Appendix B: Technical Notes — Definitions

BIRTHS

Apgar Score is a numerical expression of the condition of a newborn shortly after birth. It is the sum of points accumulated upon assessment of the heart rate, respiratory effort, muscle tone, reflex irritability, and color. The highest possible score is ten. 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 less than 5 pounds, 8 ounces (2,500 grams) as reported on the birth certificate.

DEATHS

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 prior to its first birthday.

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.

Postneonatal Death is the death of a child after 27 days of life and before its first birthday.

Postneonatal Death Rate is the number of postneonatal 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.

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.E.M. — licensed direct entry midwife.

M.D. — medical doctor.

N.D. — naturopathic doctor.

R.N. — registered nurse.

ENDNOTES

1 *Vital Statistics of the United States*, 1982, vol. 1, section 4, page 1. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, Maryland, 1986.

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

MORTALITY

Comparability Between ICD-9 and ICD-10 Codes

The *International Classification of Diseases* (ICD) codes are periodically revised to reflect progress in the identification of diseases.[1] This practice began in 1900 and occurs every 10 to 20 years. Each of these revisions has produced some breaks in the comparability of cause of death statistics.

ICD-10 has many changes from ICD-9, including: considerably greater detail for some causes (and less detail for others); shifts of inclusion in terms and titles from one category, section, or chapter to another; regrouping of diseases; new titles and sections; and modifications in coding rules. As a result, serious breaks occur in comparability for a number of causes of death. Measures of this discontinuity are essential to the interpretation of mortality trends. Comparability ratios between ICD-9 and ICD-10 have been computed for this purpose (please see the following table). Note that data tables showing cause of death information for years prior to 1999 are based on the original ICD-9 codes and have not been adjusted using comparability ratios.

Studies of the comparability between revisions of the ICD have been carried out and published since at least the fifth revision. Comparability studies, also called bridge-coding studies, involve the dual classification of a single year of mortality data, that is classifying the underlying cause of death on mortality records by the new revision and the previous revision. The key element of the comparability study is the comparability ratio, which is derived from the dual classification. It is calculated by dividing the number of deaths for a selected cause of death classified by the new revision by the number of deaths classified to the most nearly comparable cause of death using the previous revision (in this case the number of deaths identified as being attributable to a particular cause using ICD-10 codes and rules divided by the number of deaths attributed to the same cause using ICD-9 codes and rules). The resulting ratio represents the net effect of the new revision on statistics for this cause and can be used as a factor to adjust previously calculated mortality statistics.

A comparability ratio of 1.00 indicates that the same number of deaths was assigned to a particular cause or combination of causes, regardless of the revision used. A ratio showing perfect correspondence (1.00) between the two revisions does not necessarily indicate that the cause was unaffected by changes in classification and coding procedures but merely that there was no net change.

A ratio less than 1.00 results from a decrease in assignments of death to a cause in ICD-10 compared with ICD-9. A ratio of more

than 1.00 results from an increase in assignments of deaths to a cause in ICD-10 compared to the corresponding ICD-9 cause.

In regard to the magnitude of coding effects produced by rule changes, that of Rule 3 is among the most prominent. This rule is used to determine the direct sequels of causes. It states "If the conditions selected by the general principle or by Rule I or by Rule 2 is obviously a direct consequence of another reported condition, whether in Part I or Part II [of the medical certification portion of the death certificate], select this primary condition." The cause of death most affected by Rule 3 is pneumonia, which is often the consequence of another condition or injury. In ICD-10 the applicability of Rule 3 to pneumonia is broader than in ICD-9, so pneumonia is considered a consequence of a much wider range of conditions. As a result, pneumonia is much less likely to be selected as the underlying cause of death under ICD-10 than under ICD-9.

The following describes selected leading causes of death affected by changes in classification and underlying cause of death rules.

