

For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations

Before starting an infant mortality assessment, it is helpful to first understand common definitions and available resources. Because the data sources that are available will vary by state, developing an inventory of available data sets is an essential first step to starting your infant mortality assessment.

A. Data Sources

1. Sources of Infant Mortality Data

Various sources of data can be utilized in the infant mortality assessment process. Major sources include vital statistics, surveys, interviews, reports of individual events and record reviews.

Population-based data, such as vital records, have the advantage of being widely and readily available, generalizable to the population of interest and inexpensive.

Vital statistics include data for the complete population of events, and are not a sample of events. This supports more robust analysis of rare events and more detailed

Vital records, both birth and death certificates, provide the major source of data on infant mortality in the United States.

analysis of data for population subgroups. Vital records, including both birth and death certificates, provide the major source of data on infant mortality in the United States. There are laws in all states that require the registration of all live births and all deaths. Vital events are registered in the state where the event occurs, not the state of residence, though resident-specific analysis is possible. While variation exists among states, vital record registration generally includes: live births, deaths, fetal deaths, marriages, divorces, and to a lesser extent adoptions and induced terminations of pregnancy. Vital record certificates vary by state but the CDC National Center for Health Statistics and National Association for Public Health Statistics and Information Systems (NAPHSIS) provides guidance and training to the states, and encourages states to maintain a baseline of conformity for live births, deaths and fetal deaths. Vital event registries have limitations, including:

- Adding new variables to registries is resource intensive
- Data quality can be compromised by errors in reporting and missing data
- Timely data may be unavailable because of reporting delays

- **Nativity data:** Data provided from the birth certificate.
- **Mortality data:** Data provided from the death certificate. The date of birth field distinguishes infant deaths from those occurring at older ages.
- **Linked birth-infant death records:** Linkage and merging of birth certificate and death certificate data. Linked birth and death data are the first choice for every infant that dies either within a calendar year (death cohort or period file) or a specific cohort year of birth (birth cohort). While extremely valuable, it should be noted that cohort files may be a few years behind the current calendar year because of delays in collecting and reporting of events.

For infant mortality assessment, the data sources most commonly used – birth, death and linked birth/death records – are the primary focus of this guide. The specific variables indicated below are available to states that are currently using the 2003 revised birth certificate. See Resources, Section E, for a copy of both the 2003 revised birth certificate and the 2003 revised death certificate. Those states using earlier revisions should consult the Resources, Section D. ([Click here to view map of states using the 2003 revised birth certificate versus the 1989 birth certificate as of January 1, 2008](#))



For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

Table 1. Variables listed in the Revised 2003 Birth and Death Certificates:

BIRTH RECORDS	INFANT DEATH RECORDS
<p>Infant Variables</p> <ul style="list-style-type: none"> • Name • Date/time of birth • Place of birth • Sex • Weight at birth • Birth order (for multiple order births) • Plurality • Congenital defects • Apgar score • Gestational age • Method of feeding at discharge <p>Maternal Variables</p> <ul style="list-style-type: none"> • Name (maiden and marital surname) • Date of birth • Place of birth • Last menstrual period (LMP) • Height/weight (pre-pregnancy/at delivery) • Month prenatal care began/ended • Number of prenatal care visits • Past medical history • Past obstetric history • Cigarette smoking • Race/ethnicity of mother • Marital status • Parity (total number of live births) • Gravidity (total number of pregnancies) • Maternal education • Birth attendant • Providers for prenatal care • Sources of payment for delivery • Mode of delivery • WIC food provisions during pregnancy • Social Security Number (used for linking records) • Birth interval (calculated) • Pregnancy interval (calculated) • Weight gain (calculated) <p>Paternal Variables</p> <ul style="list-style-type: none"> • Race/ethnicity • Date of birth • Place of birth • Education • Presence/absence of name • Social Security Number 	<ul style="list-style-type: none"> • Date of death • Place of death • Causes of death or underlying causes • Parent information • Age at time of death • Autopsy conducted • Method of disposition

To appropriately interpret secondary data, it is necessary to understand the accuracy and completeness of data reporting and how missing data are treated. Reliability and validation studies of specific birth certificate variables have been published that provide information on the usefulness of certain measures. For example, studies have demonstrated that birth certificates are more likely to be incomplete among women of lower socioeconomic status than among women of higher socioeconomic status. Despite the limitations of vital records data, the large number of vital records available to analyze provides robust results and makes it possible to study subpopulations. A list of other possible data sources that can be useful for assessing infant mortality are provided below.

Table 2. Data Sources Useful for Examining Infant Mortality and Related Factors:

