HUMAN FERTILITY DATABASE DOCUMENTATION: U.S.A

Authors:

Ward Kingkade
Formerly US Census Bureau

Aiva Jasilioniene
Max Planck Institute for Demographic Research, Rostock, Germany
E-mail: Jasilioniene@demogr.mpg.de

Dmitri Jdanov
Max Planck Institute for Demographic Research, Rostock, Germany
E-mail: Jdanov@demogr.mpg.de

Last revision: 19 March 2026 (by Aiva Jasilioniene)

1. Organizational History of Birth Statistics

The system of vital registration in the U.S. developed gradually, starting from separate initiatives primarily at the local level in the colonial period, then gathering momentum in the 19th Century from public health concerns. Mortality held priority in terms of these interests. It was not until 1915 that a Birth Registration Area, consisting initially of 10 states, was established by the National Board of Health. A major milestone was reached in 1933, when Texas entered the Birth Registration Area, which from that point on encompassed all 48 of the constituent states of the U.S. at that time. In 1950, the organized territory of Alaska, which became a state in 1959, was included in the birth registration area (see Table 1). From that point on major efforts have been made to enhance the quality of the data and expand the dissemination of U.S. vital statistics.

Originally the development of vital statistics was the responsibility of the U.S. Census Bureau, beginning with items included in 19th Century U.S. Censuses. Starting in 1946, the division that performed this function was transferred to the U.S. Public Health Service, and designated the National Office of Vital Statistics. In 1960 this organization was merged with the National Health Survey to form the National Center for Health Statistics (NCHS), and in 1963 the Division of Vital Statistics was organized as one of 5 operating divisions of NCHS, which has continued to improve the completeness and accuracy of birth registration, to standardize measurement across the states, and to broaden the scope of items measured and disseminated in U.S. vital statistics.

Table 1: Year in which each state was admitted into the Birth Registration Area

Year of entry	State	Year of entry	State
1915	Connecticut	1922	Illinois
	Maine		Montana
	Massachusetts		Wyoming
	Michigan	1924	Florida
	Minnesota		Iowa
	New Hampshire		North Dakota
	New York	1925	West Virginia
	Pennsylvania	1926	Arizona
	Rhode Island		Idaho
	Vermont	1927	Alabama
	District of Columbia		Arkansas
1916	Maryland		Louisiana
1917	Indiana		Missouri
	Kansas		Tennessee
	Kentucky	1928	Colorado
	North Carolina		Georgia
	Ohio		Oklahoma
	Utah	1929	Hawaii*
	Virginia		Nevada
	Washington		New Mexico
1919	California	1932	South Dakota
	Oregon	1933	Texas
	South Carolina	1950	Alaska*
1920	Nebraska		
1921	Delaware		
	Mississippi		
	New Jersey		

* Alaska became a state in January, 1959 and annual births are published for it starting that year. Hawaii became a state in August, 1959, and NCHS began publishing annual births for the state of Hawaii starting in 1960. Prior to their inclusion as states, data for these "organized territories" were provided in a supplement to the National Natality Statistics volumes. Other territories, such as the Virgin Islands, have also been included in these supplements.

2. Data Availability

The number of items included on the standard birth certificate grew from 30 in 1900 to 72 in 2003. For the purposes of the Human Fertility Database we are interested in only a few of these items, namely age of mother and birth order. These items have been included from the first version of the standard certificate (1900). Prior to the 1930 revision of the standard birth certificate, the format and detail of the birth order item varied among the states. Vital Statistics data on births by single years of age and birth order have been published since 1931 and the publications are available as scanned PDF files on the website of the National Center for Health Statistics. For a few years before and during the Second World War (1939-1945) the births by age and birth order are given in 5-year detail. However, only from 1960, when Massachusetts begins reporting information on births by birth order, data on births sorted by age of the mother and birth order become available for all the states and thus cover the entire U.S. (for more details about the availability of data before 1960, please see Table 2).

The U.S. data on births for 1931 and 1932, published by single years of age and birth order, exclude the states of Colorado, Maine, Massachusetts, New Hampshire, and Rhode Island. Due to this but mainly because of the fact that it was only in 1933 when all the states were finally added to the Birth Registration Area, the data on births for the years 1931 and 1932 are not used for the calculations in the HFD. Figures on births by single years of age and birth

order for the U.S. in 1933 leave out Colorado, Massachusetts, and New Hampshire and for the years 1934-1938, such data are missing for Massachusetts and New Hampshire. These states, however, (except Massachusetts) provided other types of birth data which, by employing some special methods, allowed to estimate the distribution of births by single years of age and birth order for the whole area of the U.S (see Table 2).

