
THE NATIONAL ARCHIVES OF THE UNITED STATES
LITTERA
SCRIPTA
MANET
★ 1934 ★

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

Person	Time	Subject
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/Genetic 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

Part	Title	Forms Included
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.

Data Item ID	From	To	Data Item Name
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.

Data Item ID	Card Num	From	To	Data Item Name
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....M-6		45	47	Cerebral palsy; diplegia, atonic (1 yr, interim, 7 yr): CHILD, GENERAL, DIAGNOSES AND CONDITIONS

[illegible] of NCPP Computerized Data Items. [illegible] data items in the NCPP data [illegible] so that a data item may appear [illegible] formatted trees enables a researcher to locate a [illegible] by a word typically used by the research community, [illegible] general categories of data items. After locating a data [illegible], the researcher is directed to either Volume II, VIII or IX, 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 (08-34). Cesarean section, indication
6140....VAR	abruptio placenta (08-55). Cesarean section, indication
942..08-34	abruptio placenta. Cesarean section, indication
1739..08-55	abruptio placenta. Uterine stimulant, augmentation indication
1730..08-55	abruptio placenta. Uterine stimulant, induction indication
6172....VAR	abruptio placenta. Uterine stimulant, induction indication
1883..08-55	abruptio placentae. Bleeding before cord clamped, cause
1834..08-55	abruptio placentae. Cesarean section, indication
6040....VAR	abruptio; placenta, (yes, no, unknown)
6039....VAR	abruptio; placenta, degree
1875..08-55	abruptio; placenta.

In this example, data items are found in the variable file and in the master file on forms 08-34 and 08-55. 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 08-34 and 08-55 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 FED-2 or FED-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 XXXXAYYZZ 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.
>
> XXXXXYYZZ7340002333000000

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).

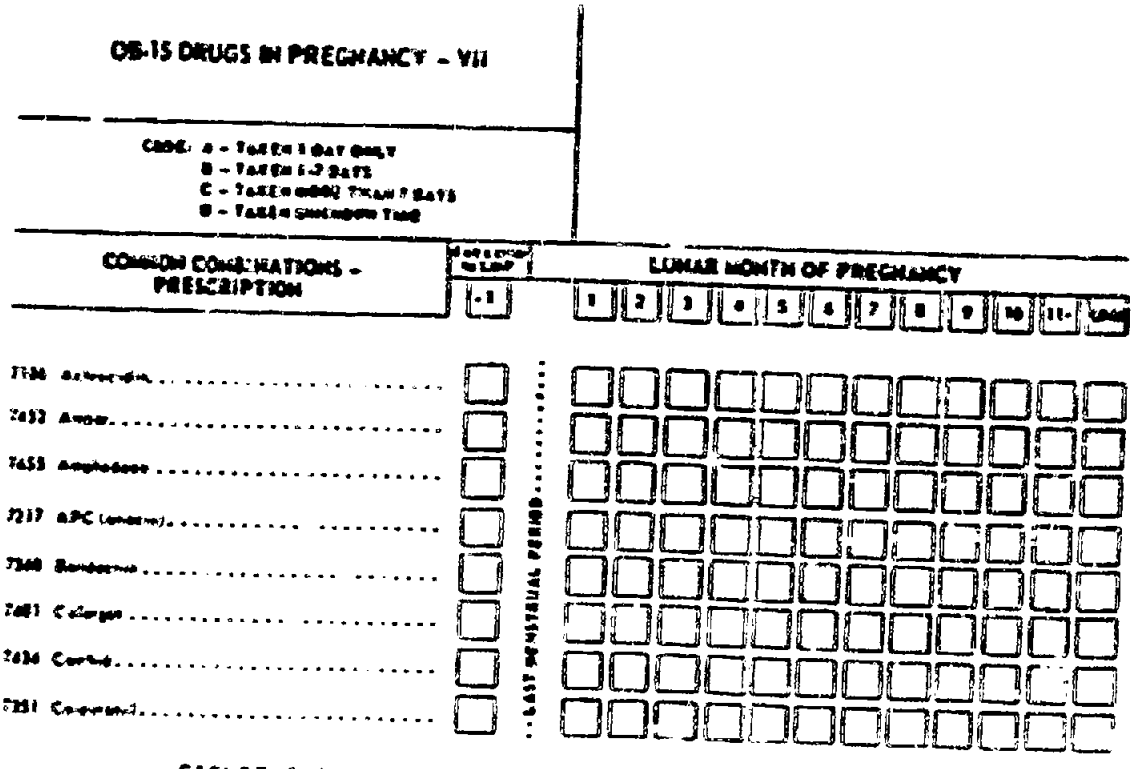


FIGURE 6.1. Drugs in Pregnancy 08-15 Example Form

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:

Person	Time	General Subject Area	Data Item ID	Data Item Name
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.