Heart Disease. The comparability ratio (CR) for this cause is 0.9858, indicating a nearly 1.5 percent net decrease in the allocation of heart disease as the underlying cause of death when using the ICD-10 classification scheme. The net decrease is a result primarily of shifts away from heart disease to other causes of death due to Rule A; under this rule, certain disorders are considered ill-defined and not reflecting the true underlying cause of death. Cardiac arrest is one such disorder. Thus, it is ignored in the selection of underlying cause of death if another more specific cause is listed on the death certificate.

Malignant Neoplasms. The CR for cancer is 1.0068, indicating considerable comparability in numbers and rates between revisions. Nevertheless, a substantial number of deaths are classified under malignant neoplasms in ICD-10 that were not classified as such under ICD-9. Most of these were classified as pneumonia in ICD-9 and were affected by the change in Rule 3 (described above). In ICD-10, the applicability of Rule 3 to pneumonia is broader than in ICD-9; that is, pneumonia is considered a consequence of a much wider range of conditions. As a result, pneumonia is much less likely to be selected as the underlying cause of death under ICD-10 than under ICD-9. In addition, some deaths shifted out of the malignant neoplasm category due to the revision. Most of these are classified in ICD-10 as HIV or, *in situ* neoplasms, benign neoplasms, and neoplasms of uncertain or unknown behavior.

Nearly all of the specified malignant neoplasm categories show some shifts of deaths into and out of the specified category. For example, because of changes in the rule governing the selection of the primary site, deaths involving cancer of the trachea, bronchus, and lung are a little less likely to be attributed to this cause. (The comparability ratio is 0.9837.) This occurred because ICD-10, in contrast to ICD-9, classifies malignant neoplasms of the lung as secondary to many other cancers. Further, when classifying deaths according to ICD-10, unlike ICD-9, selection of the primary site is not determined by order of entry on the death certificate. Thus, when

two primary sites from different organ systems are listed, the deaths are classified to C97, the category for independent (primary) multiple sites.

Alzheimer's Disease. The CR published in the previously described NCHS publication should not be applied to Oregon data. Unlike the nation, deaths assigned to this category have included both Alzheimer's disease (ICD-9 331.0) and presenile dementia (ICD-9 290.1). A study of deaths coded to ICD-9 290.1 showed that 99 out of 100 were attributable to Alzheimer's dementia and that physicians were using the terms "Alzheimer's disease" and "Alzheimer's dementia" essentially interchangeably. To provide a more realistic measure of the impact of Alzheimer's disease, both diseases were included in Oregon's "Alzheimer's Disease" category. ICD-10 eliminated the separate category for "Alzheimer's dementia"; just one code (G30) is present in the current revision.

Unintentional Injuries. With a comparability ratio of 1.0303, deaths were slightly more likely to be attributed to unintentional injuries than previously. Virtually all of this increase involves shifts from natural causes in ICD-9 to unintentional injuries in ICD-10. Most of these deaths were classified as pneumonia or cardiac arrest in ICD-9 but were coded to unintentional injuries as a consequence of the changes in Rule 3 and Rule A, respectively. The CR for the largest subset in this group, motor vehicles, is 0.9754, but the specific category with the largest difference (CR = 0.8409) is falls. This 16 percent decrease is the result of the change in the classification of unspecified fractures. In ICD-9, if the term "fracture" was listed on the death certificate without mention of an external cause, the death was classified to "Fracture, cause unspecified" (E887) within the greater "Accidental Falls" (E880-888) category. In ICD-10, a fall is not assumed to be responsible for an unspecified fracture, and the death is classified to "Exposure to Unspecified Factor," (X59), which is classified as an unintentional injury, but in a residual category, not a fall.

Intentional Self-Harm. This category (i.e., suicide) has a comparability ratio of 0.9962. The slight decline may have resulted from records pending amendment that were unable to be identified at the time of the study. Some changes in coding categories have resulted in less specific data. For example, the type of firearm used in suicide (and all other external cause categories) is no longer distinguished other than handgun vs. long gun; previously, rifles, shotguns, and military (assault) weapons were categorized individually. Further, suffocation suicides involving plastic bags are no longer identified (The number of deaths in this category was typically about the same as the number resulting from cutting and piercing injuries).