Table 2: Characteristics of data on births for the years 1933-1938

	State	Years	Age groups	Birth order
1.	USA	1933	5-year	Total
	Colorado, Massachusetts, and New Hampshire	1933	5-year	Total
	USA	1933	Total	By birth order (Massachusetts included in the category of "unknown" birth order)
	Colorado, Massachusetts, and New Hampshire	1933	Total	By birth order, except Massachusetts
	USA without Colorado, Massachusetts, and New Hampshire	1933	1-year	Total
	USA without Colorado, Massachusetts, and New Hampshire	1933	1-year	By birth order
2.	USA	1934-1938	5-year	Total
	Massachusetts and New Hampshire	1934-1938	5-year	Total
	USA	1934-1938	Total	By birth order (Massachusetts included in the category of "unknown" birth order)
	Massachusetts and New Hampshire	1934-1938	Total	By birth order, except Massachusetts
	USA without Massachusetts and New Hampshire	1934-1938	1-year	Total
	USA without Massachusetts and New Hampshire	1934-1938	1-year	By birth order

Figures on birth order for Massachusetts for the period 1939-1959 are included in the "total" and "unknown" classifications only. Births are not given by birth order because the state of Massachusetts did not require the reporting of such information.

From 1968 onwards, public use micro datasets are available from which customized tabulations can be made combining any selection of variables coded from the birth certificates.

Data used for the HFD calculations are specified in Appendix 1.

2.1 Description of the procedure used to estimate order-specific births by single years of age in 1933-1938

Table 3: Original data that were used as an input

Code	State	Years	Age groups	Birth order
NH	Colorado and New Hampshire	1933	5-year	No (total)
M	Massachusetts	1933	5-year	No (total)
NH	Colorado, and New Hampshire	1933	No (total)	Yes
US3	USA without Colorado, Massachusetts, and New Hampshire	1933	1-year	Yes
NH	New Hampshire	1934-1938	5-year	No (total)
M	Massachusetts	1934-1938	5-year	No (total)
NH	New Hampshire	1934-1938	No (total)	Yes
US3	USA without Massachusetts and New Hampshire	1934-1938	1-year	Yes

Algorithm:

"New Hampshire" (NH) stands for New Hampshire plus Colorado in 1933. Thus, the algorithm below is applicable for the entire period 1933-1938.

First, we distribute data for Massachusetts into single year age groups, birth order 1 through 5+. Since data for NH are available by 5-year age groups only, data for US3 are aggregated using the same age scale. Then a matrix for US3 and NH is prepared for the application of IPF procedure (for a detailed description of IPF, see the HFD Methods Protocol):

1) Age-specific birth data (no birth order, total) are calculated as a sum of US3 and NH by respective age groups:

$$B(x) = B^{US3}(x) + B^{NH}(x) \quad (1)$$

2) Birth data by birth order (no age, total) equals the sum of US3 and NH:

$$B_i = B_i^{US3} + B_j^{NH},\ i=1,\ 2,\ 3,\ 4,\ 5+ \quad (2)$$

3) Data for US3 are used as initial values for age-specific birth data by birth order $B_i(x)$. The application of the IPF procedure to these data produces the distribution by 5-year age groups for each birth order for US3+NH. 5-year age groups are further distributed into one-year age groups using the distribution of US3 data (for each birth order and for all birth orders):

$$B_i^{US3+NH}(x;x+1) = B_i^{US3+NH}(x;x+5)\frac{B_i^{US3}(x;x+1)}{B_i^{US3}(x;x+5)},\ i=1,\ 2,\ 3,\ 4,\ 5+,\ UNK \quad (3)$$

$$B^{US3+NH}(x;x+1) = \sum_i B_i^{US3+NH}(x;x+1) \quad (4)$$

The same procedure is used to split 5-year age groups in Massachusetts data. Since the specification by birth order is not available for Massachusetts, we add these distributed numbers only to the category of unknown birth order.

