TECHNICAL APPENDIX FROM

VITAL STATISTICS OF
THE UNITED STATES

1999

NATALITY

U.S. DEPARTMENT OF
HEALTH AND HUMAN SERVICES

CENTERS FOR DISEASE CONTROL AND PREVENTION
NATIONAL CENTER FOR HEALTH STATISTICS
Hyattsville, Maryland: March 2001

VITAL STATISTICS OF THE UNITED STATES: NATALITY, 1999
TECHNICAL APPENDIX

ACKNOWLEDGMENTS

A copy of the technical appendix may be obtained by contacting the National Center for Health Statistics, Reproductive Statistics Branch at 301-458-4111.

VITAL STATISTICS OF THE UNITED STATES: NATALITY, 1999
TECHNICAL APPENDIX

For a list of reports published by the National Center for Health Statistics contact:

Data Dissemination Branch
National Center for Health Statistics
Centers for Disease Control and Prevention
6525 Belcrest Road, Room 1064
Hyattsville, MD 20782-2003
(301) 458-4636
Internet:www.cdc.gov/nchswww/

VITAL STATISTICS OF THE UNITED STATES: NATALITY, 1999
TECHNICAL APPENDIX

VITAL STATISTICS OF THE UNITED STATES: NATALITY, 1999
TECHNICAL APPENDIX

Figure

VITAL STATISTICS OF THE UNITED STATES: NATALITY, 1999
TECHNICAL APPENDIX

Text tables

Population tables

Introduction

This document provides detailed information on the variables and the quality and completeness of the data on

the public-use file for 1999 births, published by the Centers for Disease Control and Prevention's National Center for Health Statistics (1). This report supplements the Technical notes of "Births: Final Data for 1999" (2) and provides a thorough discussion of the definitions, coding, quality and completeness of the 1999 birth data (1). In addition, this report is recommended for use with the public-use file for 1999 births, available on CD-ROM from the National Center for Health Statistics, and the tabulated data of "Vital Statistics of the United States, 1999, Volume I, Natality" (in preparation).

Definition of live birth

Every product of conception that gives a sign of life after birth, regardless of the length of the pregnancy, is considered a live birth. This concept is included in the definition set forth by the World Health Organization in 1950 and revised in 1988 by a working group formed by the American Academy of Pediatrics and the American College of Obstetricians and Gynecologists (3,4,5):

> Live birth is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which, after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered liveborn.

This definition distinguishes in precise terms a live birth from a fetal death (see the section on fetal deaths in the Technical Appendix of volume II, *Vital Statistics of the United States*). In the interest of comparable natality statistics, both the Statistical Commission of the United Nations and the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS) have adopted this definition (6,7).

History of birth-registration area

The national birth-registration area was proposed in 1850 and established in 1915. By 1933 all 48 States and the District of Columbia were participating in the registration system. The organized territories of Hawaii and Alaska were admitted in 1929 and 1950, respectively; data from these areas were prepared separately until they became States--Alaska in 1959 and Hawaii in 1960. Currently the birth-registration system of the United States covers the 50 States, the District of Columbia, the independent registration area of New York City, Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, and the Commonwealth of the Northern Mariana Islands. However, in the statistical tabulations, "United States" refers only to the aggregate of the 50 States (including New York City) and the District of Columbia.

The original birth-registration area of 1915 consisted of 10 States and the District of Columbia. The growth of this area is indicated in table 4-1. This table also presents for each year through 1932 the estimated midyear population of the United States and of those States included in the registration system.

Because of the growth of the area for which data have been collected and tabulated, a national series of geographically comparable data before 1933 can be obtained only by estimation. Annual estimates of births were prepared by P. K. Whelpton for 1909-34 (8). These estimates include adjustments for underregistration and for States that were not part of the birth-registration area before 1933.

Sources of data

Natality statistics

Since 1985 natality statistics for all States and the District of Columbia have been based on information from the total file of records. The information is received on electronic files of individual records processed by the States and provided to NCHS through the Vital Statistics Cooperative Program. NCHS receives these files from the registration offices of all States, the District of Columbia, and New York City. Information for Puerto Rico and the Virgin Islands is also received through the Vital Statistics Cooperative Program. Information for Guam is obtained from microfilm copies of original birth certificates and is based on the total file of records for all years. Data from American Samoa first became available in 1997. Data from the Commonwealth of the Northern Mariana Islands (referred to as Northern Marianas) first became available in 1998. Similar to data from Guam, the data are obtained from microfilm copies of original birth certificates and are based on the total file of records.

Birth statistics for years prior to 1951 and for 1955 are based on the total file of birth records. Statistics for 1951-54, 1956-66, and 1968-71 are based on 50-percent samples except for data for Guam and the Virgin Islands, which are based on all records filed. During the processing of the 1967 data the sampling rate was reduced from 50 percent to 20 percent. For details of this procedure and its consequences for the 1967 data see pages 3-9 to 3-11 in volume I of *Vital Statistics of the United States*, 1967. From 1972 to 1984 statistics are based on all records filed in the States submitting computer tapes and on a 50-percent sample of records in all other States.

Information for years prior to 1970 for Puerto Rico, the Virgin Islands, and Guam is published in the annual vital statistics reports of the Department of Health of the Commonwealth of Puerto Rico, the Department of Public Health of the Virgin Islands, the Department of Public Health and Social Services of the Government of Guam, and in selected *Vital Statistics of the United States* annual reports.

U.S. natality data are limited to births occurring within the United States, including those occurring to U.S. residents and nonresidents. Births to nonresidents of the United States have been excluded from all tabulations by place of residence beginning in 1970 (for further discussion see "Classification by occurrence and residence"). Births occurring to U.S. citizens outside the United States are not included in any tabulations in this report. Similarly the data for Puerto Rico, the Virgin Islands, Guam, American Samoa, and the Northern Marianas are limited to births registered in these areas.

Standard certificate of live birth

The U.S. Standard Certificate of Live Birth, issued by the Public Health Service, has served for many years as the principal means of attaining uniformity in the content of the documents used to collect information on births in the United States. It has been modified in each State to the extent required by the particular State's needs or by special provisions of the State's vital statistics law. However, most State certificates conform closely in content to the standard certificate.

The first standard certificate of birth was developed in 1900. Since then, it has been revised periodically by the national vital statistics agency through consultation with State health officers and registrars; Federal agencies concerned with vital statistics; national, State, and county medical societies; and others working in public health, social welfare, demography, and insurance. This procedure has assured careful evaluation of each item for its current and future usefulness for legal, medical, demographic, and research purposes. New items have been added when necessary, and old items have been modified to ensure better reporting or, in some cases, dropped when their usefulness appeared to be limited.

1989 revision--Effective January 1, 1989, a revised U.S. Standard Certificate of Live Birth (figure 4-A) replaced the 1978 revision. This revision provided a wide variety of new information on maternal and infant health characteristics, representing a significant departure from previous versions in both content and format. The most significant format change was the use of check boxes to obtain detailed medical and health information about the mother and child. It has been demonstrated that this format produces higher quality and more complete information than do open-ended items.

The reformatted items included "Medical Risk Factors for This Pregnancy," which combines the former items "Complications of Pregnancy" and "Concurrent Illnesses or Conditions Affecting the Pregnancy." "Complications of Labor and/or Delivery" and "Congenital Anomalies of Child" also have been revised from the open-ended format.

For each of these items at least 15 specific conditions have been identified.

Several new items were added to the revised certificate. Included are items to obtain information on tobacco and alcohol use during pregnancy, weight gain during pregnancy, obstetric procedures, method of delivery, and abnormal conditions of the newborn. These items can be used to monitor the health practices of the mother that can affect pregnancy and the use of technology in childbirth, and to identify babies with specific abnormal conditions. When combined with other socioeconomic and health data, these items provide a wealth of information relevant to the etiology of low birth weight and other adverse pregnancy outcomes.

Another modification was the addition of a Hispanic identifier for the mother and father. Although NCHS had recommended that States add items to identify the Hispanic or ethnic origin of the newborn's parents, concurrent with the 1978 revision of the U.S. Standard Certificate of Live Birth and reported data from the cooperating States since that year, the items were new to the U.S. Standard Certificate for 1989.

The 1989 revised certificate also provided more detail than previously requested on the birth attendant and place of birth. This permits a more in-depth analysis of the number and characteristics of births by attendant and type of facility and a comparison of differences in outcome. For further discussion see individual sections for each item.

Classification of data

One of the principal values of vital statistics data is realized through the presentation of rates that are computed by relating the vital events of a class to the population of a similarly defined class. Vital statistics and population statistics, therefore, must be classified according to similarly defined systems and tabulated in comparable groups. Even when the variables common to both, such as geographic area, age, race, and sex, have been similarly classified and tabulated, differences between the enumeration method of obtaining population data and the registration method of obtaining vital statistics data may result in significant discrepancies.

The general rules used to classify geographic and personal items for live births are set forth in "Vital Statistics Classification and Coding Instructions for Live Birth Records, 1999," *NCHS Instruction Manual*, Part 3a (9). This material is incorporated in the basic file layout on the CD-ROM. The instruction materials are for States to use in coding the data items; they do not include any NCHS recodes. So, the file layout is a better source of information, since it provides the exact codes and recodes that are available. The classification of certain important items is discussed in the following pages. See table A for a listing of items and the percent of records that were not stated for each State, Puerto Rico, Virgin Islands, Guam, American Samoa, and the Northern Marianas.

Classification by occurrence and residence

Births to U.S. residents occurring outside this country are not reallocated to the United States. In tabulations by place of residence, births occurring within the United States to U.S. citizens and to resident aliens are allocated to the usual place of residence of the mother in the United States, as reported on the birth certificate. Beginning in 1970 births to nonresidents of the United States occurring in the United States are excluded from these tabulations. From 1966 to 1969 births occurring in the United States to mothers who were nonresidents of the United States were considered as births to residents of the exact place of occurrence; in 1964 and 1965 all such births were allocated to "balance of county" of occurrence even if the birth occurred in a city. The change in coding beginning in 1970 to exclude births to nonresidents of the United States from residence data significantly affects the comparability of data with years before 1970 only for Texas.

For the total United States the tabulations by place of residence and by place of occurrence are not identical. Births to nonresidents of the United States are included in data by place of occurrence but excluded from data by place of residence, as previously indicated. See table B for the number of births by residence and occurrence for the 50 States and the District of Columbia for 1999.

Residence error--A nationwide test of birth-registration completeness in 1950 provided measures of residence error for natality statistics. According to this test, errors in residence reporting for the country as a whole tend to

overstate the number of births to residents of urban areas and to understate the number of births to residents of other areas. This tendency has assumed special importance because of a concomitant development--the increased utilization of hospitals in cities by residents of nearby places--with the result that a number of births are erroneously reported as having occurred to residents of urban areas. Another factor that contributes to this overstatement of urban births is the customary practice of using "city" addresses for persons living outside the city limits. Residence error should be taken into consideration in interpreting data for small areas and for cities. Both birth and infant mortality patterns can be affected.

Incomplete residence--Beginning in 1973 where only the State of residence is reported with no city or county specified and the State named is different from the State of occurrence, the birth is allocated to the largest city of the State of residence. Before 1973 such births were allocated to the exact place of occurrence.

Geographic classification

The rules followed in the classification of geographic areas for live births are contained in the instruction manual mentioned previously. The geographic code structure for 1999 is given in another manual, "Vital Records Geographic Classification, 1995," *NCHS Instruction Manual*, Part 8 is included with the documentation file on CD-ROM (1). The geographic code structure in use is based on results of the 1990 Census of Population.

United States--In the statistical tabulations, "United States" refers only to the aggregate of the 50 States and the District of Columbia. Alaska has been included in the U.S. tabulations since 1959 and Hawaii since 1960.

Metropolitan statistical areas--The metropolitan statistical areas and primary metropolitan statistical areas (MSA's and PMSA's) used in this report are those established by the U.S. Office of Management and Budget as of April 1, 1990, and used by the U.S. Bureau of the Census (10) except in the New England States.