Assault. Like suicide, this category (i.e., homicide) showed little difference between ICD-9 and ICD-10 coding; the comparability ratio was 0.9983. The reader is cautioned that this CR is applicable only to prior years' categories based on ICD-9 codes E960-E969. Under the ICD-9 classification, legal intervention (E970-E979) deaths were included in the leading cause of death category "Homicide."

They no longer are. Further, NCHS has not published a comparability ratio for legal intervention deaths because the figure calculated did not meet standards of reliability or precision.

Super MICAR

Beginning in 1993, the underlying cause of death was determined by using Super MICAR, software distributed by the National Center for Health Statistics. In the past, the underlying cause of death was determined by a nosologist using information provided on death certificates by physicians. Super MICAR applies a set of algorithms to all the causes listed on a death certificate to arrive at the underlying cause of death.

This software is being used because the number of deaths among Oregonians has increased substantially during recent years, but has not been accompanied by an increase in staff. Consequently, data availability became increasingly untimely during recent years. Instituting the Super MICAR system is resulting in more timely data.

An advantage of the Super Micar system is that all causes recorded on the death certificate are now included in the data file. We will be able to report, for example, not only the number of Oregonians who died from Alzheimer's Disease but the number of Oregonians who had the disease at the time of their death (provided it was mentioned on the certificate).

Age-adjusted Rates

The death rates in this report are not age-adjusted. (However, age- and sex-specific death rates are presented in addition to crude death rates.) Age-adjusted death rates permit the comparison of populations with disparate age structures as if the populations had similar distributions. Age-adjusted rates should be used when comparing subsets (e.g., counties and races). See the formulas section of this Appendix for instructions on calculating age-adjusted rates. Rates may also be computed on-line at the federal Centers for Disease Control (CDC) site http://wonder.cdc.gov.

We ordinarily publish age-adjusted death rates for the U.S. and Oregon (by cause) in this report, but 1999 data were unavailable from the CDC at the time of publication.

Tobacco-linked Deaths

The number of Oregonians whose deaths were linked to tobacco use are presented in the mortality section. However, the number is artificially low. This is because the role of tobacco, if any, is not routinely noted on the death certificates of Oregonians who died out-of-state. (The footnotes in the tables describe the question on the Oregon death certificate regarding tobacco use.) The potential for undercount is greatest for Oregon residents who live in counties bordering other states. A more detailed discussion can be found in *Tobacco and Oregon: A Legacy of Illness and Death*, published in 1992.

YOUTH SUICIDE ATTEMPTS

Unlike previous years, youth who threatened to commit suicide but made no physical act are not included in the total number of attempts, but are shown in a separate table (Table 8-14). Previously they had been identified in the method of attempt tables. Had they been included in the total, the count would have been 785, a new high.

Data in the youth suicide attempts section were compiled from teen suicide attempt reports and death certifications files with the Oregon Department of Human Services' Center for Health Statistics. Attempt rates are age-specific and are expressed per 100,000 of the population at risk per year. The Center for Population Research and Census was the source of the population data. Methods of attempts are classified according to the International Classification of Diseases (ICD). The name of the attempter is not recorded on attempts reported to the Center for Health Statistics.

Several problems are apparent with the data. The first is that the total number of attempts reported is low. Because Oregon is the only state to require that adolescent suicide attempts be reported, when Oregon adolescents attempt suicide in another state, the event is not reported. More significantly, although required by law, the data suggest that not all hospitals are fully cooperating with the program. It is uncertain whether reporting hospitals are using the same criteria in determining whether the patient attempted suicide. Finally, a few data items are poorly reported.

ENDNOTE

1. This description is drawn from *National Vital Statistics Report,* Vol. 49, No. 2, June 26, 2001, which includes additional detail not included here. The document is available online at:
   http://www.cdc.gov/nchs/products/pubs/pubd/nvsr/49/49-pre.htm

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

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 to provide 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.