3. Data Quality

One of the criteria for a state's inclusion in the Birth Registration Area was the achievement of at least 90% completeness of birth registration. Prior to 1933 some states were dropped from the Registration Area because they had regressed in terms of completeness of registration, then subsequently readmitted when completeness was deemed to be at 90% or higher. Completeness of registration was assessed by a variety of methodologies, all of which had their weaknesses[1]. In 1934, Pascal Whelpton showed that by a measure which took account of census undercount of infants as well as underregistration of births, many states in the Birth Registration area were at less than 90% completeness relative to the 1930 Census. Soon after, adjustments of various fertility indicators for underregistration of births emerged in official vital statistics volumes.

The 1940 Census featured the reappearance of an item on children ever born, which had not been included on decennial Censuses since 1910. The 1940 Census also included a nationwide test of completeness of birth registration in which enumerators filled out special cards for infants born between December 1, 1939 and April 1, 1940. The cards were checked against birth records to see if certificates were on file. After a series of follow-up activities and corrections, the evaluation arrived at the result that for the US as a whole, birth registration was 92.5% complete. A similar exercise was conducted as part of the 1950 census, arriving at a national estimate of 97.9 percent completeness of birth registration.

Of greatest relevance to the Human Fertility Database are the findings that according to the 1940 evaluation, completeness of registration was 75.7 percent among mothers ages 45 and over, and 89.7 percent for women ages 40-44. According to the 1950 evaluation, completeness was over 90% for women in each (5-year) age group in the reproductive ages. Completeness of registration in 1950 was over 90% in each of 7 birth order categories reported, and this applies for the most part equally well to five-year age categories[2].

Coale and Zelnick (1963) came forward with the assessment that the quality of the birth data for the Native White population was "good enough" as of 1934, without precisely explaining why and in what manner. In any case, they present their own set of undercount estimates and adjusted series of annual births, which can be compared to the official statistics provided here.

In 1960 the National Center for Health Statistics decided to discontinue adjustment for underregistration, taking explicit account of Coale's (1955) adjustments for underregistration of births and Census undercount, and noting that these suggested that the errors in birth registration and census enumeration appeared to be in the same direction. It emerged that birth rates computed by dividing unadjusted births by unadjusted populations came closer to Coale's estimated rates than the officially adjusted series in which births adjusted for under-registration are divided by unadjusted populations[3].

[1] What is said to have been the most common approach was to mail out a card to every household in the state requesting information on births in the preceding year, and check the replies indicating births against the birth register to see if a certificate was on file. The low response rate to these inquiries is an obvious liability of the method.

[2] There were two curious exceptions: 84.1% completeness of registration for 5th order births and 88.2% completeness for 6th and 7th order births *among 15-19 year olds*. It should be borne in mind that no check or correction for age misreporting was included in the scope of the 1940 and 1950 evaluations.

[3] Another element of Coale's legacy to US Vital Statistics seems to have been the introduction of intrinsic vital rates and other Lotka parameters, which first appeared in the 1962 Natality volume of Vital Statistics of the United States and continue to be reported by NCHS in various publications and releases.

4. Age of Mother

Up to and including 1988, mother's age was elicited on the standard birth certificate in terms of age at the time of the birth. The 1989 revision of the standard birth certificate included no age item other than the mother's own date of birth. Of course, not all states were in compliance with this item, nor are several as of 2006, so mother's age may be either coded as stated on the birth certificate or computed based on dates of birth of mother and the birth in question.

Prior to 1961, births for which age of mother was not ascertained were published as a separate category in the published natality statistics volumes. In 1961 and 1962 such births were allocated *pro rata* over the distribution of valid responses. Starting in 1963 age was imputed for this category of births by hot-deck imputation[4]. In 1963 records were matched to the last previously processed record with a valid response and matching age and live birth order. Beginning in 1964, live birth order was replaced by total birth order including fetal deaths. From 1964 to 1996, births associated with ages under 10 or over 50 were treated in the same manner as if age were not ascertained. In 1997 the age limits were revised to 9 and 55. As of 1989, instances in which mother's birthday was not ascertained were handled by a hot deck method analogous to that employed for unreported age.

From 1989 to 2003, the mother's directly reported age, containing one or more "unknown or not stated" categories, was included in the public use microdatasets. In these instances we present the data according to the age variable computed from dates of birth when given, with age not ascertained imputed as indicated above[5]. A primary reason for doing so is that the imputed age variable is the one NCHS presents in its official publications.

