailed information on the selection methods and sampling frame for each institution can be found in Niswander and Gordon (1972) and in Appendix A to this volume.

Selection ratios at the various hospitals ranged from 10 to 100 percent of the sampling frame. This resulted in the creation of 58,760 NINDB case numbers, each uniquely identifying the institution, the type of patient selection, and the gravida or child. Because some institutions used NINDB forms for special concurrent studies, for walk-ins, or for other purposes, several categorizations were developed to identify which cases met the study criteria. These are shown in Table 1.2 and are further described in Appendix A.

Case selection was continuously monitored. The first case was selected on January 2, 1959; the last, on December 31, 1965. The earliest delivery occurred on January 11, 1959; the last, in September of 1966.

Sample sizes by institution and ethnic group are shown in Table 1.3. The largest sample was collected at the Boston Lying-In Hospital, Boston, Massachusetts, where the sampling frame consisted of all clinic patients admitted for prenatal care. Special exclusions were unwed mothers planning to place their babies for adoption. The selection ratio was initially 50%, but was raised to 100% after two months. Nearly 90% of the Boston sample (12,000 cases) were white. Children were followed at the Children's Hospital Medical Center.

TABLE 1.2. Categorization of NCPP Cases

CATEGORY	DESCRIPTION	NUMBER OF CASES
Core	Cases Meeting General Study Criteria	56,134
Non-Core	Ancillary Cases Used By Individual Institutions	2,626
		58,760
Walk-In	Cases That Delivered on Same Day Registered	298
Cohort I	Core Cases Excluding Walk-Ins	
Cohort IA	First Study Pregnancy, Single Birth, Registered on or Before 12/31/64	37,998
Cohort IB	All Cases Registered on or Before 12/31/64	48,488
Cohort IC	First Study Pregnancy, Single Birth, Registered at Any Time	43,521
Cohort ID	All Cases Registered at Any Time	55,857 55,908*
Cohort II	Core Cases Excluding Walk-Ins and Lost to Study	
Cohort IIA	Cohort IA Minus Cases Lost to Study	37,579
Cohort IIB	Cohort IB Minus Cases Lost to Study	46,052
Cohort IIC	Cohort IC Minus Cases Lost to Study	43,073
Cohort IID	Cohort ID Minus Cases Lost to Study	53,043
Cohort IID Rev.	Cohort ID Minus Cases Lost to Study	53,039
Basic Document Cohort	Cohort IC Minus Abortions	42,878

*Cohort ID was reported in Women and Their Pregnancies as 55,908 cases; this included 51 registrants who were not pregnant and were subsequently deleted from the data base.

TABLE 1.3. Sample Size By Institution and Ethnic Group in the NCPP Population (Cohort IID)*

INSTITUTION	ETHNIC GROUP				TOTAL
	WHITE	BLACK	PUERTO RICAN	OTHERS	
Boston Lying-In Hospital	10,803	1,198	25	167	12,193
Children's Hospital, Buffalo	2,383	59	12	15	2,469
Charity Hospital, New Orleans	0	2,582	0	0	2,582
Columbia-Presbyterian Medical Center	633	876	602	27	2,138
Johns Hopkins Hospital	798	2,744	1	6	3,549
Medical College of Virginia	831	2,367	0	6	3,204
Univ. of Minnesota Hospital	2,986	19	2	140	3,147
New York Medical College	269	1,558	2,630	17	4,474
Univ. of Oregon Medical School	2,216	861	1	72	3,150
Pennsylvania Hospital	882	8,580	316	14	9,792
Providence Lying-In Hospital	2,096	672	5	49	2,822
University of Tennessee College of Medicine	22	3,501	0	0	3,523
Total	23,919	25,017	3,594	513	53,043

*From Broman et al. (1975)

The sample from Children's Hospital of the State University of New York at Buffalo was unique because it consisted of private patients referred by several obstetricians. Women who planned to move out of the area or deliver in another hospital were excluded. More than 95% of the 2500 patients selected were white.

The sampling frame at Charity Hospital, New Orleans, Louisiana, included black patients residing in Orleans Parish and assigned to the Tulane or Louisiana State University medical services in the hospital. The selection ratio varied from one in ten to one in six of the eligible patients and produced a sample of approximately 2500.

All clinic patients admitted to the Columbia-Presbyterian Hospital in New York were included in the sampling frame, selection varied between one in six and one in five patients. After April, 1962, difficulty in follow-up necessitated exclusion of patients residing outside Manhattan or the Bronx. In April of 1963 case selection was terminated. The NCPP sample of 2100, the smallest from any institution, was approximately 30% white, 40% black, and 30% Puerto Rican.

The sampling frame at Johns Hopkins Hospital, Baltimore, Maryland, consisted of all clinic patients living in metropolitan Baltimore; transients and patients referred to county clinics for obstetrical care were excluded. The selection ratio was initially 20% and was raised to 30%, 40%, and finally 100%. Nearly 80% of the 3500 patients selected were black.

Clinic patients residing within a 50 mile radius of the Medical College of Virginia, Richmond, Virginia, were initially included in the sampling frame. Later, the area was reduced to Richmond and three surrounding counties, and still later to the city itself. Excluded were white welfare cases and patients planning to put their children up for adoption. The selection ratio was 100% of the white patients and was increased from 25% to 100% of the black patients. About 75% of the 3200 patients selected were black.

The sampling frame at the University of Minnesota Hospital, Minneapolis, Minnesota, included all clinic patients, although in the first year of registration, women who were divorced, separated, widowed or unmarried before the start of their pregnancy were excluded. The selection ratio was 100% and yielded a sample of 3100. Ninety-five percent of the sample were white.

The sampling frame of all clinic patients was also used at the Metropolitan Hospital of New York Medical College in New York. An initial selection of one in 10 patients chosen from the sampling frame was gradually increased to one in six. The NCPP sample of 4500 was about 60% Puerto Rican and 35% black.

At the University of Oregon Medical School in Portland, the sampling frame again consisted of all clinic patients. Residence requirements were later restricted to certain areas within Multnomah County. Medical students' wives and clients of private adoption agencies were excluded. The sampling ratio varied from one in three to two in three. The sample of 3200 was approximately 70% white.

The second largest sample came from Pennsylvania Hospital in Philadelphia. The sampling frame was the set of all clinic patients except those planning to deliver elsewhere. The selection ratio was 100% and 90% of the nearly 10,000 patients selected were black. Children were followed at the Children's Hospital of Philadelphia.

The sampling frame at Providence Lying-In Hospital, Providence, Rhode Island, was defined as all clinic patients. The NCPP sample of 2800 consisted of about 45% of the sampling frame and was 75% white. Children were followed at the Child Study Center of Brown University.

At the University of Tennessee College of Medicine in Memphis, the sampling frame included all clinic patients living inside the city limits. Initially the patient selection ratio was one in 10, but it was raised to one in seven within six months.

In summary, of the 53,000 pregnant women registered in the NCPP, 95% were clinic patients. All were from urban areas, 64% from cities in the north-eastern U.S. Forty-five percent of the women registrants were white, 47% black, 7% Puerto Rican, and 1% from a variety of other ethnic groups. Broman et al. (1975) include an extensive analysis of the demographic characteristics of the NCPP study population which, although too extensive to reproduce here, is of considerable potential value to the researcher.

Women Lost to the Study

The important question of the characteristics of the women who dropped out of the NCPP before completion of their pregnancy is considered as a potential source of bias by Niswander and Gordon (1972). They note that 4.1% of the study registrants were lost to the study before the completion of their pregnancy.

They conclude that on the whole, fewer of the very young white or black women dropped out than might normally be expected. This difference was not consistent by collaborating center.

Among women of both races, the more highly educated mothers were lost to the study more frequently than were those of the lower educational group. Again, the disparity between the two groups was not consistent across the collaborating centers.

An excess of nulliparas and a reduced frequency of grand multiparas occurred in the group lost to study as compared with the study population. Though the trend was reasonably consistent by collaborating center, no consistency was observed in marital status of the women who were nulliparas or grand multiparas.

For the characteristics compared, some disparities were present between the lost-to-study gravidas and those of the study population. With the exception of the number of prior pregnancies, little consistency could be found in the differences by collaborating center.

Niswander and Gordon (1972) also summarize the differences between study women and women lost-to-study, with regard to the distributions of certain characteristics and perinatal mortality rates. These data are presented in Table 1.4.

Perinatal mortality rates among study women age 35 or older, of both races, were higher than rates among women in the intermediate ages. Importantly, the lost-to-study women do not show a disproportionate percent of cases in this age bracket.

Data did not indicate that education of the gravida was related to the perinatal mortality rate. Therefore, the large number of lost-to-study women in the high education group was not thought to be a biasing factor with regard to perinatal death rate.

TABLE 1.4. Comparison of Selected Characteristics of Study and Lost-to-Study Gravidas

ITEM	WHITE			BLACK		
	LOST TO STUDY GRAVIDAS		STUDY GRAVIDAS	LOST TO STUDY GRAVIDAS		STUDY GRAVIDAS
	PERCENT	PERCENT	PERINATAL MORTALITY RATE*	PERCENT	PERCENT	PERINATAL MORTALITY RATE*
AGE OF GRAVIDA (YRS)						
UNDER 18	3.4	6.3	23.5	9.9	15.9	38.0
18-34	86.2	85.9	53.2	82.0	76.9	40.7
35+	8.4	7.8	63.8	8.1	7.2	63.7
EDUCATION OF GRAVIDA (YRS)						
UNDER 9	10.5	13.1	33.2	8.2	19.0	42.7
9-11	49.8	65.7	33.6	80.7	75.9	41.9
12+	39.7	21.1	32.0	11.1	5.1	38.2
NO. OF PRIOR PREGNANCIES						
0	48.9	40.0	29.1	36.7	34.1	39.7
1-5	48.2	56.0	33.6	58.5	57.3	40.1
6+	2.9	4.0	70.1	4.8	8.6	52.3
MARITAL STATUS						
MARRIED	93.7	86.6	35.5	67.5	60.4	40.1
NOT MARRIED	6.3	13.4	32.5	32.5	39.6	44.7

* Per 1000 births (Niswander and Gordon, 1972)

White and black women having six or more prior pregnancies showed the highest perinatal mortality rates. A smaller percentage of lost-to-study women fell into this category than was the case with study women. If one were willing to assume that the babies of lost-to-study mothers had the same perinatal death rates as babies of study mothers of comparable parity, then the overall perinatal mortality rate would change from 35.1 to 34.9 for whites and would be unchanged for blacks when the lost-to-study cases were included. The impact of these cases lost-to-study would not appear to be very significant.

Similarly, the perinatal mortality rate for babies born to never-married white women was lower than that for babies born to married white women. The rate of never-married women among the cases lost to study was half the rate of the study women. An adjustment comparable to that described above would not change the overall perinatal mortality rates if the cases lost-to-study were included.

We have included this discussion of differences between the lost-to-study gravidas and study gravidas in relation to perinatal mortality because we believe it illustrates a lack of significant bias resulting from lost-to-study patients. While the results of analyses related to perinatal mortality can

not be generalized to all outcomes of interest, it is reassuring to note that loss of women from the study has apparently not been an important biasing factor with regard to this major outcome variable.

Children Lost to the Study

Problems of follow-up, inherent in all long-term, longitudinal studies, are especially present in multicenter studies conducted in a highly mobile, free society. Throughout the NCPP, a considerable effort was expended to prevent attrition of the study population and reasonable success was achieved. The number and characteristics of children lost to study from the various collaborating centers are discussed in detail in Niswander and Gordon (1972), Broman et al. (1975), and Hardy et al. (1979). Discussion directed specifically to the children included in the speech, language and hearing exams can be found in Lassman et al. (1980).

Follow-up rates for survivors of the total population of 53,042 pregnancies were 85% at one year, 75% at four years and 79% at seven years.* At three and eight years, speech, language and hearing exams were given at selected institutions; follow-up rates were 48% and 47%, respectively. The reduction in subjects for the speech, language and hearing examinations is considered below.

Hardy et al. (1979) note that children who missed one examination did not necessarily miss later ones. They state that a child's missing an evaluation led to intensified efforts to get the child to return for the next examination. While the rates for follow-up were high for the combined institutions, some of the centers were more successful than others (Niswander and Gordon, 1972). As Hardy et al. (1979) point out, however, review of rates of children with major abnormalities by institution shows reasonable consistency, indicating no obvious patterns in the loss of cases.

At Johns Hopkins University in 1962, a study of 50 consecutive infants who missed the one-year examination was conducted (Hardy et al., 1979). A pediatric-neurologist and a nurse examined each infant on a home visit. It was found that most of the children had failed to return because their mothers had other young children to care for, family illness or other problems, rather than for reasons related to the presence or absence of abnormalities of the study child.

Hardy et al. (1979) present data comparing characteristics of the children examined to characteristics of the children from the whole NCPP study population lost to follow-up at age one. This was done to evaluate the introduction of possible bias occurring as a result of loss to follow-up. Variables collected earlier and known to relate to abnormal pregnancy outcome, such as selected maternal characteristics and low birth weight, were compared for those infants receiving the one-year neurological examination and those

*Detailed data on the distributions of infants examined, death and attrition through one year are presented in Niswander and Gordon (1972). Overall, more than 85% of all children were examined at four months and a similar percentage had the neurological examination at age one year (Hardy et al., 1979).

who failed to receive it. If the frequency of these predictive variables in the two groups is similar, then there would be no evidence of bias due to missed examinations.

For the study as a whole, relatively small differences existed between infants examined and not examined at one year. For three maternal characteristics, age, parity and education, differences were slight and inconsistent. For three characteristics of the newborn, sex of the child, birthweight and Apgar score, essentially no difference existed between the groups.

In their discussion of IQ at four years of age, Broman et al. (1975) compare the total NCPP cohort, the sample they studied at the four-year exam, those lost by attrition, and those lost by death. Comparing a number of demographic and neonatal characteristics, they found the groups to be very similar.

In summary, characteristics of the subjects lost to follow-up have been found to be similar to those of the subjects examined at both the one-year and four-year evaluations. Active follow-up methods were employed to keep the loss-to-follow-up percentages low and a concerted effort was put forward to maintain the sample.

Sample Size Reduction at Later Ages: Attrition of Collaborating Institutions

In the latter part of the study, another type of problem related to the administration of the speech, language, and hearing (SLH) examinations arose.

Most of the collaborating institutions began to perform the three-year SLH tests as the children in their respective samples became of age. Some losses of study data resulted from failures to keep appointments and from uncooperative behavior by the children. The greatest loss of three-year data was caused by difficulties in providing space and staff at Boston and Philadelphia, where the largest numbers of NCPP cases were registered. These two institutions found it necessary to screen their three-year-old children for possible SLH problems and to test only the smaller number who were considered to have SLH deficits. Screening was performed by nonspecialists during home visits. As a result, only 1059 children who were given SLH tests at Boston and 460 who were tested at Philadelphia were available for analysis after culling for correctness of age and NCPP registration and completeness of data. Data from these two subsamples were analyzed by the study staff and were examined by a special group of consultants as part of an inquiry into the quality of the total NCPP data base. These data were judged suitable for inclusion within the larger SLH study subsamples because the prevalence and types of SLH problems in these subsamples did not differ significantly from those at other NCPP institutions.

The eight year SLH examination likewise was routinely scheduled for all children once they reached the prescribed age range. Boston, however, was delayed by problems of space and staff until two years after the first children registered there came of age. At midyear in 1970 the eight year examination was discontinued for administrative reasons at six of the 12

participating centers. Examinations continued at Children's Medical Center in Boston, Children's Hospital at Buffalo, Johns Hopkins Hospital in Baltimore, and the Universities of Tennessee, Minnesota and Oregon. The reduction in the size of the eight-year SLH subsample, as compared with the NCPP population, is mainly a reflection of these events.

THE EXPANDED RESEARCH PROGRAM AND RESEARCH AREA

In 1970 the Perinatal Research Committee established a number of ad hoc task forces. These task forces included the following: Basic Document Task Force; Task Force on Toxemia of Pregnancy; Task Force on Labor and Delivery; Task Force on Speech, Language and Hearing; Task Force on Physical Growth and Development; Task Force on Four and Seven-Year Data; Task Force on Congenital Malformations; and Task Force on Pathology. An Epidemiological and Statistical Advisory Committee was established, as was a Coordinating Committee for Data Analysis. This latter committee developed a master plan for data analysis to insure that the objectives of the NCPP were attained. The master plan established priorities, coordinated task force studies, and contained a consensus on methodologies from the Epidemiological and Statistical Advisory Committee.

A Comprehensive Plan for Analysis and Interpretation of Collaborative Perinatal Project Data was developed in the early 1970s. After a careful review of the objectives of the NCPP, the data available for analysis and the work in progress, it was recommended that major efforts in analysis and interpretation were needed in ten primary areas to meet the basic objectives of the project. The ten primary areas are: Cerebral Palsy; Mental Retardation; Communicative Disorders; Visual Abnormalities; Convulsive Disorders; Learning Disorders; Minimal Brain Dysfunction; Congenital Malformations; Birthweight - Gestational Age Relationships (Prematurity); Neuropathology, General Pathology and Placentology.

Implementation of the comprehensive plan was through a team of researchers for each of the ten primary areas, each team headed by a member of the professional staff of the Perinatal Research Branch. On each team was a member of the Office of Biometry staff and a member of the Perinatal Research Branch's Section for Production of Data Analysis. In each of the ten primary areas, a program plan was developed to expand on the summary statements in the comprehensive plan and to provide a detailed approach to the major components of the area and required data analysis. Monographs in book form were planned in each of the areas.

In addition to the ten primary areas, ten secondary areas of analytical focus were identified. These secondary areas are: Pregnancy Hypertension; Maternal Infection During Pregnancy; Labor and Delivery; Neonatal Hyperbilirubinemia; Maternal Anesthesia - Analgesia During Labor and Delivery; Intelligence Test Scores at Age Four; Physical Growth and Development (Birth to Seven Years); Multiple Births; Genetic and Socioeconomic Factors; Drugs Taken During Pregnancy.

In August of 1975, the Developmental Neurology Branch was created within the Neurological Disorder's Program of NINCDS. With this organizational change, the Perinatal Research Branch became the Perinatal Research Section

within the newly formed Developmental Neurology Branch. A major objective of the Developmental Neurology Branch was the completion of the Comprehensive Plan for Analysis and Interpretation of Collaborative Perinatal Project Data.

Analysis and reporting of the data from the NCPP have been carried out by a diverse group of investigators. Some have been entirely within the individual collaborative institutions, others within the Perinatal Research Branch; some studies were based on investigator-initiated grants while others were supported by individual contracts for a particular analysis. Analyses and reports in the identified primary and secondary areas were primarily carried out by NINCDS staff or through specific contracts with outside investigators.

Major books and monographs from the study include the following: The Women and Their Pregnancies (Niswander and Gordon, 1972), Blood Pressure, Edema and Proteinuria in Pregnancy (Friedman et al., 1976), Pregnancy Hypertension (Friedman and Neff, 1977), Birth Defects and Drugs in Pregnancy (Heinonen et al., 1977), Congenital Malformations in Singletons (Myriantopoulos and Chung, 1974), Congenital Malformations in Twins (Myriantopoulos, 1975), External Ear Malformations: Epidemiology, Genetics, and Natural History (Melnick and Myriantopoulos, 1979), The First Year of Life (Hardy et al., 1970), Preschool IQ: Prenatal and Early Developmental Correlates (Broman et al., 1975), Minimal Brain Dysfunction: A Prospective Study (Nichols and Chen, 1981), Early Correlates of Speech, Language, and Hearing (Lassman et al., 1980), and The Developing Human Brain. Growth and Epidemiologic Neuropathology (Gilles et al., 1983).

THE NINCDS COLLABORATIVE PERINATAL PROJECT BIBLIOGRAPHY

We have mentioned only briefly the research conducted using the data from the NCPP. For the researcher who is interested in using these data, we strongly recommend obtaining a copy of the NINCDS Collaborative Perinatal Project Bibliography. The bibliography is available from the Developmental Neurology Branch at NINCDS and lists all publications that have included data from the NCPP. Of particular usefulness is the subject index, which readily allows an investigator to identify publications by specific research areas. Over 500 publications have been based, at least in part, on the NCPP; complete citations can be found in the bibliography.

SUMMARY

In this chapter we have provided a discussion of the background, history, organization, subject selection and research areas of the NCPP. We have drawn extensively from the published works of Niswander and Gordon (1972), Broman et al. (1975), and Hardy et al. (1970), as well as from NINCDS publications and documents to provide the researcher with an introductory overview of the study.

CHAPTER 2. DATA COLLECTION METHODOLOGY

THE PRETEST: FORM DESIGN AND DEVELOPMENT

Collection of data for the NINCDS Collaborative Perinatal Project required a well established and reviewed methodology. Because of the prospective nature of the study, data were collected as soon after an event as possible. In addition, they were collected without reference to antecedent events to provide as unbiased a data base as possible.

Hardy et al. (1979) review some of the special challenges involved in NCPP data collection. They identify the following key features: (1) data collection took place over a sixteen-year time span; (2) large amounts of reliable, standardized, highly specific and detailed information on each mother-child pair were required; (3) a very large number of women, and later children, were enrolled; (4) the geographic distribution of the collaborating centers required that special attention be paid to communication; and (5) the study personnel changed over time and varied in their professional orientation.

