APPENDIX III-A

OVERVIEW OF CONFIDENCE INTERVALS

Disclaimer: This document is designed to familiarize you with confidence intervals. It is not intended to be (or to replace) a proper course in statistics. It does contain suggested rules of thumb that will help you to interpret information commonly reported in tables, figures and charts. For more detailed information, we suggest that you download our monograph "Guidelines for Statistical Analysis of Public Health Data" from http://www.ucsf.edu/fhop/fhoprep.html.

Is it Better or Worse?

If you know the value of an indicator for only one group of people (e.g. low birth weight for a county is 6.2%), you might ask, "so what?" In order to know if this value has relevance, you also need to know if the value for that group is really "different" than the value for another group. In most cases, when we talk about differences, we want to consider "statistically significant differences" between groups.

Populations and Samples

We usually use statistics to describe a characteristic of a specific group. All possible members of the group form what is called, in statistical parlance, a "population". In most cases, it is not possible to record information about every person or element of a population. Instead, one can select a smaller group from the population. This smaller group is called a "sample" of the population. One can then measure the characteristic of interest in the sample. The summary statistic, the "point estimate", used to describe the sample can also be used to describe the whole population.

However, by definition, a point estimate is always an approximate measure of the true value of the characteristic found in the population. Even with a completely representative sample, it is always an approximate measure because there will always be random error in the sample. Such random error is due to chance alone and is not related to bias in sample selection or to bias in measurement. Because of random statistical error, we calculate a range of values around the point estimate and the probability that the "true" value is within the calculated range.

Because the true value in the population is unknown, any effort to specify a range of values that includes the true value with 100% certainty means that we would have to set forth a range which is simply too broad to be useful. So, by convention, we are generally satisfied if we are 95% certain that the true value is within our specified range. In other words, we consider a statistic to be reasonably accurate if it produces the "correct" value 95 times out of 100. Conversely, we can say that it is unacceptable if chance factors lead to errors more than 5 times out of 100.

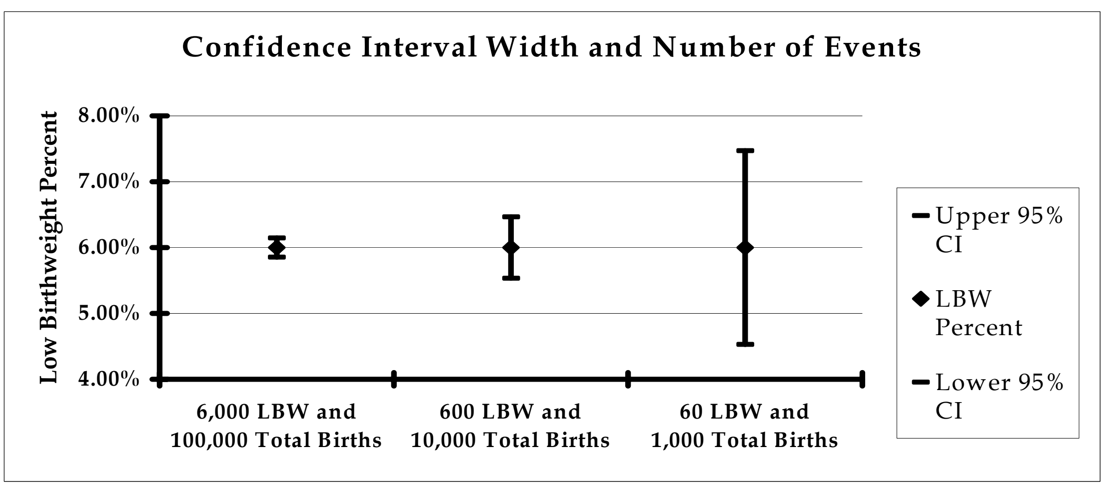
Data contained in vital records, such as the number of live births or infant deaths, represent complete counts of these occurrences. Thus, they are not influenced by sampling error. However, the National Center for Health Statistics states that "when the figures (such as the number of births or deaths) are used for analytic purposes, such as the comparison of rates over time, for different areas, or among different subgroups, the number of events that actually occurred may be considered as one of a large series of possible results that could have arisen under the same circumstances" (NCHS, Monthly Vital Statistics Report, Vol. 46, No. 6(S)2, February 26, 1998). For this reason, we continue to treat vital statistics as if they were samples from some broader population.

Confidence Limits

We can use the concept of 95% confidence to tell how well our statistic measures the true characteristic of the population. We can set limits above and below our point estimate, between which the "true" value found in the population will lie in 95 out of 100 cases. This means that, if we take 100 samples from the population, the true population value will be situated within the confidence interval for 95 out of the 100 samples. We call these limits the 95% confidence limits around the statistic. The lower limit is called the lower 95% confidence limit; the value above the summary statistic is called the upper 95% confidence limit. These limits form what is known as the 95% confidence interval.

When we create a confidence interval, it allows us to say that we are 95% confident that the true value for the population lies between the lower and upper limits we have determined.

Confidence intervals vary in width (the distance between the lower and upper limits) depending on the size of the sample. As the number of events or the denominator with which a rate is calculated grows larger, the confidence interval surrounding the rate grows narrower. A confidence interval surrounding an estimate, such as the low birth weight percent for a county, where there are, for example, 1,000 births will be wide. On the other hand, a confidence interval surrounding the low birth weight percent in a county in which there are 10,000 births will be much narrower than that of the 1,000-birth county, even if the low birth weight percents are identical in the two counties. The following figure shows how the width of a confidence interval varies by the size of the denominator from which the proportion is calculated (in this case, the number of live births). As the number of live births decreases, the confidence interval widens, even if the proportion remains the same.



Formula for Calculation of 95% Confidence Interval

The FHOP Data Templates use two types of formulas to calculate 95% confidence intervals:

For **proportions**, the 95% confidence interval around the point estimate is described by adding or subtracting:

1.96 times the square root of (p * q / n), where:
"p" is the proportion,
"q" is 1 minus the proportion, and
"n" is the number in the study sample.

For **rates**, the 95% confidence interval around the point estimate is described by adding or subtracting:

1.96 times the rate divided by the square root of the numerator (cases)

Statistical Differences

We often want to determine if the values of a statistic for two different groups really are different from each other. To do so, we cannot simply look at the value of each statistic, because there is always random error involved in sampling the populations. We can use 95% confidence intervals to compare statistics for different groups (a variety of statistical tests are available).

Thus, a confidence interval is simply a way to decide whether a mathematical difference is statistically significant. For example, we may have a statistic and a confidence interval for the infant mortality rate for a county and want to know if the rate is significantly different from the rate for the state of California. We can look to see if the confidence interval for the county's rate includes the figure for the State. If so, then we cannot conclude that the two rates are different. On the other hand, if the confidence interval for the county does not include the State's figure, then we can view the two rates as significantly different.

The data templates allow use of confidence intervals in different ways. As described above, you can compare the confidence interval for the county with the benchmark figure for the State. Also, you can compare the county with the Healthy People 2000 or 2010 objective. You can also look at the confidence interval for the county's most recent year, and look back to earlier years to assess whether the rate for the county has changed over time.