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 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, the NCHS method of distributing births where age of father was not stated in the same proportion as births where age of father was stated within each 5-year age interval of mother was used to facilitate national comparisons. NCHS uses this procedure to avoid distortion in rates that would result if the relationship between age of mother and age of father were disregarded.

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.

ENDNOTE

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*

INDUCED TERMINATIONS OF PREGNANCY

Except for incomplete reporting by providers, the data represent *all* abortions performed in Oregon during the current data year. 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 the patient's 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 who leave the state to obtain an abortion equals the number of out-of-state residents who obtain an abortion in Oregon. In formulating generalizations which 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 only as careful approximations and interpreted only within the framework of statistical safeguards developed to take sampling variability into account.

Some rates in this 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. This may

NUMBER OF FIRST-TIME ABORTIONS BY YEAR AND AGE GROUP, OREGON OCCURRENCE, 1975-1989

YEAR	AGE GROUPS					
	15-19	20-24	25-29	30-34	35-39	40-44
1975	3,470	2,751	1,331	620	296	107
1976	3,877	3,125	1,551	616	297	108
1977	3,605	2,921	1,467	650	300	107
1978	3,620	3,041	1,573	786	327	98
1979	3,821	3,149	1,552	811	289	108
1980	3,792	2,965	1,540	795	345	90
1981	3,261	2,643	1,361	760	343	96
1982	2,530	2,066	1,093	607	263	83
1983	2,340	1,976	971	519	287	67
1984	2,340	2,091	995	580	299	80
1985	2,442	2,041	915	496	324	64
1986	2,065	1,694	880	506	270	70
1987	2,375	1,926	935	584	322	83
1988	2,844	2,281	1,086	661	379	94
1989	2,801	2,453	1,245	637	415	110

greatly affect the estimation of rates. To minimize the potential bias inherent in such estimates, unknown events in some cases (Table 4-1) are assigned to the categories of analysis proportional to the distribution of known events. In this way, rates calculated for subsets (e.g., "abortions per thousand teen females") are, on average, less affected by incomplete data.

ESTIMATION OF THE CUMULATIVE PROPORTION OF FEMALES WHO HAVE EXPERIENCED AN ABORTION

This figure is estimated by tracing the abortion experience of a specific cohort of females over an extended time period. In the table on the previous page, an approximation of the "cumulative total" of first-time abortions by one of the cohorts may be obtained by summing the figures in the boxed area.

To obtain this value, it is necessary to sum the number of first-time abortions for 15- to 19-year-olds from 1975 to 1979 and those of 20- to 24-year-olds from 1980 to 1984 with those of 25- to 29-year-olds from 1985 to 1989. This provides an estimate of the numerator in the following equation:

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

The denominator may be estimated by averaging the size of the cohort during 1975-1989. Table A-1 lists the annual estimate of the number of females within each cohort. For example, in 1975 the number of 15- to 19-year-old females was estimated to be 110,334; in the next year it was 111,184. The average size of this age group from 1975 to 1979 was 112,047. Similarly, the number of 20- to 24-year-old women between 1980 and 1984 was 114,553 on average; the number of 25- to 29-year-olds averaged 111,724 between 1985 and 1989. Thus, between 1975 and 1989 the cohort of interest had an average population size of 112,775.

Substituting into the formula given above:

$$Cp = \frac{\text{Sum of First Abortions}}{N} = \frac{35.195}{112.775} = .312 \text{ or } 31.2 \text{ percent}$$

This figure approximates the proportion of females in the 25- to 29-year-old cohort who, by 1989, had *ever had an abortion*. This method of estimation assumes that factors such as deaths and migration have not altered the composition of the female population in Oregon--that is, the women who have left the state display the same characteristics as those who have moved into Oregon. It also assumes that patients with a history of previous abortions do not report the current procedure as a first abortion.

TEEN PREGNANCY

Pregnancy estimates are based upon the estimated number of teen births and induced terminations among Oregon teens; they do not include the number of fetal deaths or miscarriages (spontaneous abortions) which occur. The estimation of teen births is considered to be relatively complete and includes births to resident teens even when they occur out-of-state. The estimation of teen abortions is based on all reported abortions to teen age residents of Oregon; 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, estimates of teen abortions and teen pregnancies should be considered minimal in nature.