The collection of study data required that standardized protocols, forms and manuals be developed with the above challenges in mind. In addition, developers of the study instruments sought to: (1) include detailed and comprehensive information required for thorough etiologic studies; (2) reduce ambiguity to a minimum, assuring reproducibility and comparability of information collected over time by different examiners and institutions; and (3) simplify and standardize the processing of information at all stages of data collection and processing.

To meet these requirements, specific forms and manuals were developed through collaboration of staff members from NINDB, the participating institutions, and consultants. The Perinatal Research Committee and a number of ad hoc committees, described in Chapter 1, devoted much time and attention to all aspects of data collection and production. A large number of task forces developed data collection protocols, which, after a series of revisions, resulted in the production of pretest forms in 1957. The design of the study required that all data be collected and recorded in a uniform fashion as quickly as possible after an event. Because most of the data collection forms were structured and precoded, positive findings and responses were described more extensively in a special section of each form. Information was collected and recorded by specially trained and highly skilled interviewers and examiners using detailed instruction manuals. Standardized manuals, available for every form, provided instructions on how the form should be filled out and definitions of specific items of information, such as diseases.

When the pretest forms were introduced for evaluation by the collaborating institutions on January 1, 1958, a number of difficulties were identified. Many pregnant women resented some of the questions asked. Physicians often objected to working with structured forms that took a long time to complete and contained many questions that were thought to have little relevance to patient care. Another problem was duplication of effort; certain

hospitals continued to use their own hospital records in addition to completing the NCPP study forms. In many instances, however, the new forms were adopted by the hospital and became the official hospital record.

In parallel with the development of the forms, procedure manuals were developed to ensure uniformity of data collection. Certain instructions in the use of the protocols seemed too stringent to a few institutions. Adjustments were permitted, which in turn affected the uniformity with which the records were used across institutions.

During the pretest period, the staff of the coordinating unit at NINDB was small and only limited evaluation of the preliminary data was possible. Review at this point included an assessment of the problems encountered in the use of the data collection forms, an examination of the uniformity and adequacy with which the data were collected, and identification of population differences as shown in the reported data (Broman et al., 1975).

While the institutions were evaluating the pretest protocols, the NINDB initiated workshops and training sessions for staff concerned with data collection. Designed to assure accuracy of the examinations and observations, the sessions provided specific instructions in interviewing procedures, examination techniques and other relevant factors.

After two years of protocol development, pretesting and revision, the collection of study data began on January 1, 1959. Because some of the collaborating hospitals continued to have problems with the use of the obstetrical protocols, protocols were revised even after January 1, 1959 so that they could serve as hospital as well as research records. Timing and other aspects of forms revision are considered in the next section.

CATEGORIES OF DATA COLLECTED

Data for the NCPP were collected in a number of major categories, including: obstetrics, pediatrics, pathology, serology, socioeconomics, genetic history, psychology, speech, language, and hearing. A more complete discussion of the data categories may be found in Chapter 4. The data were collected over a 16 year period with final data collection for eight year speech, language and hearing occurring in 1974 (Figure 2.1).

All of the major data categories required multiple forms for the collection of information relevant to research variables of interest. In addition to forms required to collect study data, general and administrative forms were used to maintain subject records and to ensure smooth communication between the collaborating institutions and NINDB. The forms for these areas were developed through the collaborative activities of specialists in the individual fields. In all, 97 different forms were used. Copies of the individual forms, with the exception of a few of the administrative forms, are included as part of Volume II of this guide.

In listing the forms used in the NCPP by major category (Table 2.1), we have included the form identification codes (an abbreviation of the data category and a form number) and the titles of the forms. The major phases of the data collection process, indicated in Table 2.2, also appear in

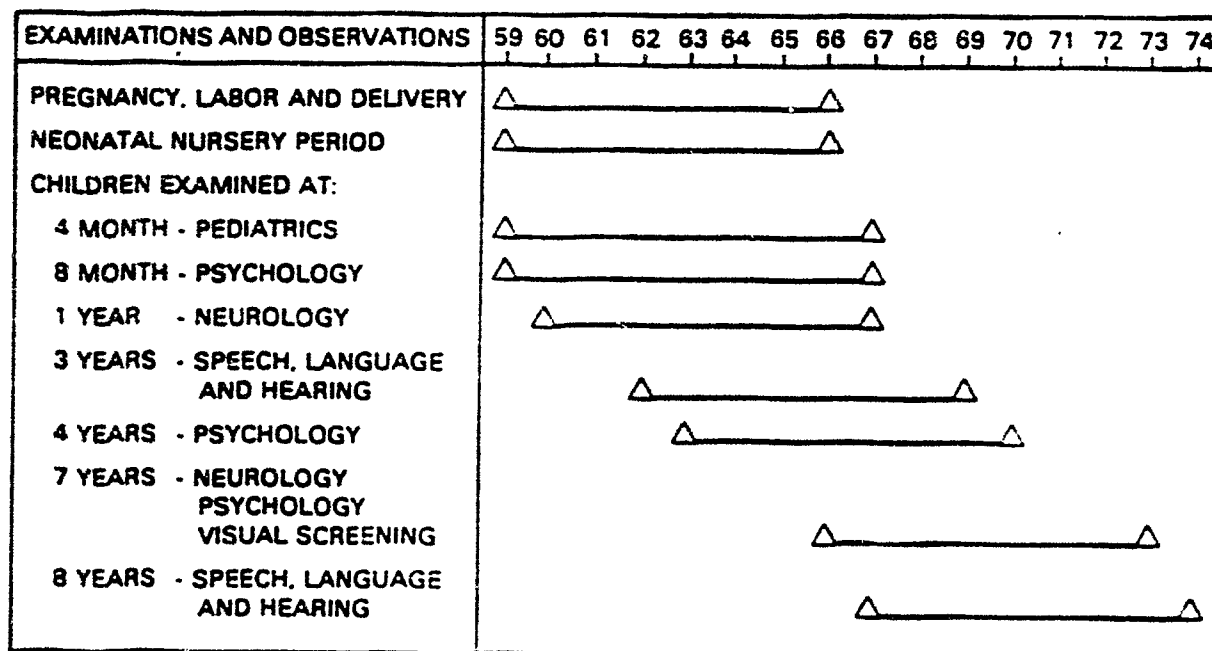


FIGURE 2.1. Data Collection for the NCPP, 1959-1974

Figure 2.1. Included here are the types of information collected at each examination and the relevant data collection forms, identified by form code.

FORM MODIFICATION AND REPLACEMENT

As noted earlier, obstetric protocols were revised because they presented problems at some institutions after the study began. Protocol revision had important implications. The new forms differed substantially from those used initially, and as a result, pregnancies registered from 1959 through the early part of 1962 were recorded on the original OB forms. Beginning in the spring of 1962, however, pregnancies were observed and recorded using the revised protocol. At the beginning of the study, family history information at registration was collected using a series of forms FHH 1-4. Subsequently, these forms were replaced by GEN 5-8 (1961) and SE-1 (1963). Table 2.3 shows a summary of form replacements. Changes in other forms, such as pediatrics and psychology, were of a lesser degree and mostly minor in nature.

All the study forms were printed by NINDB on multicopy NCR paper and distributed to the collaborating institutions. Included on each form is the revision number and the date for the form. Some revisions were cosmetic only. The color of the identification stripe may have changed, for example, or the ordering of the questions may have been changed. Other revisions involved rewording of some questions, while a third type of change involved addition of new data items. Each form and its revisions are included in Volume II. Form revisions were assigned numbers that are included on all punched cards copied onto the master data file.

TABLE 2.1. Forms Used in the NCPP by Subject Area

GENERAL AND ADMINISTRATIVE

AR-2	Administrative report for record inventory and patient followup	CP-9	Report of transmission of specimen by hospital neuropathologist
AR-3	Notice of identification change	CP-5	Continuation sheet for obstetric or pediatric forms
AR-4	Inventory of completed seven-year examinations		
AR-5	Sample maintenance data		
AR-8	Inventory of completed eight-year examinations		

OBSTETRICS

AR-1	Obstetrical administrative record	OB-33	Delivery room events
OB-2	Reproductive history	OB-34	Obstetrician's summary of labor and delivery
OB-3	History since last menstrual period	OB-35	Anesthesia record
OB-4	Gynecological history	OB-40	Optional prenatal record
OB-5	Recent medical history	OB-42	Past medical history
OB-6	Past medical history	OB-43	Initial prenatal examination
OB-7	Infectious disease and system review	OB-44	Prenatal observations
OB-8	Repeat prenatal history	OB-45	Laboratory record
OB-9	Prenatal record	OB-46	Physician's clinic record
OB-10	Return visit and laboratory record	OB-47	Summary of antepartum hospitalization (also used in reporting death for both pre- and post-partum)
OB-12	Summary of hospitalization	OB-50	Admission history
OB-15	Drugs in Pregnancy	OB-51	Admission examination, Part I, General examination
OB-30	Admitting record by obstetrician	OB-52	Admission examinations, Part II, Abdomino-pelvic examination
OB-31	Admitting examination by obstetrician	OB-55	Delivery report
OB-32	Labor room record	OB-56	Obstetric summary
		OB-57	Anesthetic agents
		OB-58	Summary of puerperium
		OB-60	Obstetric diagnostic summary

PEDIATRICS

- | | | | |
|--------|---|--------|--|
| PED-1 | Delivery room observation of the neonate | PED-12 | Summary of the first year of life after the duration summarized on PED-8 |
| PED-2 | Neonatal examination | PED-14 | Physical growth measurements |
| PED-3 | Nursery history | PED-20 | Interval medical history |
| PED-4 | Report of fetal or infant death (also used for child death) | PED-29 | Summary of medical records of illness or hospitalization (if indicated) |
| PED-5 | Results of tests and procedures done on the neonate | PED-74 | Ophthalmology consultant's report, vision screening study (if indicated) |
| PED-6 | Neonatal Neurological examination | PED-75 | Visual screening at 7 years |
| PED-7 | Summary of hospital course of neonate | PED-76 | Pediatric-neurologic examination at 7 years |
| PED-8 | Newborn diagnostic summary | ADM-86 | interdisciplinary diagnostic code one year - seven years (IDC-77) |
| PED-10 | Four-month pediatric examination | | |
| PED-11 | One-year neurological examination | | |

SEROLOGICAL AND PATHOLOGICAL STUDIES

- | | | | |
|-------|--|--------|--|
| VIR-1 | Blood samples for viral serological study | PATH-1 | Placental examination - gross |
| VIR-3 | Special rubella study (used limited time in Project) | PATH-2 | Placental examination - microscopic |
| | | PATH-3 | Autopsy protocol - complete autopsy report |

PSYCHOLOGICAL EXAMINATION AT 8 MONTHS

- | | | | |
|------|--|------|--|
| PS-1 | COLR Research form of Bayley Scale of Mental Development | PS-3 | Infant behavior profile |
| PS-2 | COLR Research form of Bayley Scale of Motor Development | PS-4 | Additional observations |
| | | PS-5 | Maternal behavior in testing situation |

SPEECH, LANGUAGE AND HEARING EXAMINATION AT 3 YEARS

- | | | | |
|-------|---|-------|-----------------------------------|
| PS-10 | Language reception | PS-14 | Speech mechanism |
| PS-11 | Language expression | PS-15 | Speech production |
| PS-12 | Auditory memory for digits and nonsense syllables | PS-16 | Additional observations |
| PS-13 | Hearing test | PS-17 | Final summary of test performance |

PSYCHOLOGICAL EXAMINATION AT 4 YEARS

- | | | | |
|-------|--|-------|--------------------------------------|
| PS-20 | Stanford-Binet Intelligence Scale Form L-M | PS-24 | Additional observations |
| PS-21 | Graham-Ernhart Block Sort Test | PS-25 | Four year psychological test summary |
| PS-22 | Motor test | PS-26 | SRA nonverbal form AH (mother) |
| PS-23 | Behavior profile | | |

PSYCHOLOGICAL EXAMINATION AT 7 YEARS

PS-30	Bender Gestalt Test (with Koppitz scoring)	PS-34	Tactile Finger Recognition Test
PS-31	Wechsler Intelligence Scale for Children	PS-35	Wide Range Achievement Test
PS-32	Auditory-Vocal Association Test (Illinois Test of Psycholinguistic Abilities)	PS-36	Behavior profile
PS-33	Goodenough-Harris Draw-A-Person Test	PS-37	Additional observations
		PS-38	Test summary

FINAL SPEECH, LANGUAGE AND HEARING EXAMINATION

PS-40	Hearing examination	PS-43	Speech mechanism
PS-41	Language comprehension	PS-44	Speech production
PS-42	Language expression	PS-45	Additional observations

SOCIOECONOMIC AND FAMILY HISTORIES

SE-1	Socioeconomic interview at registration	FHH-1	Family health history, Part I (gravida and father of baby)
GEN-5	Family history interview - outcomes from gravida's prior pregnancies and medical conditions in outcomes (at registration)	FHH-2	Family health history, Part II (family of gravida and father of baby)
GEN-6	Family history interview - family composition (at registration)	FHH-3	Family health history, Part III (household income)
GEN-7	Family history interview - health of gravida and her family (at registration)	FHH-4	Family health history - detailed health information (family of gravida and father of baby)
GEN-8	Family history interview - health of baby's father and his family (at registration)	FHH-9	Family health history review (mother at seven years)

TABLE 2.2. Final Forms Used In The NCPP
By Time of Administration

PRENATAL

<u>Registration and First Prenatal Visit</u>	<u>Subsequent Prenatal Visits</u>
Obstetric administrative record (AR-1)	Repeat prenatal history (OB-8)
Reproductive and gynecological history and history since last menstrual period (OB-2,3,4)	Prenatal observations (OB-44)
Recent and past medical history including infectious disease and system review (OB-5,6,7,42)	Laboratory record (OB-45)
Physical examination (OB-43)	Physician's clinic record (OB-46)
Socioeconomic interview (SE-1)	Blood samples for serological studies (VIR-1,3)
Family history interview including outcomes of prior pregnancies, family composition and health history of parents and their relatives (GEN-5,6,7,8)	
	Summary of Antepartum Hospitalization (OB-47)

LABOR AND DELIVERY

Repeat prenatal history and admission history (OB-8)	Delivery room events (OB-33)
Admission examination (OB-51,52)	Delivery report (OB-55)
Laboratory record (OB-45)	Obstetric summary (OB-56)
Labor room record (OB-32)	Anesthetic agents (OB-57)
	Summary of puerperium (OB-58)
	Placental Examinations (Gross and Microscopic) (Path-1,2)
	Obstetric Diagnostic Summary (OB-60)

NEWBORN

Delivery room observation (PED-1)	Results of tests and procedures (PED-5)
Neonatal examination (PED-2)	Neonatal neurological examination (PED-6)
Nursery history (PED-3)	
	Newborn Diagnostic Summary (PED-8)

FOUR MONTHS

Pediatric examination (PED-10)	Interval medical history (PED-20,29)
--------------------------------	--------------------------------------

EIGHT MONTHS

Bayley Scales of Mental and Motor Development (PS-1,2)	Physical measurements (PED-14)
Infant behavior profile, additional observations, and maternal behavior ratings (PS-3,4,5)	Interval medical history (PED-20,29)

TABLE 2.2. Final Forms Used In The NCPP
By Time of Administration (Cont.)

EIGHT YEARS

Speech, language and hearing examination with tests of language comprehension and expression, auditory discrimination, speech mechanism and production, and additional observations (PS-40,41,42,43,44,45)	Physical measurements (PED-14)
	Interval medical history (PED-20,29)

GENERAL FORMS

Administrative reports for record inventory, patient follow-up
and sample maintenance (AR-2,3,4,5,8; CP-5,9)

Report of fetal, infant, or child death (PED-4)

Autopsy report (PATH-3)

TABLE 2.3. Summary of Form Replacement

Form Number* (At End of Study)	Title of Form	Old Forms Replaced	Date of Replacement
OB-40	Prenatal Record	Page 1 of OB-40 was used to replace page 1 of OB-9 (optional, for hospital use only)	April, 1962
OB-42	Past Medical History	Replaced page 2 of OB-9	April, 1962
OB-43	Initial Prenatal Exam	Replaced pages 3 and 4 of OB-9	April, 1962
OB-44	Prenatal Observations	Replaced clinical findings from OB-10	April, 1962
OB-45	Laboratory Record	Replaced laboratory findings from OB-10	April, 1962
OB-47	Summary of Antepartum Hospitalization	Replaced OB-12	April, 1962
OB-50	Admission History	Replaced OB-30	April, 1962
OB-51	Admission Exam, Pt. 1	Replaced general findings on OB-31	April, 1962
OB-52	Admission Exam, Pt. 2	Replaced portion of OB-31 where results obstetric exams were reported.	April, 1962
OB-55	Delivery Report	Replaced pages 2,3 & 4 of OB-34	April, 1962
OB-56	Obstetric Summary	Replaced page 1 of OB-34	April, 1962
OB-57	Anesthetic Agents	Replaced OB-35	April, 1962
SE-1	Socioeconomic Interview	Card records in master file include data from FHH-1 and FHH-3. Replaced FHH-1,3.	April, 1963
GEN-5,6,7,8	Family Histories	Cards in master file include data from FHH-2 and FHH-4. Replaced FHH-2,4.	May, 1961
PED-8	Newborn Diagnostic Summary	Replaced PED-7	January, 1963

*Copies of actual forms appear in Volume II of the Guide, except OB-40 which was retained by the institutions if used

PHASES IN DATA COLLECTION AND ASSOCIATED FORMS

Prenatal Phase

When a woman was selected in accordance with the approved sampling technique by a collaborating institution, an administrative record (AR-1) with identification information was completed and submitted to the Perinatal Research Branch. At the same time, she was interviewed to complete: her past obstetrical history (OB-2); her menstrual and gynecological history (OB-4); her past medical history, including a history of X-ray exposure, drug intake, and hospitalizations (OB-5, OB-6, OB-7); and history since her last menstrual period (OB-3). This history included information regarding minor and major illnesses, X-ray exposure, visits to physicians or episodes of hospitalization that occurred during the interval between the last menstrual period and registration in the NCPP.

During one of the early visits to the prenatal clinic, socioeconomic and genetic histories of the woman and her family were obtained. Socioeconomic information (SE-1) covered such items as race, religion, place of birth, education, occupation, family income, housing density, marital status, geographic mobility, etc. Genetic and family health information included information on outcome of prior pregnancies (GEN-5), full and half siblings, twinning, consanguinity (GEN-6) and any history of blood group incompatibility, congenital malformations, motor and sensory disorders or mental retardation in the gravida, her previous offspring, her immediate family (GEN-7) or in the father of the baby and his family (GEN-8).

The obstetrician completed a detailed physical examination (OB-9 or OB-43) of the gravida. As routine laboratory tests, he ordered and recorded (OB-10 or OB-45) a hemoglobin and/or a microhematocrit, a complete urinalysis, serologic tests for syphilis, and blood typing for ABO and Rh type. A Coombs' test was performed on all Rh negative gravidas.

The patient returned to the prenatal clinic for reexamination at intervals of four weeks during the first seven months of pregnancy. This schedule changed to visits every two weeks during the eighth month, and to weekly visits following the eighth month of pregnancy. At each visit, interviews were conducted to elicit information on illnesses occurring after her most recent visit to the prenatal clinic (OB-8). In addition, the obstetrician obtained information about the presence or absence of certain intercurrent events such as bleeding, edema or trauma, observed the presence or absence of fetal heart activity, and noted presentations (OB-10). A blood pressure reading was obtained and repeat urine tests performed to identify albuminuria and glycosuria.

If the patient reported she had been examined by non-study physicians during the prenatal period, or in a special clinic, or was hospitalized for an intercurrent illness, either in the study hospital facility or an outside hospital, verification of this information was sought. The physician responsible for her care in each instance was contacted. A report (OB-12) of each episode of hospitalization was completed.

Blood samples of 20 milliliters were collected in Vacutainers at the first visit to the prenatal clinic (VIR-1; VIR forms are not on the master file and are not included in Volume II. See Volume IV, Work Files.). Repeat samples of blood were drawn according to a specific schedule; at bi-monthly intervals through pregnancy, at delivery, and finally at six weeks post partum. After proper separation, the serum from these blood samples was frozen and shipped to the Section on Infectious Diseases of the Perinatal Research Branch.

Labor and Delivery

When the gravida was admitted to the hospital for delivery or observation, her admission history (OB-30 or OB-50) was established, and a reevaluation of her physical status obtained (OB-31 or OB-51-52). During labor and delivery, she was under surveillance by a trained observer who obtained, at specified intervals, data on blood pressure, fetal heart rate, the frequency and spacing of her contractions, and any other intercurrent events, such as bleeding or meconium staining of amniotic fluid (OB-32). The observer also recorded information on the progress of labor as narrated by the obstetrician in charge. Any other remarkable events occurring during labor and delivery were documented by the observer (OB-33). Observers were usually nurses who had been trained specifically for this role. Records were completed identifying anesthetic agents administered during labor or delivery (OB-35 or OB-57).

The obstetrician, after termination of the delivery, completed the summary form of labor and delivery (OB-34 or OB-55). The placenta was placed in a plastic bag (sealed to avoid dessication) and sent to the study pathologists for examination (Path-1-2).

Newborn Phase

When the child was born, the 1-, 2-, and 5-minute Apgar scores were obtained by a person specially trained for this purpose (usually the delivery room observer). In addition, the onset of respiration relative to time of birth was recorded, as well as information on types of resuscitation (PED-1). The information collected was designed primarily to record the time and sequence of events taking place at the time of delivery and to record the functional integrity of the infant and any potentially stressful influences present immediately after birth.