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

STEP 1: FINDING THE CORRECT NUMBER

The first step is to determine how many 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 which 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 which *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 that 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 that 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 table(s) 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 that the number of events which occurred 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:

CRUDE DEATH RATE = (DEATHS/POPULATION) X 1,000

the number of people who could have died

a number chosen by vital statisticians to improve the ease of comparisons

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 pre-pubescent or post-menopausal women in the population. (The turn of the century notion that only *married* women between the age of 15 and 44 would be considered at risk of pregnancy has been abandoned for obvious reasons.)

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.

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

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." For example, a statistician will say, "We are 95% sure that the *true* infant death rate for Oregon in 1986 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 which do not reflect real changes. Consider Tillamook County's infant mortality rates for a five-year period.

TILLAMOOK COUNTY			
YEAR	BIRTHS	INFANT DEATHS	INFANT DEATH RATES
1981	324	5	15.4
1982	318	2	6.3
1983	306	4	13.1
1984	264	1	3.8
1985	266	3	11.3
1981-1985	1,478	15	10.1

The overall rate of 10.1 is quite close to the state rate for the same time period (10.2). Yet, for some years the rate is four times as high as the rate of other years simply because four 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. But, 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% 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 is too few, how many cases are sufficient to say that 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 cause-of-death item provides an excellent example in comparability:

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

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.

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 twenties to have more babies than the very young or the very old. Sex and race, as well as age, can affect rates drastically.

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

	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

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 death rates for each age 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, then 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 which 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 statistically significantly different, how can we find out *why* they are different? If the differences which we expected did not prove to be significant, is there another item which 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 1985, has chronic lower respiratory disease 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 the book to illustrate changes. And finally, the staff of the Center for Health Statistics are available for data users who need assistance.

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 one 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 which occur is not available in vital records (perhaps 10 percent of all pregnancies). Thus, a measure which excludes these outcomes provides an adequate indicator of the number of pregnancies.

Technical Notes — Formulas

GENERAL:

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

Birth rate, Oregon, 1993 = 13.7
Birth rate, Oregon, 1994 = 13.6

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

PREGNANCY:

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

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

---

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

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

---

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

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

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

---

4. $$TOTAL\ FERTILITY\ RATE = \begin{array}{l}The\ Sum\ of\ Age\text{-}Specific\ Birth\ Rates\ in \\ 5\text{-}Year\ Categories\ between\ 15\ and\ 44\end{array} X\ 5$$

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

5. $$\textit{FETAL DEATH RATIO} = \frac{\textit{Resident Fetal Deaths (20+ Weeks Gestation)}}{\textit{Resident Live Births}} \; X \; 1{,}000$$

$$\textit{Oregon, 1994, Residents} = \frac{224}{41{,}832} \; X \; 1{,}000 = 5.4$$

$$\textit{FETAL DEATH RATE} = \frac{\textit{Resident Fetal Deaths (20+ Weeks Gestation)}}{\textit{Resident Live Births + Resident Fetal Deaths}} \; X \; 1{,}000$$

$$\textit{Oregon, 1994, Residents} = \frac{224}{43{,}591 + 224} \; x \; 1{,}000 = 5.1$$

$$\textit{PERINATAL DEATH RATE} = \frac{\textit{Resident Neonatal Deaths + Resident Fetal Deaths (20+ Weeks Gestation)}}{\textit{Resident Live Births + Resident Fetal Deaths}} \; X \; 1{,}000$$

$$\textit{Oregon, 1994, Residents} = \frac{148 + 203}{41{,}566 + 203} \; X \; 1{,}000 = 8.4$$

Note: *Publications vary in the gestation cutoff for fetal deaths. In addition, some measures employ birthweight in place of gestational age. Fetal and perinatal death rates are based on 1993 year of birth.*

