National Children's Study

Teaching Database User Manual

Version 3.0

April 2019

NICHD
NCS Program Office

1. Study Background

The National Children's Study (NCS) Vanguard Study was a pilot for the planned Main Study cohort and was started in 2009. The study aimed to evaluate environmental influences on child health and development. The Vanguard Study tested different recruitment strategies in 43 counties across 31 states throughout the U.S.

The study enrolled over 5,600 birth families and followed them through 2014. NCS collected prenatal, perinatal, and early childhood data up to age 5. At protocol-specified study visits, questionnaires and interviews, neuro-psychosocial and cognitive assessments, and physical examination data were collected, along with nearly 19,000 biological and 5,500 environmental primary samples from which a sample repository of over 250,000 items was created.

2. Teaching Database

In addition to the data available for researchers, the NCS Archive provides the NCS Teaching Database comprised of a base child-level data set and several topic-related modules. This data is derived from files available within the NCS Archive and is based on real data collected during the NCS Vanguard Study.

The purpose of the NCS Teaching Database is to provide researchers a ready-made database system that can be used for various statistical models, including repeated measures analyses. It is freely available as a resource to supplement classroom examples and student academic projects. Those interested in learning more about the source data and study data collection procedures are encouraged to visit the NCS Archive website (https://ncsarchive.s-3.net) and register for full site access.

Disclaimer: The NCS Teaching Database is intended for teaching and educational purposes only. The content within the NCS base dataset and content modules is unsuitable for publication purposes. Specific programming techniques were used to make these data public downloadable files.

3. Child Dataset

Filename:	NCS_Teaching_Child_V1_1
Available Formats:	.sas7bdat, .csv

The Child dataset contains 5,604 records, each record representing a child with a unique child identifier CHILD_PIDX. The CHILD_PIDX then may be used to link to the modules described below. Each record also contains the mother's identifier MOM_PIDX that is not unique since siblings and children of multiple births exist on the Child dataset. The data topics included at child-level are birth measurements, demographics, and mother's socio-economic characteristics and demographics. Community-level characteristics are included based on the American Community Survey (ACS). To protect confidentiality, only the birth quarter and year are provided. Similarly, geographic characteristics only include the census region (midwest, northeast, south, west) and an indicator whether the participant's address is within a metropolitan area.

Several ACS characteristics have been merged to the child dataset to capture community information on citizenship, foreign-born population, income, household characteristics, educational attainment, and

race-ethnicity. To avoid disclosure of the participants' communities all ACS data have been converted to deciles (< 10 percent, 10 to < 20 percent, etc.). The codebook formats include the corresponding scores associated with each decile range.

3.1 Dataset Variables

Variable	Type	Length	Label
CHILD_PIDX	Char	9	Participant ID of the Child (Encrypted)
MOM_PIDX	Char	9	Participant ID of the Mother (Encrypted)
PSU_IDX	Char	4	Primary Sampling Unit ID (Encrypted)
RECRUITTYPE	Num	8	Mother's Recruitment Strategy
CHILD_SEX	Num	8	Child's sex
CHILD_DOB_YEAR	Num	8	Child's year of birth
CHILD_DOB_QTR	Num	8	Child's quarter within the year of birth
CHILD_RACE	Num	8	Child's race (5 categories)
CHILD_ETHNICITY	Num	8	Child's ethnicity (Hispanic or not Hispanic)
GESTATIONAL_AGE	Num	8	Child's gestational age at birth (5 categories)
BABY_WEIGHT	Num	8	Baby's birth weight (9 categories)
MULTIPLE	Num	8	Was Child part of a birth with multiple children (e.g., twins)?
SIBLINGS	Num	8	Did Child have an enrolled NCS sibling from a separate birth?
WITHDREW	Num	8	Child was withdrawn from the study
MOM_AGE	Num	8	Mother's age at Child's birth
MOM_RACE	Num	8	Race of Child's mother (5 categories)
MOM_ETHNICITY	Num	8	Ethnicity of Child's mother
MOM_MARISTAT	Num	8	Mother's marital status collected at the pregnancy visit (4 categories)
MOM_EDUCATION	Num	8	Mother's education level collected at the pregnancy visit (6 categories)
MOM_INSURANCE	Num	8	Mother's health insurance coverage collected at the pregnancy visit (5 categories)
SURVEY_LANG	Num	8	Interview language used during the birth visit
HOUSEHOLD_INCOME	Num	8	Household income collected at the pregnancy visit (4 categories)
CENSUS_REGION	Char	10	Census Region
MSA	Num	8	Address within Metropolitan Statistical Area (MSA)
CITIZENSHIP_PERC	Num	8	2009-2013 ACS 5-year estimate of non-citizen residents (percentile)
FOREIGN_BORN_PERC	Num	8	2009-2013 ACS 5-year estimate of foreign-born population (percentile)
HH_TYPE_PERC	Num	8	2009-2013 ACS 5-year estimate of population under 18 years in a non-family household (percentile)
INCOME_PERC	Num	8	2009-2013 ACS 5-year estimate of median household income (percentile)
POVERTY_PERC	Num	8	2009-2013 ACS 5-year estimate of poverty status: population below 100% federal poverty limit (percentile)
TENURE_PERC	Num	8	2009-2013 ACS 5-year estimate of owner-occupied households (percentile)
LING_ISO_PERC	Num	8	2009-2013 ACS 5-year estimate of linguistically-isolated households (percentile)
WNH_PERC	Num	8	2009-2013 ACS 5-year estimate of white non-Hispanic households (percentile)
AFAMNH_PERC	Num	8	2009-2013 ACS 5-year estimate of African-American non-Hispanic households (percentile)
ASNH_PERC	Num	8	2009-2013 ACS 5-year estimate of Asian non-Hispanic households (percentile)
HISP_PERC	Num	8	2009-2013 ACS 5-year estimate of Hispanic households (percentile)
OTHERNH_PERC	Num	8	2009-2013 ACS 5-year estimate of Other non-Hispanic households (percentile)
LESSTHANHS_PERC	Num	8	2009-2013 ACS 5-year estimate of population with less than high school education (percentile)
HSPLUS_PERC	Num	8	2009-2013 ACS 5-year estimate of high school or higher education (percentile)