The newborn child was examined (PED-2) by a pediatrician within 24 hours after delivery. Repeat pediatric (PED-2) examinations were performed at 24-hour intervals and children who remained in the hospital longer than one week were examined at weekly intervals. A newborn neurological examination (PED-6) was performed at two days of age. Observations of the child in the nursery, such as body temperature, respirations, feedings, and other intercurrent events, were made and recorded by nurses (PED-3). Determination of bilirubin was done on every child at 36 hours of age and repeated at 24-hour intervals as long as the most recent value was above 10 milligrams percent. Bilirubin determinations on premature infants were done at daily intervals until five days of age and discontinued unless the most recent value was above 10 milligrams percent. Hemoglobin and/or microhematocrit

determinations were obtained on every child at 48 hours of age (PED-5). ABO and Rh blood typing was performed at the same time, followed by a Coombs' test if the Rh factor was negative. Other laboratory determinations were done as indicated. Following discharge from the newborn nursery, a diagnostic evaluation of the data collected during the newborn period was made by a physician, who completed an extensive, structured, diagnostic summary (PED-8).

Four Months Phase

Each child was scheduled for a physical examination (PED-10) at four months of age. Concurrently, an interval history (PED-20) was obtained to establish information concerning possible visits to physicians or hospitals following discharge from the newborn nursery. If this information revealed that the child was hospitalized or was seen by a physician for anything other than routine care, a copy of the physician's and/or hospital record (PED-29) was obtained.

Eight Months Phase

A psychological examination, utilizing an early research version of the Bayley Scales of Infant Development, was administered at age 8 months to assess the child's mental development (PS-1) and fine and gross motor development (PS-2). The Bayley Scales were supplemented by observations of the child, ratings of behavior characteristics (PS-3), additional information on physical and behavioral abnormalities and hearing acuity (PS-4) and interactions of the mother with the infant (PS-5). At this time, an interview took place covering intercurrent events and medical history during the interval since the last examination (PED-20). Physical growth measurements were also taken (PED-14).

Twelve Months Phase

At one year of age, a neurological examination (PED-11) was given and an interval history (PED-20) obtained. After completion of the one-year examination, a diagnostic summary (PED-12) was completed to summarize events, illnesses and conditions that occurred or were recognized during the interval beginning with the terminal date of the PED-8 summary and ending with the PED-11 examination.

Three Year Phase

At three years of age, the child was brought back for examination of speech, language and hearing status (PS-10 through PS-17). At that time, an interval history (PED-20) was obtained and physical growth measurements were taken (PED-14). Language reception (PS-10) and expression (PS-11) were evaluated, as was auditory memory for digits and nonsense syllables (PS-12). A hearing test (PS-13) was administered during the visit. In terms of speech evaluation, both speech mechanism (PS-14) and speech production (PS-15) were examined. Additional observations, such as observable physical anomalies or unusual behaviors during the test period, were recorded (PS-16). Following the testing, a final summary of the speech, language and hearing examinations was prepared (PS-17).

Four Year Phase

At four years of age, the child was seen by a psychologist and a detailed examination conducted. This was accomplished by administration of the Stanford-Binet Intelligence Scale (PS-20), in addition to assessment of fine and gross motor development (PS-22), and a test of the child's concept formation skills (PS-21). A behavioral profile (PS-23) consisting of examiner ratings of the child's behavior during the examination was included. Additional observations of the child's appearance and behavior were recorded (PS-24). A test summary (PS-25) prepared on the child included the examiner's clinical impressions. At the time of the four-year exam, an intellectual assessment of the mother or mother-surrogate was performed (PS-26). An interval history (PED-20) and physical growth measurements (PED-14) were also obtained for the child at this time. Thereafter, interval histories (PED-20) were obtained at five and six years of age, usually by home visits.

Seven Year Phase

At age seven, the child returned for a neurological examination (PED-76), visual screening (PED-75), and an interval history (PED-20). An ophthalmologic consultation (PED-74) was also carried out when required. During the same visit, or shortly thereafter, a detailed psychological examination was also performed. The tests evaluated the child's intelligence by means of the Wechsler Intelligence Scales for Children (WISC) (PS-31) and examined the child's perceptual-visual-motor skills by means of the Bender Gestalt Test (PS-30). Other psychological tests administered at age seven included the Auditory-Vocal Association Test (Illinois Test of Psycholinguistic Abilities) (PS-32), Goodenough Harris Draw-A-Person Test (PS-33), Tactile Finger Recognition Test (PS-34), Wide Range Achievement Test (PS-35), and a behavioral profile (PS-36). Additional observations were also made of the child's appearance, movements and behavior (PS-37). Researchers recorded any attendance in a special class or school and prepared a summary of the psychological examinations and clinical impressions (PS-38).

During the seventh year of the child's life, the genetic and socio-economic information collected prenatally was brought up-to-date (FHH-9). This form was designed in such a way as to permit comparison of information at two points in time.

Subsequent to these examinations, a diagnostic summary (PED-77, ADM-86, IDC-77) was again completed for each child by NINDB staff, covering all conditions and events that were recognized or had occurred since the completion of the diagnostic summary at age one year.

Eight Year Phase

At eight years of age, the child returned to the study facility for a detailed evaluation of speech, language, and hearing status (PS-40 through PS-45). A thorough hearing examination was conducted (PS-40). Language comprehension (PS-41) and expression (PS-42) were evaluated. In terms of speech, both speech mechanism (PS-43) and speech production (PS-44) were examined. Additional observations recorded at the time of the speech, language and hearing examinations (PS-45) included observable anomalies and

aberrant behavior. An interval history (PED-20) covering intercurrent medical events and final physical growth measurements were recorded at this time (PED-14).

Supplementary Information

Data obtained during scheduled examinations in the follow-up of study children were supplemented by: (1) interval medical histories (PED-20) obtained at 18 months, 24 months, five years and six years of age, usually by home visits, and (2) summaries of medical illness or hospitalization (PED-29) for illness, injury, condition or hospitalization completed from the hospital records of the study facility, and other hospitals, clinic, private physicians, etc., where study children were seen for diagnostic purposes.

PROCEDURE AND INSTRUCTION MANUALS

All of the study forms for the Collaborative Perinatal Project were accompanied by manuals that detailed how the forms were to be completed. For each form, information was provided on the purpose of the form, general instructions, specific instructions for the completion of each item on the form, procedure for examinations and observations, and working definitions of variables. A general manual containing a history of the development of the Collaborative Perinatal Project and information on the methodology, case selection, data analysis and the administrative and organizational framework of the project was also available.

Copies of all instruction manuals are included along with the individual forms in Volume II. The manuals were revised when the forms were revised; the manuals included are those that accompanied the final version of the forms.

PROCEDURES INSTITUTED TO ENSURE CONSISTENCY AND ACCURACY OF PRIMARY DATA

As mentioned above, efforts to ensure the consistency and accuracy of the data collected throughout the study were made continuously. Staff members from the collaborating institutions met in workshops. Films were produced describing the neonatal and one-year neurological examinations and were used in training neurologists and pediatricians. An interchange of visits among personnel from the Perinatal Research Branch and the collaborating institutions occurred, with the purpose of exchanging views and standardizing examination techniques and the recording of data.

Because the NCPP encompassed multiple medical centers and involved numerous personnel with various levels of training completing many types of forms, it was obvious from the beginning that strong efforts were needed to establish and maintain uniformity of data collection. Some of the approaches used to ensure the necessary consistency have been noted. The data on pregnancy, birth and on the infant's first year were collected and processed using those approaches. With the subsequent examinations, active quality control programs were instituted. The quality control programs developed for the psychology, speech, language and hearing and seven year pediatric-neurological examinations are described in Appendix C.

CHAPTER 3. DATA PROCESSING PROCEDURES

DATA PROCESSING AT THE COLLABORATING INSTITUTIONS

All records of examinations, observations, interviews, etc., were reviewed by non-medical personnel in the collaborating institutions and compared with hospital records to check legibility, consistency, completeness, and adherence to study requirements and definitions. Discrepancies were brought to the attention of the person responsible for the completion of the form for consideration and resolution. Corrections on study forms were made in such a manner that the original information could be identified on the record. Before the forms were submitted to the Perinatal Research Branch, they were edited in detail for consistency and accuracy by medical personnel. After review and editing, copies of the completed forms were sent to the Perinatal Research Branch for key punching and data entry.

DATA PROCESSING AT THE PERINATAL RESEARCH BRANCH

Data processing procedures at the Perinatal Research Branch were designed to minimize errors and identify mistakes that might have occurred at the collaborating institutions. The data processing system included comprehensive reviews and edits at every stage. Niswander and Gordon (1972) describe the organization of the system as follows:

1. When an examination was completed and reviewed at the Center, a copy of the form was sent for data processing to the Perinatal Research Branch.
2. The form was then edited by specially trained nurses for completeness and accuracy, and was then coded.
3. Cards were punched, verified, and sent to the computer facility.
4. The next stage of processing included a screening of every column in every card for invalid codes.
5. The data on the cards were checked to determine whether they fell outside a range of levels established by the medical group responsible for that particular form. For example, the record for a child with a first breath recorded in excess of ten minutes after birth, and who was reported to be liveborn, would be questioned. Similar reviews were made for many other measurements.
6. The cards earmarked for review in this procedure were returned to the appropriate evaluations unit, which then examined the original form. If a mistake was found, the card was corrected and returned for processing. If the item was correctly recorded, it was then forwarded to the physician in charge who attempted to ascertain the reason for the unusual reading. He had two options. The first was to accept the recording as legitimate and send the data back to the processing group. The second option was to

request a review by the hospital for confirmation or rejection of the observation and a substitution of the correct observation, if known. If the observation was incorrect and no substitution was possible, the item was classed as unknown.

7. After data were processed into the computer file, frequency distributions were tabulated periodically for specific items in the file so that unusual values could be rechecked. The original forms were examined to provide a review of these unusual observations."

The same general procedure was employed for pediatric and behavioral data.

ACCURACY OF THE DATA PROCESSING

The quality of the data processing effort is reflected in the results of a study of the data processing operation described by Niswander and Gordon (1972). The case numbers of 20 NCPP registrants were selected at random; photocopies of all of the study forms filled out for these mothers and their children were requested, and computer printouts were made of the data processed from these records. Cases were selected from women registered in each of the several years of the study.

A total of 40,000 separate pieces of information in this sample was examined. In all, 34 unique errors were identified, yielding an error rate of less than one-tenth of one percent, which is very low for the large volume of records processed. A similar review of 100 cases carried out at the Johns Hopkins Center compared study information and hospital records and found a similarly small error rate.

In another study, the validity of obstetric information in the NCPP records was assessed at two hospitals. This was accomplished by comparing study records with the hospital records. At these two hospitals a sample of eight percent of the records was reviewed; the records were stratified to insure a sufficient number of normal pregnancies in the NCPP sample.

A sample of cases was drawn and the centers were requested to provide the actual hospital records for the sample. Arrangements were made for the hospital records to be reviewed by a physician. Information on demographic, prenatal, delivery, postpartum and infant characteristics was obtained from 14 different NCPP forms and records. The hospital and the NCPP forms for each selected patient were reviewed and abstracted "blind" and independently by the same physician. Forty of the most important characteristics were selected for detailed analysis. The items can be characterized as five demographic, ten prenatal, thirteen delivery, four postpartum, and eight infant.

The review showed that more information was missing from the hospital records than from NCPP forms. In addition, the NCPP forms contained more detailed information than did the hospital records.

An insignificantly small number of patients had important facts missing from both study and hospital records. The only frequent omissions in both records were the results of a failure to check "not present" or "not done"

boxes for the rare conditions and procedures. In one hospital, five important items were totally absent for a small proportion of the patients. These items were: maternal pre-conception weight and height for 1 percent of the patients; birthweight for 0.5 percent of the babies; blood pressure measurements for 1.5 percent of the patients; and postpartum temperatures for 1 percent of the patients. Only one item, the staff position of the person who delivered the baby, was completely missing for a significantly large number of patients: 19 percent in one hospital and 3.5 percent in the other.

The NCPP forms were also compared with the hospital records by computing the percent of records in which the item was present in both hospital and study records, and also the percent of the hospital records with items missing that were present in the study records. When the hospital record contained an item of information, it was generally present in the study record; when the item was absent in the hospital record, it again frequently appeared in the study record. More of the important items of information were recorded in the NCPP forms than in the hospital records.

The study concluded that the NCPP forms contained extensive and detailed information not available in the other hospital records and that the study records had a high standard of completeness in the two centers where they could be evaluated.

POOLING OF DATA FROM THE COLLABORATING INSTITUTIONS

During NINDB Perinatal Statistical Ad Hoc Committee review, one of the major questions that the committee attempted to resolve was the appropriateness of pooling information from the collaborating institutions.

The committee recognized generally that the data for white and black gravidas should not be combined. While many similarities existed between white and black women with respect to the medical and obstetrical conditions and complications they experienced, their demographic characteristics were very different. In addition, there was interest in examining differences between racial groups in pregnancy outcome, and neurological and behavioral attributes of the children. The two groups differed in mortality rates, in low birthweight rates, and in the morbidity experienced by the child from birth onward to the end of the study period.

In studying the problem of pooling data for each race across collaborating institutions, the committee found that the application of standard statistical measures of variability was not a useful way to identify biologically meaningful variation. In many instances, demonstration of statistical significance, because of the relatively large sample sizes, need not correspond to substantive significance. Medical investigators could not be assumed to consider such variation unusual or suspect.

The committee found that, as would be expected, demographic characteristics of the gravidas varied considerably among collaborating institutions. This is a basic strength of the NCPP: that a group of collaborating institutions, heterogeneous with regard to the demographic characteristics of their gravidas, show, in general, the same basic relationships of prenatal characteristics to fetal outcome.

The committee also found that antepartum characteristics, with the exception of "infections during pregnancy," were quite uniform. They found that the labor and delivery characteristics were, for the most part, fairly uniform. As might be expected, the committee found that the relative frequencies of "definite" findings were much more consistent than were "suspect" findings.

DATA FILES CONSTRUCTED

The data from the first prenatal visit through pregnancy, birth and the eight years of follow-up on the children, collected using approximately 100 study forms for each mother-child pair, were transferred from the completed study forms to computer data files. This transfer occurred throughout the 16 year span of the NCPP data collection phase and continued into subsequent substantive analysis phases. The data files documented by this user's guide include all primary data that were computerized. Not all data items were computerized. For example, clinical comments and notes recorded on study forms were usually not coded or computerized as such.

There are three major groups of data files: the master file, the variable file and the work files. The names of the files reflect historical name conventions and partially reflect their characteristics. At the time the study was conducted, punched cards were the standard data entry method. Throughout this document we refer to cards as a convenience; the actual cards were discarded after their images were archived on magnetic tape.

Master File

The master file consists of computer cards that were punched directly from completed study forms (see Volume II for specific details). These computer cards are tied to specific study forms and contain data items as defined by the definition of codes accompanying each study form. The master file consists of approximately six million unblocked, 80 column card images.

Not all study forms have corresponding computer cards in the master file. In some cases, information from a study form was never computerized; in other cases, the study forms were computerized, but appear only as work files and were not entered on the master file.

The master file has the advantage of containing most of the primary data, but it is unwieldy and complicated. Data for each case are grouped together; all cases appear in the same file. As a result, the user must search extensively for specific data items involving a number of patients, since cards of the same type are not grouped together. Each case does not have the same number or even the same types of computer cards. The file is structured so that the cards appear (if present) in the same order for each case. This structure facilitates accessing all data from a particular case at the same time and allows combining of data across study forms.

The master file requires 16 computer reels at 1600 bpi or four reels at 6250 bpi. The computer tapes are encoded in EBCDIC. The data are sequenced by case number and within a case, the cards are arranged to reflect the gathering of data through forms and subsequent revisions to the forms

(Figure 3.1). Knowledge of the relationship between card numbers and study form identification numbers is helpful in understanding where information is located.

In most instances, a researcher will access the master file only once for each research project. This is due mainly to the size of the master file and its associated cost. Accessing the master file should result in the creation of a work file that can be used in a specific research project. A researcher who must access the master file should clearly define his research project's data requirements and thoroughly understand what information must be extracted from the master file. A computer program tailored to the specific data request must be written or an extensive data base management system be designed.

Master File Card Number and NINDB Case Number Rationale

Computer cards for each NCPP study form are numbered to reflect their origin and possible revisions. Card numbers are assigned to identify the type of data (subject), the presence of multiple cards in a series, NCPP study form and form revisions. The first five digits of each card on the master file are the card number. The study forms and card numbers are given in Figure 3.1.

The first fourteen columns of each master file computer card contain the master file card number and the NINDB case number. Table 3.1 identifies the function of each of these columns.

TABLE 3.1. Derivation of Master File Card Number and NINDB Case Number.

<u>Contents</u>	<u>Columns</u>
Master File Card Number	
card identifier	1
general subject matter	2
form number	3-4
revision code	5
NINDB Case Number	
collaborating institution	6-7
type of patient selection	8
gravida identification number	9-12
order of the pregnancy	13
identifies child or gravida	14

Column 1 identifies multiple cards in a series. It contains a zero for cards unique to a particular form (that is, no other cards are present), for example OB-3, or for cards where repetitive data are contained. Cards for

OB-2 are an example of this second type; no new categories of information are included on successive cards, but previous births in excess of four must be recorded on an add-on card. For card series where data entered are unique to a card and more than one card is required to complete the series, a "1" is used to designate the first card, for example OB-5. OB-57, PATH-2 and PED-14 are exceptions to these rules.

The second digit on the card reveals the general subject matter covered by data on the card. All cards containing information pertaining to obstetrics, for example, are designated by a "3" in column 2; family histories are designated by a "5"; pathology with a "2"; pediatrics, with a "4"; and psychological testing with a "1".

Columns three and four reveal the form number. In the case of forms where old and new forms having different numbers are included together, the number of the latest form appears on the master file. This rule does not apply to data abstracted from several forms by NINCDS staff (ADM forms).

Column 5 of the card contains a revision code indicating which form or combination of forms was used in arriving at data on a particular card. A typical card will have one to three revision codes, with a zero indicating the first version of a form and "1", "2", and "3" indicating later revisions. As a rule, revision codes used on cards differ from card to card; investigators should check the definition of codes provided in Volume II to determine the meaning of revision codes used.

Each woman and child studied in the project received a unique case number (NINDB case number) composed of nine digits, recorded in columns 6 through 14 of all master file cards. The case number identified the institution, the mother and the child. The first two digits represented the collaborating institution (see Table 3.2 below). The third digit indicated the type of patient selection. A "1" was used for patients selected for the central core study; a "6" indicated that a patient had been transferred from one institution to another, and a "7" indicated that the patient was part of a special study undertaken by the collaborating institution. The fourth through seventh digits were used to identify the gravida, while the eighth digit identified the order of the pregnancy of a given gravida in the project. The ninth digit was used to identify the gravida or child of the pregnancy; "9" indicated the gravida, "0" indicated the child of a single birth, "1" indicated the first child of a multiple birth, "2" indicated the second child of a multiple birth, etc.

Variable File

The variable file was created by NINCDS to facilitate research studies based on the NCPP data. It contains 1222 explicitly defined items of primary interest to NCPP and was created mainly from the master file with a few items abstracted from internal work files.

The variable file is unique in that:

1. It is the only source where each case is associated with the specific study cohorts shown in Table 1.2 and defined in Appendix B.

2. Each computer record refers to a single case. Hence, in contrast to the master file, the variable file is easy to access and use for research analysis.
3. Individual diseases and conditions from the Obstetric Diagnostic Summary (OB-60), the Newborn Diagnostic Summary (PED-8), and the Summary of the First Year of Life After the Duration Summarized on the PED-8 (PED-12) appear on the variable file as individual data items. This information appears on the master file in coded form but not as individual data items.
4. It contains some items (such as Parity, Weight Gain During Pregnancy, etc.) which were computed from specific data items on the master file.

The variable file data items and their derivation are documented in Volume III.

TABLE 3.2. Collaborating Institutions and Their Code Number
(Columns six and seven of all master file cards.)

<p>05 - <u>Boston, Massachusetts</u> Harvard Medical School Boston Lying-In Hospital Children's Hospital Medical Center</p>	<p>50 - <u>Minneapolis, Minnesota</u> University of Minnesota Hospital Health Sciences Center</p>
<p>10 - <u>Buffalo, New York</u> University of Buffalo Children's Hospital</p>	<p>55 - <u>New York, New York</u> New York Medical College Metropolitan Hospital</p>
<p>15 - <u>New Orleans, Louisiana</u> Charity Hospital Tulane University School of Medicine Medical Center Louisiana State University</p>	<p>60 - <u>Portland, Oregon</u> University of Oregon Medical School</p>
<p>31 - <u>New York, New York</u> Columbia University College of Physicians & Surgeons Columbia-Presbyterian Medical Center</p>	<p>66 - <u>Philadelphia, Pennsylvania</u> University of Pennsylvania Pennsylvania Hospital The Children's Hospital of Philadelphia</p>
<p>37 - <u>Baltimore, Maryland</u> The Johns Hopkins University School of Medicine The Johns Hopkins Hospital</p>	<p>71 - <u>Providence, Rhode Island</u> Brown University Child Study Center</p>
<p>45 - <u>Richmond, Virginia</u> Virginia Commonwealth University Medical College of Virginia</p>	<p>82 - <u>Memphis, Tennessee</u> University of Tennessee College of Medicine Gailor Hospital</p>

Work Files

During initial analyses of NCPP data, significant efforts were undertaken to create a number of specific work files. In most cases, the data items on the work files are not direct transfers of information from either the master file or variable file. The data items are derived quantities or the result of combined search of the master file and hand review of original completed study forms. Data items found on the work files are preferred over similar data items on the master file.