Except in the New England States, an MSA has either a city with a population of at least 50,000, or a Bureau of the Census urbanized area of at least 50,000 and a total MSA population of at least 100,000. A PMSA consists of a large urbanized county, or cluster of counties, that demonstrates very strong internal economic and social links and has a population over 1 million. When PMSA's are defined, the large area of which they are component parts is designated a Consolidated Metropolitan Statistical Area (CMSA) (11).

In the New England States the U.S. Office of Management and Budget uses towns and cities rather than counties as geographic components of MSA's and PMSA's. NCHS cannot, however, use this classification for these States because its data are not coded to identify all towns. Instead, the New England County Metropolitan Areas (NECMA's) are used. These areas are established by the U.S. Office of Management and Budget (12) and are made up of county units.

Metropolitan and nonmetropolitan counties— Independent cities and counties included in MSA's and PMSA's or NECMA's are included in data for metropolitan counties; all other counties are classified as nonmetropolitan.

Population-size groups--Beginning in 1994 vital statistics data for cities and certain other urban places have been classified according to the population enumerated in the 1990 Census of Population. Data are available for individual cities and other urban places of 100,000 or more population. Data for the remaining areas not separately identified are shown in the tables under the heading "Balance of area" or "Balance of county." Classification of areas for 1982-93 was determined by the population enumerated in the 1980 Census of Population. As a result of changes in the enumerated population between 1980 and 1990, some urban places identified in previous reports are no longer included, and a number of other urban places have been added.

Urban places other than incorporated cities for which vital statistics data are shown in the tabulated data in "Vital Statistics of the United States, Natality" include the following:

C   Each town in New England, New York, and Wisconsin and each township in Michigan, New Jersey, and Pennsylvania that had no incorporated municipality as a subdivision and had either 25,000 inhabitants or more, or a population of 10,000 to 25,000 and a density of 1,000 persons or more per square mile.

C   Each county in States other than those indicated above that had no incorporated municipality within its boundary and had a density of 1,000 persons or more per square mile. (Arlington County, Virginia, is the only

county classified as urban under this rule.)

C   Each place in Hawaii with 10,000 or more population. (There are no incorporated cities in Hawaii.)

Places of less than 100,000 population are not separately identified on the public-use file because of confidentiality limitations.

Race or national origin

     Beginning with the 1989 data year, birth data are tabulated primarily by race of mother. In 1988 and prior years the race or national origin shown in tabulations was that of the newborn child. The race of the child was determined for statistical purposes by an algorithm based on the race of the mother and father as reported on the birth certificate. 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 race of the other parent. When the parents were of different races and neither parent was white, the child was assigned to the race of the father, 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 it was reported. When information on race was missing for both parents, the race of the child was considered not stated and the birth was allocated according to rules discussed on page 4 of the Technical Appendix, volume I, *Vital Statistics of the United States*, 1988. In 1989 the criteria for reporting the race of the parents did not change and continues to reflect the response of the informant (usually the mother). Beginning with the 1992 issue of *Vital Statistics of the United States, Volume I, Natality*, trend data for years beginning with 1980 have been retabulated by race of mother.

     The most important factor influencing the decision to tabulate births by race of the mother was the decennial revision of the U.S. Standard Certificate of Live Birth in 1989. This revision included many more health questions that are directly associated with the mother, including alcohol and tobacco use, weight gain during pregnancy, medical risk factors, obstetric procedures, complications of labor and/or delivery, and method of delivery. Additionally, many of the other items that have been on the birth certificate for more than two decades also relate directly to the mother, for example, marital status, education level, and receipt of prenatal care. It is more appropriate to use the race of the mother than the race of the child in tabulating these items.

     A second factor has been the increasing incidence of interracial parentage. When race is aggregated into the four categories mandated in 1977 by the Office of Management and Budget, the proportion of children born to parents of different races is 5.1 percent, more than double the percent in 1977 (2.0 percent). More than half of these births were to white mothers and fathers of another race (55 percent in 1999). There have been two major consequences of the increasing interracial parentage. One is the effect on birth rates by race. The number of white births under the former procedures has been arbitrarily limited to infants whose parents were both white (or one parent if the race of only one parent was reported). At the same time, the number of births of other races has been arbitrarily increased to include all births to white mothers and fathers of other races. Thus, prior to 1989, if race of mother had been used, birth rates per 1,000 white women in a given age group would have been higher, while comparable rates for black women and women of other races would have been lower. The other consequence of increasing interracial parentage is the impact on the racial differential in various characteristics of births, particularly in cases where there is generally a large racial disparity, such as the incidence of low birthweight. In this instance, the racial differential is larger when the data are tabulated by race of mother rather than by race of child. The same effect has been noted for characteristics such as nonmarital childbearing, preterm births, late or no prenatal care, and low educational attainment of mother.

     The third factor influencing the change is the growing proportion of births with race of father not stated, 14 percent in 1999. Although this proportion has stabilized and declined slightly in the 1990's, it is still higher than in 1979, 11 percent. The high proportion of records with the father's race not reported reflects the increase in the proportion of births to unmarried women; in many cases no information is reported on the father. These births were already assigned the race of the mother because there is no alternative. Tabulating births by race of mother provides a more uniform approach, rather than a necessarily arbitrary combination of parental races.

The change in the tabulation of births by race presents some problems when analyzing birth data by race, particularly trend data. The problem is likely to be acute for races other than white and black.

The categories for race or national origin are "White," "Black," "American Indian" (including Aleuts and Eskimos), "Chinese," "Japanese," "Hawaiian," "Filipino," and "Other Asian or Pacific Islander" (including Asian Indian). Before 1992 there was also an "other" category, which is now combined with the "Not stated" category. Before 1978 the category "Other Asian or Pacific Islander" was not identified separately but included with "Other" races. The separation of this category from "other" allows identification of the category "Asian or Pacific Islander" by combining the new category "Other Asian or Pacific Islander" with Chinese, Japanese, Hawaiian, and Filipino.

Beginning in 1992, NCHS contracted with seven States with the highest API populations to code births to additional API subgroups. The API subgroups include births to Vietnamese, Asian Indian, Korean, Samoan, Guamanian, and other API women. The seven States included in this reporting area are: California, Hawaii, Illinois, New Jersey, New York, Texas, and Washington. At least two-thirds of the U.S. population of each of these additional API groups lived in the seven-State reporting area (13). The data are available on the detailed natality tapes and CD-ROMs beginning with the 1992 data year. An analytic report based on the 1992 data year is also available upon request (14). Minnesota began reporting additional API subgroups in 1996 and Virginia began reporting in 1998. Missouri and West Virginia started reporting in 1999 for a total of 11 reporting states.

The category "White" comprises births reported as white and births where race, as distinguished from Hispanic origin, is reported as Hispanic. Before 1964 all births for which race or national origin was not stated were classified as white. Beginning in 1964 changes in the procedures for allocating race when race or national origin is not stated have changed the composition of this category. (See discussion on "Race or national origin not stated.")

If the race or national origin of an Asian parent is ill-defined or not clearly identifiable with one of the categories used in the classification (for example, if "Oriental" is entered), an attempt is made to determine the specific race or national origin from the entry for place of birth. If the birthplace is China, Japan, or the Philippines, the race of the parent is assigned to that category. When race cannot be determined from birthplace, it is assigned to the category "Other Asian or Pacific Islander."

Race or national origin not stated--If the race of the mother is not defined or not identifiable with one of the categories used in the classification (0.6 percent of births in 1999) and the race of the father is known, the race of the father is assigned to the mother. Where information for both parents is missing, the race of the mother is allocated electronically according to the specific race of the mother on the preceding record with a known race of mother. Data for both parents were missing for only 0.4 percent of birth certificates for 1999. Nearly all statistics by race or national origin for the United States as a whole in 1962 and 1963 are affected by a lack of information for New Jersey, which did not report the race of the parents in those years. Birth rates by race for those years are computed on a population base that excluded New Jersey. For the method of estimating the U.S. population by age, sex, and race excluding New Jersey in 1962 and 1963, see page 4-8 in the Technical Appendix of volume I, *Vital Statistics of the United States*, 1963.

Age of mother

Beginning in 1989 an item on the birth certificate asks for "Date of Birth." In previous years, "Age (at time of this birth)" was requested. Not all States revised this item and therefore the age of mother either is derived from the reported month and year of birth or coded as stated on the certificate. In 1999, the mother's age was reported directly by five States (Kentucky, Nevada, North Dakota, Virginia, and Wyoming) and American Samoa. From 1964 to 1996, the age of mother was edited for 10-49 years. When the age of mother was computed to be under 10 years or 50 years or over, it was considered not stated and was assigned as described below. Beginning in 1997, age of mother is edited for ages 10-54 years. When the age of mother is computed to be under 10 years or 55 years or over, it is considered not stated and was assigned as described below. A review and verification of unedited birth data for 1996 showed that the vast majority of births reported as occurring to women aged 50 years and older were to women aged 50-54 years. The number of births to women 50-54 years is too small for computing age-specific birth rates. These births have been included with births to women 45-49 for computing birth rates.

Age-specific birth rates are based on populations of women by age, prepared by the U.S. Bureau of the Census. In census years the decennial census counts are used. In intercensal years, estimates of the population of women by age are published by the U.S. Bureau of the Census in *Current Population Reports*.

The 1990 Census of Population derived age in completed years as of April 1, 1990, from the responses to questions on age at last birthday and month and year of birth, with the latter given preference. In the 1960, 1970, and the 1980 Census of Population, age was also derived from month and year of birth. "Age in completed years" was asked in censuses before 1960. This was nearly the equivalent of the former birth certificate question, which the 1950 test of matched birth and census records confirms by showing a high degree of consistency in reporting age in these two sources (15). More recently, reporting of maternal age on the birth certificate was compared with reporting of age in a survey of women who had recently given birth. Reporting of age was very consistent between the two sources (16).

Median age of mother--Median age is the value that divides an age distribution into two equal parts, one-half of the values being less and one-half being greater. Median ages of mothers for 1960 to the present have been computed from birth rates for 5-year age groups rather than from birth frequencies. This method eliminates the effects of changes in the age composition of the childbearing population over time. Changes in the median ages from year to year can thus be attributed solely to changes in the age-specific birth rates. Trend data on the median age is shown in table 1-5 of "Vital Statistics of the United States, volume 1, natality (at http://www.cdc.gov/nchs/datawh/statab/unpubd/natality/natab97.htm).

Not stated date of birth of mother– In 1999, age of mother was not reported on 0.02% of the records. Beginning in 1964 birth records with date of birth of mother and/or age of mother not stated have had age imputed according to the age of mother from the previous birth record of the same race and total-birth order (total of fetal deaths and live births). (See "Computer Edits for Natality Data, Effective 1993" NCHS Instruction Manual , Part 12, page 9; available on request from the Division of Vital Statistics.) In 1963 birth records with age not stated were allocated according to the age appearing on the record previously processed for a mother of identical race and parity (number of live births). For 1960-62 not stated ages were distributed in proportion to the known ages for each racial group. Before 1960 this was done for age-specific birth rates but not for the birth frequency tables, which showed a separate category for age not stated.

Age of father

Age of father is derived from the reported date of birth or coded as stated on the birth certificate. If the age is under 10 years, it is considered not stated and grouped with those cases for which age is not stated on the certificate. Information on age of father is often missing on birth certificates of children born to unmarried mothers, greatly inflating the number of "not stated" in all tabulations by age of father. In computing birth rates by age of father, births tabulated as age of father not stated are distributed in the same proportions as births with known age within each 5-year-age classification of the mother. This procedure is followed because, while father's age is missing in 14 percent of the birth certificates in 1999, one third of these were on records where the mother is a teenager. This distribution procedure is done separately by race. The resulting distributions are summed to form a composite frequency distribution that is the basis for computing birth rates by age of father. This procedure avoids the distortion in rates that would result if the relationship between age of mother and age of father were disregarded. Births with age of father not stated are distributed only for rates and means, not for frequency tabulations (2).

Live-birth order and parity

Live-birth order and parity classifications refer to the total number of live births the mother has had including the 1999 birth. Fetal deaths are excluded.