4. Modules

Each module contains an encrypted participant identifier (**CHILD_PIDX**) that can be used to link to the base dataset. In addition, age at the time of the interview (**CHILD_AGE** or **CHILD_ADJ_AGE**) and interview visit (**VISIT**) can be found in each module. Not all variables were collected for every child participant enrolled in NCS, and some were collected as "don't know" and "refused". All missing values, including the "don't know" and "refused", are indicated by ".M".

To link the Child dataset to a teaching Module, please refer to the following SAS code provided below:

```
PROC SORT DATA=[CHILD DATASET]; BY CHILD_PIDX; RUN;
PROC SORT DATA=[MODULE DATASET]; BY CHILD_PIDX; RUN;

DATA [NEW];
MERGE CHILD (IN=A) MODULE (IN=B);
BY CHILD_PIDX;
IF A AND B;
RUN;

-OR-

PROC SQL;
CREATE TABLE [NEW] AS
SELECT A.CHILD_PIDX, A.[CHILD VARIABLES OF INTEREST],B.[MODULE
VARIABLES OF INTEREST]
FROM
(SELECT * FROM [CHILD DATASET])AS A
INNER JOIN
(SELECT * FROM [MODULE DATASET]) AS B
ON A.CHILD_PIDX=B.CHILD_PIDX;
QUIT;
```

4.1   Anthropometric   Measures

Filename:	NCS_Teaching_Anthro_V1
Available Formats:	.sas7bdat, .csv

The Anthropometry module contains a battery of physical measures collected at the 6-month, 12-month, 24-month, and 36-month interviews. The NCS used standardized procedures and equipment to collect precise body measurement data. This file contains 11 variables and 2,795 observations from 2,204 child participants. Children in this file can have up to four observations across the four study visits (approximately 35% of children have two or more records). To account for children born prematurely, an adjusted age variable is provided (**CHILD_ADJ_AGE**) on the file. Child's weight in kilograms (kg) (**WEIGHT**); length of the upper arm (humerus) in centimeters (cm) (**UPPER_ARM_LENGTH**); and mid upper arm (**UPPER_ARM_CIRC**), waist (**WAIST_CIRC**), and head (**HEAD_CIRC**) circumference (cm) were measured at all four study visits. Recumbent length (cm) (**LENGTH**) was measured from the 6-month interviews to the 24-month interviews. Standing height (cm) (**HEIGHT**) was measured at the 24-month and 36-month interviews. Per recommendation by the CDC, Body Mass Index (kg/m^2) (**BMI**) was calculated only for children 24 months of age or older. **BMI** was calculated by dividing the weight measurement by recumbent length or standing height (converted to meters) squared.