The work files are structured similar to the variable file; they are easy to access and use. Four types of work files are available: files that contain data that are basic to NCPP, meant to augment the master file; special subject or study files; serology files, and administrative files. Eighteen separate work files are documented in Volume IV, Selected NCPP Work Files. The names of the files are given in Table 3.3.

TABLE 3.3. NCPP Work Files Documented

<u>File</u>	<u>Work File Name</u>
W1	Socioeconomic index at registration
W2	Socioeconomic index at seven years
W3	Drugs taken during pregnancy, trade names
W4	Drugs taken during pregnancy, active compounds
W5	Congenital malformations, one and seven years
W6	Cerebral palsy diagnosis
W7	Abnormalities at seven years
W8	Speech, language and hearing at eight years
W9	Toxemia classification
W10	Rupture of membranes
W11	Survey of viral, bacterial, parasitic and fungal infections during pregnancy
W12	Serological testing - complement fixation tests
W13	Serological testing for toxoplasmosis and rubella
W14	Serological testing, cord blood
W15	Serological testing, abnormalities and controls
W16	Serum specimen inventory
W17	Family linkage
W18	Visit summary

NON-COMPUTERIZED DATA

Collaborative clinical research projects by their very nature collect information that is not amenable to computerization. The NINCDS Collaborative Perinatal Project is not an exception. Computing capabilities during the active stages of the project were far less sophisticated than at present. Data in the written comments sections of study forms could not be easily

handled for computerization. Consequently, all such amplifying information is available on the microfilm only. In some cases, complete study forms are available only on microfilm. The NINCDS microfilmed all original study forms completed during the project; these records are available for research use.

Study forms available only on microfilm include:

OB-11/46	Record of Current Pregnancy/Physician's Clinical Record
OB-12/47	Summary of Antepartum Hospitalization
OB-30/50	Admitting Record/Admission History
OB-31	Admitting Examination by Obstetrician (See OB-51/52)
OB-32	Labor Room Record (See ADM-49/50/51)
PED-20	Interval Medical History
PED-29	Summary of Medical Records of Illness or Hospitalization
PED-74	Ophthalmology Consultants Report Vision Screening Study

A researcher considering use of the microfilm records should consult Chapter 5 to determine the applicable requirements and procedures.

CHAPTER 4. NCPP DATA: HIERARCHICAL CLASSIFICATION AND PERSON, TIME AND SUBJECT CATEGORIZATION

A researcher can acquire a general overview of the type of information or data items collected throughout the NCPP from reading the description of the study forms given in Chapter 2. Some researchers will find that description sufficient to determine if the NCPP data applies to their research question. Other researchers require more specific information. The purpose of this chapter is to provide that information. It is not feasible to list here all data items that are available; over 7,000 of them exist on the three types of computer files. To help a researcher obtain an idea of the contents of the NCPP data base, a hierarchical classification and a person, time, and subject categorization were constructed. The person, time, subject categorization is used extensively in Volumes II-VII and is considered an integral part of each data item name (see Chapter 5).

HIERARCHICAL CLASSIFICATION

The data from the NCPP can be placed into 20 primary data classes within the following seven subject areas: obstetrics; placenta; pediatrics; psychology; speech, language and hearing; socioeconomics; and family history (Table 4.1). Data were collected on over 4000 data items in these areas. In addition, other data items were derived from items recorded on the data collection forms. These included such simple calculations as gestational age, which was based on the date of the last menstrual period and the date of delivery, to complex scores of socioeconomic status that were based on several variables. Because of the diversity and complexity of the data, a researcher may have difficulty discovering specific information that is available.

A hierarchical classification has been developed to aid visualization of the overall scope of the collected data. Within each primary data class, the types of study information are classified into lower levels. The hierarchy is designed to allow the researcher to determine what types of information related to a specific research area are included in the data base.

The hierarchy is not intended to direct a researcher to a specific data item. The approach taken in developing the hierarchy is to organize information simultaneously according to major subject area, stage of pregnancy (for obstetric variables) or age of the child (for other variables), type of examination, and topical areas under which data were collected. This allows a researcher to determine relatively easily what general information is available concerning a particular topic.

In developing the general data classification as represented in this hierarchy, it was thought that a system that combined a time-frame dimension along with a biological or behavioral classification was potentially more useful than one that attempted to classify only on the basis of subject. For example, the inclusion of all prenatal laboratory tests under one rubric was preferable to making laboratory tests the major heading and then including prenatal, postnatal and pediatric tests under individual headings. The framework adopted also follows the design of the data collection forms and the organization of the data items in the master file.

Using a two level classification of data collected in the NCPP, a researcher can determine, at a general level, if data on certain subjects are available (Table 4.2). For example, the tests included in the four-year psychological examination are identified.

A more detailed examination of the organization of the study data items is given in Table 4.3. Here, a tertiary level of organization has been added and the data classes are referenced to the study forms on which the data were collected. With this information, a researcher can use Volume II of this guide to identify the specific questions that were asked, the way the data were recorded, the coding procedures used for the study data items, availability of data records, and the location of the variables on data tapes.

The tertiary data classes should enable a potential user of the NCPP data base to determine if the type of information required for a proposed research project is available. At a minimum it will aid a researcher in deciding if further investigation for specific data items is worthwhile. The reference given to NCPP study forms is one mechanism to help a researcher continue his search. It is not the only, or in certain circumstances, even the best way to proceed (see Chapter 6). The study form identification does not exclude the possibility that either the variable file or one of the work files may contain the relevant data items as well. A researcher is advised to consult the documentation for these files (Volumes III and IV) in addition to the study forms found in Volume II.

TABLE 4.1. Hierarchical Classification:
Primary Data Classes for the NCPP

Subject Area	Class Number	Primary Data Class
Obstetrics	1	Registration and personal information
	2	History
	3	Prenatal examinations and miscellaneous prenatal records
	4	Admission for delivery
	5	Labor
	6	Delivery and postpartum
	7	Diagnostic summary
Placenta	8	Placental examination
Pediatrics	9	Newborn
	10	Infant
	11	One to seven years
Psychology	12	Psychological examination at eight months of age
	13	Psychological examination at four years of age
	14	Psychological examination at seven years of age
Speech, Language and Hearing	15	Speech, language and hearing examination at three years of age
	16	Speech, language and hearing examination at eight years of age
Socioeconomics	17	Socioeconomics
Family History	18	Family history at time of study pregnancy
	19	Family history reviewed at the time the study child is seven years of age
Family Linkage	20	Linkage of related individuals included in the study

TABLE 4.2. Hierarchical Classification:
Secondary Data Classes for the NCPP

1. Obstetrics - registration and personal information
 - A. Identifying information
 - B. Study registration
 - C. Personal information
2. Obstetrics - history
 - A. Gynecological history
 - B. History of prior pregnancies
 - C. History since last menstrual period
 - D. Recent medical history
 - E. Past medical history
 - F. Repeat prenatal history
3. Obstetrics - prenatal examinations and miscellaneous prenatal records
 - A. Initial prenatal examination
 - B. Return prenatal examinations
 - C. Laboratory examinations
 - D. Physician's clinic record
 - E. Drugs in pregnancy
 - F. Special rubella study
 - G. Summary of antepartum hospitalizations
 - H. Visit summary
4. Obstetrics - admission for delivery
 - A. Admission history
 - B. Admission examination
5. Obstetrics - labor
 - A. Labor room record
 - B. Summary of labor
6. Obstetrics - delivery and postpartum
 - A. Delivery room events
 - B. Delivery report
 - C. Anesthetic agents
 - D. Summary of the puerperium
7. Obstetrics - diagnostic summary
 - A. Diseases/conditions - before pregnancy; during pregnancy; post partum
 - B. History of hypertension
 - C. Toxemia screen
 - D. Toxemia classification
 - E. Infections during pregnancy
8. Placental examination
 - A. Gross
 - B. Microscopic
9. Pediatrics - newborn
 - A. Delivery room observations
 - B. Neonatal examination
 - C. Nursery history (newborn period summary)
 - E. Report of fetal or infant death
 - F. Neonatal neurological examination
 - G. Newborn diagnostic summary
 - H. Summary of the hospital course of the neonate
10. Pediatrics - infant
 - A. Four-month pediatric examination
 - B. Blood samples for viral serological study
 - C. One-year neurological examination
 - D. Summary of the first year of life after the newborn period
 - E. Physical growth measurements

11. Pediatrics - one to seven years
 - A. Seven-year pediatric and neurologic exam
 - B. Seven-year visual screening and examination
 - C. Seven-year diagnostic summary
12. Psychological examination at eight months of age
 - A. Bayley Scales of Mental Development
 - B. Bayley Scales Motor Development
 - C. Infant behavior profile
 - D. Additional observations on physical and behavioral abnormalities
 - E. Maternal behavior in testing situation
13. Psychological examination at four years of age
 - A. Stanford-Binet Intelligence Scale
 - B. Graham-Ernhart Block Sort Test
 - C. Motor test
 - D. Behavior profile
 - E. Additional observations on physical and behavioral abnormalities
 - F. Psychological test summary: clinical impressions
 - G. Intellectual assessment of study mother or mother surrogate
14. Psychological examination at seven years of age
 - A. Bender Gestalt Test with Koppitz scoring
 - B. Wechsler Intelligence Scale for Children (WISC)
 - C. Auditory-Vocal Association Test
 - D. Goodenough-Harris Draw-a-Person Test
 - E. Tactile Finger Recognition Test
 - F. Wide Range Achievement Test
 - G. Behavioral profile
 - H. Additional observations on physical and behavioral abnormalities
 - I. Psychological test summary: clinical impressions
15. Speech, language and hearing examination at three years of age
 - A. Language reception
 - B. Language expression
 - C. Hearing test
 - D. Speech mechanism
 - E. Speech production
 - F. Auditory memory - digits and nonsense syllables
 - G. Additional observations
 - H. Final summary of speech, language and hearing test performance
16. Speech, language and hearing examination at eight years of age
 - A. Hearing
 - B. Language comprehension
 - C. Language expression
 - D. Speech mechanism
 - E. Speech production
 - F. Additional observations
17. Socioeconomics
 - A. Socioeconomic data at the time of the study pregnancy
 - B. Socioeconomic data reviewed at the time the child was seven years of age
18. Family history at time of study pregnancy
 - A. Outcomes from gravida's prior pregnancies
 - B. Family composition
 - C. Health of gravida and her family
 - D. Health of father of baby and his family
19. Family history reviewed at the time the study child was seven years of age
 - A. Outcome of prior pregnancies
 - B. Pregnancies since study pregnancy
 - C. Outcome of all pregnancies
 - D. Conditions in study child, parents, or siblings since birth of study child
20. Linkage of related individuals included in the study
 - A. Family linkage - mother's relationships
 - B. Family linkage - children's relationships
 - C. Family linkage - relationship groups

**TABLE 4.3. Hierarchical Classification:
Tertiary Data Classes for the NCPP.**

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
1. Obstetrics - Registration and personal information	A. Identifying information (Form AR-1)	<ol style="list-style-type: none"> 1. Study number and hospital number 2. Name, address, and telephone number
	B. Study registration (Form AR-1)	<ol style="list-style-type: none"> 1. Date registered 2. Date form initiated 3. Sampling frame patient
	C. Personal information (Form AR-1)	<ol style="list-style-type: none"> 1. First day of LMP 2. Date of birth 3. Marital status 4. Race 5. Patient status
2. Obstetrics - History (See also 7. Obstetrics - Diagnostic summary; 17. Socioeconomics; and 18. Family history at time of study pregnancy)	A. Gynecological history (Forms OB-4,9)	<ol style="list-style-type: none"> 1. Menstrual history 2. Fertility and contraceptive history
	B. History of prior pregnancies (Forms OB-2,9)	<ol style="list-style-type: none"> 1. Record of pregnancies in chronological order 2. Characteristics of prior pregnancies and their outcome
	C. History since last menstrual period (Form OB-3)	<ol style="list-style-type: none"> 1. History of symptoms, conditions, and exposures 2. Intercourse frequency 3. Smoking history
	D. Recent medical history (Forms OB-5,15)	<ol style="list-style-type: none"> 1. Illness or disability requiring confinement - prior 12 months 2. Non-confining illness or disability - prior 12 months 3. Medications or injections - prior 12 months
	E. Past medical history (Forms OB-6,7,9,42)	<ol style="list-style-type: none"> 1. Hospitalizations 2. Radiologic exams or treatments - prior 12 months 3. Other radiologic exams or treatments 4. Examinations and treatments of extremities 5. All other examinations and treatments 6. Blood and transfusions 7. Blood tests taken 8. Series of injections 9. System review 10. Surgery 11. Childhood diseases 12. Other infectious diseases 13. Parasitic diseases
	F. Repeat prenatal history (since last visit) (Form OB-8)	<ol style="list-style-type: none"> 1. Symptoms, conditions and exposures 2. Intercourse frequency 3. Smoking - cigarettes per day 4. Medications

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

3. Obstetrics - Prenatal examinations and miscellaneous prenatal records
(See also 7. Obstetrics - Diagnostic summary)

A. Initial prenatal examination
(Forms OB-9,43)

1. Weight, height, vital signs
2. General examination
3. Obstetric examination
4. X-ray pelvimetry
5. Clinical pelvic mensuration
6. Diagnostic impressions

B. Return prenatal examinations
(Forms OB-10,44)

1. Gestational age
2. Weight, blood pressure, urinalysis
3. History of symptoms, complications and fetal activity
4. Obstetric examination

C. Laboratory examinations
(Forms OB-10,45; VIR-1)

1. Virology
2. Blood type and Rh
3. Antibody tests
4. Serology
5. Blood chemistry and hematology
6. Urinalysis
7. X-ray pelvimetry and diagnostic X-ray
8. Cultures
9. Glucose tolerance tests
10. Pap smear
11. Other laboratory studies

D. Physicians clinic record
(Form OB-46)

1. Medications
2. Diagnoses and impressions
3. Signs and symptoms
4. Treatments and procedures

E. Drugs in pregnancy
(Form OB-15)

1. Date of LMP
2. Drugs taken by lunar month of pregnancy

F. Special rubella study
(Form VIR-3)

1. Exposure to rubella during study pregnancy
2. Administration of gamma globulin during study pregnancy

G. Summary of antepartum hospitalizations
(Form OB-47)

1. Place hospitalized
2. Admission impression
3. Condition of fetus at discharge
4. Condition of mother at discharge
5. Surgical procedures
6. Discharge diagnoses
7. Anesthesia given
8. Radiation exposure
9. Drug therapy
10. Laboratory work

H. Visit summary

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
4. Obstetrics - Admission for delivery	A. Admission history (Form OB-50)	<ol style="list-style-type: none"> 1. Prior pregnancies 2. Pelvic summation 3. History of labor 4. History of rupture of membranes 5. History of vaginal bleeding 6. Reason for hospital admission
	B. Admission examination (Forms OB-51,52)	<ol style="list-style-type: none"> 1. Weight and vital signs 2. General examination 3. Abdomino-pelvic examination 4. Diagnostic impressions
5. Obstetrics - Labor	A. Labor room record (Forms OB-32; ADM-49,50,51)	<ol style="list-style-type: none"> 1. Maternal vital signs 2. Fetal heart rate 3. Membranes 4. Bleeding 5. Meconium 6. Pelvic examination 7. Medications
	B. Summary of labor (Forms OB-34; OB-55)	<ol style="list-style-type: none"> 1. Onset and duration 2. Position and station 3. Rupture of membranes 4. Induction and use of uterine stimulants 5. Arrested progress of labor 6. Complications and other procedures
6. Obstetrics - Delivery and post partum	A. Delivery room events (Form OB-33)	<ol style="list-style-type: none"> 1. Timing of delivery events 2. Vital signs 3. Bleeding 4. Meconium
	B. Delivery report (Forms OB-34,55)	<ol style="list-style-type: none"> 1. Type of delivery 2. Vertex delivery procedure 3. Vertex delivery with forceps or vacuum extractor 4. Breech delivery procedure 5. Breech delivery with forceps or internal version 6. Indications for forceps, vacuum extraction or version 7. Cesarean section and other surgical procedures 8. Indications for cesarean section 9. Duration of pregnancy and birthweight 10. Umbilical cord 11. Placenta 12. Complications and other procedures 13. Fetal condition
	C. Anesthetic agents (Forms OB-35,57)	<ol style="list-style-type: none"> 1. Who administered agent and who provided information 2. Gaseous agents used 3. Intravenous agents used 4. Deepest anesthesia prior to clamping cord 5. Conduction agents used 6. Response of patient 7. Other medication
	D. Summary of the puerperium (Form OB-58)	<ol style="list-style-type: none"> 1. Post partum blood pressure - highest and lowest diastolic 2. Temperature - highest 3. Postpartum transfusions 4. Summary postpartum data and diagnoses

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

7. **Obstetrics - Diagnostic summary**
(See also 2. **Obstetrics - History**; and 3. **Obstetrics - Prenatal examinations and miscellaneous prenatal records**)

A. **Diseases/conditions - before pregnancy; during pregnancy; post partum**
(Form OB-60)

B. **History of hypertension**
(Form OB-60)

C. **Toxemia screen**
(Form OB-60)

D. **Infectious diseases during pregnancy**
(Form OB-60)

1. Cardiovascular
2. Pulmonary
3. Hematologic
4. Metabolic/endocrine
5. Venereal
6. Urinary tract
7. Gynecological
8. Neurologic/psychiatric
9. Gastrointestinal
10. Integument/appendages
11. Complications of pregnancy
12. Complications of puerperium
13. Other diseases or conditions
14. Special studies

1. Blood pressure
2. Proteinuria
3. Edema
4. Other conditions related to toxemia

1. Viral
2. Bacterial
3. Parasitic
4. Fungal
5. Etiology unknown
6. Vaccination - live attenuated

8. **Placental examination**

A. **Gross**
(Form PATH-1)

B. **Microscopic**
(Form PATH-2)

1. Size and shape
2. Umbilical cord
3. Membranes and fetal surface
4. Maternal surface
5. Cut surface
6. Multiple births
7. Abnormalities

1. Cord
2. Membranes
3. Decidua
4. Terminal villi
5. Intervillous space
6. Multiple births
7. Other abnormalities

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

9. Pediatrics - Newborn

**A. Delivery room observations
(Form PED-1)**

1. Date and time of birth
2. Race and sex
3. Birth weight
4. Timing of cord clamping, first breath and first cry
5. Suction and resuscitation procedures
6. Apgar scores
7. Physical examination (delivery room)

**B. Neonatal examination
(Form PED-2)**

1. Age at time of exam
2. Measurements
3. Respiration
4. Physical examination - system review with comments
5. Moro response
6. Motor activity
7. Tone
8. Weight
9. Dysmaturity
10. Clinical impression

**C. Nursery History (newborn
period summary)
(Form PED-3)**

1. Special conditions
2. Weight
3. Temperature
4. Feeding method
5. Activity
6. Cry
7. Abnormalities and clinical signs
8. Medications and procedures

**D. Report of fetal or infant death
(Forms PED-4; PATH-3)**

1. Sex
2. Date of delivery or date and time of death
3. Place of delivery or place of death
4. Weight and crown rump length (fetal death)
5. Birth injuries (infant death)
6. Cause of death
7. Malformations present
8. Autopsy findings

**E. Neonatal tests and procedures
(Form PED-5)**

1. Cord blood studies
2. Serum bilirubin
3. Hemoglobin
4. Hematocrit

**F. Neonatal neurological examination
(Form PED-6)**

1. Time of examination and last feeding
2. Age of child
3. Eyes
4. Movement and motor activity
5. Cry
6. Grasp
7. Jerk and ankle clonus
8. Suck
9. Response to stimulus or position
10. Reflexes
11. Tone
12. Translumination
13. Tonic neck reflex
14. Impression

**G. Newborn diagnostic summary
(Form PED-8)**

1. Summary data on abnormalities, malformations, conditions, infections and procedures by organ system
2. Specific diagnoses: suspect and definite
3. Procedures

**H. Summary of the hospital course
of the neonate
(Form PED-7)**

1. Date of birth and discharge
2. Clinical data and description of events
3. Clinical impressions

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
10. Pediatrics - Infant	A. Four-month pediatric examination (Form PED-10)	<ol style="list-style-type: none"> 1. Age 2. Weight, length, and head and chest circumference 3. Vital signs 4. Physical examination - system review with comments 5. Neurological evaluation - examination with comments 6. Cry and vocalizations 7. Maternal - child relationship evaluation 8. Impression/diagnosis
	B. Blood samples for viral serological study (Form VIR-1)	<ol style="list-style-type: none"> 1. Records, specimens obtained from abnormal infants and controls at four months of age
	C. One-year neurological examination (Form PED-11)	<ol style="list-style-type: none"> 1. Age 2. Weight, length and head circumference 3. Physical examination - system review with comments 4. Responsiveness 5. Phonation 6. Locomotor and postural development 7. Neurological evaluation - examination with comments 8. Impression
	D. Summary of the first year of life after the newborn period (Form PED-12)	<ol style="list-style-type: none"> 1. Neurologic abnormality - suspect and definite 2. Related central nervous system and skeletal conditions - other - suspect and definite 3. Abnormalities, malformations, conditions, and infections by organ system - suspect and definite 4. Procedures 5. Social and environmental conditions 6. Summary data on abnormalities, conditions and procedures
	E. Physical growth measurements (Form PED-14)	<ol style="list-style-type: none"> 1. Weight 2. Length 3. Head circumference
	F. Interval medical history (Forms PED-20,29)	<ol style="list-style-type: none"> 1. Age at history and date 2. Informant 3. Health care 4. Medical problems not treated by a physician 5. Hospitalization information 6. Summary of informant's account of medical care 7. Summary of medical records