Live-birth order indicates what number the present birth represents; for example, a baby born to a mother who has had two previous live births (even if one or both are not now living) has a live-birth order of three. Parity

indicates how many live births a mother has had. Before delivery a mother having her first baby has a parity of zero and a mother having her third baby has a parity of two. After delivery the mother of a baby who is a first live birth has a parity of one and the mother of a baby who is a third live birth has a parity of three.

Live-birth order and parity are determined from two items on the birth certificate, "Live births now living" and "Live births now dead."

Not stated birth order--Before 1969 if both of these items were blank, the birth was considered a first birth. Beginning in 1969, births for which the pregnancy history items were not completed have been tabulated as live-birth order not stated. As a result of this revised procedure, 22,686 births in 1969 that would have been assigned to the "First birth order" category under the old rules were assigned to the "Not stated" category.

All births tabulated in the "Not stated birth order" category are excluded from the computation of percents. In computing birth rates by live-birth order, births tabulated as birth order not stated are distributed in the same proportion as births of known live-birth order.

Date of last live birth

The date of last live birth was added to the U.S. Standard Certificate of Live Birth in 1968 for the purpose of providing information on child spacing. The interval since the last live birth is the difference between the date of last live birth and the date of present birth.

Beginning in 1995, NCHS ceased to collect information on the date of last live birth and thus the information on interval is only available from birth certificate data from 1968-94.

Educational attainment

Data on the educational attainment of both parents were collected beginning in 1968 and tabulated for publication in 1969 for the first time. Data on educational attainment is currently available only for the mother. Beginning in 1995, NCHS ceased to collect information on the educational attainment of the father and thus the information is available from birth certificate data only for 1969-94.

The educational attainment of the mother is defined as "the number of years of school completed." Only those years completed in "regular" schools are counted, that is, a formal educational system of public schools or the equivalent in accredited private or parochial schools. Business or trade schools, such as beauty and barber schools, are not considered "regular" schools for the purposes of this item. No attempt has been made to convert years of school completed in foreign school systems, ungraded school systems, and so forth, to equivalent grades in the American school system. Such entries are included in the category "not stated."

Women who have completed only a partial year in high school or college are tabulated as having completed the highest preceding grade. For those certificates on which a specific degree is stated, years of school completed is coded to the level at which the degree is most commonly attained; for example, women reporting B.A., A.B., or B.S. degrees are considered to have completed 16 years of school.

Education not stated--The category "Not stated" includes all records in reporting areas for which there is no information on years of school completed as well as all records for which the information provided is not compatible with coding specifications.

Births tabulated as education not stated are excluded from the computations of percents.

Marital status

National estimates of births to unmarried women are based on two methods of determining marital status. For 1994 through 1996, birth certificates in 45 states and the District of Columbia included a question about the mother's marital status. Beginning in 1997, the marital status of women giving birth in California and Nevada is determined by a direct question in the birth registration process. Beginning June 15, 1998, Connecticut discontinued inferring

the mother's marital status and added a direct question on mother's marital status to the State's birth certificate.

In the two States (Michigan and New York) which used inferential procedures to compile birth statistics by marital status in 1999, a birth is inferred as nonmarital if either of these factors is present: a paternity acknowledgment was received or the father's name is missing. In recent years, a number of States have extended their efforts to identify the fathers when the parents are not married in order to enforce child support obligations. The presence of a paternity acknowledgment therefore is the most reliable indicator that the birth is nonmarital in the States not reporting this information directly; this is now the key indicator in the nonreporting States. The inferential procedures in current use represent a substantial departure from the method used before 1980 as well as those used during the 1980's to prepare national estimates of births to unmarried women, before 1980 the incidence of births to unmarried women in States with no direct question on marital status was assumed to be the same as the incidence in reporting States in the same geographic division (17). Inferential procedures in use during the 1980's relied heavily on a comparison of the surnames of the parents and the child to infer the mother's marital status. The procedures now in use depend, as noted above, on very reliable indicators, namely a paternity affidavit or missing information on the father.

The procedures for reporting marital status in California, Nevada, New York City changed beginning January 1, 1997. The methods used to determine marital status and the impact of the procedures on the data were discussed in detail in a previous report (17).

The use of inferential marital status data together with information from a direct question represents an attempt to use related information on the birth certificate to improve the quality of national data as well as to provide data for the individual nonreporting States. An evaluation of this method and its validity for California (the largest nonreporting State until 1997) has been published (18). Because of the continued substantial increases in nonmarital childbearing throughout the 1980's, the data have been intensively evaluated by the Division of Vital Statistics, NCHS (17).

The mother's marital status was not reported in 1999 on 0.03 percent of the birth records in States reporting this information from a direct question. Marital status was imputed as "married" for these records.

When births to unmarried women are reported as second or higher order births, it is not known whether the mother was married or unmarried when the previous deliveries occurred, because her marital status at the time of these earlier births is not available from the birth record.

Rates for 1940 and 1950 are based on decennial census counts. Rates for 1955-97 are based on a smoothed series of population estimates (17,19). Because of sampling error, the original U.S. Bureau of the Census population estimates by marital status fluctuate erratically from year to year; therefore, they have been smoothed so that the rates do not show similar variations. These rates differ from those published in volumes of *Vital Statistics of the United States* before 1969, which were based on the original estimates provided annually by the U.S. Bureau of the Census. Birth rates by marital status for 1971-79 have been revised and differ from rates published before 1980 in volumes of *Vital Statistics of the United States* (see "Computation of rates and other measures").

Place of delivery and attendant at birth

The 1989 revision of the U.S. Standard Certificate of Live Birth included separate categories for freestanding birthing centers, the mother's residence, and clinic or doctor's office as the place of birth. Prior to 1989, place of birth was classified simply as either "In hospital" or "Not in hospital." Births occurring in hospitals, institutions, clinics, centers, or homes were included in the category "In hospital." In this context the word "homes" does not refer to the mother's residence but to an institution, such as a home for unmarried women. Birthing centers were included in either category, depending on each State's assessment of the facility. Beginning in 1989 births occurring in clinics and in birthing centers not attached to a hospital are classified as "Not in hospital." This change in classification may account in part for the lower proportion of "In hospital" births compared with previous years. (The change in classification of clinics should have minor impact because comparatively few births occur in these facilities, but the effect of any change in classification of freestanding birthing centers is unknown.)

Beginning in 1975 the attendant at birth and place of delivery items were coded independently, primarily to

permit the identification of the person in attendance at hospital deliveries. The 1989 certificate includes separate classifications for doctor of medicine (MD), doctor of osteopathy (DO), certified nurse midwife (CNM), other midwife, and other attendants. In earlier certificates births attended by certified nurse midwives were grouped with those attended by lay midwives. The 1989 certificate also facilitated the identification of home births, births in freestanding birthing centers, and births in clinics or physician offices.

Data for the "In hospital" category for 1975-88 include all births in clinics or maternity centers, regardless of the attendant. Data for 1975-77 published before 1980 included clinic and center births in the category "In hospital" only when the attendant was a physician. Therefore, data shown for 1975-77 published after 1980 differ from data published before 1980. As a result of this change, for 1975 an additional 12,352 births were classified as occurring in hospitals, raising the percent of births occurring in hospitals from 98.7 to 99.1. Similarly, for 1976 the number of births occurring in hospitals increased by 14,133 and the percent in hospitals raised from 98.6 to 99.1; for 1977 the increase is 15,937 and the percent in hospitals raised from 98.5 to 99.0. For 1974 and earlier the "In hospital" category includes all births in hospitals or institutions and births in clinics, centers, or maternity homes only when attended by physicians.

The "Not in hospital" category includes births for which no information is reported on place of birth. Before 1975 births for which the stated place of birth was a "doctor's office" and delivery was by a physician were included in the category ``In hospital." Beginning in 1975 these births were tabulated as "Not in hospital" and included with births delivered by physicians in this category. Although the actual number of such births is unknown, the effect of the change is minimal. In 1974, 0.3 percent of all births were delivered by physicians outside of hospitals; in 1975 this proportion was 0.4 percent.

Babies born on the way to or on arrival at the hospital are classified as having been born in the hospital. This may account for some of the hospital births not delivered by physicians or midwives.

Beginning in 1993, all in-hospital births occurring in Illinois where the attendant was classified as an "other" midwife were changed to certified nurse-midwife. This was necessary because almost all of these births were delivered by midwives certified by the American College of Nurse Midwives but because Illinois does not certify midwives, many of these births were classified as "other" midwives.

Procedures in some hospitals may require that a physician be listed as the attendant for every birth and that a physician sign each birth certificate, even if the birth is attended by a midwife and no physician is physically present. Therefore, the number of live births attended by midwives may be understated in some areas.

Birthweight

Birthweight is reported in some areas in pounds and ounces rather than in grams. However, the metric system has been used in tabulating and presenting the statistics to facilitate comparison with data published by other groups. The categories for birthweight were changed in 1979 to be consistent with the recommendations in the *Ninth Revision of the International Classification of Diseases* (ICD-9) and remain the same for the Tenth Revision of the International Classification of Diseases (ICD-10) (4). The categories in gram intervals and their equivalents in pounds and ounces are as follows:

Less than 500 grams = 1 lb 1 oz or less
500-999 grams = 1 lb 2 oz-2 lb 3 oz
1,000-1,499 grams = 2 lb 4 oz-3 lb 4 oz
1,500-1,999 grams = 3 lb 5 oz-4 lb 6 oz
2,000-2,499 grams = 4 lb 7 oz-5 lb 8 oz
2,500-2,999 grams = 5 lb 9 oz-6 lb 9 oz
3,000-3,499 grams = 6 lb 10 oz-7 lb 11 oz

3,500-3,999 grams = 7 lb 12 oz-8 lb 13 oz
4,000-4,499 grams = 8 lb l4 oz-9 lb l4 oz
4,500-4,999 grams = 9 lb 15 oz-11 lb 0 oz
5,000 grams or more = 11 lb l oz or more

The ICD-9 defines low birthweight as less than 2,500 grams. This is a shift of 1 gram from the previous criterion of 2,500 grams or less, which was recommended by the American Academy of Pediatrics in 1935 and adopted in 1948 by the World Health Organization in the *Sixth Revision of the International Lists of Diseases and Causes of Death*.

After data classified by pounds and ounces are converted to grams, median weights are computed and rounded before publication. To establish the continuity of class intervals needed to convert pounds and ounces to grams, the end points of these intervals are assumed to be half an ounce less at the lower end and half an ounce more at the upper end. For example, 2 lb 4 oz-3 lb 4 oz is interpreted as 2 lb 3 ½ oz-3 lb 4 ½ oz.

Births for which birthweight is not reported are excluded from the computation of percents and medians.

Period of gestation

The period of gestation is defined as beginning with the first day of the last normal menstrual period (LMP) and ending with the day of the birth. The LMP is used as the initial date because it can be more accurately determined than the date of conception, which usually occurs 2 weeks after the LMP.

Births occurring before 37 completed weeks of gestation are considered to be "preterm" or "premature" for purposes of classification. At 37-41 weeks gestation, births are considered to be "term," and at 42 completed weeks and over, "postterm." These distinctions are according to the ICD-9 and ICD-10 (4) definitions.

The 1989 revision of the U.S. Standard Certificate of Live Birth included a new item, "clinical estimate of gestation," that is being compared with length of gestation computed from the LMP date when the latter appears to be inconsistent with birthweight. This is done for normal weight births of apparently short gestations and very low birthweight births reported to be full term. The clinical estimate also was used if the date of the LMP was not reported. The period of gestation for 5.1 percent of the births in 1999 was based on the clinical estimate of gestation. For 97 percent of these records the clinical estimate was used because the LMP date was not reported. For the remaining 3 percent the clinical estimate was used because it was compatible with the reported birth weight, whereas the LMP-computed gestation was not. In cases where the reported birthweight was inconsistent with both the LMP-computed gestation and the clinical estimate of gestation, the LMP-computed gestation was used if it was within 5 weeks of the clinical estimate and birth weight was reclassified as "not stated." This was necessary for 336 births, less than 0.01 percent of all birth records in 1999. If the reported birthweight was inconsistent with both the LMP-computed gestation and the clinical estimate of gestation, gestation and birthweight were classified as "not stated" if the LMP-computed gestation was not within 5 weeks of the clinical estimate. These changes result in only a very small discontinuity in the data.