GEOGRAPHIC REGION	DATA SOURCE
National data only	<ul style="list-style-type: none"> • National Hospital Care Survey (NHCS; formerly National Hospital Discharge Survey and the National Hospital Ambulatory Medical Care Survey), NCHS • National Health and Nutrition Examination Survey (NHANES), NCHS
National and all states/regions	<ul style="list-style-type: none"> • National Vital Statistics, NCHS: <ul style="list-style-type: none"> • Live births • Infant deaths • Fetal deaths • Linked birth/infant death files • National Immunization Survey (NIS), NCHS
National and HHS regions	National Survey of Family Growth (NSFG), NCHS
State-specific for all states	<ul style="list-style-type: none"> • State Department of Health Vital Statistics: <ul style="list-style-type: none"> • Live births • Infant deaths • Fetal deaths • State linked birth/infant death • Behavioral Risk Factor Surveillance System (BRFSS) • Hospital discharge surveys
State-specific for selected states	<ul style="list-style-type: none"> • Health Care Utilization Project (HCUP), AHRQ • Youth Risk Behavioral Surveillance System (YRBSS), 47 states, 6 territories, 2 tribes, 22 localities, CDC • PRAMS, 40 states and New York City (approximately 78 percent of all live births in the United States), CDC • School Health Profile Surveys, 49 states, 5 territories, 20 cities (contains information on adolescent health), CDC • Birth defects registries, U.S./state-based birth defects tracking systems, 14 states, CDC • Fetal and Infant Mortality Reviews • Maternal Mortality Reviews or Pregnancy Associated Mortality Reviews • Perinatal Quality Care Collaboratives/Quality Improvement Initiatives/Neonatal Networks
State and local health social service programs	<ul style="list-style-type: none"> • Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) • Medicaid/CMS/MMIS • Family Planning • Healthy Start • Home Visiting Programs • Hospital and insurance records – important sources of information about medical service utilization <ul style="list-style-type: none"> • Date of admission • Place of residence • Duration of in-hospital stay • Diagnosis • Procedures and treatment • Birth weight • Apgar score • Cause of death
State and local review programs	<ul style="list-style-type: none"> • Fetal and infant, or child death reviews <ul style="list-style-type: none"> • Age at death • Cause of death

Note: Public use data sets exist – i.e. CDC Wonder, CPONDER, etc.

For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

2. Fetal and Infant Mortality Review (FIMR)

FIMR information complements data obtained from vital records. FIMR is a community process to assess, monitor, and improve service systems and community resources for women, infants, and families. The process begins with a fetal or infant death and can consist of data gathering from hospital records, outpatient records, social service records, and autopsy reports; an interview with the mother or family; a case review by a team of health, social service, and other experts in the community to identify issues and provide recommendations; and community action to address challenges identified by the case review team. While FIMR data cannot be used to infer causes of deaths for a given community, it can help to develop hypotheses to pursue with larger, more generalizable data sets. However, the overarching goal of the FIMR process is to provide community leaders with recommendations from case reviews so that communities can implement targeted interventions to improve community service systems and resources (see www.nfimr.org/home).

Leadership and support is provided by the National Fetal and Infant Mortality Review Program, a collaborative effort of the American College of Obstetricians and Gynecologists (ACOG) and the federal Maternal and Child Health Bureau (MCHB), Health Resources and Services Administration (HRSA).

B. Definitions

1. Infant Mortality Rates

Infant mortality refers to the death of a live-born infant within their first year of life (less than 365 days). The infant mortality rate (IMR) is the number of infant deaths per 1,000 live births for a particular time period (typically one year) and geographic population. The components of the infant mortality rate are the:

- Numerator (the number of infant deaths)
- Denominator (the number of live births)

The HP 2020 objective is to reduce the rate of all infant deaths to 6.0 infant deaths per 1,000 live births.

Infant Mortality Rate

SAMPLE PROBLEM

In 2010, **3,321** babies were born alive in Star County and of those, **24** died before reaching their first birthday. What was the infant mortality rate for Star County?

FORMULA

$$\frac{\text{\# infant deaths in 1st yr of life for a defined time period (usually one year)}}{\text{\# live births during the same time period}} \times 1,000$$

SOLUTION

$$\frac{24}{3,321} \times 1,000$$

ANSWER

7.2 infant deaths per 1,000 live births

a) Neonatal and Postneonatal Mortality

The main sub-components of the 'infant period' include:

The neonatal period (birth - 27 days of life, often referred to as the first month of life)

Neonatal Mortality Rate

$$\frac{\text{\# infant deaths between 0-27 days of life for a defined time period}}{\text{\# live births during the same time period}} \times 1,000$$

The postneonatal period (28 days - 1 year)

Postneonatal Mortality Rate

$$\frac{\text{\# infant deaths between 28-365 days of life}}{\text{\# live births during the same time period}} \times 1,000$$

In 2007, nearly two-thirds (65.5 percent) of infant deaths occurred in the neonatal period, between birth and the first 27 days after birth; and the remaining 34.5 percent occurred in the postneonatal period, between 28 and 365 days after birth (Figure 7).

For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

This difference is reflected in the 2007 neonatal mortality rate of 4.4 deaths per 1,000 live births and the postneonatal mortality rate of 2.3 deaths per 1,000 live births (Figure 8). The sum of the neonatal and postneonatal mortality rate equals the infant mortality rate. In the example we provide, the sum of the neonatal and postneonatal mortality rates differ from the infant mortality rate by 1/10th of a percent due to rounding in the individual mortality rates.

