
Appendix B: Technical notes - definitions

Births

- **Apgar score** is a summary measure of the infant's condition based on heart rate, respiratory effort, muscle tone, reflex irritability and color. The highest possible score is 10. A low Apgar score (seven or less), measured five minutes after birth, indicates the infant is at increased risk of morbidity and mortality.
- **Births to unmarried mothers ratio** is the number of births to unmarried mothers per 1,000 live births. Ratios differ from rates.
- **Crude birth rate** is the number of live births per 1,000 total population.
- **Live birth** is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which, after such a separation, breathes or shows any other evidence of life such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered live born (1).
- **Low birthweight infant** is a live born infant with a birthweight of less than 5 pounds, 8 ounces (2,500 grams) as reported on the birth certificate.
- **Birth rate per 1,000 men** is the number of births per 1,000 males in Oregon. In computing birth rates by age of father, births tabulated as age of father not stated are distributed in the same proportions as births with known age within each five-year age classification of the mother. The male birth rate is used to facilitate comparisons between Oregon and the national rate.

The National Center for Health Statistics (NCHS) uses this procedure to avoid distortion in rates resulting from the disregard of the relationship between the mothers' and fathers' age.

Deaths

- **Contributing cause** of death is defined as any significant condition that contributed to the fatal outcome, but was not related to the disease or condition directly causing death (see the underlying cause of death definition below) (2).
- **Crude death rate** is the number of deaths per 1,000 or 100,000 total population.
- **Fetal death** is death prior to the complete expulsion or extraction from its mother of a product of conception of at least 20 weeks' gestation, except where such expulsion results from a therapeutic abortion; the death is indicated by the fact that after such separation, 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.
- **Fetal death ratio** is the number of fetal deaths per 1,000 live births. Ratios differ from rates.
- **Infant death** is the death of a child within the first year of life.
- **Infant death rate** is the number of infant deaths per 1,000 live births.
- **Maternal death rate** is the number of female deaths attributed to childbirth or to complications of pregnancy or the puerperium, per 100,000 live births.
- **Neonatal death** is the death of a child within the first 27 days of life.
- **Neonatal death rate** is the number of neonatal deaths per 1,000 live births.
- **Perinatal death** is the death of a fetus after 20 weeks' gestation or the death of a live-born infant prior to the 28th day of life. Other medical literature may include different time periods.
- **Perinatal death ratio** is the number of perinatal deaths per 1,000 total live births. Ratios differ from rates.
- **Postneonatal death** is the death of a child from day 28 through 364 after birth.

- **Postneonatal death rate** is the number of postneonatal deaths per 1,000 live births.
- **Underlying cause** is the disease or injury that initiated the train of events leading directly to death, or the circumstances of the accident or violence that produced the fatal injury (2).
- **Years of potential life lost (YPLL)** is the numerical difference between a predetermined end point age, usually 75 years, and the age at death. YPLL quantifies premature deaths occurring in younger age groups.

Medical personnel - abbreviations used in tables

- C.N.M. — certified nurse midwife
- D.C. — doctor of chiropractic medicine
- D.O. — doctor of osteopathic medicine
- L.D.M. — licensed direct entry midwife
- M.D. — medical doctor
- N.D. — naturopathic doctor
- R.N. — registered nurse

Endnotes

1. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics. Vital statistics of the United States, 1982, vol. 1, section 4, page 1. Hyattsville, Maryland; 1986.
2. World Health Organization. International statistical classification of diseases and related health problems, 10th revision, 2010, vol. 10, page 33. [Internet]. 2010 [cited 2016 Jan 22]. Available from: http://apps.who.int/classifications/icd10/browse/Content/statichtml/ICD10Volume2_en_2010.pdf.

Appendix B: Technical notes - methodology

“That, sir, is the good of counting; it brings everything to a certainty, which before floated in the mind indefinitely.”

—Samuel Johnson

Induced termination of pregnancy

The induced termination of pregnancy data in this report represents nearly all abortions performed in Oregon during the current data year. Missing data is due to incomplete reporting by providers. Another consideration is the place of occurrence (Oregon) versus the mother’s place of residence (residence could be anywhere). That is, the data constitute events associated with the place of occurrence rather than the “residence data” used in estimating births. This is necessary because many abortions obtained out-of-state by Oregon residents are not reported to Oregon’s Center for Health Statistics. It reflects the great variation in abortion reporting procedures among states (e.g., some states do not record patients’ residence) as well as the fact that a comprehensive data collection network among all states, similar to that used in reporting births, does not exist in regard to abortions.

In using “occurrence” data rather than “residence” data to estimate abortion rates for Oregon residents, an implicit assumption is made that the number of Oregon residents that leave the state to obtain an abortion equals the number of out-of-state residents who obtain an abortion in Oregon. In formulating generalizations that involve trends or long-term behavioral patterns, annual totals are treated as sample values generated by ongoing social, economic or political processes and thus subject to “chance” variability. For most purposes, numbers offered in this report should be viewed as careful approximations and interpreted only within the framework of statistical safeguards developed to take sampling variability into account.

Some rates in the Induced Terminations of Pregnancy section are based on relatively few events, and for most comparisons may be used only with extreme caution—due to the chance fluctuations associated with small numbers. A small percentage of abortion reports lack certain data items.

Estimation of the cumulative proportion of females who have experienced an abortion

This estimate is computed by tracing the abortion experience of a specific cohort of females over an extended time period. In the table below, an approximation of the “cumulative total” of first-time abortions by one of the cohorts may be obtained by summing the numbers in the boxed area.