---

6. $$\textit{ABORTION RATIO} = \frac{\textit{Resident Abortions}}{\textit{Resident Births}} \; X \; 1{,}000 \; \textit{or} \; \frac{\textit{Occurrence Abortions}}{\textit{Occurrence Births}} \; X \; 1{,}000$$

$$\textit{Oregon, 1994, Occurrence} = \frac{13{,}391}{43{,}591} \; X \; 1{,}000 = 307.2$$

---

7. $$\textit{ABORTION RATE} = \frac{\textit{Resident Abortions or Occurrence Abortions}}{\textit{Female Resident Population Aged 15–44}} \; X \; 1{,}000$$

$$\textit{Oregon 1994, Occurrence with total adjusted for not stated ages} = \frac{13{,}300}{682{,}428} \; X \; 1{,}000 = 19.5$$

DEATHS:

8. (*CRUDE*) *DEATH RATE* $\frac{\textit{Resident Deaths}}{\textit{Population}} X\ 1{,}000$

Oregon, 1994 $\frac{27{,}361}{3{,}082{,}000} X\ 1{,}000$ 8.9

---

9. *INFANT DEATH RATE* $\frac{\textit{Resident Infant Deaths}}{\textit{Resident Births}} X\ 1{,}000$

Oregon, 1994 $\frac{295}{41{,}832} X\ 1{,}000$ 7.1

---

10. *NEONATAL DEATH RATE* $\frac{\textit{Resident Neonatal Deaths}}{\textit{Resident Births}} X\ 1{,}000$

Oregon, 1994 $\frac{164}{41{,}832} X\ 1{,}000$ 3.9

---

11. *POSTNEONATAL DEATH RATE* $\frac{\textit{Resident Postneonatal Deaths}}{\textit{Resident Births}} X\ 1{,}000$

Oregon, 1994 $\frac{131}{41{,}832} X\ 1{,}000$ 3.1

---

12. *CAUSE SPECIFIC DEATH RATE* $\frac{\textit{Resident Deaths Due to Specific Cause}}{\textit{Population}} X\ 100{,}000$

Oregon, 1994, *Heart Disease* $\frac{7{,}417}{3{,}082{,}000} X\ 100{,}000$ 240.7

---

13. *AGE AND SEX SPECIFIC DEATH RATE* $\frac{\textit{Resident Deaths in Age Sex Category}}{\textit{Population in Age Sex Population}} X\ 1{,}000$

Oregon, 1994, *Males Aged* 5 14 $\frac{63}{225{,}880} X\ 100{,}000$ 27.9

MARRIAGE AND DIVORCE:

$$14.\ MARRIAGE\ RATE = \frac{Marriages}{Population}\ X\ 1{,}000$$

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

$$15.\ DIVORCE\ RATE = \frac{Divorces}{Population}\ X\ 1{,}000$$

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

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:

Lower limit = 13.0 x 0.51671 = 6.7
Upper limit = 13.0 x 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:

$$\text{Lower Limit} = R - [1.96 \times R / \sqrt{N}]$$

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

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:

$$\text{Lower Limit} = 13.7 - [1.96 \times (13.7 / \sqrt{143})]$$

$$= 13.7 - [1.96 \times (13.7 / 11.96)]$$
$$= 13.7 - [1.96 \times 1.15]$$
$$= 13.7 - 2.25$$
$$= 11.5$$

$$\text{Upper Limit} = 13.7 + [1.96 \times (13.7 / \sqrt{143})]$$

$$= 13.7 + [1.96 \times (13.7 / 11.96)]$$
$$= 13.7 + [1.96 \times 1.15]$$
$$= 13.7 + 2.25$$
$$= 16.0$$

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:

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

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

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

$$1.96\sqrt{0.194}$$

= 1.96 x .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 an HSA) 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 an HSA) 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.

Each of these techniques has its 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.

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

The same logic can be used to adjust for age and/or race.

