

Washington State Department of Health

Birth Data File Technical Notes

A DATA USERS' GUIDE



DOH 422-160 07-26-2021
Center For Health Statistics

For persons with disabilities, this document is available on request in
Other formats. To submit a request, please call 1-800-525-0127
(TDD/TTY 1-800-833-6388).

TABLE OF CONTENTS

OVERVIEW:.....3

METHODOLOGY:.....3

CHANGES TO THE DATA COLLECTION SYSTEM - WHALES.....4

EFFECTS OF TRANSITION ON KEY VARIABLES4

CHANGES TO NEW FILE FORMAT5

 NEW FIELDS6

 DISCONTINUED FIELDS7

KNOWN DATA QUALITY ISSUES7

 AGE (MOTHER, FATHER).....7

 APGAR SCORE (5 MINUTES)7

 ATTENDANT/CERTIFIER CLASSIFICATION.....7

 BIRTH WEIGHT – see Historical Notes8

 BODY MASS INDEX (BMI)8

 CERTIFICATE NUMBER.....8

 CERTIFICATE TYPE (DELAYED BIRTH)9

 FACILITY OF BIRTH9

 GEOGRAPHY – CITY/COUNTY/STATE RESIDENCE & OCCURENCE9

 GESTATIONAL AGE – CALCULATED AND CLINICAL ESTIMATE11

 INFANT DEATH FLAG.....12

 INFANT LIVING AT TIME OF REPORT.....12

 MATERNAL SMOKING (NUMBER OF CIGARETTES/PACKS PER DAY).....12

 MEDICAL AND HEALTH INFORMATION SECTION.....12

 MENSES DATE.....13

 MONTH PRENATAL CARE BEGAN13

 OCCUPATION (MILHAM CODES - MOTHER, FATHER).....13

 PREVIOUS PREGNANCY HISTORY14

 RACE/ETHNIC ORIGIN (MOTHER, FATHER, CHILD).....14

 SOURCE OF PAYMENT15

 WEIGHT GAINED DURING PREGNANCY, PREPREGNANCY WEIGHT, AND DELIVERY WEIGHT15

DATA DOCUMENTATION15

 Data Dictionary.....15

 Uniform Parentage Act (UPA)16

HISTORICAL NOTES.....16

Birth Certificate and Filing System16

Age (Mother, Father)16

Alcohol use17

Attendant/Certifier Classification.....17

Birth Weight	17
Geography	17
Maternal Smoking (Number of Cigarettes/Packs per Day)	18
Medical and Health Information Section	18
Previous Pregnancy History	19
Race	20
Weight Gained During Pregnancy, Prepregnancy Weight, and Delivery Weight	20

BIRTH DATA QUALITY TECHNICAL NOTES
Washington State Department of Health
Center for Health Statistics, September 2021

OVERVIEW:

The Washington State Department of Health (WA DOH) Center for Health Statistics (CHS) makes certain that vital events data collected from Washington birth certificates are publicly available. Vital events data are primarily used to inform WA DOH state priorities and are widely used for research.

On average there are about 87,000 births per year to Washington State residents. There has been an overall drop in births in the United States (down about 4%). Washington State births are down by 2% from last year.

In 2020, there were:

- 84,259 registered births.
- 83,101 were Washington resident births.
- 82,836 were Washington occurrence births and
- 1,423 births occurred out of state to Washington residents.

The purpose of this document is to provide birth data users an overview of:

- I. Methodology/Data collection processes
- II. Changes to the data collection system
- III. Effects of transition to WHALES on key variables

METHODOLOGY:

CHS captures demographic and medical information for all births in WA State. Birth facilities are required to register a birth within 10 days of delivery (RCW 70.58.080). Information on births is collected via the Washington State Birth Filing Form ([WashingtonStateBirthFilingForm.pdf](#)) and required fields are designated in WAC 246-491-029.

The birth certificate data collection process is not streamlined. The department of health provides the birth filing form to hospitals, but hospitals are not required to use the form. Most hospitals use a series of worksheets to collect data from the mother, the medical records, and the physician. The type of worksheet used and the method of collecting data from these sources varies from hospital to hospital, as does the amount of follow up used to collect missing information.

The method for entering birth certificate data into the computer has changed over time:

- 1980-2002: See [Historical Notes](#)
- 2003: A new web-based system for entering the data replaced the Delivery Certificate Tracker (DCT). This system was called the Birth Record Real Time Registration System (BR3). The system had more data quality checks and made it easier to submit data to the state.

- 2015-2017: Washington Health and Life Events System (WHALES) was implemented in September 2015 for data collection only – the system was used for data entry but was not the database of record. All data entered in WHALES were converted back to Bedrock, and file extracts for 2015 and 2016 were pulled from Bedrock. Beginning February 2017, all data in Bedrock were converted to WHALES and WHALES became the database of record. For the 2017 birth file, data entered from January 1 through February 17 were converted from Bedrock and entered the live WHALES system. After February 17, all records were entered and stored directly into the WHALES system.

Once a birth is registered, the information is sent to the National Center for Health Statistics (NCHS) for coding including, providing coding for race and ethnicity text fields. NCHS returns the coded information to CHS within a few days.