4.1.1 Dataset Variables

Variable	Type	Length	Label
CHILD_PIDX	Char	9	Participant ID of the Child (Encrypted)
CHILD_ADJ_AGE	Num	8	Childs age at time of interview (Months) – adjusted for prematurity
VISIT	Num	8	Study Interview
WEIGHT	Num	8	Weight (kg)
LENGTH	Num	8	Recumbent length (cm) (not measured at 36-month)
HEIGHT	Num	8	Standing height (cm) (not measured at 6-month and 12-month)
BMI	Num	8	Body Mass Index (BMI) (weight(kg)/height(m)^2) (not calculated at 6-month and 12- month)
UPPER_ARM_LENGTH	Num	8	Upper arm (humerus) length (cm)
UPPER_ARM_CIRC	Num	8	Mid upper arm circumference (cm)
WAIST_CIRC	Num	8	Waist circumference (cm)
HEAD_CIRC	Num	8	Head circumference (cm)

4.2 Sleep Routine

Filename:	**NCS_Teaching_Sleep_V1**
Available Formats:	**.sas7bdat, .csv**

The Sleep Routine module contains information concerning the child's sleeping habits. Sleeping habits were collected at the 3-month, 6-month, 18-month, 24-month, 30-month, 36-month, and 42-month visit interviews. This file contains 6 variables and 18,133 observations from 5,301 participants. Children in this file may have up to seven observations across the seven study visits (approximately 91% of participants have two or more records). Data was collected on how much the child sleeps at night (**SLEEP_HRS_NIGHT**), how often the child is difficult when put to bed (**SLEEP_DIFFICULT**), and how often the child wakes up during the night (**SLEEP_THROUGH**). The variable **SLEEP_THROUGH** was not fielded at the 3-month visit.

4.2.1 Dataset Variables

Variable	Type	Length	Label
CHILD_PIDX	Char	9	Participant ID of the Child (Encrypted)
CHILD_ADJ_AGE	Num	8	Childs age at time of interview (Months) – adjusted for prematurity
VISIT	Num	8	Study Interview: 3=3-month visit, 6=6-month visit,…42=42-month visit
SLEEP_HRS_NIGHT	Num	8	How many hours does the child sleep at night?
SLEEP_DIFFICULT	Num	8	How often is the child difficult when he/she is put to bed? (5 categories)
SLEEP_THROUGH	Num	8	How often does the child wake at night? (not asked at 3-month) (5 categories)

4.3 Infant Feeding

Filename:	**NCS_Teaching_Feeding_V1**
Available Formats:	**.sas7bdat, .csv**

The Infant Feeding module contains information regarding feeding habits of the child. Infant feeding habits were collected at the 6-month and 12-month visit interviews. This file contains 9 variables and 6,115 observations from 3,918 participants. Children in this file may have two observations across the two study visits (approximately 56% of participants have two records). Information collected during both the 6-month and 12-month interviews include age when the baby stopped consuming breast milk

(**BREAST_STOP**), how often baby is fed formula (**FORMULA**), and age when baby was first fed formula on a regular basis (**FORMULA_AGE**). At the 6-month visit, information regarding if the baby was ever breastfed (**BREAST_FEED**) and if the baby is currently being breastfed (**BREAST_FEED_NOW**) was collected. At the 12-month visit, participants were asked how often the baby is fed breast milk (**BREAST_MILK**).

4.3.1 Dataset Variables

Variable	Type	Length	Label
CHILD_PIDX	Char	9	Participant ID of the Child (Encrypted)
CHILD_ADJ_AGE	Num	8	Childs age at time of interview (Months) – adjusted for prematurity
VISIT	Num	8	Study Interview: 3=3-month visit, 6=6-month visit,…42=42-month visit
BREAST_FEED	Num	8	Was baby ever breastfed? (not asked at 12-Month) 1=Yes, 2=No
BREAST_FEED_NOW	Num	8	Is baby currently being breastfed? (not asked at 12-Month) 1=Yes, 2=No
BREAST_MILK	Num	8	In the past 7 days, how often was baby fed breast milk? (total count) (not asked at 6-month)
BREAST_STOP	Num	8	How old was baby when participant completely stopped feeding baby breast milk? (in weeks)
FORMULA	Num	8	In the past 7 days, how often was baby fed formula? (total count)
FORMULA_AGE	Num	8	How old was baby when first fed formula on a daily basis? (9 categories)