Number of First-Time Abortions By Year and Age Group, Oregon Occurrence, 1991-2005						
YEAR	AGE GROUPS					
	15-19	20-24	25-29	30-34	35-39	40-44
91	2584	2678	1190	716	402	122
92	2137	2396	1067	655	380	117
93	2267	2393	1176	598	357	117
94	2370	2379	1233	693	376	135
95	2510	2486	1402	755	463	144
96	2511	2566	1416	771	468	152
97	2679	2794	1502	835	501	151
98	2525	2679	1496	786	495	175
99	2426	2776	1482	803	503	163
00	2270	2888	1499	827	487	176
01	2194	3018	1445	826	481	149
02	1840	2665	1383	836	443	181
03	1839	2575	1270	749	420	165
04	1607	2370	1232	710	396	152
05	1605	2307	1261	729	427	178

To obtain this value, it is necessary to sum the number of first-time abortions for 15- to 19-year-olds from 1991 to 1995 and those of 20- to 24-year-olds from 1996 to 2000 with those of 25- to 29-year-olds from 2001 to 2005. This provides an estimate of the numerator in the following equation:

$$\begin{array}{l}
 \text{Cumulative proportion of females} \\
 \text{who have had an abortion}
 \end{array}
 =
 \frac{\text{Total number of first time abortions} \\
 \text{among a specific cohort of females}}{\text{Number of females in cohort}}$$

The denominator may be estimated by averaging the size of the cohort during 1991 to 1995. Table A-1 lists the annual estimate of the number of females within each cohort. For example, in 1991, the number of 15- to 19-year-old females

was estimated to be 93,043; in the next year, it was 95,064.

The average size of this age group from 1991 to 1995 was 98,540. Similarly, the number of 20- to 24- year-old women between 1996 and 2000 was 104,214 on average; the number of 25- to 29-year-olds averaged 93,065 between 2001 and 2005. Thus, between 1991 and 2005 the cohort of interest had an average population size of 98,606.

Substituting into the formula given above:

$$Cp = \frac{\text{Sum of First Abortions}}{N} = \frac{32,162}{98,606} = 0.326 \text{ or } 32.6 \text{ percent}$$

This number approximates the proportion of females in the 25- to 29-year-old cohort who, by 2005, had ever had an abortion. This method of estimation assumes factors such as deaths and migration have not altered the composition of the female population in Oregon—that is, the women who left the state displayed the same characteristics as those who have moved into Oregon. It also assumes patients with a history of previous abortions do not report the current procedure as a first abortion.

Teen pregnancy

Teen pregnancy counts include live births and induced terminations of pregnancies; they do not include fetal deaths or miscarriages (spontaneous abortions).

- Birth counts include births to teens whose primary residence is in another state.
- Teen abortion counts are based on all reported abortions to teenage Oregon residents; however, because states often do not report abortions obtained within their borders to the state of residence, as occurs with vital events such as birth and death, an unknown number of Oregon teens obtain abortion services out-of-state. As a consequence, counts of Oregon resident teen abortions and pregnancies should be considered incomplete.

Furthermore, because teen abortion counts are based on “residence data,” figures given in Chapter 4 do not correspond exactly to those in Chapter 3 that are based on “occurrence data.” (See Induced Terminations of Pregnancy methodology section.) The estimation of rates requires an estimate of the size of the appropriate population. Such

Teen Birth Rates, U.S. vs. Oregon, Ages 15-19, 2008		
Race/Ethnicity	Birth Rate ¹	
	U.S.	Oregon
TOTAL*	41.5	34.0
Non-hispanic whites	26.7	26.7

¹ All rates per 1,000 females ages 15-19.
* All races and ethnicities combined.

estimates are now available for 15- to 17-year-olds and 18- to 19-year-olds for each Oregon county on an annual basis.

Rates based upon a small population increase the likelihood of variation in the data due to the influence of chance factors. For this reason, rates of teen pregnancy, birth and abortion were calculated only if each age category contained at least 50 female residents of the specified county.

Great caution must be taken in the use of pregnancy statistics associated with females under 15 years of age. This is because relatively few events are recorded each year for this group. Also, rates are based on the estimated population cohort of 10- to 14-year-old females—many of whom are physiologically not yet at risk of pregnancy. Thus, any direct comparison of rates between this group and another age group—e.g., 15- to 17-year-olds—would be inappropriate.

Demographics

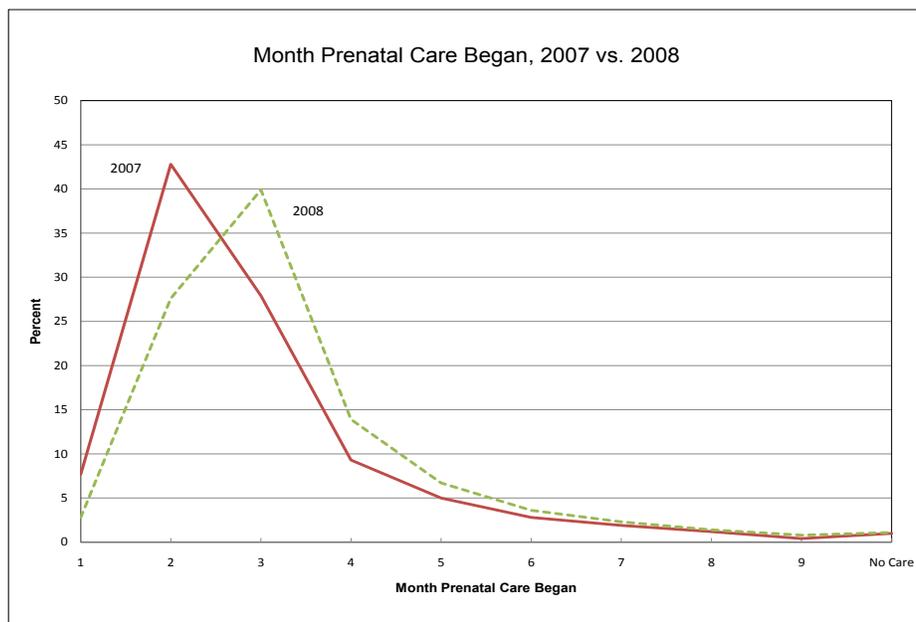
The extent to which Oregon’s demographic composition may affect its national ranking is indicated by comparisons shown in the sidebar. In 2008, Oregon’s birth rate for all teens (regardless of race or ethnic affiliation) was 7.5 percent lower than that of the United States; among all 50 states, it had the 20th lowest teen birth rate. Yet, if comparisons were made in terms of births to non-Hispanic White teens only, Oregon would have been 36th and the rate would have been 19 percent higher than that of the United States. This results from the fact that 87 percent of 15- to 19-year-old females in Oregon were non-Hispanic Whites; only 7 percent were either Hispanic or non-Hispanic African Americans. By comparison, 70 percent of the U.S. female population of that age were non-Hispanic Whites, and 26 percent were Hispanics or non-Hispanic African Americans.