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

11. Pediatrics - One to seven years

A. Seven-year pediatric and neurological examination (Form PED-76)

1. Age
2. Physical measurements
3. Blood pressure
4. Physical examination - system review
5. Neurological examination
6. Mental status
7. Intellectual status
8. Other signs, reflexes, tests, etc.
9. Neurological abnormalities
10. Abnormality on visual screening
11. Non-neurological abnormalities

B. Seven-year visual screening and examination (Forms PED-74,75)

1. Age
2. Wears glasses? (If yes - test repeated with glasses)
3. Visual acuity - each eye
4. Muscle balance
5. Color tests
6. External examination
7. Refractive error
8. Ophthalmoscopic examination
9. Diagnosis

C. Seven-year diagnostic summary (Forms ADM-86; IDC-77)

1. Forms reviewed
2. Specific conditions, diagnoses and sources

12. Psychological examination at eight months of age

A. Bayley Scales of Mental development (Form PS-1)

1. Age, sex and race
2. Scoring and diagnosis
3. Age placement on Bayley Scales (0 to 15 months)

B. Bayley Scales of Motor development (Form PS-2)

1. Scoring and diagnosis
2. Age placement on Bayley Scales (0 to 12 months)

C. Infant behavior profile (Form PS-3)

1. Orientation to objects
2. Orientation to persons
3. Activity level
4. Physical development - clinical impression
5. Mental development - clinical impression
6. Fine motor development - clinical impression
7. Gross motor development - clinical impression
8. Social/emotional development - clinical impression
9. Adequacy of examination

D. Additional observations on physical and behavioral abnormalities (Form PS-4)

1. Face, mouth, hearing, eyes
2. Comparative function of arms, hands and grip
3. Unusual muscular movements or postural adjustments
4. Deviant or stereotyped behavior
5. Specified obvious defects or anomalies

E. Maternal behavior in testing situation (Form PS-5)

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
13. Psychological examination at four years of age	A. Stanford-Binet Intelligence Scale (Form PS-20)	<ol style="list-style-type: none"> 1. Chronological age 2. Mental age 3. Intelligence Quotient (IQ) 4. Test performance on specific items 5. Adequacy of examination
	B. Graham-Ernhart Block Sort Test (Form PS-21)	<ol style="list-style-type: none"> 1. Scores by level and trial 2. Summary scores
	C. Motor test (Form PS-22)	<ol style="list-style-type: none"> 1. Gross motor 2. Fine motor 3. Dominance 4. Overall summary
	D. Behavior profile (Form PS-23)	<ol style="list-style-type: none"> 1. Orientation to testing situation 2. Orientation to examiner 3. Orientation to test materials 4. Activity 5. Communication 6. Examiner comments
	E. Additional observations on physical and behavioral abnormalities (Form PS-24)	<ol style="list-style-type: none"> 1. Face, mouth, eyes and ears 2. Unusual muscular movements or postural adjustments 3. Deviant or stereotyped behavior 4. Specified obvious defects or anomalies 5. Enrollment in nursery school 6. Examiner comments
	F. Psychological test summary: Clinical impressions (Form PS-25)	<ol style="list-style-type: none"> 1. Intelligence 2. Fine motor development 3. Gross motor development 4. Concept formation 5. Behavioral 6. Adequacy of examination 7. Overall impression 8. Examiner comments
	G. Intellectual assessment of study mother or mother surrogate - SRA non-verbal form (Form PS-26)	

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
	A. Bender Gestalt Test with Koppitz scoring (Form PS-30)	<ol style="list-style-type: none"> 1. Performance on specific figures 2. Total score and time 3. Adequacy of examination
	B. Wechsler Intelligence Scale for children (WISC) (Form PS-31)	<ol style="list-style-type: none"> 1. Verbal tests and scale scores 2. Performance tests and scale scores 3. Full scale IQ 4. Adequacy of examination
	C. Auditory-Vocal Association Test (Form PS-32)	<ol style="list-style-type: none"> 1. Scoring 2. Adequacy of examination
	D. Goodenough-Harris Draw-A-Person Test (Form PS-33)	<ol style="list-style-type: none"> 1. Scoring 2. Percentile rank 3. Adequacy of examination
	E. Tactile Finger Recognition Test (Form PS-34)	<ol style="list-style-type: none"> 1. Right hand 2. Left hand 3. Adequacy of examination
	F. Wide Range Achievement Test (Form PS-35)	<ol style="list-style-type: none"> 1. Personal data 2. Spelling test 3. Reading test 4. Arithmetic test
14. Psychological examination at seven years of age		<ol style="list-style-type: none"> 1. Separation from mother 2. Fearfulness 3. Rapport with examiner 4. Self-confidence 5. Emotional reactivity 6. Degree of cooperation 7. Level of frustration tolerance 8. Degree of dependency 9. Duration of attention span 10. Goal orientation 11. Level of activity 12. Nature of activity 13. Nature of communication 14. Assertiveness 15. Hostility
	G. Behavior Profile (Form PS-36)	
	H. Additional observations on physical and behavioral abnormalities (Form PS-37)	<ol style="list-style-type: none"> 1. Face, mouth, eyes and ears 2. Unusual muscular movements or postural adjustments 3. Deviant or stereotyped behavior 4. Specified obvious defects or anomalies 5. Additional observations 6. Enrollment in special class or school
	I. Psychological test summary: Clinical impressions (Form PS-38)	<ol style="list-style-type: none"> 1. Intelligence 2. Bender Visual Motor Production 3. Educational Achievement (WRAT) 4. Goodenough-Harris Drawing Test 5. Abstract Language Thinking (ITPA Aud Voc) 6. Tactile Finger Recognition Test 7. Behavioral 8. Overall impression

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
15. Speech, language and hearing examination at three years of age	A. Language reception (Form PS-10)	<ol style="list-style-type: none"> 1. Verbal expression 2. Alternate expression
	B. Language expression (Form PS-11)	<ol style="list-style-type: none"> 1. Verbal comprehension 2. Alternate comprehension (single word and pantomime)
	C. Hearing test (Form PS-13)	<ol style="list-style-type: none"> 1. Spondaic Word Test (verbal) 2. Spondaic Word Test (non-verbal) 3. Pure Tone Screening Test
	D. Speech mechanism (Form PS-14)	<ol style="list-style-type: none"> 1. Examination of the lips 2. Examination of the tongue 3. Examination of the soft palate 4. Diadochokinesis
	E. Speech production (Form PS-15)	<ol style="list-style-type: none"> 1. Voice 2. Articulation 3. Intelligibility of connected speech 4. Fluency of speech production
	F. Auditory memory - digits and nonsense syllables (Form PS-12)	<ol style="list-style-type: none"> 1. Recall of digits 2. Recall of nonsense syllables
	G. Additional observations (Form PS-16)	<ol style="list-style-type: none"> 1. State of child's health on day of examination 2. Observable physical anomalies 3. Unusual behavior observed during test period
	H. Final summary of speech, language and hearing test performance (Form PS-17)	<ol style="list-style-type: none"> 1. Language reception 2. Language expression 3. Hearing 4. Speech mechanism 5. Speech production 6. Global scoring 7. Auditory memory 8. Adequacy of examination 9. Referral

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
16. Speech, language and hearing examination at eight years of age	A. Hearing (Form PS-40)	<ol style="list-style-type: none"> 1. Pure tone audiometry - air conduction 2. Abnormal auditory adaptation 3. Pure tone audiometry - bone conduction 4. Discrimination test 5. Auditory memory 6. Scoring 7. Adequacy of examination
	B. Language comprehension (Form PS-41)	<ol style="list-style-type: none"> 1. Auditory verbal comprehension 2. Reading 3. Morphology - knowledge of linguistic form 4. Scoring 5. Adequacy of examination
	C. Language expression (Form PS-42)	<ol style="list-style-type: none"> 1. Connected discourse 2. Writing from dictation 3. Summary evaluation 4. Scoring 5. Adequacy of examination
	D. Speech mechanism (Form PS-43)	<ol style="list-style-type: none"> 1. Examination of the lips 2. Examination of the tongue 3. Concomitant movements present while performing 4. Examination of the soft palate
	E. Speech production (Form PS-44)	<ol style="list-style-type: none"> 1. Rate and fluency of connected speech 2. Voice 3. Intelligibility of connected speech 4. Articulation 5. Scoring
	F. Additional observations (Form PS-45)	<ol style="list-style-type: none"> 1. State of child's health on day of examination 2. Observable physical anomalies 3. General behavior aberrations observed during test period - specified
17. Socioeconomics	A. Socioeconomic data at the time of study pregnancy (Forms SE-1; FHH-1,3)	<ol style="list-style-type: none"> 1. Birthplace and education of gravida 2. Language, religion and race of gravida 3. Marital history of gravida 4. Work history of gravida 5. Household arrangement 6. Age, birthplace, education, religion and race of father of baby 7. Work history of father of baby or husband 8. Family income and number of persons supported 9. Socioeconomic index
	B. Socioeconomic data reviewed at the time the study child was seven years of age (Form FHH-9)	<ol style="list-style-type: none"> 1. Birthdate, sex and race of child 2. Residence of child 3. Socioeconomic data on foster parent, adoptive parent or guardian 4. Marital history of mother 5. Household arrangement 6. Education and employment of mother 7. Employment history of husband 8. Family income and number of persons supported 9. Socioeconomic index

Primary Data Classes

Secondary Data Classes

Tertiary Data Classes

18. Family History at time of study pregnancy (See also 2. Obstetrics - History; and 7. Obstetrics - Diagnostic summary

A. Outcomes from gravidas, prior pregnancies (Form GEN-5)

B. Family composition (Form GEN-6)

C. Health of gravida and her family (Form GEN-7)

D. Health of father of baby and his family (Form GEN-8)

1. Number and outcome of prior pregnancies
2. Prior liveborn children
3. Medical care and hospitalization of siblings
4. Summary of medical conditions in outcome pregnancies
5. Rh or other blood incompatibility
6. Congenital malformations or physical defect
7. Seizures, convulsions, epilepsy
8. Motor defects
9. Sensory defects
10. Developmental retardation
11. Inability to attend regular school
12. Description of conditions

1. Family of gravida
2. Family of father of baby

1. Physical defects/congenital malformations
2. Sensory defects
3. Diabetes
4. Seizures, convulsions, epilepsy
5. Motor defects
6. Mental retardation, special schools
7. Mental illness, nervous problems or psychiatric treatment
8. Additional diseases
9. Multiple pregnancies

1. Physical defects/congenital malformations
2. Sensory defects
3. Diabetes
4. Seizures, convulsions, epilepsy
5. Motor defects
6. Mental retardation, special schools
7. Mental illness, nervous problems or psychiatric treatment
8. Radiation

Primary Data Classes	Secondary Data Classes	Tertiary Data Classes
19. Family history reviewed at the time study child was seven years of age	A. Outcome of prior pregnancies (Form FHH-9)	<ol style="list-style-type: none"> 1. Fetal death 2. Live born 3. Summary of conditions
	B. Pregnancies since study pregnancy (Form FHH-9)	<ol style="list-style-type: none"> 1. Total number 2. Miscarriages/abortions 3. Multiple pregnancies
	C. Outcome of all pregnancies (Form FHH-9)	<ol style="list-style-type: none"> 1. Fetal death 2. Live born 3. Summary of conditions
	D. Conditions in study child, parents or siblings since birth of study child (Form FHH-9)	<ol style="list-style-type: none"> 1. Rh/blood incompatibility 2. Congenital malformations/physical defects 3. Developmental retardation 4. Child unable to attend regular school 5. Seizures, convulsions, epilepsy 6. Motor defects 7. Sensory defects 8. Diabetes 9. Mental illness, nervous problems or psychiatric treatment 10. Deaths of children 11. School and achievements of study child
20. Linkage of related individuals included in the study	A. Family linkage - Mother's relationships	
	B. Family linkage - Children's relationships	
	C. Family linkage - Relationship groups	

PERSON, TIME AND SUBJECT CATEGORIZATION

The NCPP collected information from approximately 58,000 pregnancies and included over 7,000 individual data items. In the previous section, a hierarchical classification was used to describe the general types of data items collected. This section describes an alternative categorization that was devised in conjunction with the naming of individual data items. Each data item was named and given a unique data item identification (see Chapter 6). In completing this process, it became apparent that an implicit categorization was involved: data items could be described based on the person, time, or general subject area they represented (Table 4.4).

A researcher can categorize his research project variables according to which person(s), what time(s) and what subject(s) apply. Working definitions used for each of the categories are given in Table 4.5. It is emphasized that this is only one group's attempt to define usable categories. Other groups would choose slightly different ones. A researcher should view the categories as an aid to locating specific data items and not as an end in themselves. As such, more than one combination of categories should be checked before deciding that a specific type of information is not available.

Volume VII, *Categorization of Data Items by Person, Time of Collection or Measurement and General Subject Area*, enables a researcher to use the categorization to locate individual data items of potential interest. Chapter 6 gives an example of how this is accomplished.

Volume VII is divided into three parts:

- Part A: Categorization of Data Items Organized by Person
- Part B: Categorization of Data Items Organized by Time of Measurement or Observation
- Part C: Categorization of Data Items Organized by General Subject Area.

The information contained in each part is the same, differing only in the organization of the categories. In Part A, data items are ordered by person, time, subject, while data items in Part B are ordered by time, person, subject. Data items in Part C are ordered by subject, person, time.

In part A all the computerized data items are organized by the person categories shown in Table 4.4. This allows a researcher to identify readily all the data items included in the computer files that relate to mothers, children, fathers, etc.

In Part B of Volume VII, the researcher can identify all computerized data items grouped on the basis of time of occurrence, observation, or measurement. For example, all the data items that relate to the neonate are listed under the category Neonatal. Such a categorization enables the user of the guide to determine all the information on a particular topic within a chronological category.

Finally, all the computerized variables are organized into general subject area categories in Part C of Volume VII (Table 4.4), providing an alternative classification to that developed earlier in this chapter. Unlike the example of prenatal laboratory tests given in the hierarchical classification, the classification scheme employed in Volume VII Part C includes all the clinical laboratory tests under a single heading.

SUMMARY

The data classification and categorization provided by Tables 4.1 through 4.4 of this chapter allow the guide's user to develop an understanding of the types of information collected as part of the NCPP. The hierarchical classification directs a researcher directly to study forms (Volume II) to locate specific data items. The person, time and subject categorization (Volume VII) directs a researcher to specific data items under a categorical listing and then to the source of the data item: master, variable or work file. Thus, the researcher can develop an understanding of the data available on the basis of biological-behavioral categories and stage of pregnancy/development and, in turn, identify the specific data items on the study forms and the data items in the computer files that relate to his/her research interests.

TABLE 4.4. Person, Time and Subject Categories
for the NCPP Data Items

<u>Person</u>	<u>Time</u>	<u>Subject</u>
Mother	General	Administrative
Father	Preconception	Anesthesia
Placenta	Registration	Clinical Impression
Fetus	Prenatal	Clinical Laboratory
Child	Admission	Current Pregnancy, General Information
Mother Surrogate	Intrapartum	Environmental Exposures
Family	Delivery	Events
Sibship	Post Partum	Hearing
	Neonatal	Hospitalization
	Four Month	Language
	Eight Month	Linkage
	One Year	Malformation
	Three Year	Medical Diagnoses and Conditions
	Four Year	Medical History
	Seven Year	Medications
	Eight Year	Neurological Examination
		Observations
		Pathology
		Physical Examination
		Procedures
		Psychological Examination
		Reproductive History
		Serology
		Socioeconomic
		Speech
		Vision
		Work History
		X-ray
		Summary
		Gynecological History
		Special Studies
		Family/Genetic History
		SLH Examination

TABLE 4.5. Definition of Person, Time
and Subject Categories

<u>PERSON</u>	<u>DEFINITION</u>
Mother	Study registrant bearing the "study pregnancy"; biologic mother of the "study child"; gravida.
Father	Biologic father of the study child or study pregnancy; in the case of socioeconomic data, this category may indicate either the "father of baby" (not necessarily husband of the mother) or the "husband" (not necessarily related biologically to the study child).
Placenta	The organ of metabolic and gaseous interchange between the fetus and mother; also included in this category are gross and microscopic pathologic data from examination of the umbilical cord.
Fetus	Conceptus; the product of conception including the embryonic stage, i.e., from conception to the moment of birth.
Child	Product of the study pregnancy from the moment of birth onward; study child.
Mother Surrogate	Person or persons substituting for the mother of a study child, e.g., adoptive parents, foster parents or guardian.
Family	Person or persons biologically related to the mother or father of the study child.
Sibship	Child or children having one or both of the same biologic parents as the study child; siblings; half siblings; full siblings.

TABLE 4.5. Definition of Person, Time
and Subject Categories (Cont.)

<u>TIME</u>	<u>DEFINITION</u>
General	Data with no pertinent time period or data pertaining to more than one time period.
Preconception	Data pertaining to the period prior to conception of the study pregnancy.
Registration	Data collected at the time of study mother's registration in the study.
Prenatal	Data pertaining to the period from conception of the study pregnancy to delivery of the study child.
Admission	Data collected at the time of study mother's admission to the hospital for delivery of the study child.
Intrapartum	Data pertaining to the period from admission for delivery or onset of labor to delivery of the study child.
Delivery	Data pertaining to the time period during which delivery of the study child occurred.
Post Partum	Data (pertaining to the study mother) collected during the period immediately following birth of the study child.
Neonatal	Data pertaining to the study child during the period from birth to one month of age; the majority of these data were collected prior to or at the time a study child was discharged from the hospital.
Four Month	Data collected at the time of the four month examination of the study child.
Eight Month	Data collected at the time of the eight month examination of the study child.
One Year	Data collected at the time of the one year examination of the study child.
Three Year	Data collected at the time of the three year examination of the study child.
Four Year	Data collected at the time of the four year examination of the study child.
Seven Year	Data collected at the time of the seven year examination of the study child.
Eight Year	Data collected at the time of the eight year examination of the study child.

TABLE 4.5. Definition of Person, Time
and Subject Categories (Cont.)

<u>SUBJECT</u>	<u>DEFINITION</u>
Administrative	Data pertaining to the administrative aspects of the study.
Anesthesia	Data on medications and procedures used to obtain anesthesia.
Clinical Impression	Impression of abnormality or dysfunction gained by an examiner following evaluation of clinical signs and symptoms and including a subjective component.
Clinical Laboratory	Data obtained from laboratory examination of clinical specimens.
Current Pregnancy General Information	Personal data and medically relevant information pertaining to the study pregnancy for which the mother is enrolled.
Environmental Exposures	Data on exposure to occupational or other environmental entities or hazards.
Events	Data related to a specific event, occurrence or incidence.
Hearing	Data obtained from examination and testing of hearing function.
Hospitalization	Data on specific hospital admissions or the number of hospitalizations.
Language	Data obtained from examination and testing of language function.
Linkage	Data on the genetic relationships of family members to the study mother, father or child.
Malformation	Data on the conditions in which failure of normal development has resulted in abnormal physical traits existing at the time of birth.
Medical Diagnoses and Conditions	Data on specific diagnoses or conditions obtained from past medical history or examination during the study.
Medical History	Data obtained from the study participant or medical records relevant to past or current medical diagnoses or conditions.
Medications	Data on drugs or medications used.
Neurological Examination	Data obtained from observation and physical examination of the central nervous system.
Observations	Data obtained from observations not categorized elsewhere.
Pathology	Data obtained from clinical and anatomical pathological examination.
Physical Examination	Data obtained from physical examination of the study participant.
Procedures	Data relating to specific procedures performed on the study participant prior to or during the period of enrollment in the study.
Psychological Examination	Data obtained from the psychological examinations and observations.

TABLE 4.5. Definition of Person, Time
and Subject Categories. (Cont.)

<u>SUBJECT</u>	<u>DEFINITION</u>
Reproductive History	Data pertaining to the outcome of pregnancies prior to and or during the period of enrollment in the study.
Serology	Data obtained from the laboratory examination of serum by specific immunologic methods.
Socioeconomic	Data related to the social and economic characteristics and environment of the study participant.
Speech	Data obtained from examination and observation of speech function.
Vision	Data obtained from examination of the eyes.
Work History	Data pertaining to occupation and employment prior to and during the period of enrollment in the study.
X-Ray	Data on diagnostic x rays and diagnostic or therapeutic radiological procedures.
Summary	Data presented as a summary of data collected and recorded elsewhere.
Gynecological History	Medical history specifically related to the female genital tract, reproductive physiology and endocrinology.
Special Studies	Data pertaining to participation in other special organized studies conducted during the period of enrollment in the study.
Family/Genetic History	Data on the medical histories of family members genetically related to the study child.
SLH Examination	Data obtained from the speech, language and hearing examinations not specifically or exclusively related to one of these areas.

CHAPTER 6. A RESEARCHER'S GUIDE TO THE NCPP DATA

Previous chapters give an overview of the NINCDS Collaborative Perinatal Project: how the study was conducted, how the data were processed, what type of information was collected and how requests for data are made. In this chapter, we explain how researchers can use other volumes of this guide to find both locations and documentation for each data item included on the NCPP computer files. First, we introduce concepts of data item identification and naming that are fundamental to accessing the data. Then a description of Volumes II through VII is given, followed by suggested approaches in using the volumes. Finally, example uses of the volumes are given based on hypothetical questions that might initiate access to one or more volumes.