Figure 7. Infant Deaths by Age at Death, United States, 2007

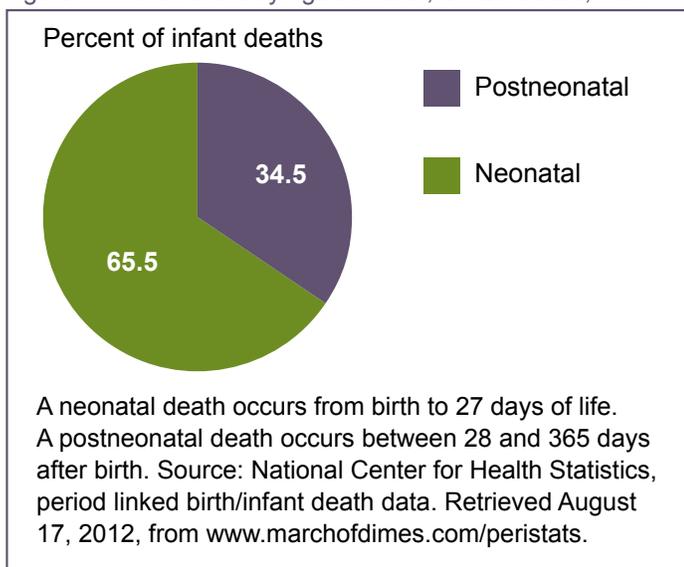
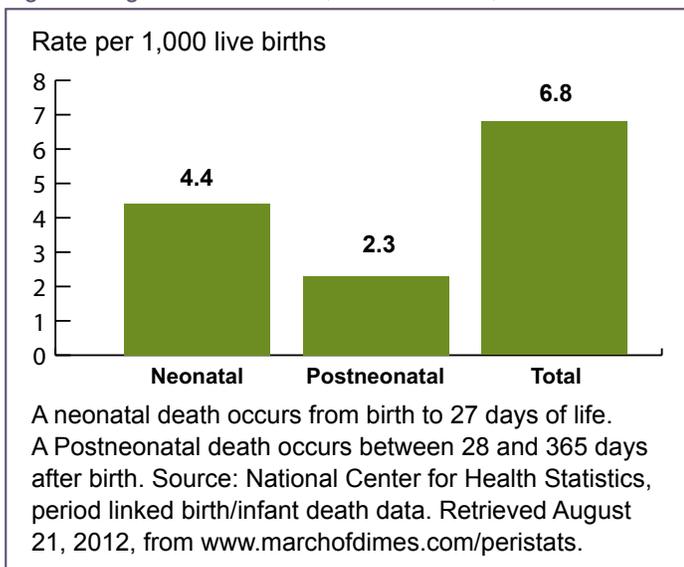


Figure 8: Age at Infant Death, United States, 2007



b) Why Look at Neonatal and Postneonatal Rates in Addition to Infant Mortality Rates?

Neonatal and postneonatal mortality are examined separately because most deaths during the neonatal period are associated with events surrounding the

prenatal period and delivery, whereas postneonatal deaths are more likely to be associated with conditions or events that arise after the delivery and may reflect environmental factors. Data on events occur during both the neonatal and postneonatal periods that contribute to the infant mortality rate, along with the cause of death, provide guidance on the types of interventions needed to decrease high infant mortality rates. For example, infants who die within the first 27 days may have experienced complications during labor and delivery, been born with congenital anomalies or been born with preterm-related complications. Likewise, if an urban center has high postneonatal mortality rates, programmatic interventions could work to reduce SIDs and other sleep-related deaths through a Safe to Sleep campaign, control exposure to infectious diseases through a hygiene or vaccination campaign, or prevent unintentional injuries through a child safety/car seat safety campaign.

c) Gestational Age and Preterm Birth Measures

Gestational age refers to the age of the fetus or newborn infant in weeks. Estimates of a newborn's gestational age based on birth certificate data are used to monitor trends in infant and maternal health and to improve our understanding of adverse pregnancy outcomes.

Term birth refers to live births between 37 and 41 completed weeks gestation. Within term births, live births that occur between 37 and 38 completed weeks are commonly referred to as Early Term.

Preterm birth refers to live births less than 37 completed weeks (<259 days) gestation. Preterm birth rates or percentages are expressed as a proportion of all live births during a particular time period and geographic population. Preterm infants are almost 10 times as likely to die in the first year of life and have a higher risk of morbidity as term infants.

Very preterm births refers to live births less than 32 completed weeks gestation. The percentage of live births that are very preterm is around 2 percent, but more than 50 percent of infant deaths in 2008 occurred among very preterm births. Very preterm infants are more than 70 times as likely to die in the first year of life as term infants.

Late preterm births refers to live births between 34 and 36 completed weeks gestation. Late preterm births comprise 70 percent of all preterm births and are the fastest growing subgroup of preterm births. Compared to term infants, late preterm infants have three times the infant mortality rate of term infants.

Disorders relating to short gestation and low birth weight, not elsewhere classified (sometimes labeled simply “prematurity/low birth weight”), are the second leading cause of infant death in the United States (www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_06.pdf). In recent years, NCHS has begun to use the classification of preterm-related causes of death, which is the leading cause of infant mortality in the United States.

Between 1981 and 2005, preterm birth rates increased more than 30 percent (9.4 percent to 12.8 percent).

There are multiple reasons that may account for these changes, including: increases in multiple births (twins, triplets and higher-order births); increases in births to women over 35 years of age; increased use of treatments for infertility; and changes in obstetric management of pregnancy. Beginning in 2006, the preterm birth rate declined from 12.8 percent to 12.2 percent in 2009, mainly as a result of quality improvement efforts to reduce elective deliveries (via induction of labor and or cesarean section) around the late preterm birth period. Despite these improvements, the rate of preterm births is higher for African Americans (17.5 percent) compared with white infants (11.1 percent). Recent data indicate that preterm rates for white infants are on the rise, with a relative increase of almost 9 percent between 1998 and 2008 (10.2 percent to 11.1 percent), while those for African Americans increased at a slower pace (4 percent) from 1998 to 2007 (17.6 percent to 18.3 percent) and then declined in 2008 (17.5 percent). Overall, health care costs for preterm babies in the United States total more than \$26 billion annually.

