

2022-2023 National Survey of Family Growth Female Pregnancy File Recode Specifications

** Recodes shown in blue font and noted with a double asterisk were defined based on final, imputed values of other recodes and have no imputation flags. Therefore, their variable type in the codebooks and file indexes is “computed in post-processing.”

Section B Recodes: Pregnancy and Birth History

OUTCOME: Pregnancy outcome

This recode assigns a single outcome code to each pregnancy, even if the pregnancy had multiple outcomes. If there were multiple outcomes for the pregnancy (e.g., 2 fetuses, with 1 born alive but other being a stillbirth), OUTCOME gives the following priority:

*live birth
induced abortion
stillbirth
miscarriage
ectopic
pregnancy*

The raw variable indicating how each pregnancy ended is BB-2 PREGOUTn, and for each pregnancy up to 5 outcomes can be reported. There were no pregnancies in the 2022-2023 NSFG data with more than 4 types of outcomes reported (PREGOUT1, PREGOUT2, PREGOUT3, PREGOUT4).

The Blaise-computed variable **prgoutcome** defined in Flow Check B-7 indicates the *primary* outcome of the pregnancy based on BB-2 PREGOUTn, and follows this hierarchy of outcomes noted above:

- live birth (code 1),
- abortion (code 2),
- stillbirth (code 3),
- miscarriage (code 4),
- ectopic pregnancy (code 5),
- current pregnancy (code 6).
- not a live birth, and not a current pregnancy (code 7, based on BB-2b HOWENDDK = no, DK, or RF)

The **prgoutcome** variable is used to define OUTCOME. Those cases that are coded 7 on prgoutcome to represent “not a live birth, and not a current pregnancy” (based on BB-2b

HOWENDDK = no, DK, or RF) are imputed on OUTCOME.

Logic for defining OUTCOME recode, based on the pregnancy file data:

If prgoutcome = 7 then impute OUTCOME;
Else OUTCOME= prgoutcome;

Note: If interested in all outcomes for a multiple-outcome pregnancy (e.g., 2 fetuses where one was born alive and one was a stillbirth), you must use raw variables BB-2 PREGOUTn (restricted-use and only available in the NCHS Research Data Center). To determine if a pregnancy resulted in a multiple birth (e.g., twins, triplets), see computed variable **bornaliv**. A multiple-gestation pregnancy may have only 1 outcome noted on BB-2 PREGOUTn if all fetuses had the same outcome, e.g., a set of twins born by vaginal delivery.

Imputation Note: Done for cases with prgoutcome = 7 “not a live birth and not a current pregnancy.” Given that all cases imputed on OUTCOME - were known to be something other than live births, all were imputed to a value of 2, 3, 4, or 5 (non-live birth outcome).

Code categories:

1 = Live birth
2 = Induced abortion
3 = Stillbirth
4 = Miscarriage
5 = Ectopic pregnancy
6 = Current pregnancy

BIRTHORD: Birth order

BIRTHORD is blank (inapplicable) if the pregnancy did not result in live birth (pregnancy file recode OUTCOME NE 1).

Otherwise, BIRTHORD is set to 1 for R’s first live birth and incremented by 1 for each additional live birth R reported. BIRTHORD reflects the respondent’s births in chronological order. No imputation is needed, but BIRTHORD values may reflect imputed values of OUTCOME and DATEND.

SAS logic:

```
data births; set pregs;  
if outcome=1; /* subset live births
```

```
*/ data birthord; set births; by
caseid;
if first.caseid then
birthord=1; else
birthord+1;
retain birthord;
label birthord="birth order";
```

Note: BIRTHORD is defined by pregnancy and does not account for multiple births. For example, if the respondent's first pregnancy resulted in twins both born alive and her second pregnancy resulted in a single live birth, her first pregnancy with twins will have BIRTHORD=1 and her second pregnancy with a single baby will have BIRTHORD=2.

Imputation Note: No imputation of BIRTHORD was needed because it is based on imputed values of OUTCOME and DATEND, but BIRTHORD may have been adjusted if any changes were made (during post-processing) to PREGORDR and the affected pregnancies include live births.

Code categories:

Blank =
inapplicable 1-nn
= birth
order

DATEND: Year when pregnancy ended

DATEND is blank (inapplicable) if the pregnancy is a current pregnancy (recode OUTCOME = 6).

