DOCUMENTATION OF THE
DETAIL NATALITY PUBLIC USE FILE
FOR 2004

**NOTE: THE RECORD LAYOUT OF THIS FILE HAS
CHANGED SUBSTANTIALLY. USERS
SHOULD READ THIS DOCUMENT
CAREFULLY.**

Public Use Data Tape Documentation - Natality Detail 2004 Data

This tape documentation was prepared in the Division of Vital Statistics by Steve Steimel of the Systems, Programming, and Statistical Resources Branch (SPSRB) and Paul Sutton, Brady Hamilton of the Reproductive Statistics Branch (RSB). Thomas D. Dunn of SPSRB was responsible for coordinating the natality documentation.

Sharon Kirmeyer of RSB prepared the Technical Appendix. The Registration Methods Section and the Data Acquisition and Evaluation Branch provided consultation to State Vital Statistics offices regarding collection of birth certificate data.

Questions on the documentation or general questions concerning the natality file should be directed to the Systems, Programming, and Statistical Resources Branch, Division of Vital Statistics, NCHS, 3311 Toledo Road, Hyattsville, MD 20782-2003 (301-458-4777).

Questions concerning the Technical Appendix or substantive questions concerning the natality data should be directed to the Reproductive Statistics Branch, Division of Vital Statistics, NCHS, 3311 Toledo Road, Hyattsville, MD 20782-2003 (301-458-4111).

Included in this document are:

1.   Introduction to the Public Use Documentation for 2004

2.   List of data elements and tape locations.

3.   Machine/File/Data Characteristics.

4.   Detail Record Layout.

5.   Geographic Code Lists for the United States and the Territories.

6.   Lists of Metropolitan Statistical Areas the United States and the
      Territories.

7.   List of Country Codes.

8.   Natality Technical Appendix for 2004.

9.   Births: Final Data for 2004.

10.   Control Total Tables for Selected Revised Certificate Items.

SYMBOLS USED IN TABLES

Symbol	Explanation
---	Data not available
...	Category not applicable
-	Quantity zero
0.0	Quantity more than 0 but less than 0.05
*	Figure does not meet standards of reliability or precision

Public Use File Documentation
Natality 2004

Introduction

United States birth data represent all births registered in the 50 States, the District of Columbia, New York City, Puerto Rico, Virgin Islands, American Samoa, and the Commonwealth of the Northern Marianas. The Center for Disease Control and Prevention's National Center for Health Statistics (NCHS) receives these data as electronic files, prepared from individual records processed by each registration area, through the Vital Statistics Cooperative Program. Information for Guam is obtained from paper copies of original birth certificates which is coded and keyed by NCHS. Data from American Samoa first became available in 1997. Data from the Northern Marianas first became available in 1998.

Birth data for the U.S. are limited to births occurring within the United States to U.S. residents and nonresidents. Births to nonresidents of the Unites States are excluded from all tabulations by place of residence. Births occurring to U.S. citizens outside of the United States are not included in this file. Birth data for Puerto Rico, Virgin Islands, Guam and American Samoa are limited to births occurring within the respective territories. For more detailed information on the 2003 Natality file *see The Technical Appendix of the Vital Statistics of the United States, 2003, vol I natality.*

The 1989 and 2003 Revisions of the U.S. Certificate of Live Birth

This data file includes data based on both the 1989 Revision of the U.S. Standard Certificate of Live Birth (unrevised) and the 2003 revision of the U.S. Standard Certificate of Live Birth in 2003 (revised). The 2003 revision is described in detail elsewhere. (See the 2003 Revision website at: http://www.cdc.gov/nchs/vital_certs_rev.htm). Pennsylvania and Washington implemented the revised certificate in 2003; five states, Idaho, Kentucky, New York (excluding New York City), South Carolina, and Tennessee implemented as of January 1, 2004. Two additional states, Florida and New Hampshire, implemented the revised birth certificate in 2004, but after January 1. The nine revised states represent 20 percent of all 2004 births; the seven revised states, for which data are available for all of 2004,

represent 14 percent of all births.   Where comparable, revised data are combined with data from the remaining 41 states and the District of Columbia. (Revised data are denoted by "R;" unrevised data are denoted by "U" in the "Rev" column of the documentation.) Where data for the 1989 and 2003 certificate revisions are not comparable (e.g., educational attainment of the mother), unrevised and revised data are given in separate fields in the data file. For the two states which revised in 2004, but after January 1, data which are not comparable across revisions are excluded from all tabulations and wherever reporting flags are used. Also see discussion of reporting flags. Data items exclusive to the 1989 Revision are shown in tables U-1 – U- 7. Selected items new to the 2003 Revision are included in this data file, see tables R-1- R-6. However, several new items are not included (e.g., infections during pregnancy, maternal morbidity). Please see section on "Incomplete National Reporting," below. For further information please contact us at births@cdc.gov or (301)458-1111.