There is not a widely accepted measurement of gestational age that is both accurate and consistent on a population level, presenting limitations to the use and interpretation of preterm birth rates. Gestational age is calculated in a variety of ways. Traditionally, gestational age has been calculated from the first day of the last normal menstrual period (LMP). However, LMP itself is not always accurately reported, especially for women with late diagnosis of pregnancy and/or late initiation of prenatal care. In cases with missing data or where preference is for another method of calculating gestational age, another commonly used measure of gestational age at the population level is the clinical or obstetrical estimate of gestational age. Studies

The HP 2020 objective is to reduce preterm births to 11.4 percent of all live births.

have reported significant differences in gestational age estimates based on the various methods and caution should be exercised when selecting your calculation methodology and when comparing results from various studies unless the same methodology is used.

d) Birth Weight Measures

Birth weight is highly correlated with gestational age, as it is the first weight of the newborn obtained following delivery. Birth weight is a strong predictor of infant mortality and morbidity and has been demonstrated to be accurately documented on the birth certificate. Vital and hospital records are the most common sources of birth weight data. Measures include:

- **Rate of low birth weight (LBW):** This indicator identifies the number of live births weighing less than 2,500 grams or 5.5 pounds, expressed as a proportion of all live births in a given place and time
 - Very low birth weight (<1,500 grams or 3.3 pounds)
 - Moderately low birth weight (1,500- 2,499 grams)

The lower the birth weight, the greater the risk for poor outcomes, including infant death. In 2007, babies born with low birth weight were about 25 times more likely to die in the first year of life compared to babies who were not born low birth weight. The risk of death was much greater for babies born very low birth weight, who were more than 100 times more likely to die before their first birthday than babies weighing 2,500 grams or more.

The HP 2020 objective is to reduce low birth weight births to 7.8 percent of all live births.

In recent years, the proportion of low birth weight infants has been generally increasing. The proportion of low birth weight infants steadily increased from 6.7 percent in 1984 to 8.3 percent in 2006. This period was followed by a decrease in the proportion of low birth weight infants to 8.2 percent in 2007 that remained relatively unchanged in 2008. The prevalence of babies born low birth weight varies by race and ethnicity. African Americans have the highest prevalence of low birth weight infants. In 2008, 13.7 percent of African American live births were low birth weight, compared with 7.2 percent of white live births. Recent data indicate that the proportion of low birth weight white infants increased 30 percent from 1990 to 2006 (5.6 percent to 7.3 percent) then decreased to 7.2 percent in 2008. Among African Americans, there was a decline in the proportion of low birth weight infants in the 1990s, followed by a 7 percent increase from 2001 to 2006 (13.1 percent to 14.0 percent) and then a decrease to 13.7 percent in 2008.

Regional variations in low birth weight deliveries have also been reported. According to 2008 data, Alabama, Louisiana and Mississippi had the highest proportion of births that were low birth weight at 10.6 percent, 10.8 percent and 11.8 percent respectively. Alternatively, the states with the lowest proportions of low birth weight deliveries were Alaska, Oregon, and Washington at 6.0 percent, 6.1 percent, and 6.3 percent, respectively.

e) Fetal Death Measures

Most state vital statistics programs use the following definition of fetal death or one very similar: “Death prior to the complete expulsion or extraction from its mother of a product of human conception, irrespective of the duration of pregnancy and which is not an induced termination of pregnancy.” State reporting requirements vary, and many require reporting around 20 weeks of gestation [www.cdc.gov/nchs/data/misc/itop97.pdf].

“The death is indicated by the fact that after such expulsion or extraction, the fetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles. Heartbeats are to be distinguished from transient cardiac contractions; respirations are to be distinguished from fleeting respiratory efforts or gasps.”

This definition is based on the definition endorsed by the World Health Organization, with additional clarifiers recommended by the AAP Committee on the Fetus and Newborn in partnership with the ACOG, Association for Neonatal Nurses, CDC and the National Institutes of Health with the exclusion statement regarding induced terminations of pregnancy. According to recommendations

made by the this same committee, fetuses that die in utero prior to 20 weeks full gestation are considered miscarriages or spontaneous abortions. The 1992 Revision of the Model State Vital Statistics Act and Regulations recommends that fetal deaths of 350 grams or more, or if weight is unknown, of 20 completed weeks gestation or more, calculated from the date the last normal menstrual period began to the date of delivery, be reported. Also, the term “stillbirth” may be used to describe these deaths at 20 weeks gestation or more.

Early fetal death: A fetal death between 20 and 27 weeks of gestation.

Late fetal death: A fetal death at or after 28 weeks of gestation.

f) Perinatal Measures

Perinatal refers to around (peri) the time of birth (natal), so it can be used to refer to the entire or parts of the period beginning before conception and through the first year of life. The following are common definitions, combining infant and fetal death time periods.

The National Center for Health Statistics uses two different definitions for perinatal mortality:

Perinatal definition I: this includes fetal deaths of 28 weeks gestation or more and infant deaths of less than seven days of age.

Perinatal definition II: this is the most inclusive definition, and includes fetal deaths of 20 weeks or more and infant deaths of less than 28 days of age.

The denominator for perinatal rates is “per 1,000 live births plus fetal deaths.”



For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

C. Basic Calculations

This section on measurement provides information on the tools needed to help interpret your data. These are the approaches that will allow you to uncover whether differences in infant mortality rates are statistically significant or, in other words, the change over time or difference between groups is unlikely to have occurred by chance alone.