The DATEND recode was initially defined as a century-month date, with values directly drawn from the Blaise-computed variable **cmpregend** (defined in Flow Check B-10).

DATEND= cmpregend

Note: For inclusion on the 2022-2023 public-use files, the CM dates for DATEND were converted to years, as done in previous public-use files. The original CM date variable INDATEND is restricted-use and available only through the NCHS Research Data Center.

Imputation Note: Done for all pregnancies where cmpregend = DK, RF, or 9997. Imputed values of OUTCOME are used to constrain imputation of DATEND (e.g., if OUTCOME is imputed to be a current pregnancy,

then DATEND is imputed to blank/inapplicable.)

Code categories for DATEND (public-use variable):

Blank = inapplicable
xxxx - 2023 = Year pregnancy ended

Code categories for INDATEND (restricted-use variable):

Blank = inapplicable
xxxx – 1488 = date (century month) pregnancy ended

AGEPREG:** **Age at pregnancy outcome**

AGEPREG is blank (inapplicable) if this is a current pregnancy (recode OUTCOME=6).

Otherwise,

AGEPREG is defined using R's date of birth and the final imputed CM value of recode DATEND:

$$\text{AGEPREG} = \text{INT}((\text{DATEND} - \text{cmbirth})/12)$$

SAS logic:

```
if outcome = 6 then agepreg = .;  
else agepreg = int((datend-cmbirth)/12);
```

Code categories:

Blank = inapplicable
xx - 49 = age at pregnancy outcome

DATECON: Year when pregnancy began

DATECON is blank (inapplicable) if the pregnancy did NOT fall in any of the categories below:

- Pregnancy ended in live birth or is a current pregnancy (recode OUTCOME = 1 or 6);
or
- Pregnancy is a 1st or 2nd order pregnancy (pregodr = 1 or 2); **or**
- Pregnancy is a recent pregnancy ending since January of 5 years prior to interview (recode DATEND GE cmjan5yr)

Otherwise, if the pregnancy met ANY of the above “applicable” criteria:

DATECON is based directly on original, non-missing values of cmprgbeg (CM date for pregnancy start as defined in Flow Check B-37).

If cmprgbeg not in(., 9997, 9998, 9999) then DATECON = cmprgbeg;

If **cmprgbeg** has missing values (sysmis or 9997, 9998, 9999 corresponding to not ascertained, refused, or don't know), then DATECON is computed as follows, using imputed values of PRGLNGTH recode:

IF recode OUTCOME=6 and cmprgbeg in(., 9997, 9998, 9999) then
 $DATECON = cmintvw - ROUND(PRGLNGTH/4.33)$.

If recode OUTCOME=1 and cmprgbeg in(., 9997, 9998, 9999) then
 $DATECON = DATEND - ROUND(PRGLNGTH/4.33)$;

Else if OUTCOME in(2 3 4 5) and ((pregodr in(1 2)) or (DATEND GE cmjan5yr)) and cmprgbeg in(., 9997, 9998, 9999) then
 $DATECON = DATEND - ROUND(PRGLNGTH/4.33)$;

Note: For inclusion on the 2022-2023 public-use files, the CM dates for DATECON were converted to years, as done in public-use files for 2015-2017 and 2017-2019. The original CM date variable INDATECON is restricted-use and available only through the NCHS Research Data Center. PRGLNGTH is also restricted-use and available only through the NCHS RDC.

Imputation Note: If either DATEND or PRGLNGTH is based on imputed values, then DATECON is based on those imputed values.

Code categories for DATECON (public-use variable)
xxxx - 2023 = Year of pregnancy conception

Code categories for INDATECON (restricted-use variable):
xxxx - 1488 = date (century month) of pregnancy conception

AGECON:** **Age at time of conception**

AGECON is blank (inapplicable) if DATECON is inapplicable.

Note: In prior NSFGs, DATECON and AGECON recodes were applicable for all pregnancies but in 2022-2023, the survey was modified to only obtain gestational length information for selected pregnancies (all live births, all 1st or 2nd pregnancies, and all pregnancies ending in or after January of 5 years prior to interview.