Classification and coding of data on Washington State vital records follow National Center for Health Statistics (NCHS) guidelines as defined in ‘Vital Statistics Instruction Manuals.’ For details see https://www.cdc.gov/nchs/nvss/instruction_manuals.htm. CHS performs routine data quality checks on an ongoing basis and especially after the close of a given calendar year as we approach the deadline for submitting the annual file to NCHS and to data users.

CHANGES TO THE DATA COLLECTION SYSTEM - WHALES

CHS began using WHALES for data entry in September of 2015, and in February 2017, the WHALES system became the database of record. Beginning in 2017, birth data files for analysis will be from WHALES, and as a result, the data extracted from the new system look different. Most field names have been updated to be more descriptive. Some variable formats and/or values have changed. There are also fields that have been added, as well as some discontinued. These changes will be reflected in the 2017 data and future releases of the Birth Statistical file. Changes to the data collection system may affect fields in the birth data. Any large gap before and after a change might be a product of the change rather than a real difference. Changes to key variables are described below.

CHS will continue to release final Birth Statistical data files annually in Microsoft Excel format, with SAS and Stata versions also available. Data files can be accessed through the Y: Drive for WA-DOH employees. Upon completion of a data sharing agreement with their agency, local health department employees can download the files from the CHS Data Files site (accessed via Secure Access Washington). The files are also available for online ordering with established data sharing agreement. The Birth Statistical file for years 1980-2016 will continue to be distributed in the old format, documented by this data dictionary, [Data Dictionary 1980-2016](#).

EFFECTS OF TRANSITION ON KEY VARIABLES

Data users utilizing statistical software’s (SAS, Stata, R, etc.) to analyze birth data will see significant changes to variable names for data fields starting with the 2017 birth file. The new variable names will not match those in the old (2016 or older) birth files. **Please Note** that any analytic programs written for pre-2017 data will have to be revised for data produced for 2017 and later.

To assist with the transition, the new [Birth Statistical File - Data Dictionary and Crosswalks](#) has the new field names formats and values, as well as the historical ones. It also includes all coded crosswalks, like the hospital facility code table.

CHANGES TO NEW FILE FORMAT

- **Birth Certificate Number:** The birth certificate number is now referred to as the birth state file number. Birth Statistical files will have an encrypted state file number labeled 'SFN Encrypt' (formerly 'certno_e').
- **Format Changes:** Examples of format changes include:
 - **Time of Birth,** will be a character field with format HH: MM.
 - **Certificate Type,** historically has been R 'Regular' D 'Delayed' O 'Out of State' and now includes B 'Foreign Born'
 - **Sex,** will include M 'Male' F 'Female' U 'Undetermined' and X 'Gender not exclusive'
 - **Mother's marital status** has historically been a yes, no or unknown question. There will be more detailed categories describing marital status i.e. 1 – 'Married to other parent listed' 2 – 'Married but not to the person listed as parent' 3 – 'Married but refuse to provide information' 4 – 'Not married and providing parent information' 5 – 'Not married now, was married during pregnancy' 6 – 'Not married and refuse to provide information' 8 – 'Not provided' U – 'Unknown'
 - **Facility Type for Birthplace** historically included 'Home- planned' and "Home Unplanned' but now only includes 1 'Hospital' 2 'Enroute' 3 'Birth Center' 4 'Clinic/Doctor's Office' 5 'Home' 6 'Other' 9 'Unknown/Not Stated'. Data on **Intended Facility** is given only if the facility where birth occurred is not the intended place of delivery.
- **Array Fields:** There were several fields that used to be provided as arrays of codes indicating birth scenarios. All the scenarios were assigned a code in each array. These will instead be individual fields with yes, no and unknown values. These include:
 - **Mother medical risk factors,** new fields like 'Gestational Hypertension' and 'Prepreg Diabetes' were formerly in 'mrf1-mrf6'
 - **Maternal infections** like 'Hep C' and 'Chlamydia', formerly in 'minfect1 – minfect7'
 - **Obstetric procedures** like 'Cervical Cerclage' and 'Tocolysis', formerly in 'obproc1-obproc4'
 - **Onset of labor conditions** like 'Prolonged Labor' and 'Ruptured Membranes', formerly in 'labons1 – labons3'
 - **Characteristics of labor** like 'Induction' and 'Steroids', formerly in 'labchar1 – labchar7'
 - **Maternal morbidities** like 'Maternal Transfusion' and 'Hysterectomy', formerly in 'mmorbid1 – mmorbid5'
 - **Infant conditions** like 'Surfactant' and 'Child Seizure', formerly in 'abcond1-abcond6' and

- **Congenital malformations** like ‘Anencephaly’ and ‘Cleft Lip’, formerly in ‘malfl – malf7’.
- **Calculated Infant Race and Ethnicity:** The 2017 file was the last year these fields were included, and they will be discontinued in all future files. The NCHS calculated child race (raceccal) and ethnicity (hispcal) fields were discontinued by NCHS in 1989 and are no longer used by NCHS.