The following approaches will be addressed:

- Measures of Occurrence
- Measures of Association: Comparing Infant Mortality Rates
- Testing the Statistical Significance of the Difference between Two Rates

1. Measures of Occurrence

To understand the occurrence of infant deaths, investigators quantify infant deaths from a defined population. Measures of infant death describe the likelihood that infants born alive with certain characteristics will die during the first year of life. Infant death measures of occurrence include:

a. Infant Mortality Rate:

As presented in the Definitions section, the infant mortality rate is a measure of the risk of dying during the first year of life, defined as the number of deaths to infants born during a certain year divided by the total number of live births for the same year, presented per 1,000 live births.

$$\frac{\text{Number of deaths among infants (born in [birth year/s])}}{\text{Total number of live births (in 2020)}} \times 1,000$$

b. Cause-Specific Infant Mortality Rate:

A measure of mortality risk, generally defined as the number of infant deaths from a specific cause of death or group of causes during a year divided by the total number of live births for the same year, presented per 100,000 live births. The sum of all of the cause-specific infant mortality rates equals the infant mortality rate for all causes combined.

$$\frac{\text{Number of deaths from causes related to prematurity and low birth weight among infants (born in [birth year/s])}}{\text{Total number of live births (in [birth year/s])}} \times 100,000$$

c. Proportionate Mortality:

A measure of the proportion of deaths due to a particular cause of death. The sum of all of the cause-specific proportionate mortality measures equals 100 percent.

$$\frac{\text{Number of deaths from causes related to prematurity and low birth weight among infants (born in [birth year/s])}}{\text{Total number of deaths from all causes among infants born (in [birth year/s])}} \times 100$$

It is important to differentiate between a cause-specific mortality rate and proportionate mortality. The denominator for a cause-specific mortality rate is all live births in a given year, while the denominator for proportionate mortality includes only all deaths.

A change in a cause-specific mortality rate seen over time, place or by other factors reflects real change that may be important for prevention and intervention strategies, whereas a change in proportionate mortality may only be reflecting a change in another cause or causes of death. For example, if there is an increase in the birth defects-specific infant mortality rate over time, this suggests the need to address birth defects; by contrast, if there is an increase in the percent of infant deaths due to birth defects, this apparent increase might be a consequence of deaths due to other causes decreasing over time. Also, you can have increases or decreases in the infant mortality rate without a change in proportionate mortality. Both cause-specific mortality and proportionate mortality provide independent information on infant mortality in your state, county or region that may be useful for program planning.

2. Standardized Rates

When populations differ from each other, such as by age or race, we can standardize the rate to remove the effect of differences between populations for comparison purposes. An adjusted rate is the same as calculating an average that is weighted to reflect the characteristics of the population of interest. Such weighting can be derived from a standard population. The following are two methods for standardizing differences in population demographic distributions: the direct and indirect method.

a. Direct Method:

Using the Direct Adjusted Rate Formula, an example is provided to calculate the infant mortality rate per 1,000 live births for State A population adjusted by the U.S. population

stratified by women of reproductive age. We multiply proportions of the standard population in each category (i.e., age groups) by the corresponding maternal age-specific mortality rates of each of the two comparison populations. These products are added up for each population. The standard or reference population may be the U.S. national population, a state population or a specific sub-population with a standardized distribution.

Direct Adjusted Rate (DAR) Formula

$$\sum_i^m c_i * r_i$$

m = number of strata

i = stratum

c = proportion of standard population in stratum i

r = rate in study population in stratum i

Direct Method Example:

Age (in years) Strata (i)	U.S. Birth Population (c)	State A Birth Population	U.S. Infant Mortality Rate / 1,000 Live Births	State A Mortality Rate / 1,000 Live Births (r)
<18	26%	22%	3	3
18-20	30%	20%	9	7
21-34	24%	28%	14	11
35-44	20%	30%	24	22
Total	100%	100%	12	12

WALK-THROUGH:

According to the table above, State A has the same infant mortality rate as the United States.

Multiply the percentage of the U.S. Birth Population in each age stratum (c) by the State A infant mortality rate (r) in the corresponding stratum. Then sum the stratum specific mortality rates. This gives you the infant mortality rate for State A by assuming that the birth population in state A had the same age distribution as the United States.

DIRECT AGE-ADJUSTED RATE FORMULA:

$$\sum_i^m c_i * r_i$$

m = number of strata

i = age stratum

c = proportion of U.S. birth population in age stratum i (in decimal form)

r = state A infant mortality rate in age stratum i

SOLUTION:

$$.26 \times 3 = 0.78/1,000+$$

$$.30 \times 7 = 2.10/1,000+$$

$$.24 \times 11 = 2.64/1,000 +$$

$$.20 \times 22 = 4.40/1,000+$$

This sums to the State A maternal age at delivery-adjusted mortality rate of 9.92 infant deaths per 1,000 live births.

If the crude mortality rate for the State A had been reported, it would be concluded that the United States and State A had the same mortality rate. After adjustment for differences in the age distribution between the state A birth population and the U.S. birth population, state A had a lower infant mortality rate (direct adjusted infant mortality rate of 9.92 per 1,000 live births).