The data on births in the public use datasets from the National Center for Health Statistics, and those in earlier publications cited here, take the form of Lexis Rectangles. That is, births in a given year are tabulated by mother's age at last birthday at the moment of the birth in question.

In the public use datasets, the detail age item has generally been constructed such that over the period 1968-1996 reported ages under 10 or over 49 were assigned an imputed value. Starting in 1997 the upper age limit was raised to 54. It has not always been the case that in a given year mothers' ages outside these limits have been reported. In addition, in more recent releases ages 10-12 or 10-14 are frequently grouped together, as are ages over 50 (see Table 4).

[4] A "hot deck" procedure is one in which the characteristic missing from the current record is assigned the value from the last record previously processed which contained a valid code for this item and matched on other characteristics deemed relevant to the determination of the characteristic in question (see Shryock and Siegel, 1975, for an account of this class of procedures). To elucidate the terminology it should be borne in mind that after Hermann Hollerith introduced the cardpunch in the 19th century as a labor-saving device, punched cards were employed in processing US Census and Vital Statistics data until this technology was replaced by key-to-disk and scanning operations. For purposes of completeness we note that a "cold deck" procedure is one in which items with missing or invalid values are assigned imputed values based on the total distribution of all valid responses (i.e. the whole "deck" of punched cards having valid entries, after the first round of processing).

[5] Beginning in 2003 the "detail age of mother" variable, including imputed ages, was replaced by one in which some grouping at the extremes of the age range is employed.

Table 4: Description of data on live births, USA, 1931–2024

Period	Age range	Age interval	Birth order	Lexis elements	Source*
1931**	12-54, UNK	1-year	1-26, UNK	squares	1
1932**	12-54, UNK	1-year	1-25, UNK	squares	1
1933-1935***	10-54, UNK	5-year	Total	squares	2
1936***	10-54, 55+, UNK	5-year	Total	squares	2
1937-1938***	10-54, 55+, UNK	5-year	Total	squares	3
1933-1935	12-54, UNK	1-year	1-27, UNK	squares	2
1936	12-55, UNK	1-year	1-27, UNK	squares	2
1937	12-55, UNK	1-year	1-27, UNK	squares	3
1938	12-55, UNK	1-year	1-25, UNK	squares	3
1933-1936***	Total	–	1-27, UNK	–	2
1937***	Total	–	1-27, UNK	–	3
1938***	Total	–	1-25, UNK	–	3
1939	10-54, 55+, UNK	5-year	1-24+, UNK	squares	4
1940-1941	10-54, UNK	5-year	1-24+, UNK	squares	4
1942	10-54, UNK	5-year	1-22+, UNK	squares	4
1943-1945	10-54, 55+, UNK	5-year	1-22+, UNK	squares	4
1946	≤14, 13-58, 59+, UNK	1-year	1-13+, UNK	squares	4
1947	≤14, 15-53, 54+, UNK	1-year	1-13+, UNK	squares	4
1948	≤14, 15-53, 54+, UNK	1-year	1-17+, UNK	squares	4
1949-1959	≤14, 15-49, 50+, UNK	1-year	1-17+, UNK	squares	4-9
1960-1963	≤14, 15-49, 50+	1-year	1-8+, UNK	squares	10
1964-1967	≤14, 15-49	1-year	1-8+, UNK	squares	10
1968-1977	10-49	1-year	1-20+, UNK	squares	11
1978-1996	10-49	1-year	1-20+, UNK	squares	11
1997-2002	10-54	1-year	1-20+, UNK	squares	11
2003	≤14, 15-54	1-year	1-20+, UNK	squares	11
2004-2005	10-12, 13-49, 50-54	1-year	1-20+, UNK	squares	11
2006	10-12, 13-49, 50-54	1-year	1-8+, UNK	squares	11
2007-2008	10-12, 13-49, 50-64	1-year	1-8+, UNK	squares	11
2009-2014	10-12, 13-49, 50+	1-year	1-8+, UNK	squares	11, 17, 18
2015-2019	10-12, 13-50+	1-year	1-6+, UNK	squares	19-23
2018-2024	10-12, 13-50+	1-year	1-8+, UNK	squares	25-28

* For the references, see the document USAref.pdf.
** These data are not used for calculations in the HFD (see Section 2).
*** Data in this configuration are also provided for some states as specified in Table 2.