The original infant race variable was calculated as follows:

When the parents were of the same race, the race of the child was the same as the race of the parents. When the parents were of different races and one parent was white, the child was assigned to the other parent’s race. When the parents were of different races and neither parent was white, the child was assigned to the father’s race, with one exception. If either parent was Hawaiian, the child was assigned to Hawaiian. If race was missing for one parent, the child was assigned the race of the parent for whom race was reported. When information on race was missing for both parents, the race of the child was considered not stated.

The original infant race and ethnicity variables were based off the pre-2003 certificate which only allowed a single race to be reported. When the Revised Standard Birth Certificate was implemented in 2003, the race and ethnicity reporting changed from single race to multiple race reporting. Since the infant race and ethnicity calculations were discontinued in 1989, there was not a corresponding update to the algorithm in 2003 by NCHS; rather, WA DOH created an algorithm to translate the new race and ethnicity categories to the old race categories so that the infant race and ethnicity fields could still be calculated. **We are discontinuing this in the 2018 Birth file.**

NEW FIELDS

The following fields have been added to the Birth Statistical file beginning with 2017 data.

New Fields	Notes
Mother Date of Birth	
Father Date of Birth	
Birthplace City and Birthplace County	Literals
Residence City and Residence County	Literals
Birthplace State	Literals
Birthplace State FIPS Code	Literals
Birthplace Country	Literals
Race Tribe 1 & 2	Literals
Race Other Asian 1& 2	Literals
Race Other Pacific Islander 1 & 2	Literals
Race Other 1 & 2	Literals
NCHS Race and Ethnicity Ecodes and Ccodes	Used to calculate Mother/Father Race

	Summary and Bridged race
NCHS Hispanic E codes and C Codes	Used to calculate Mother/Father Hispanic ethnicity
Mother & Father Bridged Race	NCHS calculated bridge race
Mother Father Hispanic Other	Literal
Number of cigarette packs smoked	During each trimester
Maternal infection information	Other infections literal

DISCONTINUED FIELDS

Data fields that have been discontinued over the years will not have placeholders included in future releases. Current fields that will be removed include:

Discontinued fields	Notes	Year
Mother Occupation Milham	formerly occm_mom	2018
Father Occupation Milham	formerly occm_dad	2018
Mother's birthplace state NCHS code	formerly bctrymom	2017
Date of last live birth	formerly llbmm4y, month and year fields remain	2017
Date of other pregnancy outcomes	formerly othmm4y, month and year fields remain	2017

KNOWN DATA QUALITY ISSUES

For issues related to files prior to 2003, please see [Historical Notes](#)

AGE (MOTHER, FATHER)

The father's age is missing from a substantial number of records, mostly because the mother is unmarried, and a paternity affidavit has not yet been filed.

APGAR SCORE (5 MINUTES)

In 2003, the 1-minute Apgar score was discontinued. At the beginning of the year, some hospitals were putting the 1-minute score in the place of the 5-minute score since they were used to putting in the 1-minute score first. This was corrected in the middle of the year. For 2003 and later, users should just use the later data for 5-minute score. The 10-minute Apgar score is collected only if the 5-minute score is less than 6, therefore there is a high number of missing's for APGAR 10.

ATTENDANT/CERTIFIER CLASSIFICATION

The birth certificate collects data on both the certifier (who attests that the child was born alive at the time, place, and date stated) and the attendant (who delivered the baby).

- Since mid-1985, the classification has changed numerous times. See [Historical Notes](#) on attendant/certifier for additional data quality notes.
- 2017: Both certifier and attendant are populated in the data.
- Between 2015 and 2017, we saw an increase in the number of births to Registered Nurses (RN's) where RN's were reported as the birth attendant.
- In 2019: After thorough investigation we found the root cause to be old license types listed for some attendants. 2019 data will show an increase in Certified Nurse Midwives (CNM's) and License Midwives (LM's) and a decrease in RN's and that will be due to this cleanup.

BIRTH WEIGHT – see [Historical Notes](#)

BIRTHPLACE TYPE:

Pre 2017, if the birth occurred in the home, (home can be the mother's residence or another person's residence) the birth certificate had two options to select i.e. "Home -Planned" or "Home - Unplanned" (see data dictionary). As of 2017, if the birth occurred only one option is available and that is "Home (planned)".

BODY MASS INDEX (BMI)

The Body Mass Index (BMI) is a measure of weight for height. The formula for calculating Body Mass Index is: $BMI = 703.1 \times (\text{pregnancy weight in lb.} / \text{square of height in inches})$. For analysis, Body Mass Index is generally grouped as follows: Underweight (<18.5), Normal (18.5 – 24.9), Overweight (25.0 – 29.9), and Obese (30.0 and above).

CERTIFICATE NUMBER

In 2007, increases in the number of babies born in the state required the Center for Health Statistics (CHS) to change the certificate numbering scheme used in previous years. The range of certificate numbers for in-state births changed from 1-88,999 to 1-199,999. Since the certificate number has always been 6 characters (plus the birth year) the field size did not expand.

Public use data files do not include the actual certificate number. Instead, they have an encrypted number. In previous years these encrypted numbers have all started with '7' to distinguish them from actual certificate numbers. Because of the expansion in certificate number range, CHS added more encrypted numbers to cover the new ranges. These new numbers start with '5' or '6,' are still distinct from actual certificate numbers, and are still 10 characters long (4-digit birth year plus 6-digit encrypted number).