For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

Indirect Method:

When data are unavailable or insufficient to directly calculate the adjusted infant mortality rate, indirect methodologies may be used. The standardized mortality ratio (SMR) is one methodology that is used to indirectly calculate and compare rates. In the indirect method example, we multiply the U.S. birth weight stratum-specific infant mortality rates (r) by the number of births at Hospital A (c) in the corresponding birth weight stratum (i). The product is the expected number of infant deaths (e) in each birth weight stratum if Hospital A has the same birth weight stratum-specific infant mortality rate as the United States. After the expected infant deaths for each birth weight stratum are calculated, they are then summed to get the total number of expected infant deaths at Hospital A if it had the same birth weight-specific mortality rates as the United States. A ratio of the observed total number of deaths at Hospital A (o) to the total number of expected deaths based on U.S. birth weight-specific infant mortality rates is then calculated. This ratio is called the standardized mortality ratio – in this example the birth weight standardized infant mortality ratio.

Standardized Infant Mortality Ratio

$$\frac{\text{Observed number of infant deaths}}{\text{Expected number of infant deaths}} \times 100$$

A SMR above 100 means the number of observed deaths is greater than what would be expected if the study population had the same probability of dying as the standard population, while a SMR of below 100 means the number of observed deaths is less than expected.

The indirect method is used when category-specific rates are unavailable or unstable because of small numbers. Caution should be used in comparing SMRs as the estimates may be weighted to different reference populations.

Indirect Method Example:

In this example, we use the indirect method since the number of births and infant deaths in birth weight categories may be small; resulting in rates of birth weight-specific mortality that are unstable and unreliable. Hospital A has the highest infant mortality rate in the country among newborns weighing <2,500 grams.

Birth Weight Stratum (i)	Live Births Hospital A (c)	U.S. Infant Mortality Rate* (r)	Expected Deaths (e)	Observed Deaths (o)
500-999 g	300	0.750	225	168
1,000-1,499 g	250	0.125	31	20
1,500-1,999 g	250	0.060	15	10
2,000-2,500 g	200	0.020	4	2
Total	1,000		275	200

*Per 1,000 live births

SOLUTION:

$$\text{Standardized Mortality Rate} = \frac{200}{275} \times 100 = 72.7\%$$

Multiply the birth weight-specific infant mortality rates of the United States by the live births in Hospital A for each birth weight stratum to obtain the expected number of infant deaths. The sum of the observed deaths (200 infant deaths) divided by the sum of expected deaths (275 infant deaths) equals a standardized infant mortality ratio of 72.7 percent. If Hospital A had the same infant mortality experience for each birth weight stratum as the United States, we would expect to see 275 infant deaths, or 75 more infant deaths than the 200 that actually took place. After adjusting for differences in birth weight distribution among live births, Hospital A has better infant survival than the national average.

Case Study:

Florida analyzed vital records data to identify geographic areas where low birth weight and infant mortality rates were statistically higher than expected after adjusting for race, marital status and education level. Florida used an indirect method to calculate adjusted rates. This analysis informed further analysis and helped the Department of Health to focus efforts in counties that had significantly higher than expected infant deaths and low birth weight births.

For further detail, see information provided in the link below.

www.doh.state.fl.us/Family/mch/docs/infant_mortality_docs/ExpectedRatesByCounty09.pdf



3. Measuring Associations: Comparing Infant Mortality Rates

Measures of association quantify the infant mortality relationship between two groups. Measures of association include:

a. Relative Risk of Infant Death: A measure comparing two infant mortality rates, defined as an infant mortality rate in one group divided by an infant mortality rate in another group or more simply, the ratio of the two groups.

$$\frac{\left(\frac{\text{Number of deaths among infants (born in 2018-2020)}}{\text{Total number of live births among infants (born in 2018-2020)}} \right)}{\left(\frac{\text{Number of deaths among infants (born in 2011-2013)}}{\text{Total number of live births among infants (born in 2011-2013)}} \right)}$$

[Click here for excel document](#)

The relative risk has been preferred for etiological studies of association as it controls for baseline differences in mortality rates when comparing a risk factor across different groups, time periods or outcomes like infant mortality. The absolute difference or attributable risk (below) has been preferred for identifying the potential to prevent a poor outcome or disease and quantifying the actual numbers in the population affected.

b. Absolute Risk Difference of Infant Death (Excess Death Rate): A measure comparing two infant mortality rates across time periods, defined as the absolute difference in infant mortality between two groups, where one is subtracted from the other (typically expressed per 1,000 live births).

$$\left(\frac{\text{Number of deaths among infants (born in 2018-2020)} \times 1,000}{\text{Total number of live-born infants (born in 2018-2020)}} \right) - \left(\frac{\text{Number of deaths among infants (born in 2011-2013)} \times 1,000}{\text{Total number of live-born infants (born in 2011-2013)}} \right)$$

[Click here for excel document](#)

The absolute risk difference, also known as the attributable risk, is useful to quantify excess deaths between two time periods or populations. Absolute measures are necessary to

evaluate the meaning of a given relative risk. For example, a relative risk of 2.00 can be obtained when comparing the ratio of infant mortality rates of 20 to 10, a difference of 10 per 1,000; or comparing infant mortality rates of 2 to 1, a difference of only 1 per 1,000. When the excess death rate is multiplied by the number of births in the time period (or population) with greater infant mortality, this results in the number of excess deaths in that period/population that could have been prevented if they had the infant mortality rate of the reference period/population. Excess deaths provide a powerful tool to communicate the impact of infant mortality rate changes over time, or of geographic/demographic infant mortality rate disparities between populations. The absolute risk difference can be used to examine the attributable fractions by cause, age at death (neonatal/postneonatal), gestational age or birth weight.