DATA ITEM IDENTIFICATION AND NAMING

The NCPP data base contains over 6700 different data items and blank filler locations on computer files. We have assigned each of these a unique identification and a terse, stylized name. Because names were chosen to facilitate use of this guide, they do not duplicate names used by NINDB during the active phase of the project. Users should consult appropriate documentation before using data items from the master, variable or work files (Volumes II, III and IV).

The data item identifiers consist of 11 characters. At the far left are four unique numbers that were assigned sequentially. The next character is always a period and is followed by up to six characters. For data items on the master file, these characters describe the data collection form from which a data item was derived; for data items on the variable (VAR) or work (WXX) files, these characters indicate the appropriate file. If the right side is less than six characters, periods are inserted as shown in these examples:

850..OB-34	an item from OB-34; on the master file
3650.PATH-3	an item from PATH-3; on the master file
5223...VAR	an item on the variable file
6340...W-10	an item on work file 10, Rupture of Membranes

We assigned the numbers sequentially as they appear in Volume V. For the master file, we followed the order in which the cards would be found within an NINDB case. All card columns are accounted for by one of our data item identifications. For the variable and work files, the numbers were assigned in the order that data items appear within a case.

We categorized each data item according to the person to whom the data refer, by the time of measurement and/or the time to which the item applies and by general type or subject area (Table 4.4). Then we assigned names to the data items using the following guidelines:

- The name and the three associated categories had to stand alone - they must describe the data item out of context.
- The first word in the data item name had to be an important or key word when all names were listed alphabetically as in

Volumes VI and VII. Thus "cry, abnormal" was used rather than "abnormal cry" because a researcher is more likely to look for this item under "C" than under "A" in an alphabetic list.

- Secondary key words were preceded with a semicolon to facilitate preparation of the permuted index. For example, "abruptio; placenta" will be found under both the "A" and "P" portion of Volume VI.
- Qualifying words are delimited by commas and will not appear as keywords in Volume VI. Thus "abruptio; placenta, degree" will not be found in the "D" section.
- If medical terminology or usage has changed since the study was conducted, modern terms may be included and will be enclosed in brackets. Thus "mongolism; [Down's syndrome]" will appear under both the "M" and "D" portions of Volume VI.
- If measurement units are associated with a data item name, they are enclosed in parentheses and placed at the end of the name as in "Birthdate (yr)."
- The categories (person, time and subject) are appended to the right of the data item name.

Definitions for each category used in naming data items are given in Table 4.5. In the computerized listings of data item names, abbreviated category names are used (Table 6.1).

Data item names thus assigned are terse and highly stylized; as we have already indicated, they are not the names used by NINDB during the active phase of the project. Our aim was to develop standardized names that would stand alone. These names are intended to facilitate a user's search for data items potentially useful in a research project. Before an item is used, a researcher should consult its complete description. For a data item from the master files, e.g., 850..OB-34, the data item should be traced to the appropriate study form, e.g., OB-34, located in Volume II. A variable file data item, e.g., 5223...VAR, is traced to Volume III, where it is defined and its original source given. A data item from a work file is traced to Volume IV for its description.

Some data items contained in the indexes may include the notation "DO NOT USE." These items are either inaccurate or an alternative data item is available that gives better information. Users will find more appropriate data items by consulting one of the indexes to the data items (Volumes, V, VI and VII).

TABLE 6.1. Abbreviations for Person, Time and Subject categories

<u>Person</u>	<u>Time</u>	<u>Subject</u>
Mother	General	Administrative
Father	Preconception	Anesthesia
Placenta	Registration	Clin. Impression
Fetus	Prenatal	Clinical Lab
Child	Admission	Current Pregnancy
M Surrogate	Intrapartum	Environ. Exposure
Family	Delivery	Events
Sibship	Post Partum	Hearing
	Neonatal	Hospitalizations
	Four month	Language
	Eight month	Linkage
	One year	Malformations
	Three year	Diag. & Cond.
	Four year	Med. History
	Seven year	Medications
	Eight year	Neurological Exam
		Observations
		Pathology
		Physical Exam
		Procedure
		Psych. Exam
		Reproductive Hist.
		Serology
		Socioecon. Info
		Speech
		Vision
		Work History
		X-ray
		Summary
		Gyn. History
		Special Studies
		Fam/Genet.: Hist.
		SLH Exam

VOLUMES OF THE GUIDE

The remaining volumes of this user's guide are of two types: documentation and indices. Volumes II, III and IV document the master file, variable file, and work files, respectively. Volume II is primary to the NCPP data and should probably be used by all researchers. It contains over 2000 pages and is subdivided into 10 separately bound parts. Volume III, Variable File, has two parts with approximately 1000 pages. Volume IV, Selected NCPP Work Files, is approximately 500 pages. Volumes V, VI and VII are indices containing all data items in the NCPP data base. These volumes consist entirely of computer generated lists. They are designed to aid a researcher in his search for data items.

Volume II, The Master File

In Volume II, Project Study Forms and Documentation of Transfer to Computerized Data Items, all forms (and their revisions), instructions for completing the forms, definition of codes (coding instructions), and computer card layouts for the master file are reproduced. In addition, we provide a summary of each form's purpose and revision history, along with a table indicating the number of records available for each revision. Data items pertaining directly to each form are listed in two separate orderings, one linking data item ID's to forms, the other linking item numbers from forms back to names of data items. All data items that can be linked to a study form are included in these tables, even items on the variable file or work files. Nevertheless, Volume II contains the primary documentation for the NCPP and, therefore, the project's master file. Volumes III and IV, which describe the variable and work files, contain descriptions for computerized data items that do not appear in the master file.

For researchers interested in pursuing data obtained from specific questions asked on forms, Volume II is essential. By comparing form revisions to determine if the wording of the question changed, an informed decision on consistency of data can be reached before a request is made. The researcher can also learn how many records exist for any one revision of a study form. Finally, computer generated listings of data items provided for each form will identify data items in the master file as well as in the variable or work files, when data from that form appear in those files. Usually data items from the variable or work files are preferred if they are available.

Because Volume II in its entirety covers over 2000 pages, the volume has been subdivided into parts according to form type and time of data collection. A description of each part appears in Table 6.2.

Volume III, Variable File

Volume III, Variable File, provides documentation on creation of variable file data items and an index of items and their locations on the variable file tape. This volume is essential only for those researchers obtaining data from the variable file. Because the variable file was created to facilitate access to key data items from the NCPP, researchers are urged to use this file whenever it can satisfy research requirements.

TABLE 6.2. Structure of Volume II Parts

<u>Part</u>	<u>Title</u>	<u>Forms Included</u>
A	Prenatal Record and Medical History	AR-1 OB-2 thru OB-12 OB-15 OB-42 thru OB-47
B	Labor and Delivery	OB-30 thru OB-34 OB-35/57 OB-50 OB-51/52 OB-55/56 OB-58 OB-60 ADM-49 thru ADM-51
C	Pathological Exams and Autopsies	PATH-1 thru PATH-3
D	Family and Socioeconomic History	FHH-1/3 FHH-2/4 FHH-9 SE-1 GEN-5 thru GEN-8
E	Neonatal Exams and Observations	PED-1 thru PED-8 ADM-44
F	Pediatric and Neurological Exams, Four Months - One Year Physical Growth Measurements, Interval History, and Summary of Illness or Hospitalization	PED-10 thru PED-12 PED-14 PED-20 PED-29
G	Pediatric Neurological Exams, Seven Years	PED-74 thru PED-76 IDC-77
H	Psychological Exams, Eight Months	PS-1 thru PS-5
I	Psychological Exams, Four Years and Seven Years	PS-20 thru PS-26 PS-30 thru PS-38
J	Speech, Language and Hearing Exams, Three Years and Eight Years (Final)	PS-10 thru PS-17 PS-40 thru PS-45

Each data item on the NCPP variable file was obtained by computer interrogation of the master file. Part A of Volume III is an index of data items contained in the variable file. A title for each data item, our data item name, a data item identification number, and the codes for the data are included in the listing. Original NINCDS titles used in the variable file listing do not correspond exactly to data item names constructed specifically for this user's guide, so both names are given. The data item identification number enables other volumes to be consulted about a specific data item. In Part B of Volume III, the logical procedure followed in creating the derived data items on the variable file are described. Each method has been numbered according to the location of the data item on the file and is ordered numerically within Part B.

An example of a data item as it appears in Part A is given below.

<u>Data Item ID</u>	<u>From</u>	<u>To</u>	<u>Data Item Name</u>
5918....VAR	1095	1098	Birth; weight (gms) BIRTHWEIGHT (GRAMS) CODES: 0001-7400 = AS GIVEN BLANK,9999 = UNKNOWN

The unique data item ID is 5918....VAR and it is on the variable file at tape locations 1095-1098. Our stylized data name is given first, followed by the NINDB data name and the codes applicable. The original source of the data item can be found by looking in Part B under either 5918....VAR or tape location 1095.

Volume IV, Selected NCPP Work Files

Volume IV, Selected NCPP Work Files, provides documentation of work files that are available and tape locations for data items contained on each file. Using this volume, investigators will be able to discover how previous researchers used the NCPP data base to create their own study files (referred to in this guide as work files). Extensive data validation and clean-up was performed on each work file during creation. Some work files, e.g., Cerebral Palsy Diagnosis, may represent review of original microfilmed study forms and classification by medical experts. The result is a concise subset of data that can be accessed easier than the master file. Computer time and cost of obtaining data are much reduced whenever a work file can be used, so researchers are encouraged to use these files whenever possible.

Four types of work files are available: files that contain data that were basic to the NCPP, meant to augment the master file (W1, W2, W3, W4); special subject or study files (W5, W6, W7, W8, W9, W17, W18); serology files (W11, W12, W13, W14, W15, W16); and an administrative file (W10). Eighteen separate work files are documented here, listed below.

W1	Socioeconomic index at registration
W2	Socioeconomic index at seven years
W3	Drugs taken during pregnancy, trade names

- W4 Drugs taken during pregnancy, active compounds
- W5 Congenital malformations, one and seven years
- W6 Cerebral palsy diagnosis
- W7 Abnormalities at seven years
- W8 Speech, language and hearing at eight years
- W9 Toxemia classification
- W10 Rupture of membranes
- W11 Survey of viral, bacterial, parasitic and fungal infections during pregnancy
- W12 Serological testing, complement fixation tests
- W13 Serological testing for toxoplasmosis and rubella
- W14 Serological testing, cord blood
- W15 Serological testing, abnormalities and controls
- W16 Serum specimen inventory
- W17 Family linkage
- W18 Visit summary

Volume V, Master Index

Volume V, Master Index to the NCPP Computerized Data Items, provides a complete listing of all data items and their locations on the master, variable and work files. All data items from all files are included in the master index, including those data items that were derived from a study form. Master file data items are listed in the first portion of the master index, arranged according to card number and column. Data items from the variable file are arranged numerically by location on the file; data items from work files are arranged according to work file ID and location within that specific file. Volume V is useful during final preparation of a data request.

Examples of information in the master index appear below. From this information, investigators can learn the data item's location on either master, variable or work file.

<u>Data Item ID</u>	<u>Card Num</u>	<u>From</u>	<u>To</u>	<u>Data Item Name</u>
3732..PED-1	1401	22	23	Birth; weight (lbs): CHILD, DELIVERY, PHYSICAL EXAM
3733..PED-1	1401	24	25	Birth; weight (oz): CHILD, DELIVERY, PHYSICAL EXAM
5918....VAR		1095	1098	Birth; weight (gms): CHILD, DELIVERY, PHYSICAL EXAM (PREFERRED)
6269....W-6		45	47	Cerebral palsy; diplegia, atonic (1 yr, interim, 7 yr): CHILD, GENERAL, DIAGNOSES AND CONDITIONS

Volume VI, Alphabetical Permuted Index

Volume VI, Alphabetical Permuted Glossary of NCPP Computerized Data Items, provides an alphabetical listing of all data items in the NCPP data files. Key phrases in a data item are permuted so that a data item may appear more than once. The Alphabetical Permuted Index enables a researcher to locate a specific data item by a name typically used by the research community; it is not well suited to locating general categories of data items. After locating a data item in the index, the researcher is directed to either Volume II, III or IV, which describe the data items in the master, variable and work files, respectively.

Volume VII, Categorization of Data Items

Volume VII, Categorization of Data Items, provides three separate listings of data items categorized by person, time and subject. This volume provides researchers with a listing of NCPP data items organized by general categories rather than by specific data items, as the data items appear in Volume VI. The procedure for using the two volumes is the same. The categorization is based on classifying a data item into three dimensions: the person the data item refers to (mother, child, etc.), the time of measurement or observation of the data item, and the general subject area of the data item (see Chapter 4). The three listings in Volume VI order the categories three separate ways, but are otherwise exactly alike. Part A orders items by person, time and then subject; Part B by time, person and then subject, and Part C by subject, person and then time. The three parts are intended to facilitate a researcher's search for the category desired.

STEPS TO FOLLOW IN USING THE GUIDE

We anticipate that users of the guide will approach the data base from a variety of viewpoints, depending on their particular research question and orientation. For this reason, we have provided a variety of indices to the data base in addition to detailed documentation on how data items arrived in the master, variable and work files. In this section, we suggest general approaches an investigator might take in considering the NCPP as a source for data. We then provide examples of how an investigator can use the guide to make a request.

If you approach the NCPP by asking how the study was conducted, when data were collected, and what data were collected on each form, then study Volume I followed by Volume II. Volume I provides an overview of the NCPP and is the best place for a researcher unfamiliar with the study to begin; Volume II documents the study forms from initial data collection to final transfer to computerized data items.

A researcher who begins assimilating information based on a study of actual data collection forms must use Volume II. In your examination of Volume II, you will want to determine how revisions of the forms may have affected data collected. If a revision did result in changes in meaning, you will want to check the number of records (cases) for particular revisions before proceeding. You must determine if the number of cases is sufficient for your use. This may not be possible without obtaining the actual data.

Volume II reveals how data were coded onto the master file and where data are located on other files as well. The final question Volume II should answer is whether you must request data from the master file, or if you can satisfy your needs with the variable file or one or more of the work files. It is not necessary to consult Volumes V, VI or VII if you find it easy to locate the data items of interest to you by going directly to the original data collection forms. You may require Volumes III and IV if some data items of interest are referenced to the variable or work files.

Another approach to finding data items in the NCPP data base is to search for information on a specific variable, e.g., a specific diagnosis or condition. In this case, the Alphabetical Permuted Index (Volume VI) is the best mechanism for locating the data item. As with all permuted indices, a researcher may have to look under several different synonyms before locating the item or determining it is not available.

Other researchers will approach the NCPP with several general topics or categories of interest in their research. Although the project study forms (Volume II) may be consulted, an alternative has been provided in Volume VII, Categorization of Data Items. A categorization of all data items is provided to assist the researcher (Chapter 4).

Finally, some researchers initially approach a data base of potential use to them by consulting a list of all data items available. The Master Index (Volume V) is constructed for just this purpose. The order of the data items in the index is the same as on the computer files. Naturally, some logical order was used in constructing the files.

EXAMPLE USES OF THIS USER'S GUIDE

To aid researchers in using this guide we have posed some hypothetical questions that might initiate access to one or more volumes. In all cases, the original data collection forms and manuals should be referred to at an early stage in the project development to understand how the data items were collected. In addition, the demographic information for defining/selecting cases and controls must be considered. Included would be factors such as age, race, parity, etc.

Example 1

A group is conducting research on the association between abruptio placenta and neurological function of the newborn. Are relevant data available from the NCPP?

One approach is to determine first what information is available on abruptio placenta. Since this is a possible name for a specific data item, our search will start in the Alphabetical Permuted Index (Volume VI). By looking under abruptio placenta, the following data items are found:

6125....VAR	abruptio placenta (OB-34). Cesarean section, indication
6140....VAR	abruptio placenta (OB-55). Cesarean section, indication
942..OB-34	abruptio placenta. Cesarean section, indication
1739..OB-55	abruptio placenta. Uterine stimulant, augmentation indication
1730..OB-55	abruptio placenta. Uterine stimulant, induction indication
6172....VAR	abruptio placenta. Uterine stimulant, induction indication
1883..OB-55	abruptio placentae. Bleeding before cord clamped, cause
1834..OB-55	abruptio placentae. Cesarean section, indication
6040....VAR	abruptio; placenta, (yes, no, unknown)
6039....VAR	abruptio; placenta, degree
1875..OB-55	abruptio; placenta.

In this example, data items are found in the variable file and in the master file on forms OB-34 and OB-55. The exact definitions of these data items are found in Volume III for the Variable File and Volume II, Part B for the master file. A data item is located in the computer listing for Volume III by using the data item identification, i.e., 6125, 6140, 6172, 6040 and 6039. Once the data item is located, the original source of the information, the derivation method and definition of codes can be determined. For data items in the master file, the researcher need only locate the sections in Volume II that describe the OB-34 and OB-55 forms. Once the appropriate section is located, a brief computer listing of all data items derived from the form (from the master file) is given. This helps locate the specific question items on the form associated with the computerized data item. A thorough researcher will determine how the question was asked, what coding instructions were used and how the response was keypunched. All of this information is available in Volume II, Part B, organized by data collection form.

The example research question also asks for information on the neurological development of the newborn. One way to proceed is to determine from Chapters 2 and 4 of this volume which forms completed for the newborn contain information on neurological development. PED-6 refers to a neonatal neurological exam, but other PED forms may apply as well. Part E of Volume II would be used to begin a search based on this information.

An alternative approach is to use the categorization of data items in Volume VII. A primary focus of the research question is the neonatal period of observation. A secondary focus is a general area(s) related to neurological development. On that basis, Part B: Categorization by Time, may be used to find all data items in the NEONATAL time category and the CHILD person category. All general subject categories for the NEONATAL, CHILD grouping are then together. The subject categories with data items are:

Administrative	Malformations	Observations
Anesthesia	Diag & Cond	Physical Exam
Clinical Lab	Med. History	Procedure
Events	Medications	X-ray
Hospitalizations	Neurological Exam	

Under neurological observations, a number of data items are available, almost all of these from PED-2 or PED-6 on the master file (Volume II, Part E). For completeness, a researcher should also check under the other subject categories, as the assignment of data items to a single subject category is sometimes difficult. Once the researcher finds the appropriate computer file containing the data items of interest, the appropriate volume documenting that computer file (II, III or IV) should be consulted for detailed information.

Example 2

"There has been considerable discussion of possible teratogenic risks associated with Bendectin use in pregnancy. We would like to identify a group of prospectively ascertained women who took Bendectin and determine if there are increased rates of congenital malformations in their offspring."

This example is similar in design to the approach taken by Heinonen et al. (1977) in their study of drugs in pregnancy using the NCPP data. We use it here to illustrate an approach to the data set utilizing specific work files.

The initial research requirement is to be able to identify a cohort of exposed women. Appropriate controls must also be identified. Since data will be provided on all registrants for the study, it will be necessary for the researcher to develop the criteria necessary for selecting controls. Data for cases and control women then must be derived from the files for analysis.

Table 2.1 (Chapter 2) shows that information on drug use in pregnancy was collected on Form OB-15. The hierarchical classification of data categories, Table 4.3, shows that Category 3, Obstetrics-Miscellaneous Prenatal Records, includes a secondary category Drugs in Pregnancy that includes information on drugs taken by lunar month of pregnancy.

Volume II indicates that Form OB-15 was completed by the Perinatal Research Branch, rather than by the collaborating institutions, and was based on review of the patient's records by Perinatal Research Branch staff. As illustrated below, drugs were coded by name (Figure 6.1). For each drug taken, information was collected on extent of drug use by lunar month of pregnancy.

An important factor relating to records of drug use in pregnancy is that the data from OB-15 were subsequently reorganized and two drug files were created (Heinonen et al., 1977). The drug files, one showing drugs taken during pregnancy by trade name, and the second showing drugs by active compounds, are described in Volume IV, Selected NCPP Work Files. The researcher can identify subjects who used Bendectin (work file 3 -- Drugs taken during pregnancy, trade names) or either of its active ingredients,

dicyclomine and doxylamine succinate (work file 4 -- Drugs taken during pregnancy, active compounds). Bendectin usage by lunar month of pregnancy and by the frequency of use is given. Definitions of codes are included in the description of the work file (Volume IV), along with the drug dictionaries. These dictionaries allow identification of both the trade names and active compounds of the drugs for which data are available.

The following example illustrates how the data on drug use are stored in work file W3. A complete description is found in Volume IV.

Subject XXXXXYYZZ took Bendectin (Code 7340) for seven days in her second lunar month of pregnancy and daily throughout her third and fourth months. She took the drug for the first 12 days of the fifth month of pregnancy.

XXXXXXYYZZ73400023330000000

Identification of outcomes of interest, once the cohort of exposed women has been identified, can involve the use of Volumes II and IV. Table 4.3 (Chapter 4) shows that information on malformations observed in the neonate is included on Form PED-8. For fetal and infant deaths, information on malformations is included on Forms PED-4 and PATH-3. Information on malformations recorded at the one-year exam is included on Form PED-12; congenital malformations recorded through age seven are included in the seven-year diagnostic summary (ADM-86 or IDC-77).