Prenatal care

In 2008, information on the timing of prenatal care was based on the difference between the date of first prenatal visit and the date of last normal menses. When the data of last normal menses is missing or invalid, the clinical estimate of gestation is used. This change has made direct comparison between 2007 data and 2008 data unreliable.

Prenatal care information based on the revised system suggests a markedly less favorable picture of prenatal care

use than data from 2007. In 2008, prenatal care began in the first month of pregnancy in 2.8 percent of births, while in 2007 prenatal care began in the first month in 7.7 percent of births. Most of this difference is likely attributable to the changes in data collection rather than changes in prenatal care utilization.



Race and ethnicity

The Center for Health Statistics began collecting multiple race and ethnicity information for decedents in 2006. Prior to 2006, Oregon's data systems were limited to a single race. In 2006, Oregon adopted the 2003 revision of the U.S. standard death certificate. Oregon now collects up to four Hispanic ethnicities and 36 races for each decedent. This change in data led to the revision of tables including race and ethnicity information in the annual report. More detailed reporting for race and ethnicity began in 2008 for birth and fetal death records.

Collection and reporting of race and ethnicity

Source of information

Birth, death and fetal death race and ethnicity information is collected about the subject of the vital record from the best available source.

Birth and fetal death —The birth mother usually provides the race and ethnicity information for birth and fetal death records, but occasionally another family member, such as the father or a grandparent, provides it. The mother is asked to identify her race and Hispanic ethnicity as well as the father's/second parent's race and ethnicity. No race or ethnicity information is collected about the child in Oregon statistical data. All statistical tables in this report present information on the mother.

Death — The informant, usually a close family member, reports race and ethnicity on the “Report of Death.” However, there are deaths where no close family member is identified, and the information is obtained from a friend, police officer or facility staff. In 2014, the informant was the spouse or domestic partner on 30.3% of records, a child of the decedent on 45.8% and a parent or sibling of the decedent on 12.6% of records. Combined, 88.7% of informants were immediate family to the decedent.

Each informant is allowed to identify the race or races and Hispanic ethnicity or ethnicities of the decedent to the best of their knowledge. Race and ethnicity are intended to be self-identified and are not defined by parentage or national origin.

Categories collected

Oregon collects up to four Hispanic ethnicities (Mexican, Cuban, Puerto Rican and Other). Hispanic ethnicities can be chosen in combination.

Oregon collects up to 36 race categories. These include: White, Black or African American, American Indian or Alaska Native, Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Other Asian, Native Hawaiian, Guamanian or Chamorro, Samoan, Other Pacific Islander, Other, and Unknown. The informant can specify up to two races if he/she selects one of the non-specific race categories, specifically, other Asian, other Pacific Islander and other. If the decedent is reported as American Indian or Alaska Native, the informant is asked to indicate up to two tribes. Enrollment or official affiliation is not required to report a tribal relationship.

Birth and fetal death — Hospitals, birthing centers and midwives are required to use a standard parent worksheet to collect information from the mother. The worksheet specifies

each category and allows for hand entry for more specificity. The worksheets can be viewed at <https://public.health.oregon.gov/BirthDeathCertificates/RegisterVitalRecords/Pages/InstructionsBirth.aspx>.

Death — Funeral service practitioners are instructed to ask open-ended questions on ethnicity and race when gathering information for the “Report of Death.” The Oregon Center for Health Statistics has provided letter-sized cardstock forms that list all race and ethnicity categories to assist the family in reporting accurately. “Other” and “Unknown” are options for both race and ethnicity.

Presentation of data

The Center for Health Statistics creates tables based on numeric codes associated with the races (including “other specify”) and ethnicity reported. The Center for Health Statistics sends record level data to the National Center for Health Statistics. The National Center for Health Statistics then processes the data to create numeric codes that are assigned to more than 300 literal race categories. This allows the coding to be standardized nationwide. An example of the detailed listing is available at www.cdc.gov/nchs/data/dvs/Appendix_E_Accessible_Race_Code_List_Update_2011.pdf.

The race codes are three digits, with the first digit representing a category and the last two digits representing a specific group. For example, white checkbox is 100, white literal is 101, Arab is 102, English is 103, French is 104, and so on through Kosovian at 134. These numeric codes are used to create the statistical tables. Considering the space available to relay information, most tables report categories based on the code’s first digit. The tables in this report present the five major race and ethnicity categories used at the Center for Health Statistics: White, African American or Black, American Indian, Asian, Pacific Islander, Hispanic (any race), and Other.

Multiple race — Although Oregon collects multiple races for each record, for deaths occurring in 2014 only 508 or approximately 1.5 percent of 34,160 resident decedents were reported as belonging to multiple races. The mean age of the decedents decreased as more racial categories were reported. Generally, younger decedents selected more race categories than their older counterparts. The mean age for

decedents with only one race indicated was 75 years, while the mean age for decedents where two or more races were reported was 52 years.

Examples of multiple race tables include 6-10 and 6-12 in Volume 2 of the annual report. In these two tables, individual decedents can be listed in more than one race category. If a decedent is listed as both White and African American/Black on the “Report of Death,” he or she would be included in the totals for both White and Black in the multiple race tables. This means the race category totals will exceed the total number of deaths in tables reporting multiple races. In tables presenting single-mention race, persons with two or more race selections are included in the “two or more races” total. Compare multiple race tables (e.g., 6-10 and 6-12) with similar single-mention race tables (e.g., 6-9 and 6-11) to determine the practical impact of this distinction.