c. Relative Risk of Cause-Specific Infant Death: A measure comparing two cause-specific infant mortality rates, defined as a cause-specific infant mortality rate in one group divided by a cause-specific infant mortality rate in another group.

$$\frac{\left(\frac{\text{Number of deaths from birth defects among infants (born in 2018-2020)} \times 1,000}{\text{Total number of live births among infants (born in 2018-2020)}} \right)}{\left(\frac{\text{Number of deaths from birth defects among infants (born in 2011-2013)} \times 1,000}{\text{Total number of live births among infants (born in 2011-2013)}} \right)}$$

[Click here for excel document](#)

d. Absolute Risk Difference of Cause-Specific Infant Death (Excess Cause-Specific Death Rate): A measure comparing two cause-specific infant mortality rates, defined as the absolute difference in cause-specific infant mortality rates between two groups where one is subtracted from the other (typically expressed per 100,000 live births).

$$\left(\frac{\text{Number of deaths from birth defects among infants (born in 2018-2020)} \times 1,000}{\text{Total number of live-born infants (born in 2018-2020)}} \right) - \left(\frac{\text{Number of deaths from birth defects among infants (born in 2011-2013)} \times 1,000}{\text{Total number of live-born infants (born in 2011-2013)}} \right)$$

[Click here for excel document](#)

For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

Cause-Specific Attributable Risk Example

In this example, absolute differences in cause-specific infant death rates for two periods/populations are calculated and used to determine the proportion of the overall Infant Mortality Rate difference that is attributable to various causes.

	Period/Population A	Period/Population B	Excess Cause-Specific Death Rate	Cause-Specific Proportion of Excess Deaths
Cause-Specific Infant Mortality Rate (per 100,000)			$(A_i - B_i)$	$(A_i - B_i) / (\sum A_i - \sum B_i)$
Preterm-related	222.6	237.6	-15.0	-18.7%
Other perinatal conditions	90.1	90.3	-0.3	-0.3%
Congenital anomalies	152.9	130.4	22.5	28.0%
SUID	136.8	92.9	43.9	54.7%
Infection	26.3	19.4	6.9	8.5%
Injury	32.9	22.5	10.3	12.9%
All other causes	69.5	57.6	11.9	14.9%
TOTAL (All causes)	731.0	650.7	80.2	100%

Period/population A has an excess of 80.2 deaths per 100,000 compared to period/population B ($731.0 - 650.7 = 80.2$). Period/population A had 242,090 births, so 194 deaths occurred that would have been prevented if period/population A had the infant mortality rate of period/population B ($80.2 \times 242,090 / 100,000 = 194$), and the Infant Mortality Rate would be lowered by 11% ($80.2/731.0$). The difference in SUID death rates between periods/populations A and B is 43.9 per 100,000 ($136.8 - 92.9 = 43.9$). Therefore, excess SUID in period/population A accounts for more than half of the total excess infant mortality in period/population A compared to B ($43.9 / 80.2 = 54.7\%$). Preterm-related mortality is actually lower in period/population A than period/population B; if the differences in all other causes were eliminated then period/population A would have lower infant mortality than period/population B. In other words, SUID, congenital anomalies, infection, injury, and other causes explain more than the total difference between period/population (summing to 119%) so the original difference would be reversed in favor of period/population A by 19%. To avoid percentages that are negative or over 100, it is also possible to determine cause-specific fractions only among causes that positively contribute to excess infant mortality. In this example, SUID accounts for 46% of positive causes of excess infant death in period/population A compared to B ($43.9 / (22.5+43.9+6.0+10.3+11.9) = 46\%$). Regardless of percentage method, SUID is the leading cause of excess infant mortality, followed by congenital anomalies, injury and infection.

This same method of excess death rates and excess death fractions can be applied to determine how much of a difference over time or between groups can be attributed to IMR differences in the neonatal versus postneonatal period.

4. Statistical Significance

(Click here to download excel document with formulas already included)

The term significant is used to indicate something important or meaningful that may warrant further discussion or exploration. This differs from the concept of statistically significant, which implies that a statistical test has been applied, and it has been determined that an observed difference is not likely to have happened by chance alone, based on a predetermined probability level (referred to as the alpha level or significance level). It is important for the investigator to differentiate and consider both statistical and non-statistical significance, because with large enough sample sizes, almost any difference can be statistically significant; while a smaller sample size may not have sufficient power to detect a statistically significant difference, but the difference is large enough to suggest programmatic action may be needed.

The following examples illustrate:

- 1) A difference that is statistically significant, but perhaps is not clinically or programmatically meaningful
~versus~
- 2) A difference that is not statistically significant, but perhaps is clinically or programmatically meaningful.

EXAMPLE 1

United States – Very Low Birth Weight (VLBW) Deliveries

	Number	Percentage
1996	53,425	1.4%
2008	61,773	1.5%

This absolute difference of 0.1% is statistically significant ($p < .05$); but consider if this statistically significant difference is significant in other ways. Some contextual questions you may ask yourself include, is a 0.1% fluctuation typical? Is the increase in 2008 the first such increase since 1996? What resources would it take to address that 0.1% increase?

EXAMPLE 2

County A – Very Low Birth Weight Deliveries

	Number	Percentage
1996	44	1.1%
2008	110	2.2%

This absolute difference of 1.1% is not statistically significant; but consider if this difference is significant in other ways. For example, it is possible that this increase in VLBW may cause a large increase in infant mortality rates. The same contextual questions as asked in example 1 apply in this case as well. Often, the investigator alone is not in a position to determine the answer of whether a difference is meaningful outside of statistical significance, which is why infant mortality investigations are best conducted as a team.