Table 1. Estimated comparability ratios for 113 selected causes of death

List number	Cause of death[1]	Number of deaths allocated according to ICD–10[2]	ICD–9[2]	Estimated comparability ratio	Standard error	Relative standard error	95 percent confidence limits Lower	Upper
001	Salmonella infections	30	37	0.8108	0.0644	7.9	0.6846	0.9370
002	Shigellosis and amebiasis	*	*	*	*	*	*	*
003	Certain other intestinal infections	*	*	*	*	*	*	*
004	Tuberculosis	653	764	0.8547	0.0172	2	0.8209	0.8885
005	Respiratory tuberculosis	518	572	0.9056	0.0201	2.2	0.8662	0.9450
006	Other tuberculosis	135	192	0.7031	0.0407	5.8	0.6233	0.7830
007	Whooping cough	*	*	*	*	*	*	*
008	Scarlet fever and erysipelas	*	*	*	*	*	*	*
009	Meningococcal infection	221	222	0.9955	0.0149	1.5	0.9663	1.0247
010	Septicemia	21,258	17,791	1.1949	0.0042	0.3	1.1867	1.2030
011	Syphilis	21	33	0.6364	0.1184	18.6	0.4043	0.8685
012	Acute poliomyelitis	*	*	*	*	*	*	*
013	Arthropod-borne viral encephalitis	*	*	*	*	*	*	*
014	Measles	*	*	*	*	*	*	*
015	Viral hepatitis	1,123	1,346	0.8343	0.0120	1.4	0.8109	0.8578
016	Human immunodeficiency virus (HIV) disease	25,089	23,586	1.0637	0.0018	0.2	1.0601	1.0673
017	Malaria	*	*	*	*	*	*	*
018	Other and unspecified infectious and parasitic diseases and their sequelae	2,865	2,607	1.0990	0.0154	1.4	1.0688	1.1291
019	Malignant neoplasms	464,688	461,544	1.0068	0.0002	0.0	1.0064	1.0072
020	Malignant neoplasms of lip, oral cavity and pharynx	5,927	6,172	0.9603	0.0040	0.4	0.9525	0.9681
021	Malignant neoplasm of esophagus	9,596	9,630	0.9965	0.0020	0.2	0.9926	1.0003
022	Malignant neoplasm of stomach	11,480	11,408	1.0063	0.0019	0.2	1.0025	1.0101
023	Malignant neoplasms of colon, rectum and anus	48,583	48,619	0.9993	0.0009	0.1	0.9975	1.0010
024	Malignant neoplasms of liver and intrahepatic bile ducts	9,732	10,102	0.9634	0.0023	0.2	0.9588	0.9679
025	Malignant neoplasm of pancreas	24,313	24,361	0.9980	0.0009	0.1	0.9963	0.9997
026	Malignant neoplasm of larynx	3,209	3,194	1.0047	0.0053	0.5	0.9943	1.0150
027	Malignant neoplasms of trachea, bronchus and lung	131,750	133,936	0.9837	0.0005	0.1	0.9827	0.9846
028	Malignant melanoma of skin	5,941	6,139	0.9677	0.0032	0.3	0.9614	0.9741
029	Malignant neoplasm of breast	38,102	37,891	1.0056	0.0010	0.1	1.0036	1.0075
030	Malignant neoplasm of cervix uteri	3,753	3,802	0.9871	0.0034	0.3	0.9805	0.9938
031	Malignant neoplasms of corpus uteri and uterus, part unspecified	5,318	5,183	1.0260	0.0040	0.4	1.0182	1.0339
032	Malignant neoplasm of ovary	11,292	11,344	0.9954	0.0016	0.2	0.9923	0.9985
033	Malignant neoplasm of prostate	30,672	30,267	1.0134	0.0015	0.1	1.0105	1.0162
034	Malignant neoplasms of kidney and renal pelvis	9,521	9,521	1.0000	0.0022	0.2	0.9957	1.0043
035	Malignant neoplasm of bladder	9,563	9,594	0.9968	0.0026	0.3	0.9916	1.0019
036	Malignant neoplasms of meninges, brain and other parts of central nervous system	10,039	10,359	0.9691	0.0025	0.3	0.9642	0.9740