For 2017 data, and moving forward, the birth certificate number is now referred to as the birth state file number. Birth Statistical files will have an encrypted state file number labeled 'SFN Encrypt' (formerly 'certno_e').

CERTIFICATE TYPE (DELAYED BIRTH)

Births registered more than 10 days after birth are called “Delayed Registrations” and are assigned a Type D Birth Certificate. These types of certificates record a very limited amount of information due to the delay between the date of birth and the date of registration. They are included here for completeness sake.

FACILITY OF BIRTH

Two things can affect the number of births in a particular facility:

- A change in facility code due to changes in licensing (e.g., a new name, new ownership)
- A change in facility characteristics such as a merger with another facility or changes in facility service area.

Thus, a given facility code may have several births in one year and then drop to zero the next year – or may have a large increase in births. Every year, the Center for Health Statistics examines reporting trends in births by facility and verifies any unusual patterns.

Midwives (particularly licensed midwives) have two series of facility codes. The ‘300’ series is for midwives delivering at the mother’s home or someone else’s private home. The ‘400’ series is for midwives who deliver at their place of business, typically a licensed birthing center. A single midwife may have two codes if s/he delivers in both places.

Beginning in 2016 all home births get the generic facility code of 810, regardless of who delivers the child. The ‘300’ and ‘400’ series codes were discontinued for all home births and are no longer used by DOH.

GEOGRAPHY – CITY/COUNTY/STATE RESIDENCE & OCCURENCE

Place of Residence

There are multiple variables in the birth file that provide information on the place of residence.

WA city and county codes (Birthplace County City WA Code formerly city_occ and Birthplace County WA Code formerly cnty_occ)

- All WA County and city codes are four-digits including leading zeros, with the first two digits being a county code and the second two being a city code. A WA city code is given if the city has at least 2,500 people. Otherwise, it is given a ‘balance of county’ code (a two-digit county code and a city code of '00'), along with other small areas in the county. A city near the cutoff point may fluctuate above and below 2,500 population and thus may have a separate code in some years and not in others. A count of zero births in a particular year may simply mean that the city did not have a separate code in that year.

- The number of births in a city may also fluctuate if the city annexes a large area. In that case, the number of births might increase substantially but the population also increases and thus the birth rate will show little change.
- Variables (Mother Residence City WA Code and Mother Residence County WA Code) showing county and city of residence or occurrence in 2017 show values with a series of '9's. These values do not appear in the previous birth files. Records with '9's in the county and/or city fields may indicate that the individual was homeless and used to be coded as a series of '0's along with out of state residents in prior years.
- The code for city of residence is based on whether or not the birth occurred within city limits. (These data are collected from the item on the birth certificate: 'Inside city limits - yes/no'.) The city of occurrence is coded to the hospital, birth center or other location where the birth occurred. The mother's residence only affects the city occurrence when it is a home birth.
- Population estimates provided by the Washington State Office of Financial Management establish which cities meet the population criteria for separate coding. Because these estimates are published in the middle of the year, changes do not appear in the birth data file until the following year. Thus, a city which first exceeds 2,500 population in 2016 would not have a separate code until 2017.

Zip Codes - Geocoded

- The zip code is determined by geocoding addresses (geozip) was added to files starting with 1988 data. The values for geozip are determined from geocoding the mother's resident address. The geocoded zip code may differ from the reported zip code for several reasons: (1) a data entry error or reporting error for the reported zip code; (2) the zip code boundary changed; (3) the geocoding process matched the address to the wrong location.
 - When working with data from 1988-2016, it is recommended that you use the geozip field in preference to the reported zip code. The reported zip code should only be used when the geozip field is missing. Zip code boundaries change over time, so caution should be used when using zip codes for births occurring in earlier years, especially prior to 1994. The number of addresses with differences between reported zip code and geocoded zip code will increase when using older files because the geocoded zip field will have different boundaries.
 - For 2017 data and going forward there will only be one residence zip code field, 'Mother Residence Zip'. This is produced by autocorrecting and geocoding the reported address, and is analogous to 'geozip' from previous years.
- ***Selecting residence or occurrence:*** All files have data for both the mother's place of residence and the place where the birth occurred. To study Washington State residents,

select state of residence ('**Mother Residence State NCHS code**') = '48'. Similarly, to study Washington State occurrences, select state of occurrence ('**Birthplace State NCHS Code**') = '48'.

GESTATIONAL AGE – CALCULATED AND CLINICAL ESTIMATE

The birth certificate provides for 2 ways of determining gestational age:

- By a clinical estimate
- By calculation from the menses date and the birth date

Clinical estimate: Has been collected since 1989 but has had limited completeness and accuracy until the last decade. Increasing evidence of greater validity of clinical estimates compared to calculated estimates, and the national availability of clinical estimates data, have prompted NCHS to use the clinical estimate as its primary measure of gestational age beginning with 2014 data.

Compared to the calculated gestational age, the clinical estimate of gestation has a very strong peak at 39 weeks ('term' birth). A 2015 study (National Vital Statistics Reports, Vol. 64 No. 5 June 1, 2015) found that the two gestational ages agree exactly for 62% of the births and within one week for 83%. Agreement is best for babies born near normal term (38-42 weeks). The clinical estimate exhibits lower levels of preterm and post term births and higher levels of births at full term.