Before 1981 the period of gestation was computed only when there was a valid month, day, and year of LMP. However, length of gestation could not be determined from a substantial number of live-birth certificates each year because the day of LMP was missing. Beginning in 1981 weeks of gestation have been imputed for records with missing day of LMP when there is a valid month and year. Each such record is assigned the gestational period in weeks of the preceding record that has a complete LMP date with the same computed months of gestation and the same 500-gram birthweight interval. The effect of the imputation procedure is to increase slightly the proportion of preterm births and to lower the proportion of births at 39, 40, 41, and 42 weeks of gestation. A more complete discussion of this procedure and its implications is presented in a previous report (20).

Because of postconception bleeding or menstrual irregularities, the presumed date of LMP may be in error. In these instances the computed gestational period may be longer or shorter than the true gestational period, but the extent of such errors is unknown.

Month of pregnancy prenatal care began

For those records in which the name of the month is entered for this item, instead of first, second, third, and so forth, the month of pregnancy in which prenatal care began is determined from the month named and the month last normal menses began. For these births, if the item "Date last normal menses began" is not stated, the month of pregnancy in which prenatal care began is tabulated as not stated.

Number of prenatal visits

Tabulations of the number of prenatal visits were presented for the first time in 1972. Beginning in 1989 these data were collected from the birth certificates of all States. Percent distributions and the median number of prenatal visits exclude births to mothers who had no prenatal care.

Apgar score

The 1- and 5-minute Apgar scores were added to the U.S. Standard Certificate of Live Birth in 1978 to evaluate the condition of the newborn infant at 1 and 5 minutes after birth. The Apgar score is a useful measure of the need for resuscitation and a predictor of the infant's chances of surviving the first year of life. It is a summary measure of the infant's condition based on heart rate, respiratory effort, muscle tone, reflex irritability, and color. Each of these factors is given a score of 0, 1, or 2; the sum of these 5 values is the Apgar score, which ranges from 0 to 10. A score of 10 is optimum, and a low score raises some doubts about the survival and subsequent health of the infant. Beginning in 1995, NCHS only collected information on the 5-minute Apgar score. Since 1991, the reporting area for the 5-minute Apgar score has been comprised of 48 States and the District of Columbia, accounting for 78 percent of all births in the United States in 1999. California and Texas did not have information on Apgar scores on their birth certificate.

Tobacco and alcohol use during pregnancy

The checkbox format allows for classification of a mother as a smoker or drinker during pregnancy and for reporting the average number of cigarettes smoked per day or drinks consumed per week. When smoking and/or drinking status is not reported or is inconsistent with the quantity of cigarettes or drinks reported, the status is changed to be consistent with the amount reported. For example, if the drinking status is reported as "no" but one or more average drinks a week are reported, the mother is classified as a drinker. If the number of cigarettes smoked per day is reported as one or more, the mother is considered a smoker. When one (or a fraction of one) drink a week is recorded, the mother is classified as a drinker. For records on which the number of drinks or number of cigarettes is reported as a span, for example, 10-15, the lower number is used. The number of drinkers and number of drinks reported on birth certificates are believed to underestimate actual alcohol use.

For 1999, information on number of cigarettes smoked per day was reported in a consistent manner for 46 States, the District of Columbia, and New York City (figure 4-A). Indiana and New York State (except for New York City) reported this information but in a format that was inconsistent with the NCHS standards. This reporting area accounted for 87 percent of all births in the U.S. in 1999. Information was not available for California and South Dakota.

Weight gained during pregnancy

Weight gain is reported in pounds. A loss of weight is reported as zero gain. Computations of median weight gain were based on ungrouped data. This item was included on the certificates of 49 States and the District of Columbia; California did not report this information. This reporting area excluding California accounted for 87 percent of all births in the United States in 1999.

Medical risk factors for this pregnancy

An item on medical risk factors was included on the 1989 birth certificate, but 2 States did not report all of the

16 risk factors in 1999. Texas did not report genital herpes or uterine bleeding, and Kansas did not report Rh sensitization.

The format allows for the designation of more than one risk factor and includes a choice of "None." Accordingly, if the item is not completed, it is classified as "Not stated."

The following definitions are adapted and abbreviated from a set of definitions compiled by a committee of Federal and State health statistics officials for the Association for Vital Records and Health Statistics (21).

Definitions of medical terms:

Anemia--Hemoglobin level of less than 10.0 g/dL during pregnancy or a hematocrit of less than 30 percent during pregnancy.

Cardiac disease--Disease of the heart.

Acute or chronic lung disease--Disease of the lungs during pregnancy.

Diabetes--Metabolic disorder characterized by excessive discharge of urine and persistent thirst; includes juvenile onset, adult onset, and gestational diabetes during pregnancy.

Genital herpes--Infection of the skin of the genital area by herpes simplex virus.

Hydramnios/oligohydramnios--Any noticeable excess (hydramnios) or lack (oligohydramnios) of amniotic fluid.

Hemoglobinopathy--A blood disorder caused by alteration in the genetically determined molecular structure of hemoglobin (for example, sickle cell anemia).

Hypertension, chronic--Blood pressure persistently greater than 140/90, diagnosed prior to onset of pregnancy or before the 20th week of gestation.

Hypertension, pregnancy-associated--An increase in blood pressure of at least 30 mm Hg systolic or 15 mm Hg diastolic on two measurements taken 6 hours apart after the 20th week of gestation.

Eclampsia--The occurrence of convulsions and/or coma unrelated to other cerebral conditions in women with signs and symptoms of pre-eclampsia.

Incompetent cervix--Characterized by painless dilation of the cervix in the second trimester or early in the third trimester of pregnancy, with prolapse of membranes through the cervix and ballooning of the membranes into the vagina, followed by rupture of membranes and subsequent expulsion of the fetus.

Previous infant 4,000+ grams--The birthweight of a previous live-born child was over 4,000 grams (8 lbs 13 oz).

Previous preterm or small-for-gestational-age infant--Previous birth of an infant prior to term (before 37 completed weeks of gestation) or of an infant weighing less than the 10th percentile for gestational age using a standard weight-for-age chart.

Renal disease--Kidney disease.

Rh sensitization--The process or state of becoming sensitized to the Rh factor as when an Rh-negative woman is pregnant with an Rh-positive fetus.

Uterine bleeding--Any clinically significant bleeding during the pregnancy, taking into consideration the stage of pregnancy; any second or third trimester bleeding of the uterus prior to the onset of labor.

Obstetric procedures

This item includes six specific obstetric procedures. Birth records with "Obstetric procedures" left blank are considered "not stated."   Data on obstetric procedures were reported by all States and the District of Columbia in 1999.

The following definitions are adapted and abbreviated from a set of definitions compiled by a committee of Federal and State health statistics officials for the National Association for Public Health Statistics and Information Systems (NAPHSIS), formerly the Association for Vital Records and Health Statistics (21).

Definitions of medical terms:

Amniocentesis--Surgical transabdominal perforation of the uterus to obtain amniotic fluid to be used in the detection of genetic disorders, fetal abnormalities, and fetal lung maturity.

Electronic fetal monitoring--Monitoring with external devices applied to the maternal abdomen or with internal devices with an electrode attached to the fetal scalp and a catheter through the cervix into the uterus, to detect and record fetal heart tones and uterine contractions.

Induction of labor--The initiation of uterine contractions before the spontaneous onset of labor by medical and/or surgical means for the purpose of delivery.

Stimulation of labor--Augmentation of previously established labor by use of oxytocin.

Tocolysis--Use of medications to inhibit preterm uterine contractions to extend the length of pregnancy and therefore avoid a preterm birth.

Ultrasound--Visualization of the fetus and placenta by means of sound waves.

Complications of labor and/or delivery

The checkbox format allows for the selection of 15 specific complications and for the designation of more than 1 complication where appropriate. A choice of "None" is also included. Accordingly, if the item is not completed, it is classified as "not stated."

All States and the District of Columbia included this item on their birth certificates in 1999. However, Texas did not report all of the complications. Texas did not report anesthetic complications or fetal distress.

The following definitions are adapted and abbreviated from a set of definitions compiled by a committee of Federal and State health statistics officials (21).

Definitions of medical terms:

Febrile--A fever greater than 100 degrees F. or 38 C. occurring during labor and/or delivery.

Meconium, moderate/heavy--Meconium consists of undigested debris from swallowed amniotic fluid, various products of secretion, excretion, and shedding by the gastrointestinal tract; moderate to heavy amounts of meconium in the amniotic fluid noted during labor and/or delivery.

Premature rupture of membranes (more than 12 hours)--Rupture of the membranes at any time during pregnancy and more than 12 hours before the onset of labor.

Abruptio placenta--Premature separation of a normally implanted placenta from the uterus.

Placenta previa--Implantation of the placenta over or near the internal opening of the cervix.

Other excessive bleeding--The loss of a significant amount of blood from conditions other than abruptio placenta or placenta previa.

Seizures during labor--Maternal seizures occurring during labor from any cause.

Precipitous labor (less than 3 hours)--Extremely rapid labor and delivery lasting less than 3 hours.

Prolonged labor (more than 20 hours)--Abnormally slow progress of labor lasting more than 20 hours.

Dysfunctional labor--Failure to progress in a normal pattern of labor.

Breech/malpresentation--At birth, the presentation of the fetal buttocks rather than the head, or other malpresentation.

Cephalopelvic disproportion--The relationship of the size, presentation, and position of the fetal head to the maternal pelvis prevents dilation of the cervix and/or descent of the fetal head.

Cord prolapse--Premature expulsion of the umbilical cord in labor before the fetus is delivered.

Anesthetic complications--Any complication during labor and/or delivery brought on by an anesthetic agent or agents.

Fetal distress--Signs indicating fetal hypoxia (deficiency in amount of oxygen reaching fetal tissues).

Abnormal conditions of the newborn

This item provides information on eight specific abnormal conditions. More than one abnormal condition may be reported for a given birth or "None" may be selected. If the item is not completed it is tabulated as "not stated."

This item was included on the birth certificates of all States and the District of Columbia in 1999. However, four areas did not include all conditions. Nebraska and Texas did not report birth injury, New York City did not report assisted ventilation less than 30 minutes or assisted ventilation of 30 minutes or more, and Wisconsin did not report fetal alcohol syndrome.

The following definitions are adapted and abbreviated from a set of definitions compiled by a committee of Federal and State health statistics (21).

Definitions of medical terms:

Anemia--Hemoglobin level of less than 13.0 g/dL or a hematocrit of less than 39 percent.

Birth injury--Impairment of the infant's body function or structure due to adverse influences that occurred at birth.

Fetal alcohol syndrome--A syndrome of altered prenatal growth and development occurring in infants born of women who consumed excessive amounts of alcohol during pregnancy.

Hyaline membrane disease/RDS--A disorder primarily of prematurity, manifested clinically by respiratory distress and pathologically by pulmonary hyaline membranes and incomplete expansion of the lungs at birth.

Meconium aspiration syndrome--Aspiration of meconium by the fetus or newborn, affecting the lower respiratory system.

Assisted ventilation (less than 30 minutes)--A mechanical method of assisting respiration for newborns with respiratory failure.

Assisted ventilation (30 minutes or more)--Newborn placed on assisted ventilation for 30 minutes or longer.

Seizures--A seizure of any etiology.

Congenital anomalies of child

The data provided in this item relate to 21 specific anomalies or anomaly groups. It is well documented that congenital anomalies, except for the most visible and most severe, are incompletely reported on birth certificates (22). The completeness of reporting specific anomalies depends on how easily they are recognized in the short time between birth and birth-registration. Forty-nine States and the District of Columbia included this item on their birth certificates (New Mexico did not). This reporting area included 99 percent of all births in the United States in 1999. The format allows for the identification of more than one anomaly including a choice of "None" should no anomalies be evident. The category "not stated" includes birth records for which the item is not completed.