Other table conventions include reporting Hispanic as a separate category in most tables that include race or ethnicity. This means records with Hispanic ethnicity are removed from the single-mention race categories in most tables. Persons of Hispanic ethnicity may belong to any race category (or categories). Footnotes in tables presenting race and ethnicity indicate when records with Hispanic ethnicity reported are removed from the race categories. These tables will also include “Non-Hispanic Single Mention Race” as a header title. There are two primary reasons for this reporting convention. First, many Hispanic individuals identify their race as “Other” (in 2008, 77.3 percent of decedents with other or unknown race were Hispanic). Second, “Non-Hispanic White” is often used as a reference category when doing statistical analysis, allowing the information contained in the tables to be used as an effective reference group.

Tobacco

National Healthy People 2020 objective (1)

Percentage of infants whose mothers did not use tobacco during pregnancy (self-reported).

Year 2020 target:	98.6 %
2007:	89.6 %

Women who smoke when pregnant have a far higher incidence of low birthweight babies than nonsmokers. Low birthweight infants experience more serious health problems, including increased rates of infant mortality. In 2008, the Oregon infant mortality rate during the first 27 days of life (neonatal) was 51.8 per 1,000 live births for low birthweight (less than 2,500 grams) infants compared to 0.7 per 1,000 for infants with birthweights of 2,500 grams or more. Women who smoked had a low birthweight rate of 84.7 per 1,000 live births, compared to 57.1 per 1,000 among women who did not smoke. One of nine mothers (11.8 %) reported using tobacco during pregnancy, a proportion that is among the lowest observed in the last 20 years. (See sidebar 2-D, page 2-8.) The percentage of tobacco use among unmarried women was nearly four times that of married women (22.9 % vs. 5.6 %). The highest percentage of tobacco use during pregnancy in 2008 was among unmarried mothers aged 20–24 and unmarried mothers aged 25–29 (24.7% and 24.3% respectively). Generally, the percentage of mothers who reported smoking during pregnancy decreased with age regardless of marital status. The lowest percentage of smokers was reported for married mothers aged 35–39 (2.9 %). (See Figure 2-5.)

Endnotes

1. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, Healthy People 2020: <http://www.healthypeople.gov/2020/topics-objectives/topic/maternal-infant-and-child-health/objectives>. Accessed Jan 25, 2016.

Appendix B: Technical notes - step-by-step instructions

“Through and through the world is infested with quantity: To talk sense is to talk quantities. It is no use saying the nation is large—How large? It is no use saying that radium is scarce—How scarce? You cannot evade quantity. You may fly to poetry and music, and quantity and number will face you in your rhythms and your octaves.”

—Alfred North Whitehead

DEATHS
 INFANT DEATHS
 NEONATAL DEATHS
 POSTNEONATAL DEATHS
 FETAL DEATHS
 LOW BIRTHWEIGHT INFANTS
 PREGNANCIES
 INDUCED ABORTIONS
 MARRIAGES
 ANNULMENTS
 DIVORCES

Data users are diverse, including public health officials evaluating a program by using death data, demographers projecting school enrollments with birth data and business people deciding to open a formal-wear shop based on marriage data. Many of these users have a thorough

knowledge of statistics. But others find the entire subject matter confusing and intimidating. For either group, a misunderstanding of what vital statistics mean can lead to wrong conclusions. Therefore, this section is included as an overview of how to use vital statistics. It is addressed to the person looking at vital events for the first time, but the experienced user may also find a review helpful.

Step 1: Finding the correct number

The first step is to determine how many instances of a particular vital event took place during the year. This involves asking two questions:

Which event or events are appropriate?

This may not be as simple as it sounds. For one thing, examining more than one type of event may be required. For example, someone concerned with teenage pregnancies will have to consider the number of induced abortions as well as the number of births that occur among teens. Taken together, they provide a useful measure of the number of pregnancies (1).

Deciding which events to use is important since sometimes the choice of one event over another can easily lead to different conclusions. To determine which events are appropriate, read the “Technical notes: Definitions” section. The narratives also contain useful examples.

Who should be counted?

If you are a hospital planner who is deciding to expand or contract delivery services, you want to count the number of births that occurred in your area, regardless of where the parents live. If you are projecting school enrollment, you want to count only how many children will potentially be residing in your area. Fortunately, vital events are usually reported so both of these data needs can be met.

Occurrence data:

The event (the death, birth, marriage, etc.) actually took place in the geographic region indicated (either Oregon or a particular county). The person participating in the event may have lived in Podunk, New York.

Residence data:

The person involved in the event lived in the geographic region mentioned, but the event itself may have taken place anywhere in the United States or Canada. In other words, a resident of Marion County who died in an accident while on vacation in Michigan has been added to the Marion County resident death figure.

When in doubt about which type of data to use, resident figures are usually the best choice. Most birth and death data are published by residence, which means comparisons with other states or the United States as a whole will be easier. Exceptions to this rule are listed in the individual sections.

Once the right event has been determined and the choice between occurrence and residence data has been made, the statistician can find the correct figures in the tables in this book. If the needed table is not listed, contact the Center for Health Statistics for more information.

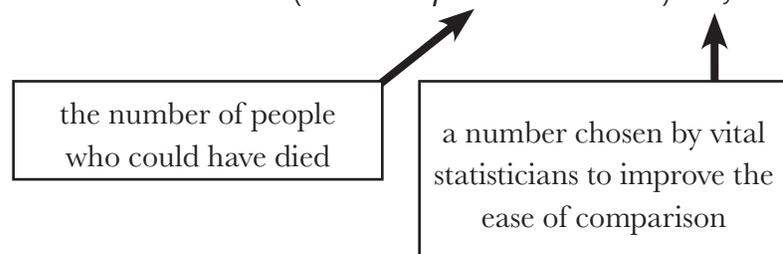
Step 2: Making the number meaningful with rates and ratios

In many instances, simply knowing the number of events is not sufficient. For example, we know more people died in Multnomah County than in Wheeler County because Multnomah County has a much larger population. But what is the likelihood of dying in each county?

