

Fragile Families and Child Health - Medical Record Data For FF Mothers and Focal Children

November 2011

The `ffch_a3_v2withIH52011_Y9.dta` file contains medical records data (from the birth hospitalization) for Fragile Families (FF) mothers and focal children, merged with data files from all FF core and in-home studies (including fathers' surveys) that are available to date.

File overview

The file contains observations for all 4,898 births in the FF study. There are medical record data for 3,684 of those births (mothers and focal children). It is recommended that users change the maximum number of variables to 10,000. This can be done using the 'set maxvar' command in Stata. In the file, the medical record data appear first, followed by FF baseline data (mothers, fathers), FF 1-year data (mothers, fathers), FF 3-year data (mother core, father core, in-home), and then FF 5-year data (mother core, father core, in-home). The prefixes "ih3" and "ih5" for in-home three-year and five-year waves were added to the variable names for the in-home survey. Data from the Year 9 wave, including core mother and father, child, and teacher surveys, as well as the home visit activities, are also now included on this file.

A variable named `ffmrflag` equals 1 if there are medical record data for a given case and equals 0 if there are no medical record data for that case. When there are no medical record data for a given case, it was for 1 of 3 reasons: (1) The hospital did not permit us to abstract records or there were too few cases for it to be financially feasible to collect data at that hospital (38%), (2) the mother refused consent (33%), or (3) the records could not be located in the hospital (29%).

The family id # `idnum` was replaced with a random id # named `ffchid`. The data are sorted by `ffchid`. The city of interview variable `mlcity` was replaced with a random variable `ffcity`, which ranges between 121 and 140 (the cities are labeled `city A` through `city T`). There is no relationship between the `mlcity` and `ffcity` variables. The file also includes dummy variables (`hosp1-hosp61`) for each of the 61 hospitals in which medical record data were collected. There is no relationship between the hospital id numbers in the FF baseline survey and the hospital dummy variables in this file. Random noise has been introduced into the data to protect confidentiality. The noise should have no impact on statistical analyses.

Within the medical records data, constructed variables appear first. Those are followed by raw data from the mother's and infant's medical records. A list of the constructed variables appears below. The instrument (abstraction form) that was used to collect the medical record data is attached to this memo for reference. Some items on the abstraction form, including names, dates, open-ended responses, notes, and ICD 9 billing codes, are not included on this data file and are marked accordingly on the attached data collection instrument.

Variable naming convention

Variable names are a maximum of 8 characters long. The first two characters are the "ch" prefix, indicating that the variable is from the Fragile Families and Child Health study. The last six characters are either the variable name as shown on the data abstraction form, or an abbreviated description of the constructed variable.

Examples:

chc4 is question C4 on the data abstraction form (does the mother have a history of unwanted pregnancy?)
chdrugp is a constructed variable indicating whether or not the mother used illicit drugs during pregnancy

Constructed variables

Several constructed variables are included on the file. These variables incorporate information from a variety of sources, including dates, open-ended responses, and ICD 9 billing codes which are not included in this data file. The following constructed variables are included:

Health care

chbstay Number of days infant stayed in the hospital (birth hospitalization)
chmstay Number of days mother stayed in the hospital post-partum
chtripnc Trimester prenatal care began

Mother's reproductive history

chlstbth # of months since mother's last live birth
chdeliv Method of delivery - vaginal or c-section
chprevpr # of previous pregnancies mother had (EXCLUDES focal child)
chprevd # of mother's previous term and preterm deliveries (> 20 weeks gest. age)
chstillb # of mother's previous deliveries that were stillbirths
chnolive # of mother's previous pregnancies that ended as abortions, miscarriages, or stillbirths
chprvabr Any previous induced/spontaneous/therapeutic abortions/miscarriages/stillbirths
chnumabr # of previous induced/spontaneous/therapeutic abortions/miscarriages/stillbirths

