



# STATISTICAL PRIMER

State Center for Health Statistics • 1908 Mail Service Center • Raleigh, NC 27699-1908 • 919/733-4728  
[www.schs.state.nc.us/SCHS/](http://www.schs.state.nc.us/SCHS/)

No. 14

April 2000

## Matched Live Birth/Infant Death Files

by

Paul A. Buescher

### Introduction

Every year the State Center for Health Statistics creates a matched birth/infant death file. Each death certificate for an infant (under one year of age) is matched to that baby's certificate of live birth. This matched file allows analysis of infant mortality by characteristics that are available on the birth certificate, such as birth weight, mother's age, and prenatal care participation. The matched birth/infant death file is often used to produce the numerator for a rate or percentage, with the file of all live births in the same year used to produce the denominator. The table to the right presents infant mortality rates by birth weight, mother's education, and mother's smoking during pregnancy. None of these characteristics are available on the death certificate, so this type of analysis is possible only if the matched birth/infant death file is used.

#### Table: Infant deaths per 1,000 live births for babies born in 1993-1997 by selected birth certificate characteristics (North Carolina residents)

##### *Birth weight*

Less than 1,500 grams .....	293.6
1,500-2,499 grams .....	17.3
2,500 grams or more .....	3.1

##### *Mother's education*

Under 9 years .....	10.1
9-11 years .....	12.8
12 years or more .....	8.6

##### *Smoking during pregnancy*

Yes .....	13.0
No .....	8.7

**ALL BIRTHS** ..... 9.5



These rates show that babies of low birth weight, those whose mothers had low education, and those whose mothers smoked during pregnancy had higher infant mortality rates. This does not necessarily demonstrate cause and effect. For example, mothers who smoke are likely to have other characteristics that are associated with an increased risk of infant mortality (such as low education or being unmarried). So the higher infant mortality rate among babies of mothers who smoked during pregnancy may be attributable to these other risk factors, in addition to the effects of smoking *per se*. A multivariable analysis to predict infant mortality that includes risk factors (control variables) in addition to smoking would allow a better estimate of the independent effect of maternal smoking. This could be accomplished by first creating a file of all live births for a time period, with the infant death information appended for those births where the baby died prior to one year of age.

Death certificates are completed for babies who were born alive (and thus also have a certificate of live birth). Babies born dead (fetal deaths) are reportable to the State of North Carolina only if the gestational age was 20 weeks or more, in which case a Report of Fetal Death should be completed. These fetuses do not have a birth certificate and are not included in the matched birth/infant death file.

In creating the matched birth/infant death files, a confidential “detailed” version is created which has personal identifiers, and a “statistical” file with shortened records and no personal identifiers is created for public use. This Statistical Primer describes some changes in the format of the matched birth/infant death files produced by the State Center for Health Statistics, to inform the users of these data.

## **Format of the File for 1997 and Earlier Years**

Since 1959, the State Center for Health Statistics has created a matched birth/infant death file based on year of birth. This is sometimes called a “birth cohort” file. Babies **born** in a given year who subsequently die as an infant are included in the file. Babies born in 1997, for example, were followed through 1997 and 1998 to determine which ones died during the first year of life. One advantage of using the birth cohort approach is that the entire file of live births in the year becomes an exact “population at risk” to use as the denominator for infant mortality measures. In fact, “infant mortality **risk**” is technically the appropriate term to use when infant mortality measures are calculated using a matched birth cohort file. Most infant mortality measures published by the State Center for Health Statistics consist of infants who died in a given year in the numerator and babies born in the same year in the denominator. Some of the babies who died in the year could have been born in the previous year. A measure calculated in this fashion is more appropriately referred to as an “infant mortality **rate**.”