AGECON is defined using R's date of birth and the final imputed value of recode DATECON:

AGECON = INT[(DATECON – cmbirth)/12]

Code categories:

xx - 49 = age at pregnancy conception

PMARPREG:** **Whether pregnancy ended before R’s 1st marriage (premarital)**

PMARPREG is blank (inapplicable) if this is a current pregnancy (recode OUTCOME = 6).

Otherwise:

PMARPREG is defined based on final imputed values of preg file recode DATEND and resp file recodes FMARITAL and MARDAT01.

PMARPREG=1 (yes) if: --R has never been married (FMARITAL=5), or
--pregnancy recode DATEND < MARDAT01

PMARPREG=2 (no) if: DATEND >= MARDAT01

SAS logic:

if outcome = 6 then pmarpreg = .;
else if fmarital = 5 or datend < mardat01 then pmarpreg = 1;
else if datend GE mardat01 then pmarpreg = 2;

Note: *If users wish to limit to pregnancies to “ever-married” respondents, they should subset cases with FMARITAL NE 5.*

Code categories:

Blank = inapplicable
1 = yes (pregnancy ended before 1st marriage, including never-married)
2 = no (pregnancy ended in same month as or later than 1st marriage)

BFEEDWKS: **Duration of breastfeeding in weeks**

Note: *Because duration of breastfeeding can be affected by multiple births and because multiple births were relatively rare in the NSFG sample, this recode is defined ONLY for singleton gestations. However, this algorithm could be used for all births.*

BFEEDWKS is blank (inapplicable) if:

-- this is a current pregnancy or it did not result in live birth (recode OUTCOME NE 1);
-- it was a multiple birth (Blaise-computed variable **bornaliv > 1**);

- this singleton child was born earlier than January of 5 years before the interview (recode DATEND LT cmjan5yr);
- the child died at age younger than 1 year old (BC-18 AGEDIED = 1);
- this child is alive but stopped living with R at age younger than 1 year old (BC-20 AGELEFT = 1).

OTHERWISE, for all applicable singleton live births where the child was born since cmjan5yr and lived with R for at least for the first year of life:

BFEEDWKS=995 if R reported never breastfeeding this baby (BC-21 ANYNURSE = NO)

BFEEDWKS=994 if R is still breastfeeding this baby (BC-23 QUITNURS = 5 (no) or BC-24a AGEQTNUR_N=996)

Otherwise, if BC-21 ANYNURSE = YES:

BFEEDWKS is based on Blaise-computed variable **ageqtnur**, the child's age in months when R stopped breastfeeding this child, as defined in Flow Check B-36).

For all cases where **ageqtnur** has a valid value (not DK/RF, nor "not ascertained), the **ageqtnur** value was converted from months to weeks:

$$BFEEDWKS=ROUND(AGEQTNUR*4.33)$$

Note: BFEEDWKS was redefined as a categorical variable for inclusion on the 2022-2023 NSFG public-use file, as in 2017-2019. INBFEEDWKS, available in the NCHS Research Data Center, shows the exact number of weeks as originally defined.

Imputation Note: For all imputation of BFEEDWKS in weeks, no imputed value should be greater than the smallest value among 48 months and kidage (in months). Put another way, they can never be imputed to have breastfed longer than the child's age, AND no imputed value should exceed 48 months (which equates to 208 weeks).

Imputation needed for cases with:

- *DK/RF on questions determining if this singleton child lived at least for 1 year with R (BC-16 LIVEHERE, BC-18 AGEDIED, or BC-20 WHENLEFT).*
- *DK/RF on questions determining duration of breastfeeding (BC-21 ANYNURSE, BC-23 QUITNURS, BC-24a AGEQTNUR_N, or BC-24b AGEQTNUR).*

Code categories for BFEEDWKS (public-use variable):

- Blank = Inapplicable
- 0 = Never breast-fed this child
- 1 = Breastfed less than 1 week
- 2 = Breastfed 1-13 weeks (LE 3 mos)
- 3 = Breastfed 14-26 weeks (>3/LE 6 mos)
- 4 = Breastfed 27-52 weeks (>6/LE 12 mos)
- 5 = Breastfed 53-104 weeks (>1/LE 2 yrs)
- 6 = Breastfed 105+ weeks (>2 yrs)
- 7 = Still breastfeeding this child

Code categories for INBFEEDWKS (restricted-use variable):