OB-15 DRUGS IN PREGNANCY - VII														
CODE: A - TAKEN 1 DAY ONLY B - TAKEN 1-7 DAYS C - TAKEN MORE THAN 7 DAYS D - TAKEN UNKNOWN TIME		LUNAR MONTH OF PREGNANCY												
COMMON COMBINATIONS - PRESCRIPTION	4 wks prior to LMP	1	2	3	4	5	6	7	8	9	10	11	UNIC	
	-1													
7106 Achloridin.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7432 Amber.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7435 Amphetese.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7217 APC (enacin).....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7340 Bendectin.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7401 Colofret.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7434 Comb-d.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7351 Co-pyrrol.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

FIGURE 6.1. Drugs in Pregnancy OB-15 Example Form

When the investigator reviews the material in Volume II, he will note that specific diagnostic codes are recorded for each condition. The most appropriate way to identify malformation cases is to use a special work file (W5) that was constructed to include congenital malformations recorded at one and seven years rather than the master file.

Using the work file W5, fully documented in Volume IV, the researcher can identify all cases of malformations diagnosed at, or prior to, one and seven years. Thus, appropriate diagnoses can be selected for analysis or the occurrence of malformations of all types can be compared between groups of women who differ in antecedent exposure to some substance of interest, in this case, Bendectin. The details of coding and data retrieval procedures may be determined by studying Volume IV.

In summary, this example illustrates how the work files can be used to address a research question. Women who took a specific drug, identified by either trade name or active compound, may be ascertained from work files W3 or W4. The occurrence of congenital malformations in the offspring of these women can be determined from work file W5. The process by which the data were collected are presented in Volume II and the construction and coding of the work files is detailed in Volume IV.

Example 3

"I have a graduate student who is determining if there is any association between neonatal anoxia and language development. What relevant information is available from the NCPP data files?"

This example illustrates how general questions can be addressed using the volumes of this guide and how specific relevant variables can be identified. It illustrates how this guide can be reviewed to develop and clarify a research hypothesis.

The most logical approach to this question is to go to the alphabetical permuted index (Volume VI) and determine if anoxia appears as an entry. The following would then be found:

<u>Person</u>	<u>Time</u>	<u>General Subject Area</u>	<u>Data Item ID</u>	<u>Data Item Name</u>
Child	Neonatal	Events	5611....VAR	Anoxia, presumed etiology of conditions

Reference to the variable file description in Volume III shows the following for this variable.

<u>Data Item ID</u>	<u>From</u>	<u>To</u>	<u>Data Item Name</u>
5611....VAR	785		Anoxia, presumed etiology of conditions PRESUMED ANOXIA (codes:987)

Review of the information in Volume III regarding the source of this variable reveals that it comes from PED-8, the Newborn Diagnostic Summary. Reference to the PED-8 form and manual in Volume II allows us to determine how this information was collected and coded.

It is important to note that this category relates not to the presence of anoxia but rather to impressions concerning the etiology of previously coded conditions, as these relate to presumed anoxia (Volume II - Manual for the PED-8 Form). Thus we cannot identify cases of anoxia by use of the alphabetical index, but only those cases where anoxia is the presumptive etiologic mechanism for another condition.

We include this example because it illustrates the necessity of a thorough approach to the documentation of the NCPP data base. Anoxia does not appear in the data base as a coded diagnosis or condition. Thus, to examine an association between anoxia and a subsequent outcome requires the identification of indications of anoxia, the approach we now present.

The researcher might begin by reviewing the hierarchical classification of data categories (Table 4.3 in Chapter 4). Under the category Pediatrics-Newborn, we find that delivery room observations are recorded on form PED-1; these observations include timing of cord clamping, first breath and first cry, suction and resuscitation procedures and Apgar scores. Reference to form PED-1 in Volume II will identify all the specific items relevant to these topics; from these, the information regarding anoxia can be determined. In addition, the location of these variables in the data files can be determined from Volume II. In this example, many of the relevant variables are included in the variable file. Other information of interest is recorded only in the master file. For example, one indication of anoxia would be the administration of oxygen in the delivery room. In the variable file (Volume III), the following information is recorded:

<u>Data Item ID</u>	<u>From</u>	<u>To</u>	<u>Data Item Name</u>
5403....VAR	573		Oxygen administered, open PROCEDURES-OPEN OXYGEN CODES: 0= Not used 1 = Used 9 = Unknown
5404....VAR	574		Oxygen or air administered, positive pressure PROCEDURES-POSITIVE PRESSURE CODES: 0 = Not used 1 = Used 9 = Unknown

In the master file, the following information is recorded on oxygen administration:

<u>Data Item ID</u>	<u>Card Num</u>	<u>From</u>	<u>To</u>	<u>Data Item Name</u>
3828..PED-1	3401	32	32	Oxygen administered, open: CHILD, DELIVERY, PROCEDURE
3829..PED-1	3401	33	34	Oxygen administered, open, age begun (min): CHILD, DELIVERY, PROCEDURE
3830..PED-1	3401	35	36	Oxygen administered, open, duration (min): CHILD, DELIVERY, PROCEDURE
3831..PED-1	3401	37	37	Oxygen or air administered, positive pressure: CHILD, DELIVERY, PROCEDURE
3832..PED-1	3401	38	39	Oxygen or air administered, positive pressure, age begun (min): CHILD, DELIVERY, PROCEDURE
3833..PED-1	3401	40	41	Oxygen or air administered, positive pressure, duration (min): CHILD, DELIVERY, PROCEDURE

We are now in a position to redefine our research question in terms of specific indicators of anoxia. In addition, we could develop an approach to anoxia that is based on the presence of combinations of specific conditions such as Apgar scores, oxygen administration, length of time to first cry, etc.

In this hypothetical example, the outcome variable of interest is language development. Examination of Table 4.1 reveals that information on language was collected at three years and eight years of age. Table 4.3 shows that data were collected on language expression as part of both examinations (PS-11, PS-42). Data were collected on language reception as part of the three-year exam (PS-10) and on language comprehension (PS-41) as part of the eight-year exam. The specific test items are described in Volume II, along with information as to how the data were collected.

To identify the data items available readily, the researcher is referred to Volume VII, where data items are categorized by person, time and subject (see Table 4.4). We are interested in the subject area language and the child at three (or eight) years. The entry for three years is shown below.

As can be seen, information from the three-year language exam is found on the master file only, with the exception of two summary scores that are also included on the variable file.

LANGUAGE
CHILD
THREE YEAR

2933..PS-10	Comprehension, non-verbal, matching
2932..PS-10	Comprehension, non-verbal, pantomime
2930..PS-10	Comprehension, non-verbal, picture
2931..PS-10	Comprehension, non-verbal, word, object
2928..PS-10	Comprehension, verbal, action word
2927..PS-10	Comprehension, verbal, familiar object
2929..PS-10	Comprehension, verbal, space
2934..PS-11	Expression, verbal, naming objects
2937..PS-11	Expression, verbal, relevance
2935..PS-11	Expression, verbal, sentence length
2936..PS-11	Expression, verbal, sentence structure
2940..PS-11	Expression, verbal, summary score
2939..PS-11	Expression, verbal, use of pronouns
2938..PS-11	Expression, verbal, word order
2915..PS-17	Language expression, summary, final
5928....VAR	Language expression, summary, final
2914..PS-17	Language reception, summary, final
5927....VAR	Language reception, summary, final
2942..PS-11	Non-verbal; expression, objects
2941..PS-11	Non-verbal; expression, picture

This example illustrates how an investigator can make effective use of Volume VII. The investigator who has a particular subject area in mind can identify the specific variables for which information is available. In this case, the subject area was language development; the researcher must identify which of the study variables most closely relate to his/her concept of language development. We have shown how Volume II can be used to identify the data items collected and how Volume VII can be used to determine the availability of specific variables in the NCPP data sets.

SUMMARY

These examples provide an introduction to how the volumes of the user's guide can be used in developing a study based on data from the NCPP. It is likely that different researchers will approach the data from different perspectives. That is the reason different approaches have been utilized in the examples.

It is important to recognize that in planning a research project based on the NCPP data, the definition of case and control populations, or exposed and unexposed cohorts, needs to include demographic characteristics and also characteristics unique to the study itself such as the institutions' subject selection procedures (Appendix A) and the definition of standard cohorts (Appendix B). In addition, the procedures for the collection and coding of data items for the study, as detailed in Volume II, should be reviewed and evaluated in lieu of the researcher's data needs. It is through a thorough understanding of the data collection process and the data set that the NCPP data can be most effectively utilized.

CHAPTER 7. REFERENCES

- Broman, S.H., P.L. Nichols and W.A. Kennedy. 1975. Preschool IQ Prenatal and Early Developmental Correlates. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey. Distributed by the Halsted Press Division of John Wiley and Sons, New York.
- Friedman, E. A., ed. 1976. Blood Pressure, Edema and Proteinuria in Pregnancy. Prog. Clin. Biol. Res., Vol. 7, Alan R. Liss, Inc., New York.
- Friedman, E. A. and R. K. Neff. 1977. Pregnancy Hypertension. PSG Publishing Company, Littleton, Massachusetts.
- Gilles, F. H., A. Leviton and E. C. Dooling. 1983. The Developing Human Brain. Growth and Epidemiological Neuropathology. John Wright - PSG, Inc., Boston,
- Hardy, J.B., J.S. Drage, and E.C. Jackson. 1979. The First Year of Life. The Collaborative Perinatal Project of the National Institute of Neurological and Communicative Disorders and Stroke. The Johns Hopkins University Press, Baltimore.
- Heinonen, O.P., D. Sloan, and S. Shapiro. 1977. Birth Defects and Drugs in Pregnancy. Publishing Sciences Group, Inc., Littleton, Massachusetts.
- Lassman, F.M., R.O. Fisch, D.K. Vetter, and E.S. LaBenz. 1980. Early Correlates of Speech, Language, and Hearing. PSG Publishing Company, Littleton, Massachusetts.
- Melnick, M. and N. C. Myrianthopoulos. 1979. External Ear Malformations: Epidemiology, Genetics, and Natural History. Birth Defects Original Article Series, Vol. XV, No. 9.
- Myrianthopoulos, N. C. 1975. Congenital Malformations in Twins: Epidemiologic Survey. Birth Defects Original Article Series, Vol. XI, No. 8.
- Myrianthopoulos, N. C. and C. S. Chung. 1974. Congenital Malformations in Singletons: Epidemiologic Survey. Birth Defects Original Article Series, Vol. X, No. 11.
- National Institute of Health. 1982. Policies and Practices for Storing, Retrieving, Accessing, Retaining and Disposing of Records in the System. Federal Register, Volume 47, No. 198, pp. 45785-45786. (Cited in text as 47 FR 198, October 13, 1982.)
- Nichols, P. L. and T-C. Chen. 1981. Minimal Brain Dysfunction: A Prospective Study. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.
- Niswander, K.R., and M. Gordon. 1972. The Women and Their Pregnancies. U.S. Department of Health, Education and Welfare, Public Health Service, National Institutes of Health. Published by the U.S. Government Printing Office, Washington, D.C.

APPENDIX A
SUBJECT SELECTION

SUBJECT SELECTION

This appendix includes a table of information on the case selection process for each collaborating institution. The tables include the date sample selection began and dates any major changes in the study selection procedures were initiated. The sampling frame criteria are defined and exclusions from the sampling frame, or from the sample, are identified. Estimates of the frame size, method of sampling and expected number of cases per year are included.

Study Selection Procedures At Boston Lying-In Hospital (NINDB Institution Number 05)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
January 1959	All clinic patients admitted	Unwed referrals from Florence Crittenden Home	None	2000	50% using hospital unit number - 1000 -	None
March 1959	(a)	Unwed referrals from Florence Crittenden Home	None	2000	100% using hospital unit number - 2000 -	None
May 1960	(a)	As above and walk-ins	None	2000	100% using hospital unit number - 2000 -	None
December 1960	(a)	As above and unwed mothers who plan to give up their babies	None	2000	100% using hospital unit number - 2000 -	None
June 1962	(a)	As above and unwed mothers who plan to give up their babies	None	2000	100% using hospital unit number - 2000 -	Diabetic mothers who made weekly visits to Joslin Clinic

(a) All clinic patients admitted.

Study Selection Procedures At Children's Hospital — Buffalo (NINDB Institution Number 10)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
October 1960	Private patients seen by 13 participating obstetricians	Walk-ins, women who do not intend to deliver at Children's Hospital or who plan to move out of area	None	Unknown	Average of two patients per month, randomly selected, per participating obstetrician - 312 -	Courtesy cases at request of participating obstetricians
August 1961	Private patients seen by 12 participating obstetricians	(a)	None	Unknown	Average of two patients per month, randomly selected, per participating obstetrician - 288 -	Courtesy cases at request of participating obstetricians
November 1961	Private patients seen by 4 participating obstetricians	(a)	None	Unknown	100% of eligible women seen by 3 obstetricians & 50% of those seen by the remaining obstetrician - 400 -	None
May 1962	Private patients seen by 4 participating obstetricians	(a)	Patients who terminate pregnancy within 7 days of initial visit defined as walk-ins	Unknown	100% of eligible women seen by 3 obstetricians & 50% of those seen by the remaining obstetrician - 400 -	None
December 1963	Private patients seen by 4 participating obstetricians	(a)	(b)	Unknown	100% of eligible women seen by 4 obstetricians - 450 -	None
January 1964	Private patients seen by 6 participating obstetricians	(a)	(b)	Unknown	100% of eligible women seen by 6 obstetricians - 900 -	None

(a) Walk-ins, women who do not intend to deliver at Children's Hospital or who plan to move out of area.

(b) Patients who terminate pregnancy within 7 days of initial visit defined as walk-ins.

Study Selection Procedures At Charity Hospital (NINDB Institution Number 15)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
March 1960	All black patients residing in Orleans Parish and assigned to Charity Hospital services of Tulane or LSU	None	None	4300	10% systematic All OPD numbers ending in zero - 430 -	None
May 1960	(a)	Walk-ins	None	4300	10% systematic All OPD numbers ending in zero - 430 -	None
April 1963	(a)	Walk-ins	None	3800	1 out of each 8 patients - 450 -	None
July 1963	(a)	Walk-ins	None	3000	1 out of each 6 patients - 500 -	None

(a) All black patients residing in Orleans Parish and assigned to Charity Hospital services of Tulane or LSU

Study Selection Procedures At Columbia Presbyterian Hospital (NINDB Institution Number 31)

<u>Date Initiated</u>	<u>Sampling Frame Criteria</u>	<u>Exclusions From Sampling Frame</u>	<u>Special Exclusions From Sample</u>	<u>Annual Sampling</u>		<u>Special Studies</u>
				<u>Frame Size</u>	<u>Method and Expected Cases</u>	
January 1959	All clinic patients admitted	None	None	3000	Every sixth case - 500 -	None
April 1, 1960	All clinic patients admitted	None	None	3000	Every fifth case - 600 -	None
April 14, 1960	All clinic patients admitted	None	None	3600	Every sixth case - 600 -	None
May 1960	All clinic patients admitted	Walk-ins	None	3600	Every sixth case - 600 -	None
April 1962	Clinic patients admitted residing in Manhattan and the Bronx only	Walk-ins	none	2500	Every fifth case - 500 -	None
April 1963			Case Selection Terminated			

Study Selection Procedures At John Hopkins Hospital (NINDB Institution Number 37)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
January 1959	All clinic patients who live within Metropolitan Baltimore (approx. 25 mile radius)	Transients and patients admitted who are referred to county clinics for obstetrical care	None	1800	20% of gravida based on history numbers ending in 5 or 8 - 360 -	Non-selected women who: 1. are 15 yrs. or under 2. are 40 yrs. or over 3. are diabetic at registration 4. have history of 4 or more fetal losses 5. have history of malformed child
September 1959	(a)	(b)	None	1600	30% of gravida based on history numbers ending in 5, 6, or 8 - 480 -	Non-selected women who: 1. are 15 yrs. or under 2. are 40 yrs. or over 3. are diabetic at registration 4. have history of 3 or more fetal losses 5. have history of malformed child
May 1960	(a)	As above and walk-ins	None	1600	30% of gravida based on history numbers ending in 5, 6, or 8 - 480 -	(c)
April 1962	(a)	As above and walk-ins	None	1400	40% of gravida based on history numbers ending in 5, 6, 8 or 9	(c)
January 1964	(a)	As above and walk-ins	None	1400	100% of gravida in sampling frame	(c)
December 1964	Case Selection Terminated					

- (a) All Clinic patients who live within Metropolitan Baltimore (approx. 25 mile radius)
 (b) Transients and patients admitted who are referred to county clinics for obstetrical care.
 (c) Non-selected women who:
 1. are 15 yrs. or under
 2. are 40 yrs. or over
 3. are diabetic at registration
 4. have history of 3 or more fetal losses
 5. have history of malformed child

Study Selection Procedures At Medical College of Virginia (NINDB Institution Number 45)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
January 1959	Clinic patients residing within a 50-mile radius of Richmond, Virginia	1. White welfare cases 2. Any woman who indicates that her child is up for adoption	Occasionally exclude patients who refuse to enter the study for personal or other reasons	1200	Every fourth black patient. All white patients - 500 -	Non-selected women in the first trimester of pregnancy
May 1960	Clinic patients residing within a 50-mile radius of Richmond, Virginia	Walk-ins	(b)	1200	Every fourth black patient. All white patients - 500 -	(c)
September 1960	Clinic patients residing in the city of Richmond and the surrounding counties of Chesterfield, Hanover and Henrico	1. Wards of the state who are in correctional institutions 2. Patients who plan to put their children up for adoption	(b)	950	Every second black patient. All white patients - 600 -	(c)
January 1962	Clinic patients residing in the city of Richmond only	(a)	None	1200	Every second black patient. All white patients - 600 -	(c)
August 1962	Clinic patients who are "city residents" of the city of Richmond	(a)	None	780	2 of every 3 black patients, all white patients - 550 -	(c)
October 1962	Clinic patients who are "city residents" of the city of Richmond	(a)	None	500	100%	None

(a) 1. Wards of the state who are in correctional institutions

2. Patients who plan to put their children up for adoption

(b) Occasionally exclude patients who refuse to enter the study for personal or other reasons

(c) Non-selected women in the first trimester of pregnancy

Study Selection Procedures At University of Minnesota (NINDB Institution Number 50)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
January 1959	All clinic patients admitted	<ol style="list-style-type: none"> 1. Women who were never married 2. Those divorced, widowed, or seperated before the start of the current pregnancy 3. Those registered for the first time after 246 days of pregnancy (as determined by the obstetrician)(a) 	Occasional patient because of language difficulty	300	100% - 300 -	None
October 1959	All clinic patients admitted	Eliminated #3 except that walk-ins are still excluded(a)	(b)	330	100% - 330 -	None
December 1959	All clinic patients admitted	Eliminated #1 and #2 therefore walk-ins only exclusions(a)	(b)	500	100% - 500 -	None

(a) Booth Memorial Referrals considered as "For Delivery Only"

(b) Occasional patient because of language difficulty

Study Selection Procedures At New York Medical College (NINDB Institution Number 55)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
February 1959	All clinic patients admitted	Patients admitted for delivery only	Occasional refusal to enter program	5000	10 cases per week which is approximately 10% - 500 -	None
January 1960	(a)	Patient admitted for delivery only	(b)	5000	Every ninth case - 555 -	None
March 1960	(a)	Patients admitted for delivery only	(b)	5000	Every eighth case - 625 -	None
May 1960	(a)	Walk-ins (also see above)	(b)	5000	Every eighth case - 625 -	None
June 1960	(a)	Walk-ins (also see above)	(b)	5000	Every seventh case - 715 -	None
January 1961	(a)	Walk-ins (also see above)	(b)	5000	Every sixth case - 835 -	None

Follow-up Terminated July 1970

-
- (a) All clinic patients admitted
 (b) Occasional refusal to enter program

Study Selection Procedures At University of Oregon (NINDB Institution Number 60)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
March 1959	All clinic patients admitted	None	None	1500	Every third case - 500 -	Non-selected repeaters
June 1959	(a)	(b)	None	1500	Every third case - 500 -	Non-selected repeaters
March 1960	(c)	(b)	None	750	Two of every three cases - 500 -	Non-selected repeaters
May 1960	(c)	As above and Walk-ins	None	750	Two of every three cases - 500 -	Non-selected repeaters and sample of non-frame Walk-ins
July 1960	(c)	As above and Walk-ins	None	1000	Every other case - 500 -	Non-selected repeaters and sample of non-frame Walk-ins
April 1961	(d)	As above and Walk-ins	None	1000	42% of grvida based on Month of birth; Feb., April, August, Oct., Dec. - 417-	None
January 1962	(d)	As above and Walk-ins	None	1000	54% of grvida based on birth on odd numbered days - 540 -	None
January 1964	(d)	As above and juvenile detention home, Ballington Home and jail patients	Repeat cases who refuse will be excluded from sampling frame	1000	54% of grvida based on birth on odd numbered days - 500 -	First trimester registrants not selected as study cases.

(a) All clinic patients admitted who reside within Portland mailing area

(b) Special Clinic: 1. Clients of private adoption agencies - 2. Medical students wives.

(c) All clinic patients admitted who reside within Multnomah County west of 122nd Street

(d) All clinic patients admitted who reside within Multnomah County west of 122nd Street.