In order to answer this question, statisticians calculate rates. This means the number of events is compared to the population for which that event could have occurred, and the figure is then standardized to some number (such as 1,000 or 100,000) for convenience.

Here is an example:

$$\text{CRUDE DEATH RATE} = (\text{DEATHS}/\text{POPULATION}) \times 1,000$$



The more specifically a statistician can define the “population at risk” (the denominator or bottom part of the formula), the more meaningful the rate is. For example, the crude birth rate, which compares the number of births to the population, is not nearly as informative as the fertility rate, which uses only the number of women of childbearing age (15-44) for comparative purposes. The fertility rate is not distorted by changes in the number of men or prepubescent or post-menopausal women in the population. (The turn of the century notion that only married women between the ages of 15 and 44 would be considered at risk of pregnancy has been abandoned for obvious reasons.)

When calculating rates and ratios, great care must be taken to make certain the appropriate time periods, geographical boundaries and populations are used.

Unfortunately we do not always have the correct denominator for the equation. In these situations a substitute is used. For example, how many people are at risk of getting divorced? The number of married people is only available for census years. As a substitute, the crude divorce rate is calculated using the total population

regardless of marital status. In other situations, the event is simply compared to another related number. For instance, the abortion ratio compares the number of abortions to the number of births. This is easier and more accurate than trying to determine the true denominator, which is the total number of pregnant women.

Step 3: Comparing two or more numbers

Numbers are more meaningful when they are converted into rates and ratios. But problems can arise when rates or ratios are compared for different geographical areas, different time periods, or different categories such as men versus women.

Chance variation

Statisticians expect a certain amount of chance variation and have methods to take this into account. The confidence interval uses the number of cases and their distributions to determine what the rate “really is.” Hypothetically, a statistician will say, “We are 95% sure the true infant death rate for Oregon in 2008 was 9.47 ± 0.97 ; that is, it lies somewhere between 8.50 and 10.44.” If two rates have overlapping confidence intervals, then the difference between them may be due to this chance variation. In other words the difference is not statistically significant.

When comparing rates and ratios, differences should be tested for statistical significance. Formulas are listed in the next section of this chapter.

Small numbers

Chance variation is a common problem when the numbers being used to calculate rates are extremely small. Large swings often occur in the rates that do not reflect real changes. Consider Clatsop County’s infant mortality rates for a five-year period.

CLATSOP COUNTY			
YEAR	BIRTHS	INFANT DEATHS	INFANT DEATH RATES
2001	380	1	2.63
2002	432	6	13.89
2003	367	6	16.35
2004	397	2	5.04
2005	411	1	2.43
2001-2005	1,987	16	8.1

Clatsop County's five-year infant death rate is 8.1, which is 2.5 percentage points higher than the state rate (5.6). Yet, for some years Clatsop's rate is more than six times as high as the rate of other years simply because five additional infants died. Public health officials would waste a good deal of energy reacting to these annual rates.

Many rates based on small numbers are published in this book because readers demand them. However, anyone preparing to make important decisions based on these rates should be wary. Consider this rule of thumb: A rate based on 20 cases has a 95 percent confidence interval about as wide as the rate itself (i.e., the interval for a rate of 50 is between 25 and 75). Even large differences between two rates based on 20 cases or less are probably not statistically significant.

If 20 are too few, how many cases are sufficient to say a true difference exists? Unfortunately, we have no easy rules for this. To be safe, the vital statistician should always try to combine several years of data or consolidate geographical areas. Confidence intervals should be calculated, and differences should be tested for statistical significance.

Changes in measurement

Another problem is that the numbers being compared have not always been based on the same type of measurement. Definitions, population estimates, certificates and coding procedures change from time to time as the need arises. This can create "artificial" differences and can disguise "real" differences. The following cause-of-death item provides an excellent example in comparability:

It appears that the incidence of hypertensive disease increased. But actually, a new coding scheme resulted in more deaths being coded as due to hypertensive disease.

During the late 1970s, approximately 80 to 85 people died each year due to hypertensive disease.	Rate = 3.3 per 100,000 population
In 1979, 250 people died from this cause.	Rate = 9.8 per 100,000 population

Taking age, sex, and race into account

Mr. G.C. Whipple noted in 1923 that, “We might find that the death rate of bank presidents was higher than that of newsboys; but this would not be because of different occupations, but because of different ages.” We expect older people to die at a higher rate than younger people. We also expect people in their 20s to have more babies than the very young or the very old. Sex and race, as well as age, can drastically affect rates.

When comparing two places or two points in time, it is necessary to take these influencing characteristics into account.

To the right is an example.

The crude death rate increased between 1950 and 1960 from 9.1 to 9.5 deaths per 1,000 population.

But, an examination of the age-specific death rates for each

	1950	1960
Crude death rate	9.1	9.5
Age-specific death rates		
0-4	5.9	5.7
5-14	0.6	0.4
15-24	1.5	1.1
25-44	2.4	2.1
45-64	11.1	10.6
65+	58.4	56.8

group indicates that all these rates decreased. This apparent contradiction is explained by the fact that in 1960 a larger proportion of the population was older. Because the risk of death is higher in older persons, the crude death rate increased.

Before comparing two places or two time periods, always compare the population characteristics first. If discrepancies are noted in any relevant variables, the rates should be adjusted or standardized in order to make the comparisons free of differences in the structure of the populations. The formulas for doing this are listed in the following section.

Step 4: Analyzing the data

The first three steps have been fairly mechanical:

- (1) = Choose the correct events and the correct group to determine the number of events that took place for the geographical areas and time periods.
- (2) = Calculate the rates.
- (3) = Compare these rates to determine if the differences are statistically significant.

NOW the vital statistician must begin to ask the difficult questions. If we find that two rates are significantly different, how can we find out why they are different? If the differences we expected did not prove to be significant, is there another item that perhaps is masking an actual difference? Frequently, the statistician has to refine the research question and begin all over again.