Furthermore, because estimates of abortion for teens are based on "residence data," figures given in Chapter 4 do not correspond exactly to those in Chapter 3, which 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 estimates are now available for 15- to 17-year-olds and 18- to 19-year-olds for each county on an annual basis. Because estimated rates based on a small population may vary greatly due to chance factors, rates of teen pregnancy, birth, and abortion were calculated for these age groups only if there were 50 or more female residents of the appropriate age group in the county.

Similarly, rates for 15- to 19-year-olds were calculated whenever a county had 50 or more female residents in this age group.

Great caution must be taken in the use of pregnancy statistics associated with females under 15 years of age. This is due to the fact that relatively few events are recorded each year for this group. Also, rates are based on the estimated population cohort of 10-14 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 1990, Oregon's birth rate for all teens (regardless of race or ethnic affiliation) was nine percent lower than that of the U.S. and, among all 50 states, it had the 24th 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 U.S. This results from the fact that 87 percent of 15- to 19-year-old females in Oregon were non-Hispanic whites and only seven 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.

TEEN BIRTH RATES, U.S. VS. OREGON, AGES 15-19, 1990		
RACE/ETHNICITY	BIRTH RATE „	
	U.S.	OREGON
TOTAL †	59.9	54.8
NON-HISPANIC WHITES	42.5	50.6
„ ALL RATES PER 1,000 FEMALES.		
† ALL RACES AND ETHNICITIES COMBINED.		

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.

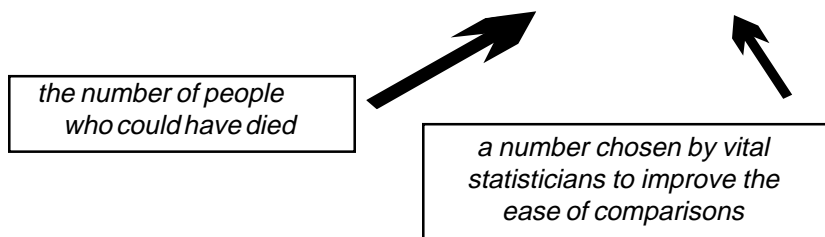## STEP 2: MAKING THE NUMBER MEANINGFUL WITH RATES AND RATIOS

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:

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.

	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

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 this report to illustrate changes. And finally, the staff of the Center for Health Statistics are available for data users who need assistance.

ENDNOTE

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. Nevertheless, 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* = $\dfrac{Resident\ Births}{Population}$ X 1,000

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

---

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

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

---

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

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

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

---

4. *TOTAL FERTILITY RATE* = *The Sum of Age-Specific Birth Rates in 5-Year Categories between* 15 *and* 44 X 5

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

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

$$\textit{Oregon, 1994, Residents} = \dfrac{224}{41{,}832} \times 1{,}000 = 5.4$$

$\textit{FETAL DEATH RATE} = \dfrac{\textit{Resident Fetal Deaths (350+ grams Birthweight)}}{\textit{Resident Live Births + Resident Fetal Deaths}} \times 1{,}000$

$$\textit{Oregon, 1994, Residents} = \dfrac{224}{43{,}591 + 224} \times 1{,}000 = 5.1$$

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

$$\textit{Oregon, 1994, Residents} = \dfrac{148 + 203}{41{,}566 + 203} \times 1{,}000 = 8.4$$

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.*

6. $\textit{ABORTION RATIO} = \dfrac{\textit{Resident Abortions}}{\textit{Resident Births}} \times 1{,}000 \ \textit{or} \ \dfrac{\textit{Occurrence Abortions}}{\textit{Occurrence Births}} \times 1{,}000$

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

7. $\textit{ABORTION RATE} = \dfrac{\textit{Resident Abortions or Occurrence Abortions}}{\textit{Female Resident Population Aged 15-44}} \times 1{,}000$

$$\begin{array}{l}\textit{Oregon 1994, Occurrence}\\ \textit{with total adjusted}\\ \textit{for not stated ages}\end{array} = \dfrac{13{,}300}{682{,}428} \times 1{,}000 = 19.5$$

DEATHS:

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

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

---

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

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

---

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

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

---

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

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

---

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

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

---

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

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

MARRIAGE AND DIVORCE:

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

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

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

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

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:

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

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

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:

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

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

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

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

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 <u>not</u> 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 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.

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.

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

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

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, <u>Statistical Notes for Health Planners</u>, No. 2. Health Resources Administration, Washington, D.C., July 1976.

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