4.4 Child Health

Filename:	**NCS_Teaching_ChildHealth_V1**
Available Formats:	**.sas7bdat, .csv**

The Child Health module contains information regarding frequently occurring child health conditions. Data for this module were collected at the 6-month, 12-month, 18-month, and 24-month visit interviews. This file contains 11 variables and 14,445 observations from 5,189 unique children. Children in this file may have up to four observations across the four study visits. Approximately 87% of the children have data from at least two interviews; approximately 33% have data from all four visits. The variable **VISIT** defines the visit interview. Across all four visits, variables collected are **CHILD_AGE** (child's age in months at time of interview), **VISIT_WT** (child's weight in pounds at most recent well-child visit or checkup), **CHILD_HEALTH** (caregiver's characterization of child's health), **RESPIRATORY** (in the last three months – since birth for 6-month visit – has child had wheezing or whistling in the chest?), and **FEVER** (in the last three months – since birth for 6- month visit – how many days has child had a fever over 101 degrees, not relating to receiving i m m u n i z a t i o n s ?).

At the 6-month and 12-month visits, caregivers were asked in the last three months (since birth for the 6-month visit) has child had diarrhea or vomiting (**GASTRO**) and in the last three months (since birth for 6-month visit) has child had an ear infection (**EAR_INFECTION**). At the 18-month and 24-month visits, caregivers were asked in the last three months, has child had frequent or repeated diarrhea (**DIARRHEA**) and in the last three months, has child had three or more ear infections (**EAR_INFECTION_FREQ**).

4.4.1 Dataset Variables

Variable	Type	Length	Label
CHILD_PIDX	Char	9	Participant ID of the Child (Encrypted)
CHILD_AGE	Num	8	Childs age at time of interview (Months)

VISIT	Num	8	Study Interview
VISIT_WT	Num	8	What was child's weight at most recent well-child visit or checkup (lbs)?
CHILD_HEALTH	Num	8	How would you categorize child's health?
GASTRO	Num	8	In the last 3 months (since birth for 6 month visit), has child had diarrhea or vomiting?
DIARRHEA	Num	8	In the last 3 months, has child had frequent or repeated diarrhea? (asked at 18 and 24 month visits)
EAR_INFECTION	Num	8	In the last 3 months (since birth for 6 month visit), has child had an ear infection?
EAR_INFECTION_FREQ	Num	8	In the last 3 months, has child had 3 or more ear infections? (asked at 18 and 24 month visits))
RESPIRATORY	Num	8	In the last 3 months (since birth for 6 month visit), has child had wheezing or whistling in the chest?
FEVER	Num	8	In the last 3 months (since birth for 6 month visit), how many days has child had a fever over 101 degrees, not related to receiving immunizations?

4.5 Mother's Pregnancy Health

Filename:	NCS_Teaching_MomPregHealth_V1
Available Formats:	.sas7bdat, .csv

The Mother's Pregnancy Health module contains information regarding frequently monitored health conditions prior to and during pregnancy. Data for this module were aggregated across multiple pregnancy visit interviews. If a mother reported a condition on any of the pregnancy visit interviews, then the flag for the condition was assigned a "Yes". This file contains 19 variables for 4,314 mothers, each mother having only one record. When linking to a child file, a mother may link to multiple children in the case of twins or multiple pregnancies.

This file contains the mother's self-reported health status (**HEALTH**) and body mass index (BMI) based on self-reported height and a pre-pregnancy weight. BMI is stored on the file as both a continuous variable (**BMI**) and a categorical variable (**BMI_CAT**) having six categories. If a woman has ever been diagnosed with high blood pressure (**HIGHBP_NOTPREG**) when not pregnant, hypothyroidism (**THYROID**), and asthma (**ASTHMA**) then these are reported on the file. The conditions reported during the pregnancy include diabetes (**DIABETES**), high blood pressure (**HIGHBP_PREG**), preeclampsia (**PREECLAMPSIA**), early labor or premature labor (**EARLY_LABOR**), anemia or low blood count (**ANEMIA**), bladder or kidney infection (**KIDNEY**), severe nausea or vomiting known as hyperemesis (**NAUSEA**), Rh disease or isoimmunization (**RH_DISEASE**), protein in the urine (**URINE**), infection of the vagina or bacterial vaginosis (**VAGINOSIS**), and infection with bacteria called Group B strep (**GROUP_B**). Additionally the current smoking status of the mother is included (**CIG_NOW**).