Consider the researcher who asks, “Since 2005, has chronic lower respiratory disease (CLRD) posed a greater risk to Oregonians?” If the researcher looked at the overall rate, the answer would be “yes,” but closer examination reveals that the death rate for males has declined. It is among women that the rate has moved sharply upward, reflecting their increased smoking prevalence during recent decades. This gender dichotomy would need to be addressed in a study of CLRD fatalities.

Help

Several sources of help are available. Many of the widely used rates and ratios are presented in the Quick Reference section, and narratives and figures are included throughout this report to illustrate changes. And finally, Center for Health Statistics’ staff are available to help data users.

Endnotes

1. A more complete and accurate estimate of pregnancies based on outcomes would include: (1) births; (2) fetal deaths (stillbirths); (3) induced abortions; and (4) spontaneous abortions (miscarriages). However, fetal deaths occur in less than 1 percent of all pregnancies and are relatively constant in relation to births (see the Fetal and Infant Mortality chapter in Volume 2) and the number of miscarriages that occur is not available in vital records. Nevertheless, a measure that excludes these outcomes provides an adequate indicator of the number of pregnancies.

Appendix B: Technical notes - formulas

GENERAL:

$$\text{PERCENT CHANGE} = \frac{\text{New Data} - \text{Old Data}}{\text{Old Data}} \times 100$$

$$\text{Birth rate, Oregon, 1993} = 13.7$$

$$\text{Birth rate, Oregon, 1994} = 13.6$$

$$\text{Percent change} = \frac{13.6 - 13.7}{13.7} \times 100 = -0.7\%$$

PREGNANCY:

$$1. \text{ (CRUDE) BIRTH RATE} = \frac{\text{Resident Births}}{\text{Population}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{41,832}{3,082,800} \times 1,000 = 13.6$$

$$2. \text{ AGE-SPECIFIC BIRTH RATE} = \frac{\text{Resident Births To Mothers in Age Category}}{\text{Female Population in Age Category}} \times 1,000$$

$$\text{Oregon, 1994, Age 20-24} = \frac{10,999}{104,718} \times 1,000 = 105.0$$

$$3. \text{ FERTILITY RATE} = \frac{\text{Resident Births to Mothers Aged 15-44}}{\text{Female Population Aged 15-44}} \times 1,000$$

NOTE: Some publications use the following: $\frac{\text{All Resident Births}}{\text{Female Population Aged 15-44}}$

$$\text{Oregon, 1994} = \frac{41,659}{682,428} \times 1,000 = 61.0$$

$$4. \text{ TOTAL FERTILITY RATE} = \left(\text{The Sum of Age Specific Birth Rates in 5-Year Categories between 15 and 44} \right) \times 5$$

$$\text{Oregon, 1994} = 5 (51.3 + 105.0 + 115.4 + 78.5 + 30.2 + 6.0) = 1,932.0$$

$$5. \text{ FETAL DEATH RATIO} = \frac{\text{Resident Fetal Deaths (350+ grams Birthweight)}}{\text{Resident Live Births}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{224}{41,832} \times 1,000 = 5.4$$

$$6. \text{ FETAL DEATH RATE} = \frac{\text{Resident Fetal Deaths (350+ grams Birthweight)}}{\text{Resident Live Births} + \text{Resident Fetal Deaths}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{224}{43,591 + 224} \times 1,000 = 5.1$$

$$7. \text{ PERINATAL DEATH RATE} = \frac{\text{Resident Neonatal Deaths} + \text{Resident Fetal Deaths (350+ grams Birthweight)}}{\text{Resident Live Births} + \text{Resident Fetal Deaths}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{148 + 203}{41,566 + 203} \times 1,000 = 8.4$$

Note: Publications vary in the definition of fetal deaths. In addition, some measures employ gestational age in place of birthweight. Fetal and perinatal death rates are based on year of birth.

$$8. \text{ ABORTION RATIO} = \frac{\text{Resident Abortions}}{\text{Resident Births}} \times 1,000 \text{ or } \frac{\text{Occurrence Abortions}}{\text{Occurrence Births}} \times 1,000$$

$$\text{Oregon, 1994, Occurrence} = \frac{13,392}{43,591} \times 1,000 = 307.2$$

$$9. \text{ ABORTION RATE} = \frac{\text{Resident Abortions or Occurrence Abortions}}{\text{Female Resident Population Aged 15-44}} \times 1,000$$

$$\begin{aligned} \text{Oregon 1994, Occurrence} \\ \text{with total adjusted} \\ \text{for unknown ages} \end{aligned} = \frac{13,300}{682,428} \times 1,000 = 19.5$$

DEATHS:

$$10. \text{ (CRUDE) DEATH RATE} = \frac{\text{Resident Deaths}}{\text{Population}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{27,361}{3,082,000} \times 1,000 = 8.9$$

$$11. \text{ INFANT DEATH RATE} = \frac{\text{Resident Infant Deaths}}{\text{Resident Births}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{295}{41,832} \times 1,000 = 7.1$$

$$12. \text{ NEONATAL DEATH RATE} = \frac{\text{Resident Neonatal Deaths}}{\text{Resident Births}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{164}{41,832} \times 1,000 = 3.9$$

$$13. \text{ POSTNEONATAL DEATH RATE} = \frac{\text{Resident Postneonatal Deaths}}{\text{Resident Births}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{131}{41,832} \times 1,000 = 3.1$$

$$14. \text{ CAUSE-SPECIFIC DEATH RATE} = \frac{\text{Resident Deaths Due to Specific Cause}}{\text{Population}} \times 100,000$$

$$\text{Oregon, 1994, Heart Disease} = \frac{7,417}{3,082,000} \times 100,000 = 240.7$$

$$15. \text{ AGE AND SEX-SPECIFIC DEATH RATE} = \frac{\text{Resident Deaths in Age-Sex Category}}{\text{Population in Age-Sex Population}} \times 1,000$$

$$\text{Oregon, 1994, Males Aged 5-14} = \frac{63}{225,880} \times 100,000 = 27.9$$