Calculated age: The gestational age in weeks is calculated by subtracting the date of last normal menses from the birth date, dividing by 7 and truncating the result to eliminate decimal places. If the menses day is missing but the month and year are present, a value of '15' is used for the day.

Prior to 2005, if the menses month and/or year were missing or the calculated gestational age was beyond a reasonable range (<18 or >45 weeks), the gestational age was estimated from the child's birth weight.

Currently, if the gestational age cannot be calculated because of missing menses dates or the calculated age is out of range, the clinical estimate is used. If the clinical estimate is also out of range or unknown, the calculated age is unknown. This change makes the Washington State data consistent with data published by NCHS.

For 1980-88, the birth certificate did not collect the clinical estimate. For these years, only the gestational age calculated from the menses date is included in this field. In all other cases, the calculated age is unknown.

For data analysis, **NCHS recommends using the clinical gestational age rather than the calculated estimate.** The gestation flag field identifies which ages were calculated from the menses date and which were imputed from the clinical estimate.

INFANT DEATH FLAG

The infant death flag is set to blank for the current year of data. The infant death flag is set to 'X' if the infant died by one year of age. This information comes from linking infant death records with birth certificate information. This flagging is done post annual birth file release i.e. after one year has elapsed.

INFANT LIVING AT TIME OF REPORT

This item was added to the birth certificate in 2003. A check against death records showed that many of the records marked 'No' did not have a matching death certificate. These records should have been marked 'Y', indicating that the infant was still alive. This problem has been corrected starting with the 2005 data. It was not possible to go back and correct the 2003 and 2004 data. Therefore, for these 2 years, the infant death flag should be used instead to study birth data for infants who have died.

MATERNAL SMOKING (NUMBER OF CIGARETTES/PACKS PER DAY)

This item has undergone wording and placement changes over time. Use caution in doing any trend analysis which spans the change.

- 1984-2003: See [Historical Notes](#)
- 2003: Item revised to collect average number of cigarettes/packs per day three months before pregnancy and by trimester during pregnancy, but placement not changed

MEDICAL AND HEALTH INFORMATION SECTION

This section refers to maternal risk factors, maternal infections, obstetric procedures, characteristics of labor and congenital anomalies of the newborn.

This section has changed considerably over time. Because of these changes, a particular code or box number does not necessarily mean the same thing from year to year. For example, a code '6' for method of delivery means 'Version & Extractions' for 1980-83, 'Repeat C-Section' for 1984-88, 'Forceps' for 1989-2002, and does not have a definition starting in 2003. The Center for Health Statistics has provided standard definitions for the items in this section, to improve comparability among facilities. In studying trends for a particular condition, make sure to select the correct codes for each year (see Data Dictionary).

Changes in individual reporting practices and definitions and in the data collection process affect data in this area. For example:

- **Delivery Method Final:** In 2017, Cesarean Trial Labor Unknown (Delivery Method Final = 6) was added.
- **Surfactant Use:** Surfactant is a substance given to premature infants to help lungs mature. In 2017, there was an increase number of reported use at UW due a reviewed workflow that better captures the actual use of surfactant.

- **Abnormal Conditions of the Newborn:** includes an item for assisted ventilation immediately following delivery. For 2003 and 2004 births occurring at Sacred Heart Hospital, Spokane, this item was overestimated. The box was checked if any little whiff of oxygen was given to the baby. The data have been corrected starting in 2005.
- **1989 -2003 :** See [Historical Notes](#)

MENSES DATE

This item consists of three parts - month, day, and year. Facilities may report part of the date, most often month and year, but not day. The discussion under 'Gestational age' compares gestational ages calculated from the menses date with clinical estimates of gestational age.

MONTH PRENATAL CARE BEGAN

Month prenatal care began has shown improvement in recent years but still has a high percent of unknowns. As of 2003, this field is calculated from the menses date and the date of first prenatal care visit. Because the dates are harder to get, this item now has more unknowns than in the past.

The birth certificate prenatal care data still reliably reflect the prenatal care experience of mothers. This finding came from assessing the quality of medical and health data from the 2003 Birth certificate. This study compared quality of selected medical and health data from two states by comparing birth certificate data with information abstracted from hospital records. Agreement for month and day of first prenatal care ranged from 66.5% to 79.6% for the two states.

Implementation of 2003 birth certificate revision created the need to calculate the month of pregnancy when prenatal care began. The Washington State and the National Center for Health Statistics did this calculation independently, and the numbers differ considerably. Most of the difference is related to the way in which the calculation converts fractional months to whole months. The WACHS formula uses the standard way of converting fractions to whole numbers. The NCHS formula assumes that any fractional value greater than the exact month means that care began in the next month. Thus, for example, a calculated month of 2.1 would be the 2nd month by the WACHS formula and the 3rd month by the NCHS formula. Since most states have now converted to the 2003 U.S. birth certificate we have added the field, '**NCHS Month Prenatal Care Began**' using the NCHS method to the data file for comparability with other revised states. **CHS encourages the use of 'NCHS Month Prenatal Care Began' as all states have now converted to using the same variables.**

OCCUPATION (MILHAM CODES - MOTHER, FATHER)

The Milham occupation codes were developed in Washington State to manually code mother's and father's occupation in 1980-91. Starting in 1992 the computer assigns Milham codes based on the occupation literals where there is an exact match. Records where there is no computer match are not edited or coded, thus only about 93-97% of these records have Milham codes. These codes are not necessarily comparable to Census Bureau occupation and industry codes. **2018 was the last year these fields were included. They have been discontinued.**

PREVIOUS PREGNANCY HISTORY

This includes total prior pregnancies, live births now living and now dead, fetal deaths and other terminations.