Metropolitan statistical areas

Metropolitan statistical areas in this file are based the 1994 Office of Management and Budget (OMB) definition effective July 1, 1994. This definition has been used to define metropolitan statistical areas for natality files since 1994.   A listing of these 311 Metropolitan Statistical Areas (MSA's), Primary Metropolitan Statistical Areas (PMSA's), and New England County Metropolitan Areas (NECMA's) is included in this documentation. The 18 Consolidated Metropolitan Statistical Areas (CMSA's) are also included. In June of 2003 The OMB substantially revised the methodology for classifying and coding metropolitan areas in the United States. NCHS plans to convert to the new classification scheme with the release of 2005 natality data.

Confidentiality

To minimize the risk of disclosure of individual or institutional information NCHS public use data files do not contain the actual day of the birth or the dates of birth of the mother or father. Geographic detail is also restricted; only counties and cities of 100,000 or more population based on the 2000 Census are identified.

Incomplete National Reporting: Selecting Reporting Areas for the
2004 natality file

The use of reporting flags

As a result of the delayed, phased transition to the 2003 Standard Certificate of Live

Birth, the 2004 natality file includes data for reporting areas that use the 2003 revision of

the U.S. Standard Certificate of Live Birth (revised) and data for reporting areas that use

the 1989 Standard Certificate of Live Birth (unrevised). Although many data items are

comparable across certificate revisions and are available for the entire United States,

many items have more limited reporting areas. For example, information on pre-

pregnancy and gestational diabetes, a revised data item, is available only for 7 States for

2004; information on amniocentesis, an unrevised item, not included on the revised

certificate, is available for 41 States and New York City. Reporting flags were developed

to help the user more readily identify reporting areas for items with less than national

reporting. The national reporting area is defined as the 50 States, The District of

Columbia, and New York City; (NYC is an independent reporting area from New York

State). Reporting flags are available for most items on the file. Positions for reporting

flags are noted along with each data item in the file layout.

Translating "blanks"

In the 2004 natality file, for data items which are not common or comparable across

certificate revisions, births to residents of a revised state occurring in an unrevised state,

and births to residents in an unrevised state occurring in a revised state, are represented

by "blanks." Blanks should be treated as "unknowns" for tabulations.

In sum, the correct use of reporting flags and translation of blanks will result in an accurate tally of births for items with incomplete national reporting. For an example of SAS code that may be used to incorporate the correct use of reporting flags and the translation of blanks see below.

Example of SAS code using reporting flags (and translating blanks)

The example below is for the revised prenatal care item. Prenatal care data based on the revised Certificate are not considered comparable to data based on the unrevised Certificate, and are presented separately (see also Births: Final Data for 2004). Accordingly, use of the reporting flag for this item will produce 2004 data for the month prenatal care began for the 7 revised States which had implemented the revised Certificate by January 1st 2004.   Data for states which implemented the revised certificates after January 1 (Delaware and Florida), are excluded from tabulations when reporting flags are used.

Sample SAS program

```
01   DATA work;
02        INFILE 'c:nat04us.dat' LRECL=1500;
03        INPUT
04             restatus 138
05             precare 245-246
06             f_mpcb 668;
07
08        /*Exclude foreign residents*/
09        IF restatus NE 4;
10        /*Select reporting area*/
11        IF f_mpcb=1;
12        /*Convert blanks to unknown*/
13        IF precare=. THEN precare=99;
14
15   PROC FREQ;
```

```
16        TABLE precare;
17   RUN;
```

In this example, "restatus" is used to exclude births to foreign residents (this is standard practice for all NCHS tabulations). Also in this example, blanks are represented by numeric values SAS code = (.). However, for some items in the file, e.g., obstetric procedures, blanks are represented by character values for which the SAS code is empty quotes (' ').

Alternative to the use of reporting flags

The use of reporting flags provides a relatively quick, accurate way to select records for all areas reporting comparable data for a given item in a particular year. However, should a limited reporting area be needed, specific state(s) of residence may also be selected, or unselected. This approach may be useful, for example, in trend analysis where reporting areas have changed overtime. See Table A in the "Technical Appendix" for state-specific information on reporting areas.

This approach may also be used to limit the reporting area to only states reporting multiple or single race data. See "Technical Appendix" Table A footnote 2 for the multiple race reporting area.

Example of SAS code using state of residence (and translating blanks)

```
01   DATA work;
02        INFILE 'c:nat04us.dat' LRECL=1500;
03        INPUT
04             restatus 138
05             xmrstate $ 107-108
```

```
06              precare 245-246;
07
08          /*Exclude foreign residents*/
09          IF restatus NE 4;
10          /*Select reporting area*/
11          IF xmrstate in ('ID','KY','NY','PA','SC','TN','WA');
12          /*Convert blanks to unknown*/
13          IF precare=. THEN precare=99;
14
15      PROC FREQ;
16          TABLE precare;
17      RUN;
```