- Blank = Inapplicable
- 000 = Breastfed for less than 1 week
- 001 = Breastfed for 1 week
- 002 = Breastfed for 2 weeks
- 003 = Breastfed for 3 weeks etc.
- 994 = Still breastfeeding this child
- 995 = Never breast-fed this child

Section E Recodes:

Marital/Cohabiting Status for Pregnancies; Wantedness of Pregnancies

RMAROUT3: Informal marital status at pregnancy outcome – 3 categories

Note: This new recode replaces RMAROUT6 that was specified in past NSFGs using information from Sections B and C on pregnancy, marriage, and cohabitation dates. RMAROUT3 is based primarily on EG-18b COHPEND that is part of a series of questions asked about marital and cohabiting status with the father of the pregnancy

RMAROUT3 is blank (inapplicable) if the pregnancy did NOT fall in any of the categories below:

- Pregnancy ended in live birth (recode OUTCOME = 1); **or**
- Pregnancy is a 1st or 2nd order completed pregnancy ((pregodr = 1 or 2) and (prgoutcome NE 6)); **or**
- Pregnancy is a recent pregnancy ending since January of 5 years prior to interview (recode DATEND GE cmjan5yr)

Otherwise, if the pregnancy met ANY of the above “applicable” criteria:

RMAROUT3 = EG-18b COHPEND.

Else if 0<cmmarrch<9998 (cmmarrch not sysmis nor dk/rf) then:
If recode DATEND GE cmmarrch (CM date of current marriage, as defined in Flow Check C-7a) then RMAROUT3=1. (EG-18b COHPEND is not asked when the pregnancy ended during the current marriage.)

Imputation note: Imputation needed if EG-18b COHPEND =DK/RF and (DATEND was not within current marriage or inputs to determining whether DATEND was within current marriage are missing). RMAROUT3 cannot be imputed to married or cohabiting if evrmarry NE 1 or evrcohab NE 1.

Code categories:

- Blank = Inapplicable
- 1 = Married to father of pregnancy
- 2 = Living with father of pregnancy, but not married
- 3 = Neither married nor cohabiting with father of pregnancy

RMARCON3: Informal marital status at pregnancy conception – 3 categories

Note: This new recode replaces RMARCON6 that was specified in past NSFGs using information from Sections B and C on pregnancy, marriage, and cohabitation dates. RMARCON3 is based primarily on EG-18a COHPBEG that is part of a series of questions asked about marital and cohabiting status with the father of the pregnancy

RMARCON3 is blank (inapplicable) if the pregnancy did NOT fall in any of the categories below:

- Pregnancy ended in live birth or is a current pregnancy (recode OUTCOME = 1 or 6); **or**
- Pregnancy is a 1st or 2nd order pregnancy (pregodr = 1 or 2); **or**
- Pregnancy is a recent pregnancy ending since January of 5 years prior to interview (recode DATEND GE cmjan5yr)

Otherwise, if the pregnancy met ANY of the above “applicable” criteria:

RMARCON3=EG-18a COHPBEG

Else if 0<cmmarch<9998 (cmmarch not sysmis nor dk/rf) then:

If recode DATECON GE cmmarrch (CM date of current marriage, as defined in Flow Check C-7a) then RMARCON3=1. (EG-18a COHPBEG is not asked when the pregnancy began during the current marriage.)

Imputation note: Imputation needed when EG-18a COHPBEG = DK/RF, and (DATECON

was not within current marriage or inputs to determining whether DATECON was within current marriage are missing). RMARCON3 cannot be imputed to married or cohabiting if evrmarry NE 1 or evrcohab NE 1.

Code categories:

- Blank = Inapplicable
- 1 = Married to father of pregnancy
- 2 = Living with father of pregnancy, but not married
- 3 = Neither married nor cohabiting with father of pregnancy

NEWWANTR: Detailed Wantedness of Pregnancy – Respondent

Note: This recode is comparable to recode WANTRESP in prior cycles except that it includes two categories of “too soon.” WANTRESP is still provided to users, defined based on final, imputed values of NEWWANTR.