037	Malignant neoplasms of lymphoid, hematopoietic and related tissue	44,715	44,530	1.0042	0.0012	0.1	1.0019	1.0064
038	Hodgkin's disease	1,021	1,036	0.9855	0.0089	0.9	0.9680	1.0030
039	Non-Hodgkin's lymphoma	17,924	18,326	0.9781	0.0018	0.2	0.9745	0.9817
040	Leukemia	16,600	16,405	1.0119	0.0019	0.2	1.0083	1.0155
041	Multiple myeloma and immunoproliferative neoplasms	9,099	8,763	1.0383	0.0030	0.3	1.0324	1.0443
042	Other and unspecified malignant neoplasms of lymphoid, hematopoietic and related tissue	*	*	*	*	*	*	*
043	All other and unspecified malignant neoplasms	51,182	45,492	1.1251	0.0021	0.2	1.1210	1.1292
044	In situ neoplasms, benign neoplasms and neoplasms of uncertain or unknown behavior	9,263	5,532	1.6744	0.0164	1.0	1.6422	1.7067
045	Anemias	3,059	3,200	0.9559	0.0077	0.8	0.9409	0.9710
046	Diabetes mellitus	48,636	48,242	1.0082	0.0011	0.1	1.0060	1.0103
047	Nutritional deficiencies	3,215	2,763	1.1636	0.0165	1.4	1.1312	1.1960
048	Malnutrition	2,607	2,665	0.9782	0.0151	1.5	0.9487	1.0078
049	Other nutritional deficiencies	608	98	6.2041	0.5961	9.6	5.0358	7.3724
050	Meningitis	592	584	1.0137	0.0136	1.3	0.9871	1.0403
051	Parkinson's disease	10,404	10,392	1.0012	0.0028	0.3	0.9956	1.0067
052	Alzheimer's disease	29,707	19,121	1.5536	0.0071	0.5	1.5398	1.5675
053	Major cardiovascular diseases	796,919	798,435	0.9981	0.0002	0.0	0.9977	0.9985
054	Diseases of heart	615,564	624,405	0.9858	0.0002	0.0	0.9854	0.9863
055	Acute rheumatic fever and chronic rheumatic heart diseases	2,446	2,980	0.8208	0.0089	1.1	0.8034	0.8382
056	Hypertensive heart disease	17,322	21,577	0.8028	0.0028	0.3	0.7973	0.8083
057	Hypertensive heart and renal disease	2,170	2,027	1.0705	0.0160	1.5	1.0392	1.1019
058	Ischemic heart diseases	466,459	466,935	0.9990	0.0002	0.0	0.9985	0.9994
059	Acute myocardial infarction	178,125	180,169	0.9887	0.0003	0.0	0.9880	0.9893
060	Other acute ischemic heart diseases	2,667	2,638	1.0110	0.0117	1.2	0.9880	1.0340
061	Other forms of chronic ischemic heart disease	285,667	284,128	1.0054	0.0004	0.0	1.0046	1.0062

See footnotes at end of table.

From: Anderson RN, Minino AM, Hoyert DL, Rosenberg HM. Comparability of cause of death between ICD-9 and ICD-10: Preliminary estimates. National vital statistics reports; Vol. 49, No. 2. Hyattsville, Maryland: National Center for Health Statistics. 2001.

REFERENCES:

1. US Department of Health & Human Services, Public Health Service, Centers for Disease Control and Prevention, October 1999. The original materials are available on-line at http://www.cdc.gov/nchs/products/training/phd-osp.htm.

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

National Center for Health Statistics: Infant Mortality, by J. C. Kleinman, Statistical Notes for Health Planners, No. 2. Health Resources Administration, Washington, D.C., July 1976.

National Center for Health Statistics: Mortality, by J. C. Kleinman, Statistical Notes for Health Planners, No. 3. Health Resources Administration, Washington, D.C., July 1977.