Anthropometric measures

chmhtcm Mother's height (cm)
chppwtkg Mother's pre-pregnancy weight (kg)
chadwtkg Mother's admission weight (kg)
chwtgnkg Mother's weight gain during pregnancy (kg)
chadqwtg Constructed – adequate weight gain during pregnancy
chobese Diagnosed as obese (pre-pregnancy) in chart and/or by pre-pregnancy Body Mass Index (BMI)
chwgtcat 5-level pre-pregnancy weight variable (underweight, normal weight, overweight, obese, morbidly obese)
chbmi Mother's pre-pregnancy BMI
chbthwtg Constructed - baby's birthweight (gm)
chblen Baby's length (cm)
chbhead Baby's head circumference (cm)
chhighbw Birthweight > 4000 gm
chlowbw Birthweight < 2500 gm (low birthweight, or LBW)
chpondex Baby's ponderal index = w/h³
chpondcat Constructed – Child's ponderal index (low, medium, high)
chgstage Constructed - Gestational age - using clinical assessment when available, pediatric assessment otherwise
chpterm Constructed - Baby was preterm (< 37 wks)
chvpterm Constructed - Baby was very preterm (< 32 wks)

Mental health and drug use

- chdrugp Constructed – mother used drugs during pregnancy, from all possible sources in the medical record
- chsmkp Constructed – mother smoked cigarettes during pregnancy, from all possible sources in the medical record
- chalcp Constructed – mother used alcohol during pregnancy, from all possible sources in the medical record
- chmhprob Constructed – mother had pre-pregnancy diagnosis of mental illness, from all possible sources in medical record

Abnormal infant health conditions (variables chinfc01 through chinfc16)

The coding of infants' abnormal health conditions was conducted by an outside pediatric consultant who was directed to glean information from all relevant sections of the medical record abstraction form as well as one-year maternal reports of child disability to identify infant health conditions and to assign to all conditions a number between 1 and 16 according to the grid shown below. The rows in the grid indicate how likely it is, based on current medical knowledge, that the mother's prenatal behavior caused the condition. The columns indicate the severity of the condition (i.e., how disabling and long-term the condition is thought to be). Many infants had more than one condition, and are thus assigned multiple codes.

Coding Grid for Infant Health Conditions

	Severity			
	High	Medium	Low	Unknown
Not Behavior Related	1	2	3	4
Possibly Behavior Related	5	6	7	8
Likely Behavior Related	9	10	11	12
Not Enough Information To Determine if Behavior Related	13	14	15	16

Missing values

Some mothers and children were missing information for items in the data abstraction form because the information could not be found in the charts. Most missing values of this type are represented with Stata's internal missing code, but some variables use 999 (unknown) or 997 (not applicable). For variables that specify the # of times a condition occurred (e.g., che5a - # of previous LBW, preterm, or small for gestational age infants) a -8 (N/A) value is used when it is known from a previous variable that the mother never had the condition.

Additional information

- A few variables were dropped from the core and in-home data files and replaced with constructed variables. For example, variables cm4b_age and cf4b_age were added to represent the child's age at time of FF 5-year interview and interview dates were dropped.
- Some of the items on the medical records abstraction form are not available to public users. Those items are marked with ** on the attached form. For example, since it was not always clear from the medical records whether an abortion was induced or spontaneous, che1 and che6 were dropped from the file. Those two variables was combined into a single constructed variable chnumabr.
- The variable chh5 should be used for the child's gender. All other child gender variables were dropped from this data set.
- For the focal child's birthweight, the only variables included in this data set are chbthwtg (constructed from the medical records) and cm1bthwtg (constructed from the mother's baseline survey). Both are measured in grams.

A special note about gravidity and parity

Gravidity and parity are characterized by the variables chd4g, chd4p, chd5t, chd5p, chd5l, and constructed variables chprevpr, chndeliv, chdeliv, and chstill.

Gravidity refers to the number of times a woman has been pregnant, and *parity* refers to the number of pregnancies that led to a birth beyond 20 weeks gestational age.

A more specific designation of pregnancy outcomes divides them into term and preterm deliveries, abortuses, and living children. This is known as the TPAL (**T**erm-**P**reterm-**A**abortus-**L**iving) designation. Abortuses include spontaneous, induced, and therapeutic abortions. A multiple gestation counts as one pregnancy and one delivery, but may increase the number of living children by more than one.

Example: A (hypothetical) woman in the Fragile Families study has just given birth. In her medical record, the gravidity, parity and TPAL history are recorded prior to the birth of the FF focal child. Thus, the FF focal child counts as a pregnancy, but not as a term or preterm delivery; it is a pregnancy without outcome at the time the reproductive history was recorded in the chart. Supposing this hypothetical FF mother has previously given birth to one set of preterm twins and one term infant, and has had two miscarriages. Her reproductive history would be recorded as follows:

Gravidity	5 (includes the FF focal child, so gravidity = T + P + A + 1)
Parity	2 (equals T + P)
T	1
P	1
A	2
L	3 (includes the twin)