The following definitions are adapted and abbreviated from a set of definitions compiled by a committee of Federal and State health statistics officials (21).

Definitions of medical terms:

Anencephalus--Absence of the cerebral hemispheres.

Spina bifida/meningocele--Developmental anomaly characterized by defective closure of the bony encasement of the spinal cord, through which the cord and meninges may or may not protrude.

Hydrocephalus--Excessive accumulation of cerebrospinal fluid within the ventricles of the brain with consequent enlargement of the cranium.

Microcephalus--A significantly small head.

Other central nervous system anomalies--Other specified anomalies of the brain, spinal cord, and nervous system.

Heart malformations--Congenital anomalies of the heart.

Other circulatory/respiratory anomalies--Other specified anomalies of the circulatory and respiratory systems.

Rectal atresia/stenosis--Congenital absence, closure, or narrowing of the rectum.

Tracheo-esophageal fistula/Esophageal atresia--An abnormal passage between the trachea and the esophagus; esophageal atresia is the congenital absence or closure of the esophagus.

Omphalocele/gastroschisis--An omphalocele is a protrusion of variable amounts of abdominal viscera from a midline defect at the base of the umbilicus. In gastroschisis, the abdominal viscera protrude through an abdominal wall defect, usually on the right side of the umbilical cord insertion.

Other gastrointestinal anomalies--Other specified congenital anomalies of the gastrointestinal system.

Malformed genitalia--Congenital anomalies of the reproductive organs.

Renal agenesis--One or both kidneys are completely absent.

Other urogenital anomalies--Other specified congenital anomalies of the organs concerned in the production and excretion of urine, together with organs of reproduction.

Cleft lip/palate--Cleft lip is a fissure of elongated opening of the lip; cleft palate is a fissure in the roof of the mouth. These are failures of embryonic development.

Polydactyly/syndactyly/adactyly--Polydactyly is the presence of more than five digits on either hands and/or feet; syndactyly is having fused or webbed fingers and/or toes; adactyly is the absence of fingers and/or toes.

Club foot--Deformities of the foot, which is twisted out of shape or position.

Diaphragmatic hernia-- Herniation of the abdominal contents through the diaphragm into the thoracic cavity usually resulting in respiratory distress.

Other musculoskeletal/integumental anomalies--Other specified congenital anomalies of the muscles, skeleton, or skin.

Down's syndrome--The most common chromosomal defect with most cases resulting from an extra chromosome (trisomy 21).

Other chromosomal anomalies--All other chromosomal aberrations.

Method of delivery

The birth certificate contains a checkbox item on method of delivery. The choices include vaginal delivery, with the additional options of forceps, vacuum, and vaginal birth after previous cesarean section (VBAC), as well as a choice of primary or repeat cesarean. When only forceps, vacuum, or VBAC is checked, a vaginal birth is assumed. In 1999 this information was collected from the birth certificates of all States and the District of Columbia.

Several rates are computed for method of delivery. The overall cesarean section rate or total cesarean rate is computed as the proportion of all births that were delivered by cesarean section. The primary cesarean rate is a measure that relates the number of women having a primary cesarean birth to all women giving birth who have never had a cesarean delivery. The denominator for this rate is the sum of women with a vaginal birth excluding VBACs and women with a primary cesarean birth. The rate for vaginal birth after previous cesarean (VBAC) delivery is computed by relating all VBAC deliveries to the sum of VBAC and repeat cesarean deliveries, that is, to women with a previous cesarean section. VBAC rates for first births exist because the rates are computed on the basis of previous pregnancies, not just live births.

Hispanic parentage

Concurrent with the 1978 revision of the U.S. Certificate of Live Birth, NCHS recommended that items to identify the Hispanic or ethnic origin of the newborn's parents be included on birth certificates and has tabulated and evaluated these data from the reporting States. The 1989 revision of the U.S. Standard Certificate of Live Births includes items to identify the Hispanic origin of the parents. All 50 States and the District of Columbia reported Hispanic origin of the parents for 1999. In 1989 Louisiana, New Hampshire, and Oklahoma did not report this information; in 1990 New Hampshire and Oklahoma did not report, and in 1991-92 New Hampshire did not report Hispanic origin.

In computing birth and fertility rates for the Hispanic population, births with origin of mother not stated are included with non-Hispanic births rather than being distributed. Thus, rates for the Hispanic population are underestimates of the true rates to the extent that the births with origin of mother not stated (1.2 percent in 1999) were actually to Hispanic mothers. The population with origin not stated was imputed. The effect on the rates is believed to be small.

Quality of data

Although vital statistics data are useful for a variety of administrative and scientific purposes, they cannot be correctly interpreted unless various qualifying factors and methods of classification are taken into account. The factors to be considered depend on the specific purposes for which the data are to be used. It is not feasible to discuss all the pertinent factors in the use of vital statistics tabulations, but some of the more important ones should be mentioned.

Most of the factors limiting the use of data arise from imperfections in the original records or from the impracticability of tabulating these data in very detailed categories. These limitations should not be ignored, but their existence does not lessen the value of the data for most general purposes.

Completeness of registration

An estimated 99 percent of all births occurring in the United States in 1999 were registered; for white births registration was 99.4 percent complete and for all other births, 98.6 percent complete. These estimates are based on the results of the 1964-68 test of birth-registration completeness according to place of delivery (in or out of hospital) and race and on the 1999 proportions of births in these categories. The primary purpose of the test was to obtain current measures of registration completeness for births in and out of hospital by race on a national basis. Data for States were not available as they had been from the previous birth-registration tests in 1940 and 1950. A detailed discussion of the method and results of the 1964-68 birth-registration test is available (23). A more recent test has not been conducted.

The 1964-68 test has provided an opportunity to revise the estimates of birth-registration completeness for the years since the previous test in 1950 to reflect the improvement in registration. This has been done using registration completeness figures from the two tests by place of delivery and race. Estimates of registration completeness for four groups (based on place of delivery and race) for 1951-65 were computed by interpolation between the test results. (It was assumed that the data from the more recent test are for 1966, the midpoint of the test period.) The results of the 1964-68 test are assumed to prevail for 1966 and later years. These estimates were used with the proportions of births registered in these categories to obtain revised numbers of births adjusted for underregistration for each year. The overall percent of birth-registration completeness by race was then computed. Data adjusted for underregistration for 1951-59 have been revised to be consistent with the 1964-68 test results and differ slightly from data shown in annual reports for years before 1969. For these years the published number of births and birth rates for both racial groups have been revised slightly downward because the 1964-68 test indicated that previous adjustments to registered births were slightly inflated. Because registration completeness figures by age of mother and by live-birth order are not available from the 1964-68 test, it must be assumed that the relationships among these variables have not changed since 1950.

Discontinuation of adjustment for underregistration, 1960--Adjustment for underregistration of births was discontinued in 1960 when birth registration for the United States was estimated to be 99.1 percent complete. This removed a bias introduced into age-specific rates when adjusted births classified by age were used. Age-specific rates are calculated by dividing the number of births to an age group of mothers by the population of women in that age group. Tests have shown that population figures are likely to be understated through census undercounts; these errors compensate for underregistration of births. Adjustment for underregistration of births, therefore, removes the compensating effect of under enumeration, biasing the age-specific rates more than when uncorrected birth and population data are used. (For further details see page 4-11 in the Technical Appendix of volume I, Vital Statistics of the United States, 1963.)

The age-specific rates used in the cohort fertility tables are an exception to the above statement. These rates are computed from births corrected for underregistration and population estimates adjusted for under enumeration and misstatement of age. Adjusted birth and population estimates are used for the cohort rates because they are an

integral part of a series of rates, estimated with a consistent methodology. It was considered desirable to maintain consistency with respect to the cohort rates, even though it means that they will not be precisely comparable with other rates shown for 5-year age groups.

Completeness of reporting

Interpretation of these data must include evaluation of item completeness. The percent "not stated" is one measure of the quality of the data. Completeness of reporting varies among items and States. See table A for the percent of birth records on which specified items were not stated.

Quality control procedures

As electronic files are received at NCHS, they are automatically checked for completeness, individual item code validity, and unacceptable inconsistencies between data items. The registration area is notified of any problems. In addition, NCHS staff review the files on an ongoing basis to detect problems in overall quality such as inadequate reporting for certain items, failure to follow NCHS coding rules, and systems and software errors. Traditionally, quality assurance procedures were limited to review and analysis of differences between the NCHS and registration area code assignments for a small sample of records. In recent years, as electronic birth registration became prevalent, this procedure was augmented by analyses of year to year and area to area variations in the data. These analyses are based on preliminary tabulations of the data that are cumulated by state on a year to date basis each month. All differences that are judged to have consequences for quality and completeness are investigated by NCHS. In the review process, statistical tests are used to call initial attention to differences for possible follow-up. As necessary, registration areas are informed of differences encountered in the tables and asked to verify the counts or to determine the nature of the differences. Missing records (except those permanently voided) and other problems detected by NCHS are resolved and corrections transmitted to NCHS in the same manner as for those corrections identified by the registration area.

Random variation and significance testing for natality data

The number of births reported for an area is essentially a complete count, since more than 99 percent of all births are registered. While this number is not subject to sampling error, it may be affected by nonsampling errors such as mistakes in recording the mother's residence or age during the registration process.

When the number of births is used for analytic purposes the number of events that <u>actually</u> occurred can be thought of as one in a large series of possible results that <u>could have</u> occurred under the same circumstances. When considered in this way, the number of births is subject to random variation. The probable range of values may be estimated from the actual figures according to certain statistical assumptions.

The **confidence interval** (CI) is the range of values for the number of births, birth rates, or percent of births that you could expect in 95 out of 100 cases. The **confidence limits** are the end points of this range of values (the highest and lowest values). Confidence limits tell you how much the number of events or rates could vary under similar circumstances.

Confidence limits for numbers, rates, and percents can be estimated from the actual number of events. Procedures differ for rates and percents and also differ depending on the number of births on which these statistics are based. Below are detailed procedures and examples for each type of case.

95 percent Confidence Interval: 100 or more births

When the number of events is greater than 100, the data are assumed to be approximately normally distributed. Formulas for 95-percent confidence limits are:

Lower limit = B ! (1.96 x √B)

Upper limit   =   B + (1.96 x ∕B)
where:
B = the number of births

Example

Suppose the number of first births to white women 40-44 years of age was 14,108. The 95-percent confidence limits for this number would be:

Lower limit   =   14,108 - [1.96 x ∕14,108]
                 =   14,108 - 233
                 =   13,875

Upper limit =   14,108 + [1.96 x ∕14,108]
                 =   14,108 + 233
             =   14,341

This means that the chances are 95 out of 100 that the actual number of first births to white women 40-44 years of age would lie between 13,875 and 14,341.

95 percent Confidence Interval: 1-99 births

When the number of births is less than 100 and the rate is small, the data are assumed to follow a Poisson probability distribution. Confidence limits are estimated using the following formulas:
Lower limit = B x L
Upper limit = B x U
        where:
            B   =   the number of births
            L   =   the value in Table C that corresponds to the number B, using the 95 percent CI column
            U   =   the value in Table C that corresponds to the number B, using the 95 percent CI column

Example

Suppose the number of first births to American Indian women 40-44 years of age was 47. The confidence limits for this number would be:

        Lower limit =   *B* x *L*
                    =   47 x 0.73476
                 =   35

Upper limit =   *B* x *U*
             =   47 x 1.32979
             =   63

This means that the chances are 95 out of 100 that the actual number of first births to American Indian women 40-44 years of age would lie between 35 and 63.

Computing confidence intervals for rates

The same statistical assumptions can be used to estimate the variability in birth rates. Again, one formula is used for rates based on numbers of events less than 100, and another formula for rates based on numbers of 100 or greater. For our purposes, assume that the denominators of these rates (the population estimates) have no error. While this assumption is technically correct <u>only</u> for denominators based on the census which occurs every 10 years, the error in intercensal population estimates is usually small, difficult to measure, and therefore not considered.