These items have undergone several changes since 1980, primarily in the amount of detail (gestational age breakdowns) available for specific outcomes. Refer to the Data Dictionary to see what data are available each year. See [Historical Notes](#) for related coding issues.

RACE/ETHNIC ORIGIN (MOTHER, FATHER, CHILD)

Department of Health provides race and ethnicity information using a series of check boxes in the WHALES system and on the birth filing form. These race and ethnicity options are consistent with the multiple options provided on the U.S. Standard Certificate of Birth (2003). Race and ethnicity are required fields (i.e. an option has to be selected even if the option is ‘unknown’). Parents/Informants may select as many race categories and as many Hispanic sub-categories as appropriate.

Starting with 2003, the file has three race fields and two Hispanic fields for the mother and father. DOH manually corrects spelling errors for literals and checks for valid entries against a master list of “other” race and ethnic subgroups which we update on an ongoing basis. All birth records are forwarded to NCHS where the checkboxes and textboxes are assigned codes that are further processed to give us the individual race variables e.g. (‘Race White’, ‘Race Black’) and derived variables like ‘Bridge Race’ and ‘Race Summary Code’. The imputation procedure to derive bridge race categories is described in detail at https://www.cdc.gov/nchs/data/dvs/Multiple_race_documentation_5-10-04.pdf

NCHS Bridge Race: This is the multiple race data bridged back to a single race by NCHS. The bridging methodology developed by NCHS bridges the multiple-race group population counts to the four single-race categories specified in the revised 1997 federal Office of Management and Budget (OMB) standards. See https://www.cdc.gov/nchs/nvss/bridged_race.htm for details on the bridging process. Bridged race (**Mother Bridge Race**) data are in the old mother’s (**race_mom**), father’s (**race_dad**), and calculated child’s race fields. Bridged Race data are useful when studying trends.

Hispanic NCHS Bridge: The mother’s and father’s race/Hispanic Origin should be used exactly as reported on the Birth Filing Form. Use ‘Hispanic NCHS Bridge’ when it is important to know the country to which the parents have ancestral connection.

Race Summary Code: This is the multiple race data edited by NCHS to remove duplicate entries. See https://www.cdc.gov/nchs/data/dvs/Multiple_race_documentation_5-10-04.pdf for details (this is in .pdf format; you will need to download the free Adobe Acrobat reader to open it). The Center for Health Statistics (CHS) has done further recoding to group the races into the five basic groups and to create single fields for the mother’s and father’s races with a code for each possible race combination. Use these data to study multiple races. In 2016,

about 4% of mothers reported having more than one race. This percentage agrees well with the value of 3.0% based on 2010 Census data.

PLEASE NOTE: Race Summary information is incomplete for out of State Births to Washington Residents for the 2017 Birth File. This was due to an import issue in the new WHALES database. The issue was corrected for starting with the 2018 file.

Race codes – “C” and “E”: The “C” codes are associated with the race literals provided by the parent/informant. The C code matches up with the location of the literal. E.g. the first Native American literal is RACE16. The corresponding code would be located at RACE16C. There are eight “C” codes that match up to eight possible literals, RACE16C through RACE23C. “E” codes include all codes for check boxes and literals, if it is a literal (such as checking Native American) then an E code is assigned. There can be up to eight codes; RACE1E through RACE8E.

Parent and child Hispanic: The check box data were edited by NCHS to take account of ‘other’ entries where codeable detail was provided. These edited data are in the old mother’s (**hisp_mom**), father’s (**hisp_dad**) and calculated child’s Hispanic fields. The Hispanic data do not need to be bridged because only one category was selected.

PLEASE NOTE: Hispanic Ethnicity information is incomplete for out of State Births to Washington Residents for the 2017 Birth File. This was due to an import issue in the new WHALES database. The issue was corrected starting with the 2018 file.

SOURCE OF PAYMENT

Before 2003, the birth certificate did not collect information on the source of payment. The item was added in 2003. The variable went through code changes in 2017 where the code 4 “Indian Health Services” is now “Other Government”.

WEIGHT GAINED DURING PREGNANCY, PREPREGNANCY WEIGHT, AND DELIVERY WEIGHT

In the early version of the electronic data system, a hospital could not directly enter weights of more than 250 lbs. They had to enter 250 lbs. and put the true weight in a comment field. As this comment field was not often used, there was a cluster of births at 250 lbs. Even in the data system weights of over 300 lbs. still had to be entered in the comment field so there may be a clustering of weights at 300 lbs. The upper limit was changed to 350 lbs. in 2003.