4.5.1 Dataset Variables

Variable	Type	Length	Label
MOM_PIDX	Char	9	Participant ID of the Mother (Encrypted)
HEALTH	Num	8	Participant's personal description of current general health status
BMI	Num	8	Participant's Pre-Pregnancy Body Mass Index (BMI)

BMI_CAT	Num	8	Participant's Pre-Pregnancy Body Mass Index (BMI) – Categorical variable
THYROID	Num	8	Has Participant ever been told by a doctor or other health care provider that they had hypothyroidism (under active thyroid)?
HIGHBP_NOTPREG	Num	8	Has Participant ever been told diagnosed with hypertension or high blood pressure when Participant was not pregnant?
ASTHMA	Num	8	Has Participant ever been told by a doctor or other health care provider that they had asthma?
DIABETES	Num	8	Has Participant ever been told by a doctor or other health care provider that they had Diabetes?
HIGHBP_PREG	Num	8	At any time during pregnancy, has a doctor or other ealth are provider told Participant they have High Blood Pressure?
PREECLAMPSIA	Num	8	At any time during pregnancy, has a doctor or provider other health care told Participant that they have preeclampsia?
EARLY_LABOR	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have early or prematu re labor?
ANEMIA	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have anemia or low blood count?
KIDNEY	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have a bladder or kidney infection?
NAUSEA	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have severe nausea or vomiting (hyperemesis)?
RH_DISEASE	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have Rh disease or isoimmunization?
URINE	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have protein in their urine?
VAGINOSIS	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have an infection of the vagina with bacteria (bacterial vaginosis)?
GROUP_B	Num	8	At any time during pregnancy, has a doctor or other health care provider told Participant they have an infection with bacteria called Group B strep?
CIG_NOW	Num	8	Does Participant currently smoke cigarettes?

4.6 Developmental Milestones

Filename:	**NCS_Teaching_Development_V1**
Available Formats:	**.sas7bdat, .csv**

The Developmental Milestones module contains information regarding important child developmental achievements in the first year of life. Data for this module were collected at the 3-month, 6-month, 9-month, and 12-month visit interviews. This file contains 23 variables and 10,915 observations from 4,130 unique children. Children in this file may have up to four observations across the four study visits; thus, this is a child-visit level data set. Approximately 79% of the children have data from at least two interviews; approximately 29% have data from all four visits. The variable (**VISIT**) defines the visit interview. The variable (**CHILD_AGE**) is the child's age in months at the time of the interview. Outliers with respect to age were excluded: only age values within one month (+/-) of the visit age were included – this represents the great majority of all cases. All other variables are Yes/No answers to developmental milestones such as smiling, waving goodbye, holding head up, rolling over, standing and walking.

The majority of the variables were asked at all four visits. A few variables were asked at only select visits; e.g., (**STAND**), (**STAND_ALONE**), and (**WALK**). These exceptions are noted in the variables' labels.

4.6.1 Dataset Variables

Variable	Type	Length	Label
CHILD_PIDX	Char	9	Participant ID of the Child (Encrypted)
VISIT	Num	8	Study Interview
CHILD_AGE	Num	8	Child's age at time of interview (Months)
EYES_FOLLOW	Num	8	Does child follow you with his/her eyes?
SMILE	Num	8	Does child smile at you when you smile at him/her?
REACH_TOY	Num	8	Does child try to get a toy that is out of reach?
FEED	Num	8	Does child feed him/herself a cracker or cereal?
WAVE	Num	8	Does child wave goodbye?
REACH_FOOD	Num	8	Does child reach for toys or food held to him/her? (not asked at 9 mo.)
SWITCH_HANDS	Num	8	Does child move a toy or block from one hand to the other?
PICKUP	Num	8	Does child pick up a small object like a Cheerio or a raisin?
HOLD	Num	8	Does child hold two toys or blocks at a time, one in each hand?
TURN_TO_SOUND	Num	8	Does child turn towards a sound? (not asked at 9 mo.)
TURN_TO_SPEAKER	Num	8	Does child turn towards someone when they're speaking?
MAKE_SOUNDS	Num	8	Does child make sounds as though he/she is trying to speak?
SAY_MAMA_DADA	Num	8	Does child say mama or dada?
HEADUP	Num	8	Does child keep head steady when sitting or held up?
ROLL_TO_BACK	Num	8	Does child roll over from stomach to back? (not asked at 9 mo.)
ROLL_TO_STOMACH	Num	8	Does child roll from back to stomach?
SITUP	Num	8	Does child sit up by him/herself? (not asked at 3 mo.)
STAND	Num	8	Does child stand while holding onto something? (not asked at 3 mo.)
STAND_ALONE	Num	8	Does child stand alone without holding onto something? (not asked at 3 and 6 mo.)
WALK	Num	8	Does child walk by him/herself without holding onto something? (not asked at 3 and 6 mo.)