NEWWANTR is blank (inapplicable) if the pregnancy did NOT fall in any of the categories below:

- Pregnancy ended in live birth or is a current pregnancy (recode OUTCOME = 1 or 6);
or
- Pregnancy is a 1st or 2nd order pregnancy (pregodr = 1 or 2); **or**
- Pregnancy is a recent pregnancy ending since January of 5 years prior to interview (recode DATEND GE cmjan5yr)

Otherwise, if the pregnancy met ANY of the above “applicable” criteria:

If EG-10 TIMINGOK=3 then NEWWANTR=1.

Else if TIMINGOK=1 then do:

If EG-11 TOOSOON_N NE DK/RF (<998) THEN DO:

If TOOSOON_P = 2 then TOOSOON = TOOSOON_N*12

Else if TOOSOON_P = 1 then TOOSOON = TOOSOON_N

End;

If TOOSOON LT 24 then NEWWANTR=3

Else if TOOSOON GE 24 then NEWWANTR=4

End;

Else if TIMINGOK=2 then NEWWANTR=2.

Else if TIMINGOK=4 then NEWWANTR=5.

Else if EG-6 WANTBOLD=5 or EG-7 PROBBABE=5 then NEWWANTR=6.

Else if PROBBABE=6 or PROBBABE=DK then NEWWANTR=7.

Note: Unlike WANTRESP in the 1995 through 2011-2019 data files, this recode does not take

into account the confirmation question EG-9 WANTBLD2, formerly asked of ages 15-19 but no longer asked, as of 2022. The number of respondents who changed their responses as a result of this confirmation question was too small to make a difference on the recode.

Imputation Note: If EG-6 WANTBOLD=1 (yes), then NEWWANTR cannot be imputed to 5 (unwanted). If set for imputation due to missing values on TOOSOON_P or TOOSOON_N, imputation must be constrained to be 3 or 4. In other words, if TIMINGOK=1 then imputation must be constrained to be 3 or 4.

Code categories:

Blank = inapplicable
1= Later, overdue
2= Right time
3= Too soon: by less than 2 years
4= Too soon: by 2 years or more
5= Didn't care, indifferent
6= Unwanted
7= Don't know, not sure

WANTRESP:** Wantedness of Pregnancy – Respondent

WANTRESP is blank (inapplicable) if the source recode NEWWANTR is inapplicable.

WANTRESP is defined based on NEWWANTR values as shown in SAS logic below.

SAS logic:

```
if newwantr = . then wantresp = .;  
else if newwantr = 1 then wantresp = 1; /* later, overdue */  
else if newwantr = 2 then wantresp = 2; /* right time */  
else if newwantr in(3,4) then wantresp = 3; /* too soon */  
else if newwantr = 5 then wantresp = 4; /* didn't care, indifferent */  
else if newwantr = 6 then wantresp = 5; /* unwanted */  
else if newwantr = 7 then wantresp = 6; /* don't know, not sure */
```

Code categories:

1= Later, overdue
2= Right time
3= Too soon, mistimed
4= Didn't care, indifferent
5= Unwanted
6= Don't know, not sure

WANTPART: Wantedness of pregnancy -- Respondent's Partner (father of pregnancy)

WANTPART is blank (inapplicable) if the pregnancy did NOT fall in any of the categories below:

- Pregnancy ended in live birth or is a current pregnancy (recode OUTCOME = 1 or 6);
or
- Pregnancy is a 1st or 2nd order pregnancy (pregodr = 1 or 2); **or**
- Pregnancy is a recent pregnancy ending since January of 5 years prior to interview (recode DATEND GE cmjan5yr)

Otherwise, if the pregnancy met ANY of the above “applicable” criteria:

If EG-17 TIMOKHP=1 then WANTPART=3.

Else if TIMOKHP=2 then WANTPART=2.

Else if TIMOKHP=3 then WANTPART=1.

Else if TIMOKHP=4 then WANTPART=4.

Else if EG-16 HPWNOLD=5 then WANTPART=5. Else if HPWNOLD=6 then WANTPART=6.

Note: If users wish to retain the “don’t know” and “prefer not to answer” values on HPWNOLD for analysis, the imputed values on this recode can be replaced by the original HPWNOLD value for these cases.

Imputation Note: If HPWNOLD=1 (yes), then WANTPART cannot be imputed to 5 (unwanted).

Code categories:

Blank = inapplicable

1= Later, overdue

2= Right time

3= Too soon, mistimed

4= Didn’t care, indifferent

5= Unwanted

6= Don’t know, not sure