95 percent Confidence Interval: 100 or more births

In this case, use the following formula for the birth rate R based on the number of births B:

$$\textit{Lower limit} \ = \ R \ ! \ [\ 1.96 \ x \ R \ / \ \sqrt{B) } \ ]$$
$$\textit{Upper limit} \ = \ R \ + \ [\ 1.96 \ x \ R \ / \ \sqrt{B) } \ ]$$

where:
R = rate (births per 1,000 population)
B = the number of births

Example

Suppose the first birth rate for white women 40-44 years of age was 1.55 per thousand, based on 14,108 births in the numerator. Therefore, the 95-percent confidence interval would be:

Lower limit = 1.55 - [1.96 x (1.55 / $\sqrt{14,108}$)]
= 1.55 - .026
= 1.52

Upper limit = 1.55 + [1.96 x (1.55 / $\sqrt{14,108}$)]
= 1.55 + .026
= 1.58

This means that the chances are 95 out of 100 that the actual first birth rate for white women 40-44 years of age lies between 1.52 and 1.58.

95 percent Confidence Interval: 1-99 births

When the number of events in the numerator is less than 20, an asterisk is shown in place of the rate because there were too few births to compute a statistically reliable rate. When the number of events in the numerator is greater than 20 but less than 100, the confidence interval for a rate can be estimated using the two formulas which follow and the values in the 95 percent CI column of Table C.

$$\textit{Lower limit} \ = \ R \ x \ L$$
$$\textit{Upper limit} \ = \ R \ x \ U$$

where:
R = rate (births per 1,000 population)
L = the value in Table C that corresponds to the number B in the numerator of the rate
U = the value in Table C that corresponds to the number B in the numerator of the rate

Example

Suppose that the first birth rate for American Indian women 40-44 years of age was 0.54 per thousand, based on 47 births in the numerator. Using Table C:

    Lower limit =   0.54 x 0.73476   =   .40
    Upper limit =   0.54 x 1.32979   = .72

This means that the chances are 95 out of 100 that the actual first birth rate for American Indian women 40-44 year of age lies between .40 and .72.

Computing confidence intervals for Hispanic subgroups

<u>Tables 6, 8, 9, and 14 in "Births: Final Data for 1999" and tables 1-4 and 1-12 in Vital Statistics of the United States, part 1 Natality</u> show birth and fertility rates for Mexicans, Puerto Ricans, Cubans, and "Other" Hispanics. Population estimates are derived from the U.S. Bureau of the Census' Current Population Survey and adjusted to resident population control totals. As a result, the rates are subject to the variability of the denominator as well as the numerator. For these Hispanic subgroups only (not for all origin, total Hispanic, total non-Hispanic, non-Hispanic white, or non-Hispanic black populations), the formulas above would be substituted by the following formulas:

Approximate 95 percent Confidence Interval: 100 or more births

When the number of events in the numerator is greater than 100, the confidence interval for the birth rate can be estimated from the following formulas:

For crude and age-specific birth rates,

$$\textit{Lower limit}: \quad R - 1.96 \left(R \left(\sqrt{ \left(\frac{1}{B} \right) + f \left(a + \frac{b}{P} \right) } \right. \right.$$

$$\textit{Upper limit}: \quad R + 1.96 \left(R \left(\sqrt{ \left(\frac{1}{B} \right) + f \left(a + \frac{b}{P} \right) } \right. \right.$$

where

R = rate (births per 1,000 population).
B = total number of births upon which rate is based
f = factor that depends on whether the population estimate is based on demographic analysis or CPS and the number of years used, equals 0.670 for single year.
a and b are single year averages of the 1998 and 1999 CPS standard error parameters; a equals -0.000238 and b equals 7,486 (24,25).
P = total estimated population upon which rate is based

Example

Suppose that the fertility rate of Cuban American women 15-44 years of age was 51.2 per thousand based on 13,088 births in the numerator and an estimated resident population of 255,399 in the denominator. The 95 percent confidence interval would be:

$$Lower\ limit = 51.2 - 1.96 * 51.2 * \sqrt{\left(\frac{1}{13,088}\right) + 0.670\left[-0.000238 + \left(\frac{7,486}{255,399}\right)\right]}$$

$$= 51.2 - 1.96 * 51.2 * \sqrt{0.000076405 + (0.670 * 0.029073)}$$

$$= 51.2 - 1.96 * 51.2 * \sqrt{0.019555}$$

$$= 51.2 - 1.96 * 51.2 * 0.13984$$

$$= 37.17$$

$$Upper\ limit = 51.2 + 1.96 * 51.2 * \sqrt{\left(\frac{1}{13,088}\right) + 0.670\left[-0.000238 + \left(\frac{7,486}{255,399}\right)\right]}$$

$$= 51.2 + 1.96 * 51.2 * \sqrt{0.000076405 + (0.670 * 0.029073)}$$

$$= 51.2 + 1.96 * 51.2 * \sqrt{0.019555}$$

$$= 51.2 + 1.96 * 51.2 * 0.13984$$

$$= 65.23$$

This means that the chances are 95 out of 100 that the actual fertility rate of Cuban American women 15-44 years of age lies between 37.17 and 65.23.

Approximate 95 percent Confidence Interval: 1-99 births

When the number of events in the numerator is less than 20, an asterisk is shown in place of the rate. When the number of events in the numerator is greater than 20 but less than 100, the confidence interval for the birth rate can be estimated using the formulas which follow and the values in Table C.

For crude and age-specific birth rates,

$$Lower:\quad R \ (\ L \ (1\&a' \ .96, \ B) \ (\left(1\&2.576\sqrt{f\left(a \ \% \ \frac{b}{P}\right)}\right)$$

$$Upper:\quad R \ (\ U \ (1\&a' \ .96, \ B) \ (\left(1\%2.576\sqrt{f\left(a \ \% \ \frac{b}{P}\right)}\right)$$

where

R = rate (births per 1,000 population).
B = total number of births upon which rate is based.
L = the value in Table C that corresponds to the number B, using the 96 percent CI column
U = the value in Table C that corresponds to the number B, using the 96 percent CI column
f = factor that depends on whether the population estimate is based on demographic analysis or CPS and the number of years used, equals 0.670 for single year.
a and b factors are CPS standard error parameters. (see previous section on 95 percent confidence interval for 100 or more births for description and specific values)
P = total estimated population upon which rate is based.

Example

Suppose that the birth rate of Puerto Rican American women 45-49 years of age was 0.4 per thousand, based on 35 births in the numerator and an estimated resident population of 87,892 in the denominator. Using Table C, the 95 percent confidence interval would be:

$$\text{Lower limit} = 0.4 * 0.68419 * \left(1 - 2.576\sqrt{0.670\left(-0.000238 + \left(\frac{7,486}{87,892}\right)\right)}\right)$$

$$=0.4*0.68419*(1-2.576\text{\textbf{/}}.056906)$$
$$= 0.4 *0.68419*(1-2.576*0.23855)$$
$$= 0.4 * 0.68419 * 0.38549$$
$$= 0.1$$

$$\text{Upper limit} = 0.4 * 1.41047 * \left(1 + 2.576\sqrt{0.670\left(-0.000238 + \left(\frac{7,486}{87,892}\right)\right)}\right)$$

$$=0.4*1.41047*(1+2.576\text{\textbf{/}}.056906)$$

$$=0.4*1.41047*(1+2.576*0.23855)$$
$$= 0.4 * 1.41047 * 1.61451$$
$$= 0.9$$

This means that the chances are 95 out of 100 that the actual birth rate of Puerto Rican American women 45-49 years of age lies between 0.1 and 0.9.

Note: In the formulas above, the confidence limits are estimated from the nonsampling error in the number of births, the numerator, and the sampling error in the population estimate, the denominator. A 96 percent standard error is computed for the numerator and a 99 percent standard error is computed for the denominator in order to compute a 95 percent confidence interval for the rate.

Computing 95 percent Confidence Intervals for percents

In many instances we need to compute the confidence intervals for percents. Percents derive from a binomial distribution. As with birth rates, an asterisk will be shown for any percent which is based on fewer than 20 births in the numerator. We easily compute a 95-percent confidence interval for a percent when the following conditions are met:

$$B \ x \ p \ \$5 \quad \text{and}$$
$$B \ x \ q \ \$5$$

where:
B = number of births in the denominator
p = percent divided by 100
q = $1 - p$

For natality data, these conditions will be met except for very rare events in small subgroups. If the conditions are *not* met, the variation in the percent will be so large as to render the confidence intervals meaningless. When these conditions are met the 95-percent confidence interval can be computed using the normal approximation of the binomial. The 95-percent confidence intervals are computed by the following formulas:

$$\text{Lower limit} = p - \left(1.96 * \sqrt{p * \frac{q}{B}} \right)$$

$$\text{Upper limit} = p + \left(1.96 * \sqrt{p * \frac{q}{B}} \right)$$

where:
B = number of births in the denominator
p = percent divided by 100
q = $1 - p$

Example

Suppose the percent of births to Hispanic women in Alabama that were to unmarried women was 23.0 percent. This was based on 310 births in the numerator and 1,345 births in the denominator. First we test to make sure we can use the normal approximation of the binomial:

$$1,345 \ x \ .230 = 309$$
$$1,345 \ x \ (1 - .230)$$
$$1,345 \ x \ .770 = 1,036$$

Both 309 and 1,036 are greater than 5 so we can proceed. The 95-percent confidence interval would be:

$$\text{Lower limit} = 0.23 - \left(1.96 * \sqrt{0.23 * \frac{0.77}{1,345}} \right)$$

$$= 0.23 - 0.022$$

$$= 0.208 \ \text{or} \ 20.8 \ \text{percent}$$

$$\text{Upper limit} = 0.23 + \left(1.96 * \sqrt{0.23 * \frac{0.77}{1,345}} \right)$$

$$= 0.23 + 0.022$$

$$= 0.252 \text{ or } 25.2 \text{ percent}$$

This means that the chances are 95 out of 100 that the actual percent of births in Alabama to Hispanic women that are to unmarried women lies between 20.8 and 25.2 percent.

Significance testing

Both rates are based on 100 or more events

When both rates are based on 100 or more events, the difference between the two rates is considered statistically significant if it exceeds the statistic in the formula below. This statistic equals 1.96 times the standard error for the difference between two rates.

$$1.96 \sqrt{\frac{R_1^2}{N_1} \% \frac{R_2^2}{N_2}}$$

where:
R_1 = the first rate
R_2 = the second rate
N_1 = the first number of births
N_2 = the second number of births

If the difference is **greater** than this statistic, then the difference would occur by chance less than 5 times out of 100. If the difference is **less** than this statistic, the difference might occur by chance more than 5 times out of 100. We say that the difference is not statistically significant at the 95-percent confidence level.

Example

Is the first birth rate for black women 40-44 years of age (1.08 per 1,000) significantly lower than the comparable rate for white women (1.55)? Both rates are based on more than 100 births (1,535 for black women and 14,108 for white women). The difference between the rates is 1.55 - 1.08 = .47. The statistic is then calculated as follows:

$$1.96 \sqrt{\frac{1.08^2}{1,535} \% \frac{1.55^2}{14,108}}$$

$$= 1.96 \text{ x } \sqrt{[(1.166/1,535 + 2.403/14,108)]}$$
$$= 1.96 \text{ x } \sqrt{(.00076+0.00017)}$$
$$= 1.96 \text{ x } \sqrt{.00093}$$
$$= 1.96 \text{ x } .03$$
$$= .06$$

The difference between the rates (.47) is greater than this statistic (.06). Therefore, the difference is statistically significant at the 95-percent confidence level.

Significance Testing for Hispanic Subgroups

Tables 6, 8, 9, and 14 in "Births: Final Data for 1999" and tables 1-4 and 1-12 in "Vital Statistics United States, volume 1 natality" showing birth and fertility rates based on population estimates derived from the U.S. Bureau of the Census' Current Population Survey and adjusted to resident population control totals, the formula above would be substituted by the formula which follows.