As of 2003, weight gain is no longer collected as a separate item. It is calculated from the prepregnancy weight and weight at delivery.

DATA DOCUMENTATION

Data Dictionary

The new [Birth Statistical File - Data Dictionary and Crosswalks](#) includes the new layout starting with the 2017 file, as well as columns indicating the former field names, formats, and labels.

Uniform Parentage Act (UPA)

With the implementation of the new UPA law on January 1st 2019, the data dictionary may change in its future iterations. Any field related to mother and father have the potential to change. We would provide more information as it becomes available. See [Uniform Parentage Act](#) for more information about the law.

HISTORICAL NOTES

This section includes data notes and issues related to information collected prior to 2003.

The Birth certificate filing form and birth filing system have undergone numerous changes over the years. This section documents historical changes that have occurred with the data collection systems, forms and variables.

Birth Certificate and Filing System

The method for entering birth certificate data into the computer has changed:

- 1980-1991: Hospitals or birth attendants complete a paper certificate form and send it to the local health jurisdiction, which then sends it to the state Department of Health, where the data are coded (if necessary) and keyed.
- 1992: Hospitals or birth attendants begin to use an electronic birth certificate program called the Delivery Certificate Tracker (DCT) to enter birth certificate records, which are then sent directly to the Department of Health. Computer programs now electronically code many items.
- Mid-1996: First revision of the DCT
- 1999: Second revision of the DCT

Age (Mother, Father)

Ages are computed from the parent's birth date and the child's date of delivery. Prior to 1989, ages were collected directly from the mother. A comparison of data before and after the change showed that ages calculated from birth dates are consistent with ages collected directly from the mother and that there is no substantial increase in missing data as a result of collecting the more detailed birth dates.

Some facilities still used the old birth certificate data collection form for 1989-1990, and, therefore, did not collect mother's DOB. As a result, mother's DOB and father's DOB data field may be blank for some records for those years.

Alcohol use

Alcohol use data are substantially underreported on birth certificates. A national telephone survey found that only about 5% of women who drink during pregnancy report it on the birth certificate. Because of poor reporting, the 2003 birth certificate no longer collects alcohol use data.

Attendant/Certifier Classification

From mid-1985 to 2016, the file has codes for the classification or title of the certifier and the attendant (e.g., MD, licensed midwife). However, the attendant's data were given only if s/he was different from the certifier. To analyze complete data by attendant type, use the following guidelines:

- Before mid-1985: Use the attendant classification.
- Mid-1985 and after: Select the attendant classification if given. Otherwise, use the certifier classification.

Coding issues for this item are:

- 1987: Changed the meaning of three codes (02, 04, and 05). See the Data Dictionary for details.
- 1996: Incorrectly coded certified midwives (code 05) as nurses (code 06). The data were corrected for major facilities by using a list of certified midwives to reset the code. However, the codes may still be incorrect for some of the smaller facilities.
- 1998: Improved the coding of midwives by using the name of the midwife to assign a code when the title was not given
- 1999: Increase in the number of births with a hospital administrator given as the birth attendant. In these cases the hospital administrator was the certifier but the attendant classification was missing. According to the rules the hospital administrator thus becomes the birth attendant. The Center for Health Statistics is working with facilities to correct this reporting problem. As of 2003, the problem has nearly been eliminated.

Birth Weight

Birth weight is given as grams on the data file. Many scales weigh babies in pounds and ounces, which the computer converts to grams. In 1980-91, even if the weights were reported in grams, they were converted to pounds and ounces at data entry, then reconverted to grams for data analysis. Since one ounce is equivalent to 28 grams, these converted gram weights will cluster in multiples of 28 grams and will probably not be the same as the gram weights originally reported. Starting in 1992, weights in grams were directly entered into the computer, so that values which are not multiples of 28 will be found for these years.

Geography

County coding: In earlier years, the county of residence was based on reporting by the mother. As of 1997, the mother's residence county was also coded by the DOH (Department of Health)

Standard Process for Matching and Geocoding, which uses a variety of matching maps and software to assign a county based on the mother's residence address. The county assigned by the geocoding software differs from the county of residence reported by the mother for a small number of records (< 0.5% of all births). In most instances where differences are found, the geocoded county is correct and, in those instances, it is used in place of the reported county. For 2017 data and forward, county coding is based on autocorrecting and geocoding the reported mother's residence address.

Infant Death Flag

The birth-infant death record linkage program has become more thorough since the mid-1980's. Therefore, some infant deaths in the early 1980's may not be flagged. Interpretation of the infant death data in those years should be done with extreme care.

Maternal Smoking (Number of Cigarettes/Packs per Day)

This item has undergone wording and placement changes over time. Note that data may not be comparable before and after the change. Use caution in doing any trend analysis which spans the change.

- 1984-88: Used wording 'Smoking at any time during the pregnancy' and placed in the middle section of the certificate, which the mother generally completes from a worksheet.
- 1989: Changed wording to 'Tobacco use during pregnancy' (which could include smokeless tobacco) and relocated to the bottom of the certificate, which is generally completed by the physician. The percentage of missing data increased from 4% in 1984 to 13% in 1989, possibly as a result of this change.
- 1992: Changed back to original wording and placement on the certificate

Medical and Health Information Section

This section includes maternal risk factors, maternal infections, obstetric procedures, characteristics of labor and congenital anomalies of the newborn.

