New Jersey Department of Health & Senior Services
Cancer Epidemiology Services
P.O. Box 369, Trenton, NJ 08625-0369
(609) 588-3500
www.state.nj.us/health

Fact Sheet: Explanation of Standardized Incidence Ratios

The Standardized Incidence Ratio (SIR)

A Standardized Incidence Ratio (SIR) is used to determine if the occurrence of cancer in a relatively small population is high or low. An SIR analysis can tell us if the number of observed cancer cases in a particular geographic area is higher or lower than expected, given the population and age distribution for that community.

The SIR is obtained by dividing the observed number of cases of cancer by the "expected" number of cases. The expected number is the number of cases that would occur in a community if the disease rate in a larger reference population (usually the state or country) occurred in that community. Since cancer rates increase strongly with age, the SIR takes into account whether a community's population is older or younger than the reference population.

How an SIR Is Calculated

The expected number is calculated by multiplying each age-specific cancer incidence rate of the reference population by each age-specific population of the community in question and then adding up the results. If the observed number of cancer cases equals the expected number, the SIR is 1. If more cases are observed than expected, the SIR is greater than 1. If fewer cases are observed than expected, the SIR is less than 1.

Examples:

60 observed cases / 30 expected cases: the SIR is 60/30 = 2.0
Since 2.0 is 100% greater than 1.0, the SIR indicates an excess of 100%.

45 observed cases / 30 expected cases: the SIR is 45/30 = 1.5
Since 1.5 is 50% greater than 1.0, the SIR indicates an excess of 50%.

30 observed cases / 30 expected cases: the SIR is 30/30 = 1.0
A SIR of 1 would indicate no increase or decrease.

15 observed cases / 30 expected cases: the SIR is 15/30 = 0.5
Since 0.5 is 50% less than 1.0, a SIR=0.5 would indicate a decrease of 50%.

Testing if the Difference Between Observed and Expected is Due to Chance

Differences between the observed and expected number of cases may be due to random fluctuations in disease occurrence. The farther the observed number is from the expected number, the less likely it is that chance variation can explain the difference.

A confidence interval (CI) is calculated around an SIR to determine how likely it is that the number of observed number of cases is high or low by chance. If the confidence interval includes 1.0, then the difference between the observed and expected number of cases is likely to have occurred by chance (i.e. to be due to random fluctuations in the data). If the confidence interval does not include 1.0, then the difference between the observed and expected number of cases is not very likely to have occurred by chance.

Examples:

SIR=1.15; 95% CI=0.95, 1.35
One can be 95% confident that the true SIR falls between 0.94 and 1.19. Since the 95% CI contains 1, this estimate of the SIR is not statistically significantly elevated.

SIR=1.11; 95% CI=1.03, 1.19
One can be 95% confident that the true SIR falls between 1.04 and 1.19. Since the 95% CI does not contain 1, this estimate is statistically significantly elevated. One can be 95% confident that the true SIR is at least 1.04, which represents at least a 4% increase.

Statistics, such as SIRs, generated with higher numbers are more likely to show a statistically significant increase or decrease if a true difference does in fact exist. In contrast, small numbers make it particularly difficult for statistical analyses to yield useful or valid information.

Example:

A 20% increase in an SIR derived from Observed =12 and Expected =10 is not statistically significant (SIR=1.2, 95% CI=0.62-2.10)

A 20% increase in an SIR derived from Observed =12,000 and Expected =10,000 is statistically significant (SIR=1.2, 95% CI=1.8-2.2)

The 95% CI as a test for statistical significance may still lead to results that that are due to chance alone. By definition, if a SIR is statistically significantly elevated with 95% confidence, there is still a five percent chance that the increase is due to chance alone. If multiple analyses are done, we further increase the likelihood that some statistically significant results are due to random variation. For example, if a SIR analysis was performed for 20 cancer sites among 5 geographic areas, (20 X 5=100 analyses), we could expect that 5 out of the 100 specific results might be statistically significantly high or low due to chance alone.