## **Reasons to Change the Format**

One problem with using the birth cohort approach is that a year must pass after the end of the year of birth to account for all infant deaths that could have occurred to babies born in that year. This in effect delays the production of the file and associated statistics for one year. Mainly for this reason, the National Center for Health Statistics has produced since 1995 a matched birth/infant death file based on year of death. This is sometimes referred to as a “period” file, in contrast to the birth cohort file. All official linked-file statistics of the National Center for Health Statistics are now produced from this matched year-of-death file. The National Center continues to produce an annual

matched birth cohort file for some customers, at a substantially later date. The matched year-of-death file can be completed relatively soon after the end of the calendar year of occurrence. In North Carolina the live birth and death files are now closed six to eight months after the end of the calendar year.

Another factor is that beginning January 1, 1999 all causes of death on North Carolina death certificates were coded according to the 10<sup>th</sup> revision of the International Classification of Diseases (ICD-10). Deaths occurring from 1979 through 1998 were coded according to the 9<sup>th</sup> revision (ICD-9). There were major changes in the coding scheme from ICD-9 to ICD-10. If we produced the 1998 matched birth/infant death file using the birth cohort approach, the babies who died in 1998 would have ICD-9 causes of death, while the babies who died in 1999 would have ICD-10 causes. Mixed codes in the same file would considerably complicate analyses.

For these reasons, the State Center for Health Statistics has decided to produce the matched birth/infant death file based on year of death beginning with the 1998 file year.

## **Format of the File for 1998 and Later Years**

Starting with 1998, the North Carolina matched birth/infant death file will include all babies who died in that year. For 1998, this means that most of the babies were born in 1998, but some were born in 1997. Therefore all of the dates of death in the new files will be in the same calendar year, but the dates of birth will span two years. Eighty-nine percent of the babies who died in 1998 were born in 1998 and only 11 percent were born in 1997. One reason for this distribution is that approximately one-half of all infant deaths occur in the first day of life. When computing infant mortality **rates** using the matched birth/infant death file based on year of death, all live

births occurring in the same year are used for the denominators. There is a close (but not exact) correspondence since the large majority of the babies dying in a year were born in the same year.

A birth cohort file can still be created by combining two matched year-of-death files. For example, a 1998 matched birth cohort file could be created by combining the 1998 and 1999 matched year-of-death files and selecting only those records where the year of birth was 1998.

The calculation of multi-year rates may just involve combining several annual data files, but variations may be encountered. For example, in the future, a 1999-2003 birth cohort file could be created by combining the 1999 through 2004 matched year-of-death files and selecting those records where the date of birth was in the 1999-2003 period. On the other hand, a 1993-1997 year-of-death file could be created by combining the existing 1992 through 1997 matched birth cohort files and selecting those records where the date of infant death was in the 1993-1997 period. Multi-year rates that include the 1998 year-of-death file will have to be carefully constructed. This is because there is some overlap between the 1997 birth cohort file and the 1998 year-of-death file. Babies who were born in 1997 but died in 1998 are included in both files, and for most purposes this overlap will have to be eliminated.

The number of infant deaths in the new matched year-of-death files will not correspond exactly to published numbers of North Carolina resident infant deaths for several reasons. Infant death records are included in the matched file only if a successful match is made with a birth certificate record (approximately 99 percent of infant death records are successfully matched to a birth certificate). Also, some delayed records may be included in the matched file that were not included in the regular death certificate file. Further, babies

who were not North Carolina residents at the time of death but were residents at the time of birth may be included in North Carolina infant mortality statistics if residence at birth from the matched file is used. The matched infant death file should be used only when you need information from both the birth and infant death portions of the file. Investigators interested only in the causes of death for infants, for example, should use the regular death file (selecting only infants), which will produce results that match published values.

## **Other Changes**

In addition to changing the matched birth/infant death file from a birth cohort to a year-of-death orientation, a number of other changes are and will be occurring in the North Carolina vital statistics files. Record layouts are changing substantially because of Y2K compliance and conversion to the ICD-10 cause-of-death coding system. New birth and death certificates are scheduled to be adopted in 2003 and this will result in major file changes. We will provide documentation of these changes to the users of our vital statistics data files.

---

Department of Health and Human Services  
State Center for Health Statistics  
1908 Mail Service Center  
Raleigh, N.C. 27699-1908