When both rates are based on 100 or more events, the difference between the two rates is considered statistically significant if it exceeds the statistic in the formula below. This statistic equals 1.96 times the standard error for the difference between two rates.

$$1.96 * \sqrt{R_1^2 * \left[\left(\frac{1}{B_1}\right) + f\left(a + \frac{b}{P_1}\right)\right] + R_2^2 * \left[\left(\frac{1}{B_2}\right) + f\left(a + \frac{b}{P_2}\right)\right]}$$

If the difference is greater than this statistic, then the difference would occur by chance less than 5 times out of 100. If the difference is less than this statistic, the difference might occur by chance more than 5 times out of 100. We say that the difference is not statistically significant at the 95-percent confidence level.

Example

Suppose the birth rate for Puerto Rican mothers 15-19 years of age (R_1) is 80.6, based on 11,978 births and an estimated population of 148,673, and the birth rate for Cuban mothers 15-19 years of age (R_2) is 27.1, based on 997 births and an estimated population of 36,782. Using the above formula, the z score is computed as follows:

$$1.96 * \sqrt{80.6^2 * \left[\left(\frac{1}{11,978}\right) + 0.670\left(-0.000238 + \frac{7,486}{148,673}\right)\right] + 27.1^2 * \left[\left(\frac{1}{997}\right) + 0.670\left(-0.000238 + \frac{7,486}{36,782}\right)\right]}$$

$$1.96 * \sqrt{6,496.36 * [0.000083486 + 0.670(-0.000238 + 0.050352)] + 734.41 * [0.0010030 + 0.670(-0.000238 + 0.20352)]}$$

$$1.96 * \sqrt{(6,496.36 * 0.033660) + (734.41 * 0.13720)}$$

$$1.96 * \sqrt{218.67 + 100.76}$$

$$1.96 * 17.87$$

$$= 35.03$$

Since the difference between the two rates of 53.5 is greater than the value above, the two rates are statistically significant at the 0.05 level of significance.

One of the rates is based on fewer than 100 cases

To compare two rates, when one or both of those rates are based on less than 100 cases, you first compute the confidence intervals for both rates. Then you check to see if those intervals overlap. If they **do** overlap, the difference is not statistically significant at the 95-percent level. If they **do not** overlap, the difference is indeed "statistically significant."

Example

Is the first birth rate for American Indian women 40-44 years of age (.54 per 1,000) significantly lower than the comparable rate for white women (1.55)? The rate for American Indian women is based on 47 events whereas the rate for white women is based on 14,108 events. The rate for American Indian women is based on less than 100 events; therefore, the first step is to compute the confidence intervals for both rates.

	Lower Limit	Upper Limit
American Indian women	0.40	0.72
White women	1.52	1.58

These two confidence intervals do not overlap. Therefore, the first birth rate for American women 40-44 is significantly lower (at the 95-percent confidence level) than the comparable rate for white women.

Testing differences between two percents

When testing the difference between two percents, both percents must meet the following conditions:

$$B \times p \geq 5 \quad \text{and}$$
$$B \times q \geq 5$$

where:

B	=	number of births in the denominator
p	=	percent divided by 100
q	=	$1 - p$

When both percents meet these conditions then the difference between the two percents is considered statistically significant if it exceeds the statistic in the formula below. This statistic equals 1.96 times the standard error for the difference between two percents.

$$1.96 \sqrt{p\,(1-p)\,\left(\frac{1}{B_1} + \frac{1}{B_2}\right)}$$

$$p \ ' \ \frac{B_1 \ p_1 \% B_2 \ p_2}{B_1 \% B_2}$$

where:

    B_1 = the number of births in the denominator for the first percent
    $B2$ = the number of births in the denominator for the second percent
    p_1 = the first percent divided by 100
    $p2$ = the second percent divided by 100

Example

Is the percent of births to Hispanic women that were to unmarried women higher in Alaska (28.8 percent) than in Alabama (23.0). The number in the denominator was 593 in Alaska and 1,345 in Alabama. The necessary conditions are met for both percents (calculations not shown). The difference between the two percents is .288 - .230 = .058. The statistic is then calculated as follows:

$$1.96 \ \sqrt{(.2477) \ (.7523) \ (.0024)}$$

$$= \ \ 1.96 \ \text{x} \ \sqrt{.000447}$$
$$= \ \ 1.96 \ \text{x} \ .021$$
$$= \ \ .042$$

The difference between the percents (.058) is greater than this statistic (.042). Therefore, the difference is statistically significant at the 95-percent confidence level.

Computation of rates and other measures

Population bases

The rates shown in this report were computed on the basis of population statistics prepared by the U.S. Bureau of the Census. Rates for 1940, 1950, 1960, 1970, 1980, and 1990 are based on the population enumerated as of April 1 in the censuses of those years. Rates for all other years are based on the estimated midyear (July 1) population for the respective years. Birth rates for the United States, individual States, and metropolitan areas are based on the total resident populations of the respective areas. Except as noted these populations exclude the Armed Forces abroad but include the Armed Forces stationed in each area. The resident population of the birth- and death-registration States for 1900-32 and for the United States for 1900-99 is shown in table 4-1. In addition, the population including Armed Forces abroad is shown for the United States. Table D shows the sources for these populations.

In both the 1980 and 1990 censuses, a substantial number of persons did not specify a racial group that could be classified as any of the White, Black, American Indian, Eskimo, Aleut, Asian, or Pacific Islander categories on the census form (26). In 1980 the number of persons of "other" race was 6,758,319; in 1990 it was 9,804,847. In both

censuses, the large majority of these persons were of Hispanic origin (based on response to a separate question on the form), and many wrote in their Hispanic origin, or Hispanic origin type (for example, Mexican, Puerto Rican) as their race. In both 1980 and 1990, persons of unspecified race were allocated to one of the four tabulated racial groups (white, black, American Indian, Asian or Pacific Islander), based on their response to the Hispanic origin question. These four race categories conform with the 1979 edition of OMB Directive 15 which mandates that race data must contain at least these 4 categories. These categories are also more consistent with the race categories in vital statistics.

In the allocation of unspecified race was carried out using cross-tabulations of age, sex, race, type of Hispanic origin, and county of residence. Persons of Hispanic origin and unspecified race were allocated to either white or black, based on their Hispanic origin type. Persons of "other" race and Mexican origin were categorically assumed to be white, while persons in other Hispanic categories were distributed to white and black *pro rata* within the county-age-sex group. For "other-not-specified" persons who were not Hispanic, race was allocated to white, black, or Asian and Pacific Islander, based on proportions gleaned from sample data. The 20-percent sample (respondents who were enumerated on the longer census form) provided a highly detailed coding of race, which allowed identification of otherwise unidentifiable responses with a specified race category. Allocation proportions were thus established at the State level, which were used to distribute the non-Hispanic persons of "other" race in the 100-percent tabulations.

In 1990 the race modification procedure was carried out using individual census records. Persons whose race could not be specified were assigned to a racial category using a pool of "race donors," which was derived from persons of specified race and the identical response to the Hispanic origin question within the auspices of the same Census District Office. As in 1980, the underlying assumption was that the Hispanic origin response was the major criterion for allocating race. Unlike 1980, persons of Hispanic origin, including Mexican, could be assigned to any racial group, rather than white or black only, and the non-Hispanic component of "other" race was allocated primarily on the basis of geography (District Office), rather than detailed characteristic.

The means by which respondent's age was determined were fundamentally different in the two censuses; therefore, the problems that necessitated the modification were different. In 1980 respondents reported year of birth and quarter of birth (within year) on the census form. When census results were tabulated, persons born in the first quarter of the year (before April 1) had age equal to 1980 minus year of birth, while persons born in the last 3 quarters had age equal to 1979 minus year of birth.

In 1990 the quarter year of birth was not reported on the census form, so that direct determination of age from year of birth was impossible. In 1990 census publications age is based on respondents' direct reports of age at last birthday. This definition proved inadequate for postcensal estimates, because it was apparent that many respondents had reported their age at time of either completion of the census form or interview by an enumerator, which could occur several months after the April 1 reference data. As a result, age was biased upward. Modification was based on a respecification of age, for most individual respondents, by year of birth, with allocation to first quarter (persons aged 1990 minus year of birth) and last three quarters (aged 1989 minus year of birth) based on a historical series of registered births by month. This process partially restored the 1980 logic for assignment of age. It was not considered necessary to correct for age overstatement and heaping in 1990, because the availability of age and year of birth on the census form provided elimination of spurious year-of-birth reports in the census data before modification occurred.

Populations for 1999--The population of the United States by age, sex, race, and Hispanic origin is shown in the Census Bureau report United States population estimates, by age, sex, race, and Hispanic origin: 1990 to 1999. Washington, DC: U.S. Census Bureau. Internet release, April 11, 2000. http://www.census.gov/population/estimates/nat_90s_1.html.

Populations for 1998--The population of the United States by age, sex, race, and Hispanic origin is shown in the Census Bureau report United States population estimates, by age, sex, race, and Hispanic origin: 1990 to 1998. Washington, DC: U.S. Bureau of the Census. Internet release, June 4, 1999. Http://www.census.gov/population/www/estimates/uspop.html.

Populations for 1997--The population of the United States by age, sex, race, and Hispanic origin is shown in

the Census Bureau report United States population estimates, by age, sex, race, and Hispanic origin: 1990 to 1997. PPL-91R.U.S. Bureau of the Census. Rounded populations are consistent with U.S. Bureau of the Census file NESTV97. Washington: U.S. Department of Commerce. 1998.

Populations for 1996--The population of the United States by age, sex, race, and Hispanic origin is shown in the Census Bureau report, United States population estimates by age, sex, race and Hispanic origin: 1990 to 1996. U.S. Bureau of the Census. PPL-57. Washington: U.S. Department of Commerce. 1997.

Populations for 1995--The population of the United States by age, sex, race, and Hispanic origin is shown in the Census Bureau report, United States population estimates by age, sex, race and Hispanic origin: 1990 to 1995. U.S. Bureau of the Census. Census file RESDO795, PPL-41. Washington: U.S. Department of Commerce. 1996.

Populations for 1994--The population of the United States by age, sex, race, and Hispanic origin is shown in the Census Bureau report, United States population estimates by age, sex, race and Hispanic origin: 1990 to 1994. U.S. Bureau of the Census. PPL-21. Washington: U.S. Department of Commerce. 1995.

Populations for 1993--The population of the United States by age, sex, race and Hispanic origin is tabulated from Census file RESO793.

Populations for 1992--The population of the United States by age, sex, race and Hispanic origin is tabulated from census file RESPO792.

Populations for 1991--The population of the United States by age, race, and sex is shown in *Current Population Reports*, Series P-25, Number 1095. Monthly population figures were published in *Current Population Reports*, Series P-25, Number 1097.

Populations for 1990--The population of the United States by age, race, and sex, and the population for each State is shown *in Current Population Reports*, Series P-25, Number 1095. The figures have been modified as described above. Monthly population figures were published in *Current Population Reports*, Series P-25, Number 1094.

Population estimates for 1981-89--Birth rates for 1981-89 (except those for cohorts of women) have been revised, based on revised population estimates that are consistent with the 1990 census levels, and thus may differ from rates published in volumes of *Vital Statistics of the United States* for these years. The 1990 census counted approximately 1.5 million fewer persons than had earlier been estimated for April 1, 1990. The revised estimates for the United States by age, race, and sex were published by the U.S. Bureau of the Census in *Current Population Reports*, Series P-25, Number 1095. Population estimates by month are based on data published in *Current Population Reports*, Series P-25, Number 1094 and unpublished data. Unpublished revised estimates for States were obtained from the U.S. Bureau of the Census.

Populations for 1980--The population of the United States by age, race, and sex, and the population for each State are shown in tables 4-2 and 4-3 of volume I, *Vital Statistics of the United States*, 1980. The figures by race have been modified as described above. Monthly population figures were published in *Current Population Reports*, Series P-25, Number 899.

Population estimates for 1971-79--Birth rates for 1971-79 (except those for cohorts of women) have been revised, based on revised population estimates that are consistent with the 1980 census levels, and thus may differ from rates published in volumes of *Vital Statistics of the United States* for these years. The 1980 census counted approximately 5.5 million more persons than had earlier been estimated for April 1, 1980 (27). The revised estimates for the United States by age, race, and sex were published by the U.S. Bureau of the Census in *Current Population Reports*, Series P-25, Number 917. Population estimates by month are based on data published in *Current Population Reports*, Series P-25, Number 899. Unpublished revised estimates for States were obtained from the U.S. Bureau of the Census.