5. Birth Order

Live births by birth order and mother's age have been reported in US Vital Statistics publications since 1931. The quality of measurement of birth order has received much less attention than has been devoted to mother's age. In several respects the birth order data invite skepticism. In particular, births of 5-7th order to women in their teens (see footnote 3) defy credibility, while reported birth orders of 20 or more are difficult to accept for mothers of any age. The 1968 public use dataset contained births of reported orders through 54. NCHS' practice has been to recode the birth order item by grouping the extremely high order births into categories such as 8+ and to include these recoded items on the public use dataset alongside the original item exhibiting implausible values.

Prior to 1969, when the items on live births and still births were both unreported, the current birth was assigned the order of 1. Beginning in 1969 such births were assigned an explicit "unknown" code in the public use datasets. We have retained this category so that the user may develop his/her own imputation procedure.

It should be noted that NCHS is tasked to work with US state and territorial statistical offices in an effort to standardize reporting practices. The registration of vital events is handled by the

respective states and territories, which are not always in compliance with NCHS' recommendations.

6. Births by Month

Births by calendar year and month have been tabulated for the period 1931-2021 Only births occurring in the Birth Registration Area have been included, which means that not all states appear before 1933. Births by month exhibit seasonality, and seasonally adjusted fertility rates appear in the more recent Vital Statistics published volumes.

7. Census Parity Data

Data on Children Ever Born have been collected in a number of the 20th Century Censuses: 1910, 1940, 1950, 1960, 1970, 1980, and 1990. The quality and detail of the information published from these censuses has varied widely (see Table 5).

An important aspect of the Census data on parity is that they were obtained on a sample basis (typically the Census "long form"). Therefore, they are subject to sampling variability. Moreover, the sample data are inflated so that table sums and marginals come to the complete 100% count totals for the relevant categories. Caution needs to be exercised when combining or comparing tabulations involving parity and complete count items (e.g. marital status).

The information from the last two Censuses which included a Children Ever Born item (1980 and 1990) is not much help. In these censuses data on parity, where present, are presented for three 10-year age categories (see US Bureau of the Census 1983, US Bureau of the Census 1993).

Census 1970 featured a whole volume devoted to fertility. The contents include a table (in single-year detail) on children ever born to white and "negro" ever-married women. For the US as a whole, data in five year detail are obtainable for single (never-married) and ever-married women. Generally, data on single women are available with less detail on parity than in the case of ever married women.

Census 1960 featured the racial dichotomy of "White" vs. "Nonwhite", making for two categories which exhaust the population. It contains single-year detail for ever married whites and nonwhites and data tabulated for all marital status categories in 5-year age detail.

The 1950 Census is practically the same as Census 1960. The table of all marital statuses (in 5-year detail) by children ever born involves a slightly more limited age range than the tables applying to White and Nonwhite ever married women in single year detail.

The 1940 Census Volume on Differential Fertility includes data from the 1910 Census as well. The best age detail from this volume appears to be 5-year groups, and the birth order data appear restricted to ever married women.

The problem with the U.S. Censuses is that the tabulations either exclude the parity distribution of never married women or give these data by 5-year age categories.

Table 5: Description of data: Women by parity, USA, population censuses

Census date	Age range	Age intervals	Parity	Marital status	Race	Data source*
15.04.1910	15-74	5-year	0-10+	ever married	all races	12:Ch.1, Table 2
01.04.1940	15-74	5-year	0-10+	ever married	all races	12:Ch.1, Table 1
01.04.1950	15-59	1-year	0-4, 5/6, 7/9, 10+	ever married	white & non-white	13:Ch. 3, Tables 8 and 10
	15-59	5-year	0-4, 5/6, 7/9, 10+	all marital statuses	white & non-white	13:Ch. 3, Tables 16 and 17
01.04.1960	15-85+	1-year, with 5-year after 44	0-4, 5/6, 7+	ever married	white & non-white	14:Ch. 2, Tables 4 and 5
	15-50+	5-year	0-4, 5/6, 7+	all marital statuses	white & non-white	14:Ch. 2, Tables 16 and 17
01.04.1970	15-85+	1-year, with 5-year after 44	0-4, 5/6, 7+	ever married	white & “negro”	15:Ch. 1, Tables 4 and 5
	15-65+	5-year	0-4, 5/6, 7+	ever married	all races	15:Ch. 1, Table 8

* For the references, see the document USAref.pdf.

