Appendix F: Confidence Intervals

Confidence intervals have been reported in the data tables and as error bars in the graphs for all measures in the report. Confidence intervals indicate the reliability of the measure. A more thorough description of statistical reliability may be found in Appendix G of this report.

Although the confidence interval concept draws from the scientific literature on sampling theory, it is also relevant when measures have been calculated from the entire population. In public health, we typically draw on data over a finite time period. Health events do not occur at regularly-spaced intervals. Even though the underlying risk for a health outcome might be stable, the measurable health events, such as infant mortality, occur at random intervals. Thus, when we measure a health event over an arbitrary time period, such as a calendar year, the measurement is taken from a sample in time. Therefore, each calculated rate (whether based upon survey data or count data) is an estimate, and confidence intervals define a range in which the true score (which would represent everyone at all times) would lie.

The 95% confidence interval indicates the range of values within which the statistic would fall 95% of the time if the researcher were to calculate the statistic (e.g., a percentage or rate) from an infinite number of samples of the same size drawn from the same base population. It is typically expressed as the "plus or minus" term, as in the following example:

> "The percentage of those polled who said they would vote for George W. Bush was 47%, plus or minus 2%."

In public health practice, the casual user may think of a confidence interval as the range of probable true scores. The following statements are a logical extension of this thinking.

Observed measure:

- The infant mortality rate for Utah from 1998 to 2003 was 5.2 infant deaths per 1,000 births, with a 95% confidence interval of 0.3.

Logical corollaries:

- This means that the statistic has a 95% confidence interval range from 4.9 to 5.5.
- Thus, if we assume this is a valid measure of infant mortality, there is a very high probability (95%) that the true score lies between 4.9 and 5.5 infant deaths per 1,000 births.
- This means that our best estimate for the underlying risk in the entire Utah population is 5.2 infant deaths per 1,000 births, but that the true risk might lie somewhere between 4.9 and 5.5.

The confidence interval may be used to ascertain whether a measure in a given community is statistically significant, that is, whether the difference is statistically higher or lower than the overall state rate. For example, the motor vehicle crash (MVC) death rate among Utah's American Indian/Alaska Native population was 52.4 per 100,000 population, with a confidence interval that ranged from 23.9 to 80.8. The lower limit of the 95% confidence range (23.9) is greater than the overall state rate of 15.4 deaths per 100,000 population. Therefore, it can be said that the MVC death rate in Utah's American Indian/Alaska Native population is higher than the state rate, and that the difference is statistically significant. Please note, however, that a difference can be meaningful without being statistically significant. The point estimate (in this example, 52.4) is still our best estimate of the underlying risk. We need to be mindful of the confidence interval, but we should not be overdependent on it in interpreting the results.

Appendix F: Confidence Intervals

For most of the measures in this report the 95% confidence intervals are symmetric and centered at the estimated rate (calculated as 1.96 times the standard error). However, in the context of public health measures it makes sense that confidence bounds be equal to or greater than zero, since the rates are nonnegative values. Occasionally when the proportion is close to zero or there are a small number of events, the lower bound is less than zero. Sometimes when the proportion is near 100%, the upper bound is greater than 100%. For these cases we needed to consider asymmetric distributions that constrain the lower and upper bounds to values that lie above zero or below 100%.

The following methods were applied in order to estimate the confidence bounds in these circumstances.

Score Method- applied when estimates were zero or 100%.

Vollset, S.E. (1993). Confidence intervals for a binomial proportion. *Statistics in Medicine* 12, 809-824.

Braner Method- applied to survey data when symmetric confidence bounds were less than zero or greater than 100%.

Braner, M. (2001, January). *Confidence intervals for an age adjusted rate*. Paper presented at the CDC/ATSDR 8th Biennial Symposium on Statistical Methods, Atlanta, GA.

Inverse Gamma Distribution- applied to count data with non-zero counts.

Anderson RN, Rosenberg HM. *Age Standardization of Death Rates: Implementation of the Year 2000 Standard*. National vital statistics reports; vol 47 no. 3. Hyattsville, Maryland: National Center for Health Statistics. 1998.