Data Item ID	From	To	Data Item Name
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:

Data Item ID	From	To	Data Item Name
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:

Data Item ID	Card Num	From	To	Data Item Name
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)

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	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: Frame Size	Annual Sampling: Method and Expected Cases	Special Studies
January 1959	All clinic patients who live within [illegible] politan [illegible] (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 Frame Size	Annual Sampling Method and Expected Cases	Special Studies
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	80%	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 58)

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	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 Frame Size	Annual Sampling Method and Expected Cases	Special Studies
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 gravida based on Month of birth; Feb., April, August, Oct., Dec. - 417-	None
January 1962	(d)	As above and Walk-ins	None	1000	54% of gravida 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 gravida 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 (NPIDB Institution Number 71)

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 clinic patients admitted	None	None	1188	(a)	None
May 1960	All clinic patients admitted	Walk-ins	None	1188	(a)	None
July 1960	All clinic patients admitted	Walk-ins	None	1188	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)

Date Initiated	Sampling Frame Criteria	Exclusions From Sampling Frame	Special Exclusions From Sample	Annual Sampling		Special Studies
				Frame Size	Method and Expected Cases	
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: 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 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 PBB Studies — BA **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 — SR Cohort SR 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 PBB Studies — BC Cohort BC 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	
1334-3134	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 — 80 Cohort 80 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 ADA4-44	
1334-3334	AND	9334	1355-7355	AND 9355		1461-3401	0144
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 abnormals, 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 proceoures 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.

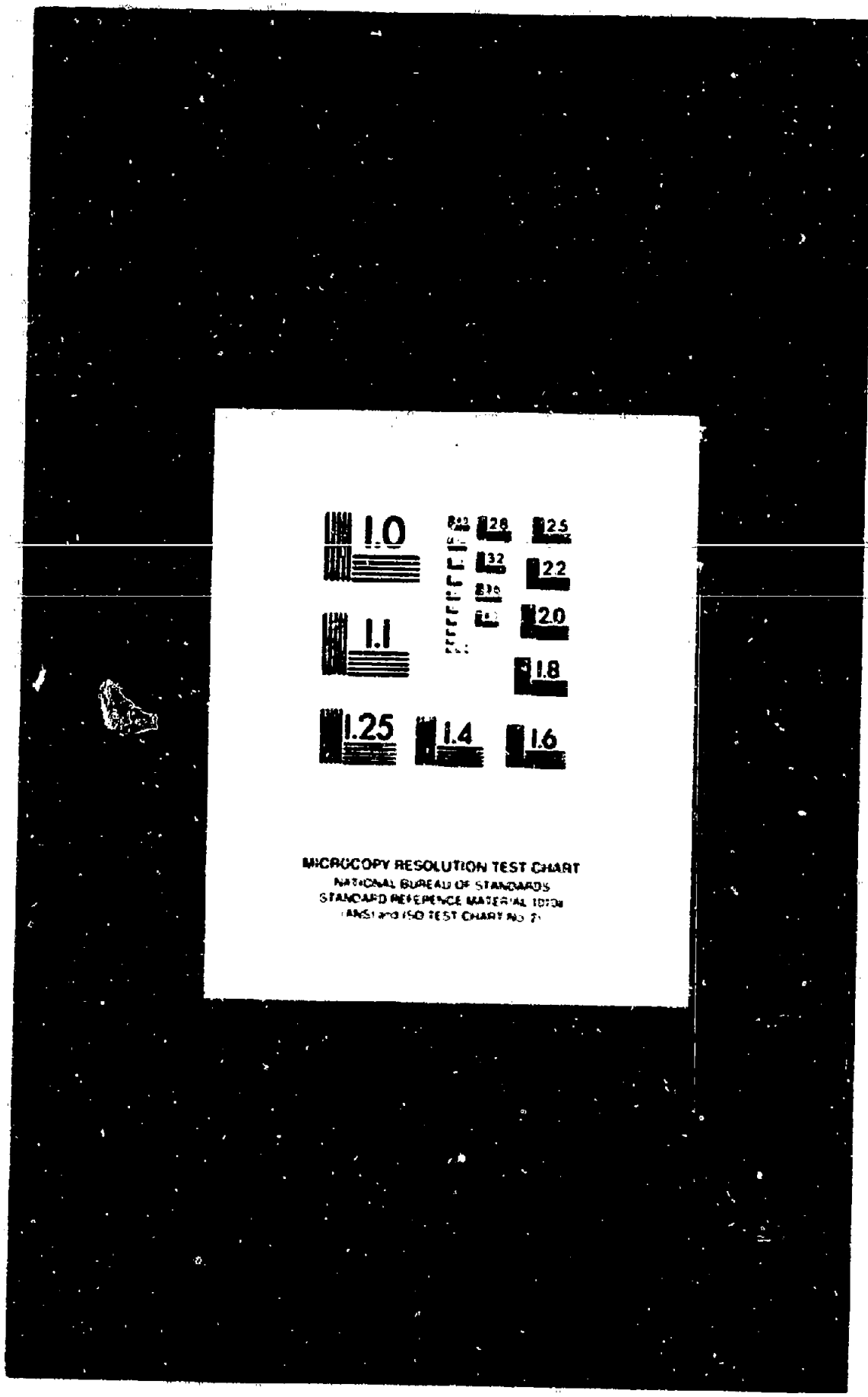
The results of each series of trials were tabulated item by item (after coding for clinical significance). A summary of these findings was sent to each project director, along with a report from the pediatric editing section

(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.


1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
2.5
2.2
2.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
ANSI and ISO TEST CHART No. 2

CONTINUED ON NEXT FICHE