MARRIAGE AND DIVORCE:

$$16. \text{ MARRIAGE RATE} = \frac{\text{Marriages}}{\text{Population}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{25,194}{3,082,000} \times 1,000 = 8.2$$

$$17. \text{ DIVORCE RATE} = \frac{\text{Divorces}}{\text{Population}} \times 1,000$$

$$\text{Oregon, 1994} = \frac{15,844}{3,082,000} \times 1,000 = 5.1$$

Beginning with 1998 data, the following methodology is being used for calculating confidence intervals and statistical significance. This explanation is paraphrased from *"Public Health Data: Our Silent Partner"*, a training manual from the Public Health Practice Program Office of the National Center for Health Statistics.¹

CALCULATING CONFIDENCE INTERVALS FOR RATES:

Confidence limits for rates based on less than 100 events

When the number of events in the numerator is less than 100, the confidence interval for a rate can be estimated using the two formulas which follow and the values in Table B-1.

Lower Limit = R x L

Upper Limit = R x U

where:

R = the rate

L = the value in Table B-1 that corresponds to the number N in the numerator of the rate

U = the value in Table B-1 that corresponds to the number N in the numerator of the rate

Example: Confidence limits for rates based on less than 100 events

In Baker County, the teen pregnancy rate for 10- to 17-year-old teens in 1998 was 13.0 per thousand, based on 12 live births in the numerator. Using Table B-1:

$$\text{Lower Limit} = 13.0 \times 0.51671 = 6.7$$

$$\text{Upper Limit} = 13.0 \times 1.7468 = 22.7$$

This means that the chances are 95 out of 100 that the pregnancy rate in Baker County for teens 10-17 lies between 6.7 and 22.7 per 1,000. So if there were 100 counties like Baker County, the teen pregnancy rate would be expected to lie between 6.7 and 22.7 per 1,000 in 95 of these counties.

TABLE B-1.
 Values of L and U for calculating 95% confidence limits for the numbers of events
 and rates when the number of events is less than 100.

N	L	U	N	L	U	N	L	U
1	0.02532	5.57164	34	0.69253	1.3974	67	0.77499	1.26996
2	0.1211	3.61234	35	0.69654	1.39076	68	0.77654	1.26774
3	0.20622	2.92242	36	0.70039	1.38442	69	0.77806	1.26556
4	0.27247	2.5604	37	0.70409	1.37837	70	0.77955	1.26344
5	0.3247	2.33367	38	0.70766	1.37258	71	0.78101	1.26136
6	0.36698	2.17658	39	0.7111	1.36703	72	0.78244	1.25933
7	0.40205	2.06038	40	0.71441	1.36172	73	0.78384	1.25735
8	0.43173	1.9704	41	0.71762	1.35661	74	0.78522	1.25541
9	0.45726	1.89831	42	0.72071	1.35171	75	0.78656	1.25351
10	0.47954	1.83904	43	0.7237	1.34699	76	0.78789	1.25165
11	0.4992	1.78928	44	0.7266	1.34245	77	0.78918	1.24983
12	0.51671	1.7468	45	0.72941	1.33808	78	0.79046	1.24805
13	0.53246	1.71003	46	0.73213	1.33386	79	0.79171	1.2463
14	0.54671	1.67783	47	0.73476	1.32979	80	0.79294	1.24459
15	0.55969	1.64935	48	0.73732	1.32585	81	0.79414	1.24291
16	0.57159	1.62394	49	0.73981	1.32205	82	0.79533	1.24126
17	0.58254	1.6011	50	0.74222	1.31838	83	0.79649	1.23965
18	0.59266	1.58043	51	0.74457	1.31482	84	0.79764	1.23807
19	0.60207	1.56162	52	0.74685	1.31137	85	0.79876	1.23652
20	0.61083	1.54442	53	0.74907	1.30802	86	0.79987	1.23499
21	0.61902	1.52861	54	0.75123	1.30478	87	0.80096	1.2335
22	0.62669	1.51401	55	0.75334	1.30164	88	0.80203	1.23203
23	0.63391	1.50049	56	0.75539	1.29858	89	0.80308	1.23059
24	0.64072	1.48792	57	0.75739	1.29562	90	0.80412	1.22917
25	0.64715	1.4762	58	0.75934	1.29273	91	0.80514	1.22778
26	0.65323	1.46523	59	0.76125	1.28993	92	0.80614	1.22641
27	0.65901	1.45495	60	0.76311	1.2872	93	0.80713	1.22507
28	0.66449	1.44528	61	0.76492	1.28454	94	0.8081	1.22375
29	0.66972	1.43617	62	0.76669	1.28195	95	0.80906	1.22245
30	0.6747	1.42756	63	0.76843	1.27943	96	0.81	1.22117
31	0.67945	1.41942	64	0.77012	1.27698	97	0.81093	1.21992
32	0.684	1.4117	65	0.77178	1.27458	98	0.81185	1.21868
33	0.68835	1.40437	66	0.7734	1.27225	99	0.81275	1.21746

Confidence limits for rates based on 100 or more events

In this case, use the following formula for the rate (R) based on the number of events (N):

$$\text{Upper Limit} = R + [1.96 \times R / \sqrt{N}]$$

where:

R = the rate (birth rate, mortality rate, teen pregnancy rate, etc.)

N = the number of events (births, deaths, teen pregnancy, etc.)