8. Definition of a live birth in the U.S.

At least as far back as 1940, the US followed the conventional WHO definition of a live birth, namely:

Live birth refers to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life - e.g. 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 live born.

9. Revision history

Changes with the March 2026 revision:

Data for 2024 were added. There are some changes in fertility rates for 2009-2023 as compared to the 2025 HFD data release, which result from revised population exposure estimates (for more details regarding these revisions, see Andreeva 2026).

Changes with the March 2025 revision:

Data for 2023 were added.

Changes with the January 2025 revision:

Data for 2022 were added. (2023 data on births are available in the input file; see section “Input Data” on the HFD country page for the USA). There are minor changes in data on births and fertility rates for 2019-2021 as compared to the release of March 2023 that are due to the revised population exposures.

Changes with the March 2023 revision:

Data for 2021 were added.

Changes with the June 2022 revision:

Data for 2020 were added. There are some changes in the 2009-2019 data on births and fertility rates as compared to the release of April 2021. The changes appeared because the new inter-censal population estimates were used in the calculations.

Changes with the April 2021 revision:

Data for 2019 were added. There are some changes in the 2018 birth count estimates and fertility rates as compared to the data released on September 24, 2020. These changes are related to adjustments in the HMD population and death count data, which we use to produce the population exposure estimates in the HFD (see the HMD background and documentation file for the USA at http://www.mortality.org/hmd/USA/InputDB/USAcom.pdf).

Changes with the September 2020 revision:

Data for 2018 were added. There are some changes observable in the birth count estimates for 2009-2017 as well as in the calculated fertility rates and other fertility indicators, respectively. The changes are caused by the revision of the post-censal population estimates used in the HMD (see the HMD background and documentation file for the USA at http://www.mortality.org/hmd/USA/InputDB/USAcom.pdf).

Changes with the September 2017 revision:

There are some changes in the birth count estimates as well as the calculated fertility indicators for 2013-2014 as compared to the data release as of 20 June 2016, which were produced by changes in the HMD population estimates (see the HMD background and documentation file for the USA at http://www.mortality.org/hmd/USA/InputDB/USAcom.pdf).

Changes with the June 2016 revision:

There are some changes in the birth count estimates for 2009-2013 as compared to the previous data release as of 18 January 2016. The changes appeared due to modified HMD population estimates (the HMD used new population estimates, revised by the US Census Bureau).

References:

Andreeva, M. (2026). About Mortality Data for United States. HMD background and documentation revised Magali Barbieri on 03.02.2026. Available at https://www.mortality.org/File/GetDocument/hmd.v6/USA/Public/InputDB/USAcom.pdf

Coale, A.J. (1955). The Population of the United States in 1950 Classified by Age, Sex, and Color. *Journal of the American Statistical Association*, March 1955, pp. 16-54.

Coale, A.J. and M. Zelnick (1963). *New Estimates of Fertility and Population in the United States.* Princeton University Press.

Hentzel, A.M. (1997). *History and Organization of the Vital Statistics System*. National Center for Health Statistics.

Shryock, H.S. and J.S. Siegel (1975). *The Methods and Materials of Demography.* 3rd Printing, US Government Printing Office.

US Bureau of the Census (1983). 1980 Census of Population. Characteristics of the Population. General Summary. United States Government Printing Office.

US Bureau of the Census (1993). 1990 Census of Population. Social and Economic Characteristics. United States. United States Government Printing Office.

Whelpton, P.K. (1934). The Completeness of Birth Registration in the United States. *Journal of the American Statistical Association*, No. 186. June, 1934.