Population estimates for 1961-69--Birth rates for 1961-69 are based on revised estimates of the population and thus may differ slightly from rates published before 1976. The revised estimates used in computing these rates were published in *Current Population Reports*, Series P-25, Number 519. The rates for 1961-64 are based on revised estimates of the population published in *Current Population Reports*, Series P-25, Numbers 321 and 324 and may differ slightly from rates published in those years.

Population estimates for 1951-59--Final intercensal estimates of the population by age, race, and sex and total

population by State for 1951-59 are shown in tables 4-4 and 4-5 of volume I, *Vital Statistics of the United States*, 1966. Beginning with 1963 these final estimates have been used to compute birth rates for 1951-59 in all issues of *Vital Statistics of the United States.*

Net census undercounts and overcounts

The U.S. Bureau of the Census has conducted extensive research to evaluate the coverage of the U.S. population (including undercount, overcount, and misstatement of age, race, and sex) in the last five decennial censuses 1950, 1960, 1970, 1980, and 1990. These studies provide estimates of the national population, that were not enumerated or over enumerated in the respective censuses, by age, race, and sex (27-29). The report for 1990 (30) includes estimates of net under enumeration and over enumeration for age, sex, and racial subgroups of the national population, modified for race consistency with previous population counts as described in the section "Population bases."

These studies indicate that there are differential coverages in the censuses among the population subgroups; that is, some age, race, and sex groups are more completely enumerated than others. To the extent that these estimates of overcounts or undercounts are valid, that they are substantial, and that they vary among subgroups and geographic areas, census miscounts can have consequences for vital statistics measures (28). However, the effects of undercounts in the census are reduced to the extent that there is underregistration of births. If these two factors are of equal magnitude, rates based on unadjusted populations are more accurate than those based on adjusted populations because the births have not been adjusted for underregistration.

The impact of net census miscounts on vital statistics measures includes the effects on levels of the rates and effects on differentials among groups.

If adjustments were made for persons who were not counted in the census of population, the size of the denominators would generally increase and the rates would be smaller than without an adjustment. Adjusted rates for 1990 can be computed by multiplying the reported rates by ratios of the 1990 census-level population adjusted for the estimated net census miscounts, which are shown in table E. A ratio of less than 1.0 indicates a net census undercount and would result in a corresponding decrease in the rate. A ratio in excess of 1.0 indicates a net census overcount and would result in a corresponding increase in the rate.

Enumeration of white females in the childbearing ages was at least 97 percent complete for all ages. Among black women, the undercount ranged up to 5 percent. Generally, females in the childbearing ages were more completely enumerated than males for similar race-age groups.

If vital statistics measures were calculated with adjustments for net census miscounts for each of these subgroups, the resulting rates would have been differentially changed from their original levels; that is, rates for those groups with the greatest estimated overcounts or undercounts would show the greatest relative changes due to these adjustments. Thus the racial differential in fertility between the white and the ``All other'' population can be affected by such adjustments.

Cohort fertility tables

The various fertility measures shown for cohorts of women are computed from births adjusted for underregistration and population estimates corrected for under enumeration and misstatement of age. Data published after 1974 use revised population estimates prepared by the U.S. Bureau of the Census and have been expanded to include data for the two major racial groups. Heuser has prepared a detailed description of the methods used in deriving these measures as well as more detailed data for earlier years (31). These tables for current years are available at http://www.cdc.gov/nchs/datawh/statab/unpubd/natality/natab97.htm.

Parity distribution--The percent distribution of women by parity (number of children ever born alive to mother) is derived from cumulative birth rates by order of birth. The percent of zero-parity women is found by subtracting the cumulative first birth rate from 1,000 and dividing by 10. The proportions of women at parities one through six are found from the following formula:

Percent at N parity =((cum. rate, order N) - (cum. rate, order N + 1))/10

The percent of women at seventh and higher parities is found by dividing the cumulative rate for seventh-order births by 10.

Birth probabilities--birth probabilities indicate the likelihood that a woman of a certain parity and age at the beginning of the year will have a child during the year. Birth probabilities differ from central birth rates in that the denominator for birth probabilities is specific for parity as well as for age.

Total fertility rate

The total fertility rate is the sum of the birth rates by age of mother (in 5-year age groups) multiplied by 5. It is an age-adjusted rate because it is based on the assumption that there are the same number of women in each age group. The rate of 2,075.0 in 1999, for example, means that if a hypothetical group of 1,000 women were to have the same birth rates in each age group that were observed in the actual childbearing population in 1999, they would have a total of 2,075.0 children by the time they reached the end of the reproductive period (taken here to be age 55 years), assuming that all of the women survived to that age.

Seasonal adjustment of rates

The seasonally adjusted birth and fertility rates are computed from the X-11 variant of Census Method II (32). This method of seasonal adjustment used since 1964 differs slightly from the U.S. Bureau of Labor Statistics (BLS) Seasonal Factor Method, which was used for Vital Statistics of the United States, 1964. The fundamental technique is the same in that it is an adaptation of the ratio-to-moving-average method. Before 1964 the method of seasonal adjustment was based on the X-9 variant and other variants of Census Method II. A comparison of the Census Method II with the BLS Seasonal Factor Method shows the differences in the seasonal patterns of births to be negligible.

Computations of percents, percent distributions, and medians

Births for which a particular characteristic is unknown were subtracted from the figures for total births that were used as denominators before percents, percent distributions, and medians were computed. The percent of records with missing information for each item is shown by State in table A. The median number of prenatal visits also excludes births to mothers who had no prenatal care. Computations of the median years of school completed and the median number of prenatal visits were based on ungrouped data. The median age of mother is computed from birth rates in 5-year age groups which eliminates the effects of changes in the age composition of the childbearing population over time. The procedures for distributing not stated age of father in order to compute mean ages are described in the section "age of father." An asterisk is shown in place of any derived statistic based on fewer than 20 births in the numerator or denominator.

References

1. National Center for Health Statistics. natality public use tape and CD-ROM. Hyattsville, Maryland: National Center for Health Statistics. Annual products.
2. Ventura SJ, Martin JA, Curtin SA, Menacker F, Hamilton BE. Births: Final data for 1999. National vital statistics reports; vol 49 no 1. Hyattsville, Maryland: National Center for Health Statistics. 2001.
3. Department of Health and Human Services. Model State Vital Statistics Act and Regulations, 1992 Revision. Publication no. (PHS). 94-1115.
4. World Health Organization. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, based on the recommendations of the Tenth Revision Conference, 1987. Geneva: World Health Organization. 1992.
5. American College of Obstetricians and Gynecologists. Guidelines for Perinatal Care, ed. 2, Washington, DC. 308-324. 1988.
6. Kowaleski J. State definitions and reporting requirements for live births, fetal deaths, and induced terminations of pregnancy (1997 revision). Hyattsville, Maryland: National Center for Health Statistics. 1997.
7. Statistical Office of the United Nations. Handbook of vital statistics systems and methods/Department of International Economic and Social Affairs. Studies in methods. Series F, no. 35, ST/ESA/STAT/SER.F35. New York: United Nations. 1985.
8. National Office of Vital Statistics. Births and birth rates in the entire United States, 1909 to 1948. Vital Statistics--Special reports; vol 33 no 8. Washington: Public Health Service. 1950.
9. Division of Vital Statistics. Instruction manual part 3a: Classification and coding instructions for birth records, 1999. Hyattsville, Maryland: National Center for Health Statistics. http://www.cdc.gov/nchs/data/3amanual/pdf.
10. U.S. Bureau of the Census. Population of metropolitan areas and component geography: 1990 and 1980 (6-30-90 definitions). 1990 CPH-L-10. Washington: U.S. Department of Commerce. 1991.
11. U.S. Department of Commerce. Metropolitan statistical area classification. Federal Register; vol 45 no 2. Washington: U.S. Government Printing Office, 956-62. 1980.
12. U.S. Office of Management and Budget. Standard metropolitan statistical areas. Rev. ed. Washington: U.S. Government Printing Office, 89-90. 1975.
13. U.S. Bureau of the Census. 1990 Census of Population. General population characteristics; (1990 CP-1-1). Washington: U.S. Department of Commerce. 1992.
14. Martin JA. Birth characteristics for Asian or Pacific Islander subgroups, 1992. Monthly vital statistics report; vol 43 no 10, suppl. Hyattsville, Maryland: National Center for Health Statistics. 1995.
15. Schachter J. Matched record comparison of birth certificate and census information in the United States, 1950. Vital statistics--Special Reports; vol 47 no12. Washington: Public Health Service. 1962.
16. Schoendorf KC, Parker JD, Batkhan LZ, Kiely JL. Comparability of the birth certificate and 1988 maternal and infant health survey. Vital and health statistics reports; series 2 no 116. Hyattsville, Maryland: National Center for Health Statistics. 1993.
17. Ventura SJ, Bachrach CA. Nonmarital Childbearing in the United States, 1940-99. National vital statistics reports; vol 48 no 16. Hyattsville, Maryland: National Center for Health Statistics. 2000.
18. Berkov B. An evaluation of California's inferred birth statistics for unmarried women. National Center for Health Statistics. Vital Health Stat 2(97). 1985.
19. Ventura SJ. Births to unmarried mothers, United States, 1980-92. National Center for Health Statistics. Vital Health Stat 21(53). 1995.
20. Taffel S, Johnson D, Heuser R. A method for imputing length of gestation on birth certificates. National Center for Health Statistics. Vital Health Stat 2(93). 1982.
21. Brockert JE, Stockbauer JW, Senner JW, et al. Recommended standard medical definitions for the U.S. Standard Certificate of Live Birth, 1989 revision. Paper presented at annual meeting of the Association for the Vital Record and Health Statistics. June 1990.

22. Watkins ML, Edmonds L, McClearn A, et al. The surveillance of birth defects: The usefulness of the revised U.S. standard birth certificate. Am J Public Health 86(5):731-34. 1996.

23. U.S. Bureau of the Census. Test of birth-registration completeness, 1964 to 1968. 1970 census of population and housing; PHC (E)-2. Evaluation and Research Program. Washington: U.S. Department of Commerce. 1973.

24. U.S. Bureau of the Census. Money income in the United States: 1998 (with separate data on valuation of noncash benefits). Current population reports, consumer income. Series P60-206. Washington: U.S. Government Printing Office. 1999.

25. U.S. Bureau of the Census. Money income in the United States: 1999 (with separate data on valuation of noncash benefits). Current population reports, consumer income. Series P60-209. Washington: U.S. Government Printing Office. 2000.

26. U.S. Bureau of the Census. U.S. population estimates, by age, sex, race, and Hispanic origin, 1980-91. Current population reports; series P-25, no 1095. Washington: U.S. Department of Commerce. 1993.

27. U.S. Bureau of the Census. Coverage of the national population in the 1980 census by age, sex, and race. Preliminary estimates by demographic analysis. Current population reports; series P-23, no 115. Washington: U.S. Department of Commerce. 1982.

28. U.S. Bureau of the Census. Estimates of coverage of the population by sex, race, and age--Demographic analysis. 1970 census of population and housing; PHC (E)-4. Evaluation and Research Program. Washington: U.S. Department of Commerce. 1974.

29. U.S. Bureau of the Census. Developmental estimates of the coverage of the population of States in the 1970 census-demographic analysis. Current population reports; series P-23, no 65. Washington: U.S. Department of Commerce. 1977.

30. Robinson JG, Ahmed B, Das Gupta P, et al. Estimation of population coverage in the 1990 United States census based on demographic analysis. JASA 88(423):1061-71. 1993.

31. Heuser R. Fertility tables for birth cohorts by color: United States, 1917-73. Washington: National Center for Health Statistics. 1976.

32. U.S. Bureau of the Census. The X-11 variant of the Census Method II Seasonal Adjustment Program. Technical paper; no 15, 1967 rev. Washington: U.S. Department of Commerce. 1967.