Example: Confidence limits for rates based on 100 or more events

In Jackson County, the teen pregnancy rate for teens 10-17 was 13.7 in 1998 based on 143 pregnancies. Therefore, the confidence interval would be:

$$\begin{aligned} \text{Lower Limit} &= 13.7 - [1.96 \times (13.7 / \sqrt{143})] \\ &= 13.7 - [1.96 \times (13.7 / 11.96)] \\ &= 13.7 - [1.96 \times 1.15] \\ &= 13.7 - 2.25 \\ &= 11.5 \end{aligned}$$

$$\begin{aligned} \text{Upper Limit} &= 13.7 + [1.96 \times (13.7 / \sqrt{143})] \\ &= 13.7 + [1.96 \times (13.7 / 11.96)] \\ &= 13.7 + [1.96 \times 1.15] \\ &= 13.7 + 2.25 \\ &= 16.0 \end{aligned}$$

So if there were 100 counties like Jackson County with similar populations, the teen pregnancy rate would be expected to lie between 11.5 and 16.0 per 1,000 in 95 of these counties.

DETERMINING STATISTICAL SIGNIFICANCE FOR RATES:

If the difference between two rates would occur due to random variability less than 5 times out of 100, then we say that the difference is statistically significant at the 95% level. Otherwise the difference is not statistically significant.

Computing statistical significance when at least one of the rates is based on fewer than 100 events

To compare two rates, when one or both rates are based on fewer than 100 events, compute the confidence intervals for both rates. If the intervals overlap, the difference is not statistically significant.

Example: comparing rates when one is based on fewer than 100 events

Baker County teen pregnancy rate for age 10-17

Lower Limit = 6.7

Upper Limit = 22.7

Jackson County teen pregnancy rate for age 10-17

Lower Limit = 11.5

Upper Limit = 16.0

The confidence intervals overlap - the interval for Jackson County is entirely within the range of the interval for Baker County. Therefore, the difference between the teen pregnancy rate for age 10-17 in Baker County and the rate for Jackson County is not statistically significant.

Computing statistical significance when both rates are based on 100 or more events

When both rates are based on 100 or more events, calculate the difference between the two rates by subtracting the lower rate from the higher rate. The difference is considered statistically significant if it exceeds 1.96 times the standard error for the difference between the two rates.

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

where:

R_1 = the first rate

R_2 = the second rate

N_1 = the first number

N_2 = the second number

If the difference is greater than the statistic, the difference would occur by chance less than 5 times out of 100. The difference is statistically significant at the 95 percent confidence level.

If the difference is less than the statistic, the difference might occur by chance more than 5 times out of 100. The difference is not statistically significant at the 95 percent confidence level.

Example: comparing rates when both are based on 100 or more events

The teen pregnancy rate for Oregon teens age 10-17 in 1997 was 18.0 and the comparable rate for 1998 was 17.2. Both rates are based on more than 100 pregnancies (3,197 in 1997 and 3,176 in 1998). The difference between the rates is $18.0 - 17.2 = 0.8$. The statistic is calculated as follows:

$$1.96 \sqrt{\frac{18.0^2}{3,197} + \frac{17.2^2}{3,176}}$$

$$1.96 \sqrt{\left(\frac{324}{3,197} + \frac{295.84}{3,176}\right)}$$

$$1.96 \sqrt{(0.101 + 0.093)}$$

$$1.96 \sqrt{0.194}$$

$$= 1.96 \times .44$$

$$= 0.86$$

The difference between the rates (0.8) is less than this statistic (0.9). Therefore, the difference is not statistically significant. A difference of 0.8 between these two rates might occur by chance more than 5 times out of 100.

CALCULATING RATES ADJUSTED FOR SEX/AGE/RACE:

When comparing rates and ratios, the influences of sex, age, and race differences in the populations must be taken into account. Comparing many different age-sex-race specific rates can be cumbersome. The following techniques are used by vital statisticians to summarize these rates into one number.

The *direct adjusted rate* applies each of the specific rates for a particular population (such as a county or a Health Service Area) to a standard population distribution (such as the state).

The *standard mortality ratio* compares the number of deaths for a particular population (such as a county or a Health Service Area) to the number of deaths which would be expected if some standard set of rates (such as the state or the U.S. rates) had occurred.²

Both of these techniques have their advantages and disadvantages. The easiest to calculate is the direct adjusted rate. The following example shows how to adjust a county's death rate for sex so that it may be compared to the state rate.

$$\frac{\left[\frac{\text{county male deaths}}{\text{county male population}} \times \text{state male population} \right] + \left[\frac{\text{county female deaths}}{\text{county female population}} \times \text{state female population} \right]}{\text{TOTAL STATE POPULATION}} \times 1,000$$

The same logic can be used to adjust for age and/or race. An example for the weights used for age-adjustment can be found in Table B-2:

TABLE B-2 2000 US STANDARD POPULATION		
Age	2000 US standard million	2000 US standard population (Census P25-1130)
0	13,818	3,794,901
0-4	55,317	15,191,619
5-9	72,533	19,919,840
10-14	73,032	20,056,779
15-19	72,169	19,819,518
20-24	66,478	18,257,225
25-29	64,529	17,722,067
30-34	71,044	19,511,370
35-39	80,762	22,179,956
40-44	81,851	22,479,229
45-49	72,118	19,805,793
50-54	62,716	17,224,359
55-59	48,454	13,307,234
60-64	38,793	10,654,272
65-69	34,264	9,409,940
70-74	31,773	8,725,574
75-79	26,999	7,414,559
80-84	17,842	4,900,234
85+	15,508	4,259,173
Total	1,000,000	274,633,642

Reference

1. U.S. Department of Health & Human Services, Public Health Service, Centers for Disease Control and Prevention. Public health data: Our silent partner [Internet]. 1999 Oct; [cited 2016 Jan 22]. Available from: www.cdc.gov/nchs/products/training/phd-osp.htm.
2. For more information, please see U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics. Direct standardization (age-adjusted death rates). 1995 March; [cited 2016 Jan 22]. Available from: www.cdc.gov/nchs/data/statnt/statnt06rv.pdf.

For further information about calculating confidence intervals and adjusting rates, see:

1. J. C. Kleinman. Infant mortality. Statistical notes for health planners, No. 2. Washington, D.C.: Health Resources Administration; 1976 July.
2. J. C. Kleinman. Mortality. Statistical notes for health planners, No. 3. National Center for Health Statistics: by, Health Resources Ad.