This section has changed considerably over time. Changes in individual reporting and definitions affect data in this area.

- In 1989, the positioning of the 'none' box for all of these items was scrambled, so they weren't all lined up across the top of the section, as they had been previously. Thus it was harder to check 'none' straight across for all items. Possibly because of this change, the number of certificates marked 'none' decreased in 1989.
- Before 1996, the DCT did not allow 'unknown' to be entered for any item in this section, so 'none' was often entered instead. The new DCT made it possible to enter 'unknown,' which may have increased the number of unknowns, at the expense of responses coded 'none.' Thus, the change from 'none' to 'unknown' is difficult to interpret. The data are different but it is not clear whether they are better or worse.

- Washington State birth certificates may overestimate two conditions because of reporting practices: Rh sensitization: Hospitals may be reporting Rh incompatibility rather than Rh sensitization (which is rare). Other excessive bleeding: Hospitals may also be misinterpreting the definition of excessive bleeding. Both of these conditions are much more often reported on Washington State birth certificates, compared to US figures.
- Starting in 1999, placenta previa decreased in frequency. Training efforts around the new birth certificate system provided clearer definitions of what should be included with the various items. Thus placenta previa may have been over reported in the past.

The current item 'Risk factors in this pregnancy' was called 'Medical risk factors for this pregnancy' in 1989-2002, 'Complications during or related to pregnancy' in 1984-88, and was two separate items, 'Complications related to pregnancy' and 'Complications not related to pregnancy' in 1980-83. Data for 'Complications related to pregnancy' are included in the 'Risk factor' field and 'Complications not related to pregnancy' are in a separate field.

The item 'Congenital anomalies of child' was an open-ended question in 1980-83, which means that hospital personnel had to write in the appropriate condition. Because open-ended questions generally have less complete reporting than check-box questions, the 1980-83 data for this item may not be comparable to later data.

For 1992-1995 infant deaths due to congenital anomalies, the corresponding birth data were examined to see if the anomaly had been reported on the birth certificate. Reporting rates ranged from 11% for circulatory/respiratory system anomalies to 64% for anencephalies. These percentages give an estimate of how well congenital anomalies are reported on birth certificates. 'Abnormal Conditions of the Newborn' includes an item for 'Assisted ventilation immediately following delivery.'

Previous Pregnancy History

This item has undergone several changes since 1980 and the related coding issues for this item are:

- 1985-88: Did not collect data for the Date of Last Fetal Death. The certificate only collected Date of Last Other Pregnancy Outcome, which could either be a fetal death or an induced termination. For births where the mother had no induced terminations but reported at least one prior fetal death, use the Date of Last Other Pregnancy Outcome as the Date of Last Fetal Death.
- 1989: Edited number of previous pregnancies so that if the sum of the detailed outcomes equals the total number of prior pregnancies specified, then any blanks in the detailed outcomes are replaced with '0'. This edit substantially decreased the number of unknowns caused by facilities not putting '0' where applicable. Users can perform a similar edit in earlier years to reduce the number of unknowns.
- 1999: Eliminated the total number of prior pregnancies as a separate item. It is now computed from the detail supplied. The data are now edited so that if at least one field contains an entry, then any blank field is converted to zero.

Race

The following changes have been made to the collection and coding of race and Hispanic Origin since 1980. Note that data may not be comparable before and after the change. Use caution in doing any trend analysis which spans the change.

- Mid-1988: Expanded race code 'C' (which previously had been used only for Mexican/Chicano) to include all Hispanics
- 1988: Added a separate item for Hispanic Origin to the birth certificate.
- 1990: Changed race coding in accordance with NCHS guidelines. If more than one race was given on the birth certificate, assign the first race listed as the person's race. Before 1990, the first non-white race was used. This change increased the number of births coded as "White."
- 1992: Expanded race codes to provide more detail on Asian Pacific American subgroups (see Data Dictionary). However, a large number of records still report just 'Asian' as the race. Thus, birth rates for Asian subgroups may substantially underestimate the actual rates.
- 1992-2002: Allow the mother to designate a race/Hispanic origin for her child.
- 1996: Modified Hispanic origin coding to comply with National Center for Health Statistics (NCHS) guidelines. Use the person's race or place of birth to provide more specificity on country of origin. See the NCHS instruction manuals cited above for more information.
- 1998: Added specific codes for two Hispanic categories previously classified "Other Hispanic" (code 5): "Boricua" now "Puerto Rican" (code 2), and "La Raza" now "Central or South American" (code 4).
- 2003: Added check boxes for race. Parents may now report more than one race. Discontinued child's designated race/Hispanic origin. See note below.

Weight Gained During Pregnancy, Prepregnancy Weight, and Delivery Weight

Before 2003, the birth certificate asked for the weight gain during pregnancy. The weight gain item had a high percent missing because many hospitals do not have this information on their data collection worksheets. The 1996 revision of the DCT allowed for negative weight gain (weight loss), which in prior years had to be coded '000.'