No.residence requirements for patients admitted for repeat study of pregnancies

Study Selection Procedures At Pennsylvania Hospital (NINDB Institution Number 66)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
January 1959	All clinic patients admitted	Unregistered emergency delivery	None	1200	100% - 1200 -	None
August 1959	All clinic patients admitted	Unregistered emergency deliveries and patients who, at initial contact, state they will deliver elsewhere	None	1200	100% - 1200 -	None
May 1960	All clinic patients admitted	Walk-ins (also see above)	None	1500	100% - 1500 -	None

Study Selection Procedures At Providence Lying-in Hospital (NINDB Institution Number 71)

<u>Date Initiated</u>	<u>Sampling Frame Criteria</u>	<u>Exclusions From Sampling Frame</u>	<u>Special Exclusions From Sample</u>	<u>Annual Sampling</u>		<u>Special Studies</u>
				<u>Frame Size</u>	<u>Method and Expected Cases</u>	
March 1960	All clinic patients admitted	None	None	1100	(a)	None
May 1960	All clinic patients admitted	Walk-ins	None	1100	(a)	None
July 1960	All clinic patients admitted	Walk-ins	None	1100	50% - 550 -	None

(a) A varying proportion of actual clinic total will be selected so as to obtain a fixed number of new cases each week. A average clinic load per day for the past 3 weeks is used in determining the sampling level for the following week.

- 500 -

Study Selection Procedures At University of Tennessee (NINDB Institution Number 82)

<u>Date Initiated</u>	<u>Sampling Frame Criteria</u>	<u>Exclusions From Sampling Frame</u>	<u>Special Exclusions From Sample</u>	<u>Annual Sampling</u>		<u>Special Studies</u>
				<u>Frame Size</u>	<u>Method and Expected Cases</u>	
October 1959	All clinic patients admitted who reside within the city limits of Memphis, Tennessee	None	None	6000	10% - 600 -	None
March 1960	All clinic patients admitted who reside within the city limits of Memphis, Tennessee	None	None	4200	Every seventh case - 600 -	None
May 1960	All clinic patients admitted who reside within the city limits of Memphis, Tennessee	Walk-ins	None	4200	Every seventh case - 600 -	None

APPENDIX B

STANDARD COHORT DEFINITIONS

STANDARD COHORT DEFINITIONS

This appendix includes the criteria used for deriving the standard study cohorts from the total population of women enrolled in NCPP. Text Table 1.2 presented a brief description of the cohort categories and the number of cases in each; here we present detailed information on the requirements for inclusion in the standard cohorts. The information is based on the data for each case in the master file. To interpret this information, the researcher is referred to Volume II, where individual computer card images can be found and the definitions of the codes are given. In the variable file, described in Volume III, all subjects are coded by standard cohort membership based on the criteria presented here. The information on the standard cohorts is included here and in Chapter 1 to provide an understanding of why differing numbers of cases appear in different publications and places in this guide.

Cohort I
Core Cases Excluding Walk-Ins

Standard Cohort — 1A **First Study Pregnancy, Single Birth,
Registered on or Before 12/31/64**

1. Core Study Cases: 9001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. First Study Pregnancy: 0001 card, col. 13, code 1
3. Single Births: Presence of any card (1401-3401), col. 14, code 0. If none of these cards present, then 8400 card, col. 14, code 0.
4. All cases Registered in 1964 or Earlier: 0001 card, cols. 38-43, codes 010159-123164.
5. Exclude from the above cases:
Walk-ins: 0001 card, col. 76 code 1
Non-Core Cases: 0001 card, col. 80, code 7

Standard Cohort — 1B **All Cases Registered on or Before 12/31/64**

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80.
2. All Cases Registered in 1964 or Earlier: 0001 cards, cols. 38-43, codes 010159-123164.
3. Exclude from the above cases:
Walk-ins: 0001 card, col. 76 code 1
Non-Core Cases: 0001 card, col. 0, code 7

Standard Cohort — 1C **First Study Pregnancy, Single Birth, Registered at Any Time**

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. First Study Pregnancy: 0001 card, col. 13, code 1
3. Single Births: Presence of any card (1401-3401), col. 14, code 0. If none of these cards present, then 0844 card, col. 14, code 0.
4. Exclude from the above cases:
Walk-ins: 0001 card, col. 76 code 1
Non-Core Cases: 0001 card, col. 80, code 7

Standard Cohort — 1D **All Cases Registered at Any Time**

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. Exclude from the above cases:
Walk-ins: 0001 card, col. 76 code 1
Non-Core Cases: 0001 card, col. 80, code 7

Cohort II
Core Cases Excluding Walk-Ins and Lost to Study

Standard Cohort for PRB Studies — IIA Cohort IA Minus Cases Lost to Study

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. First Study Pregnancy: 0001 card, col. 13, code 1
3. Single Births: Presence of any card (1401-3401), col. 14, code 0. If none of these cards present, then 0844 card, col. 14, code 0.
4. All cases Registered in 1964 or Earlier: 0001 card, cols. 38-43, codes 010159-123164.
5. Exclude from the above cases:
 - a. Walk-ins: 0001 card, col. 76 code 1
 - b. Non-Core Cases: 0001 card, col. 80, code 7
 - c. Lost to Study: see below for classification

Classification of Women Lost to Study

		OB-34 OR OB-55			AND	PED-1 OR ADM-44	
1334-3334	AND	9334	1355-7355	AND	9355	1401-3401	0844
All Cards Not Present		Not Present	All Cards Not Present		Not Present	All Cards Not Present	Not Present
All Cards Not Present		Present	All Cards Not Present		Present	All Cards Not Present	Not Present
Any Card Present		N/A	Any Card Present		N/A	Only a 14012 Card Present or all Cards Not Present(a)	Not Present
All Cards Not Present		N/A	All Cards Not Present		N/A	Only a 14012 Card Present or all Cards Not Present(a)	Not Present

(a) And Absence of all cards 1402-4402, 1403-2403, 1408-5408, 0405, 1410-3410, 1101-3101, 1411-3411, 1412-4412, 1110-2110, 1120-4120, 1406-4406, 0407, 1414 (all)

Standard Cohort for Summary Record File — IIB Cohort IB Minus Cases Lost to Study

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. All Cases Registered in 1964 or Earlier: 0001 card, cols. 38-43, codes 010159-123164
3. Exclude from the above cases:
 - a. Walk-ins: 0001 card, col. 76 code 1
 - b. Non-Core Cases: 0001 card, col. 80, code 7
 - c. Lost to Study: see below for classification

Classification of Women Lost to Study

		OB-34 OR OB-55			AND	PED-1 OR ADM-44	
1334-3334	AND	9334	1355-7355	AND	9355	1401-3401	0844
All Cards Not Present		Not Present	All Cards Not Present		Not Present	All Cards Not Present	Not Present
All Cards Not Present		Present	All Cards Not Present		Present	All Cards Not Present	Not Present
Any Card Present		N/A	Any Card Present		N/A	Only a 14012 Card Present or all Cards Not Present(a)	Not Present
All Cards Not Present		N/A	All Cards Not Present		N/A	Only a 14012 Card Present or all Cards Not Present(a)	Not Present

(a) And Absence of all cards 1402-4402, 1403-2403, 1408-5408, 0405, 1410-3410, 1101-3101, 1411-3411, 1412-4412, 1110-2110, 1120-4120, 1406-4406, 0407, 1414 (all)

Standard Cohort for PRB Studies — IIC Cohort IC Minus Cases Lost to Study

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. First Study Pregnancy: 0001 card, col. 13, code 1
3. Single Births: Presence of any card (1401-3401), col. 14, code 0. If none of these cards present, then 0844 card, col. 14, code 0.
4. Exclude from the above cases:
 - a. Walk-ins: 0001 card, col. 76 code 1
 - b. Non-Core Cases: 0001 card, col. 80, code 7
 - c. Lost to Study: see below for classification

Classification of Women Lost to Study

		OB-34 OR OB-55			AND	PED-1 OR ADM-44	
1334-3334	AND	9334	1355-7355	AND	9355	1401-3401	0844
All Cards Not Present		Not Present	All Cards Not Present		Not Present	All Cards Not Present	Not Present
All Cards Not Present		Present	All Cards Not Present		Present	All Cards Not Present	Not Present
Any Card Present		N/A	Any Card Present		N/A	Only a 14012 Card Present or all Cards Not Present ^(a)	Not Present
All Cards Not Present		N/A	All Cards Not Present		N/A	Only a 14012 Card Present or all Cards Not Present ^(a)	Not Present

(a) And Absence of all cards 1402-4402, 1403-2403, 1408-5408, 0405, 1410-3410, 1101-3101, 1411-3411, 1412-4412, 1110-2110, 1120-4120, 1406-4406, 0407, 1414 (all)

Standard Cohort for PRB Studies — IID Cohort ID Minus Cases Lost to Study

1. Core Study Cases: 0001 card, col. 8 codes 1 or 2 and no code 7 in col. 80
2. Exclude from the above cases:
 - a. Walk-ins: 0001 card, col. 76 code 1
 - b. Non-Core Cases: 0001 card, col. 80, code 7
 - c. Lost to Study: see below for classification

Classification of Women Lost to Study

		OB-34 OR OB-55			AND	PED-1 OR ADM-44	
1334-3334	AND	9334	1355-7355	AND	9355	1401-3401	0844
All Cards Not Present		Not Present	All Cards Not Present		Not Present	All Cards Not Present	Not Present
All Cards Not Present		Present	All Cards Not Present		Present	All Cards Not Present	Not Present
Any Card Present		N/A	Any Card Present		N/A	Only a 14012 Card Present or all Cards Not Present(a)	Not Present
All Cards Not Present		N/A	All Cards Not Present		N/A	Only a 14012 Card Present or all Cards Not Present(a)	Not Present

(a) And Absence of all cards 1402-4402, 1403-2403, 1408-5408, 0405, 1410-3410, 1101-3101, 1411-3411, 1412-4412, 1110-2110, 1120-4120, 1406-4406, 0407, 1414 (all)

APPENDIX C

QUALITY CONTROL PROGRAMS

QUALITY CONTROL PROGRAMS

This appendix provides information regarding quality control programs developed for the psychology, speech, language and hearing, and seven year pediatric-neurological examinations. The programs are described briefly and results of the quality control assessments presented.

PSYCHOLOGY

An important consideration in testing is reliability or reproducibility of results. Reliability can be affected by the care with which an examination is administered. Concern with reliability led to the establishment of high standards for examiners in the NCPP: persons who were concerned not only with the quality of the assessment, but who also had the skills necessary to judge this quality, to use alternate items when an item was spoiled, and to take the time necessary with each subject to obtain an adequate examination (Broman et al. 1975). All examinations were edited by a second examiner, who checked for scoring and tabulation accuracy, and then edited a yet second time when they were received at the Perinatal Research Branch.

The most frequent means of assessing reliability in psychological testing is the test-retest method, where a test is readministered and a correlation between two sets of scores obtained. Quality control programs for the psychology examinations at ages four and seven consisted of test-retest examinations of a sample of the children. For the four year examinations, 144 children were retested for the various sections of the four-year psychology battery. The children were retested by an examiner from another institution after an interval of approximately three months. The retests were observed and independently scored by the original examiners. Reliability was determined by the correlations between the original test and the retest and also by correlations between the results obtained by the retest-examiner and the original examiner, who scored the retest independently as an observer.

For the Stanford-Binet Intelligence Scale, a test-retest correlation (reliability coefficient) of $r = 0.83$ was obtained. This reliability coefficient is an indicator of both temporal stability and examiner agreement. Concurrent interobserver reliability on the Stanford-Binet was very high, $r = 0.98$. Additional information on the Stanford-Binet test-retest correlations is given by Broman et al. (1975).

The test-retest correlation for the Graham-Ernhart Block Sort Test in the four year exam was 0.43. The retest-observer correlation for interobserver reliability was 0.98.

Percent of agreement for test-retest in the "pass category" for the three gross motor tests ranged from 91% to 95%. In the "fail category" for these tests the percent agreement was much lower (ranging from 15% to 41%) due to the fact that most of the children who fail these tests the first time pass the second. Corresponding figures for agreement between retester and observer are considerably higher, but low agreement in the "fail category" for one test, line walk, indicated some scoring problems for that test.

For the four fine motor tests, percent agreement for test-retest in the pass category ranged from 92% to 97%. However, percent agreement in the fail category ranged from 12% to 61%. Retest-observer agreement was very high for all the fine motor tests (above 90%) with the exception of bead stringing, where the agreement for the fail category was 67%.

For the overall behavior rating on the four year examination, agreement was extremely low between test and retest for the "suspect" category (6%) and moderately low between retest and observer for this category (31%). Most of the children rated as suspect in behavior the first time were rated as normal the second time. A similar pattern holds for the overall test impression ratings with agreement of 21% and 42%, respectively, for the suspect category.

Similar test-retest examinations were conducted for the seven year psychological test battery. Here a total of 418 subjects were retested. Test-retest correlations showed satisfactory reliability for all tests except one (Tactile Finger Recognition). Retest-observer correlations showed excellent agreement.

In the seven-year battery, test-retest correlations for the full scale WISC IQ was 0.86, the WISC verbal and performance IQ correlations being 0.83 and 0.74 respectively. Arranged in descending order, other test-retest correlations were 0.92 for the reading subtest of the Wide Range Achievement Test, 0.86 for the spelling subtest and 0.80 for the arithmetic subtest, 0.84 for the Auditory-Vocal Association test, 0.68 for the Koppitz error score on the Bender-Gestalt Test, 0.64 on the Goodenough-Harris Draw-A-Person Test standard score and 0.38 on the Tactile Finger Recognition Test.

For the overall behavior rating at seven years, none of the children sampled were rated as abnormal, though some were rated as suspect. Of those testing as normal, 94% remained in the normal category at retest, but only 29% of the children in the suspect category remained there. Agreement between retest and observer in the suspect behavior category was 50%.

For the overall test impression on the seven year psychology examination, 94% of the normals remained so classified at retest. Of the abnormal, 58% remained in the abnormal category. Of the suspects, 43% remained in that classification. Agreement was higher between retest and observer.

The data from the psychology Inter-Institutional Quality Control Trials indicate quite close agreement between observers when standard instruments are used for evaluation. For the overall behavior rating and overall test impression, however, more variation between observers was noted. Although agreement was high on the categorization of the normal children, there was considerable variation for children thought by the initial examiner to be in the "suspect" category.

SPEECH, LANGUAGE AND HEARING TESTS

The reliability of the speech, language and hearing examinations is reported by Lassman et al. (1980). For the three year speech, language and hearing examinations, three standardization and training workshops were held for the examiners, emphasizing uniformity in test administration and

objectivity in scoring. Instructions to the children had to be memorized. At the first workshop, the test manual was reviewed item by item and standard procedures were demonstrated and discussed, using three-year-old subjects. The second workshop focused on appropriateness of facilities and equipment, logistics and operational problems. "Adequacy of examination" was considered, along with review of scoring criteria for less than fully cooperative children. The third workshop was aimed at newer examiners, many of whom had not attended the previous workshops. Demonstrations of test administration were held and considerable attention was given to scoring standards. At all workshops, the necessity of maintaining standard procedures was emphasized.

Both intramural and extramural quality control procedures were instituted. An extramural quality control program was instituted that involved visits of representatives between institutions. The speech, language and hearing examination was administered by a visiting examiner and the results compared with those obtained earlier by the local examiner. Testing techniques were discussed and standard procedures were emphasized. Reliability was checked within institutions by retesting a randomly selected sample of 5% of the children within a month of the first examination. The results were so similar that this program was terminated after three months.

The eight-year examination was expected to be more reliable than the three-year examination because of the older age group and the use of standard measuring instruments. A standardization and training conference for all examiners and supervisors included an item-by-item review of the eight-year exam, covering the rationale, the specific intent of the test items, proper administration, and uniform test procedures and scoring. The test was demonstrated with eight-year-old subjects, and scoring criteria were evaluated by means of group scoring of the demonstrated tests.

Interexaminer and interinstitutional variability was examined by Dr. Paul LaBenz (PRB, NINCDS), who studied the results obtained by 12 examiners (six pairs) at six institutions. Thirty subjects per examiner were randomly selected from a pool of 1148 children. Analysis of variances tests were carried out on performance on auditory discrimination, vocabulary, reading and articulation tests. On all four subtests, the variance contributed by examiners was not significant, but significant differences did exist between institutions.

Quality control studies between institutions were conducted for the eight year speech, language and hearing examination in the same way as was described earlier for the three year exam. Children were retested by visiting examiners within a month or two of their initial testing by a local examiner. Three times each year a small sample of children was randomly selected and recalled for reexamination. Each child recalled was examined and scored by the visiting examiner, with the original examiner scoring the new test as an observer. All retesting and observing was done without reference to the original test.

A sample of 110 children with test-retest scores on the three-year examination was studied, as was a sample of 325 for the eight-year exam. The results were categorized into three sets of tables:

1. Test/Retest - comparison of the score of the original examiner with the score of the visiting examiner.
2. Retest/Observer - Comparison of the score of the visiting examiner with the second score of the original examiner.
3. Test/Observer - Comparison of the first score of the original examiner with his second score.

Examiner differences were considered to be a source of variability in the testing situation and in the scores. Designers of the study thought that with a large group of examiners, any differences would be averaged out in data analysis (Lassman et al., 1980); for this reason, they decided not to discard data from any of the examiners. The interested reader is referred to Lassman et al. (1980) for a discussion of the role of interinstitutional and interexaminer agreement in the selection of speech, language and hearing variables for analysis.

SEVEN YEAR PEDIATRIC-NEUROLOGICAL EXAMINATION

An interinstitutional quality control exchange program was established in 1966 to "measure and increase the degree of reproducibility of data" as collected by the collaborating institutions. Three quality control trial series were scheduled for each year; each institution was to participate in each trial series as either "host" or "visitor." The seven-year examinations were begun in 1966 and as they progressed, each host institution was asked to provide to the Perinatal Research Branch a list of all children who received a seven-year pediatric-neurologic examination (PED-76) during a two-month interval before the trial. Also, all children seen in the previous two months with abnormalities marked on the "impression" item of the examinations were listed. From these lists, Perinatal Research Branch personnel randomly selected 11 children (the number that could be examined and discussed in a two day session) to be retested.

The retesting procedure was as follows. Each of the 11 patients recalled for neurological examination was seen by a physician from the visiting institution. No medical information was provided in this setting. The original examiner was present as an observer at the re-examination; no communication was permitted between the two physicians. Following each retest using the PED-76, a discussion was held involving the original tester, the visitor, and a Perinatal Research Branch representative trained to assist in assessing and interpreting the differences occurring in examination methods and findings, and pointing out incorrect examination techniques. Each child was retained for possible re-examination while the PED-76 protocols were compared. Controversial clinical findings were retested, but protocols were not changed. Copies of all exam forms were collected, and the differences itemized on a special form for later tabulation.

On the sixth round (series) of trials, a similar program was set up for repeating the visual screening exams (usually done by nurses) and maternal interval histories (usually done by lay interviewers) on the same children recalled for the PED-76 re-examination. In the early set of rounds, the children recalled were picked randomly and included a large number of

"normals." As the program progressed, an effort was made to choose mostly neurologically suspicious or abnormal children, but care was taken to include a few normals each time.

The quality control trials were supervised by physicians from the Perinatal Research Branch, who collected the exams, listed the differences, and then classified each item difference on the exams according to a scheme. For each general examination item (PED-76 items 11 through 115):

- A. No difference existed, the examiners' report for this item was the same or within defined limits, or the difference was only because of an acute condition, which was expected to resolve itself.
- B. A difference was reported, but it was considered to be within the realm of acceptable variation and not statistically significant.
- C. No significant difference, yet the finding was indicated in the diagnostic impression.
- D. Differences were observed that were important for the item, but these did not lead to differences in the diagnostic impressions.
- E. Differences were observed in the item and these differences led to differences in the diagnostic impression.

Similarly, labels were attached to the differences in diagnostic impressions using the following scheme (PED-76 items 118-120):

- F. No difference existed.
- G. Findings and/or their interpretation have led to a minor difference within the impression.
- H. Findings and/or their interpretation have led to a major difference within the impression.

The eye re-examinations were checked for differences in visual acuity, proper referral (for further testing by an ophthalmologist), for significant deficiency in visual acuity, excessive far-sightedness, muscle balance, color vision, and proper referral for non-visual acuity abnormality. The findings were recorded and tabulated. Errors were pointed out to the visual screeners involved.

The interval history forms were compared for differences in illness and conditions reported and medical service referrals. The differences were tabulated. Errors were pointed out to the interviewers.

The results of each series of trials were tabulated item by item (after coding for clinical significance). A summary of these findings was sent to each project director, along with a report from the pediatric editing section

(which did a quarterly tabulation of all errors and omissions discovered on all PED-76 and PED-75 examinations processed during that quarter at the Perinatal Research Branch).

The following discussion is restricted to differences in classification on item 118, "Neurological Abnormalities." One thousand forty-five re-examinations were conducted. Of these, 563 retests were designated as falling into category F above, 424 as G, and 58 as H. This last group represents the number that were considered to represent major differences in neurological impression - 5.5%. A review of the 58 individual cases by the Pediatric Neurology Staff at the Perinatal Research Branch found that the disparate cases could be grouped into the following categories:

Ophthalmological Differences	19
Mental Dullness	12
Behavior Disturbance or Hyperactivity	4
Minor Neurological Differences	16
Major Neurological Differences	4
Missing Exam	1

The eye differences were explained by the fact that many ocular problems are either transient or very subtle in nature. The mental illness differences were non-specific in nature, and to a large extent were based on differences in populations of study children. The differences of behavior and minor neurological differences were not considered to be of major significance in this sort of test-retest situation. That left four cases out of 1045 that were considered to indicate a significant level of error in major neurological differences (0.38% of total sample). The data support the conclusion that few serious neurologically related events were missed by the interviewers.