5. Methods for Determining Statistical Significance

a. Using Confidence Intervals to Determine Statistical Significance

A quick and easy way to interpret significant differences in rates or ratios is to use confidence intervals. They can be set at different levels, such as 95 percent and 99 percent. A confidence interval is a range of values for a variable of interest (e.g., an infant mortality rate constructed so that the range has a specified probability of including the real value of the variable). A 95 percent confidence interval would mean that if you were to repeat the analysis 100 times, you would expect the real value to be contained within the confidence interval 95 times.

Ratios: Confidence intervals (CI) are often applied to analyses where statistics are compared between two groups as a ratio. For example, if the infant mortality rate is 20.0 per 1,000 live births in Town A and 10.0 in Town B, the ratio is 2.0. Assuming that the 95 percent confidence interval around this ratio (95 percent CI) is 0.8 - 2.6, to determine if this is statistically significant, several questions must be asked:

1) Does the confidence interval include the number “1”?

If the answer is “yes”, then the ratio is not statistically significant, and the rates are not statistically significantly different from each other. In the above example, the 95 percent CI of 0.8 – 2.6 includes 1. This means that the infant mortality rates of Town A (20 per 1,000 live births) and Town B (10 per 1,000 live births) are not statistically significantly different from each other. *Note: Sample size affects confidence interval calculations. Smaller sample sizes increase the standard deviation, thereby increasing the width of the confidence interval.*

2) How wide is the confidence interval? Very little faith should be placed in confidence intervals that are too wide, such as a 95 percent confidence interval of 1.2 – 78.0. This confidence interval suggests a large amount of variance, but since it does not include “1” it is still statistically significant. You would have more faith in the difference being real if the 95 percent CI were narrower and not so close to 1 (e.g., 5.1 – 8.2). *Note: When the confidence interval is extremely wide, check the sample size.*

b. Using the Standard Error for Determining Statistically Significant Differences Between Two Rates

When you want to directly compare two rates, the standard error of the difference between two rates can be used to determine statistical significance. The difference between the two proportions may be regarded as statistically significant at the 95 percent confidence level if the difference exceeds 1.96 standard errors, as defined in the formula below. Or stated another way, the probability that the observed decline was due to random variation is less than .05, or 5 percent.

For Those Just Starting - Standing on Common Ground: Data Sources, Definitions, Basic Calculations *continued*

SAMPLE PROBLEM

There is a decline in the North Carolina infant mortality rate from 10.2 infant deaths per 1,000 live births (.01018 as a proportion) in the period 1991-1993 to rate of 8.2 (.00818 as a proportion) in 201-2003. Is this change statistically significant at the 95% confidence level?

Standard Error Formula:

$$\sqrt{\left(\frac{p1q1}{n1} + \frac{p2q2}{n2}\right)}$$

p1 = infant mortality rate in 1991-1993
 p2 = infant mortality rate in 2001-2003
 qn = 1-p1
 q2=1-p2
 n = number of live births

This formula will also work for testing the significance of a difference between two percentages (which are just rates per 100) and for rates expressed per 10,000 or per 100,000.

*Try from the beginning to calculate infant mortality rates and standard error:

SAMPLE PROBLEM

In 1991-1993, there were 3,132 infant deaths (d1) out of 307,567 live births (n1) and in 2001-2003 there were 2,892 infant deaths (d2) out of 353,711 live births (n2).

Hints:

First Step (Calculate p1 and p2):

p1 = d1/ n1
 p2 = d2/ n2

Second Step (Apply Standard Error Formula):

$$\sqrt{\left(\frac{p1q1}{n1} + \frac{p2q2}{n2}\right)}$$

p1 = rate 1
 p2 = rate 2
 qn = 1-pn
 d = # infant deaths
 n = # live births

ANSWER

The calculation of 1.96 standard errors of the difference for this example is:

$$\sqrt{\left(\frac{.01018 \times .98982}{307567} + \frac{.00818 \times .99182}{353711}\right)}$$

= .000236

Explanation:

Since the difference between the two proportions, 0.002 (i.e., 0.01018 - 0.00818), exceeds 1.96 standard errors of the difference (i.e., 0.000236), we can say that the decline in the infant mortality rate is statistically significant at the 95% confidence level. With much smaller numbers, for example for the infant mortality rate in a specific birth weight and cause-of-death category, or in a specific county within a state, a change of this magnitude might not be statistically significant.

[\(Click here for full solution with explanation\)](#)

*A Note about Statistically Significant Findings

Despite the utility of statistical significance in interpreting differences, it does not provide all of the answers. With the growing number of accessible analytic computer software packages, it has become increasingly easy to generate statistical tests and associated p values or confidence intervals. However, many statisticians and epidemiologists have raised concerns about placing too much weight on statistical significance, particularly for very small changes in rates. When deciding on whether or not to act on a particular change in rates, other policy, program and clinical significance of a difference should also be considered.

An Alternative Method for Identifying Significant Differences between Rates

We provide here a few relatively simple methods for identifying significant ratios and differences in rates. There are, however, alternative ways for determining significant differences between ratios and between rates. The NCHS recommendations for determining statistical significant differences between infant mortality rates can be found in the Technical Notes of its infant mortality reports, such as www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_05.pdf.