APPENDIX 1
INPUT DATA USED FOR HFD CALCULATIONS

BIRTHS

Period	Type of data	Age scale	Birth order	RefCode(s)
1933-1938*	Annual number of live births by age of the mother and birth order (Lexis squares)	10,11,…,59,60, unknown	1,2,…,4,5+, unknown, total	3
1939	Annual number of live births by age of the mother and birth order (Lexis squares)	10-14, 15-19… 55+, unknown, total	1-24+, unknown, total	4
1940-1941	Annual number of live births by age of the mother and birth order (Lexis squares)	10-14, 15-19… 50-54, unknown, total	1-24+, unknown, total	4
1942	Annual number of live births by age of the mother and birth order (Lexis squares)	10-14, 15-19… 50-54, unknown, total	1-22+, unknown, total	4
1943-1945	Annual number of live births by age of the mother and birth order (Lexis squares)	10-14, 15-19…55+, unknown, total	1-22+, unknown, total	4
1946	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 59+, unknown, total	1-13+, unknown, total	4
1947	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 54+, unknown, total	1-13+, unknown, total	4
1948	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 54+, unknown, total	1-17+, unknown, total	4
1949-1959	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 50+, unknown, total	1-17+, unknown, total	4–9
1960-1963	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 50+, total	1-8+, unknown, total	10
1964-1967	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 49, total	1-8+, unknown, total	10
1968-1977	Annual number of live births by age of the mother and birth order (Lexis squares)	10, 11… 49, total	1-20+, unknown, total	11
1978-1996	Annual number of live births by age of the mother and birth order (Lexis squares)	10, 11… 49, total	1-20+, unknown, total	11
1997-2002	Annual number of live births by age of the mother and birth order (Lexis squares)	10, 11… 54, total	1-20+, unknown, total	11
2003	Annual number of live births by age of the mother and birth order (Lexis squares)	≤14, 15… 54, total	1-20+, unknown, total	11
2004-2005	Annual number of live births by age of the mother and birth order (Lexis squares)	10-12, 13… 49, 50-54, total	1-20+, unknown, total	11
2006	Annual number of live births by age of the mother and birth order (Lexis squares)	10-12, 13… 49, 50-54, total	1-8+, unknown, total	11
2007-2008	Annual number of live births by age of the mother and birth order (Lexis squares)	10-12, 13… 49, 50-64, total	1-8+, unknown, total	11
2009-2014	Annual number of live births by age of the mother and birth order (Lexis squares)	10-12, 13,…,49, 50+, total	1-8+, unknown, total	11, 17–18
2015-2019	Annual number of live births by age of the mother and birth order (Lexis squares)	10-12, 13,…,49, 50+, total	1-6+, unknown	19–23

BIRTHS *(continued)*

Period	Type of data	Age scale	Birth order	RefCode(s)
2020-2024	Annual number of live births by age of the mother and birth order (Lexis squares)	10-12, 13,…,49, 50+, total	1-8+, unknown	25-28
1933-2024	Annual number of live births by month	total	total	2–11, 17–24, 26–28

*These are birth estimates. For more details about the estimation methods as well as characteristics of the initial data, please see Section 2.

FEMALE POPULATION: Exposure by age and year of birth

Female exposure population by calendar year, age, and year of birth (Lexis triangles) is estimated using data on population size and deaths from the Human Mortality Database[6], which is available at http://www.mortality.org or http://www.humanmortality.de.

Data for 1959 in the HMD refer to the territory of the United States including Hawaii, whereas the birth counts data used in the HFD include the state of Hawaii since 1960 only. However, the numbers are very small: in 1959 Hawaii would contribute to the total number of births in the USA with 0.4%. Therefore, the mismatch between the numerator and the denominator (i.e. having the data for Hawaii included in the population exposure estimates and excluding it from the birth counts data) does not introduce a significant bias as far as the computations are concerned.

[6] Data on births in the Human Mortality Database for the period 1933-1959 have been adjusted (by NCHS) for under-registration.

APPENDIX 2
AREA CODING USED IN THE INPUT DATA FILES

Codes	Description
1	Territory of the US including all the 50 states (1960 onwards)
2	Territory of the US excluding Alaska and Hawaii (years 1933-1958)
3	Territory of the US excluding Hawaii (year 1959)
4	The US birth registration area in 1931 and 1932, excluding the territories of Colorado, Maine, Massachusetts, New Hampshire, and Rhode Island
5	The US birth registration area in 1933 excluding the territories of Colorado, Massachusetts, and New Hampshire
6	The US birth registration area in 1934-1938 excluding the territories of Massachusetts and New Hampshire
7	State of Colorado (year 1933)
8	State of Massachusetts (years 1933-1938)
9	State of New Hampshire (years 1933-1938)