

## **Explanation of a Standardized Incidence Ratio (SIR) And 95% Confidence Interval**

In order to evaluate cancer incidence a statistic known as a standardized incidence ratio (SIR) was calculated for each cancer type. An SIR is an estimate of the occurrence of cancer in a population relative to what might be expected if the population had the same cancer experience as some larger comparison population designated as “normal” or average. Usually, the state as a whole is selected to be the comparison population. Using the state of Massachusetts as a comparison population provides a stable population base for the calculation of incidence rates. As a result of the instability of incidence rates based on small numbers of cases, SIRs were not calculated when fewer than five cases were observed.

Specifically, an SIR is the ratio of the observed number of cancer cases to the expected number of cases multiplied by 100. An SIR of 100 indicates that the number of cancer cases observed in the population evaluated is equal to the number of cancer cases expected in the comparison or “normal” population. An SIR greater than 100 indicates that more cancer cases occurred than expected and an SIR less than 100 indicates that fewer cancer cases occurred than expected. Accordingly, an SIR of 150 is interpreted as 50% more cases than the expected number; an SIR of 90 indicates 10% fewer cases than expected.

Caution should be exercised, however, when interpreting an SIR. The interpretation of an SIR depends on both the size and the stability of the SIR. Two SIRs can have the same size but not the same stability. For example, an SIR of 150 based on 4 expected cases and 6 observed cases indicates a 50% excess in cancer, but the excess is actually only two cases. Conversely, an SIR of 150 based on 400 expected cases and 600 observed cases represents the same 50% excess in cancer, but because the SIR is based upon a greater number of cases, the estimate is more stable. It is very unlikely that 200 excess cases of cancer would occur by chance alone.

To determine if the observed number of cases is significantly different from the expected number or if the difference may be due solely to chance, a 95% confidence interval (CI) was calculated for each SIR. A 95% CI assesses the magnitude and stability of an SIR. Specifically, a 95% CI is the range of estimated SIR values that has a 95% probability of including the true SIR for the population. If the 95% CI range does not include the value 100, then the study population is significantly different from the comparison or “normal” population. “Significantly different” means there is less than 5% percent chance that the observed difference is the result of random fluctuation in the number of observed cancer cases.

For example, if a confidence interval does not include 100 and the interval is above 100 (e.g., 105-130), then there is statistically significant excess in the number of cancer cases. Similarly, if the confidence interval does not include 100 and the interval is below 100 (e.g., 45-96), then the number of cancer cases is statistically significantly lower than expected. If the confidence interval range includes 100, then the true SIR may

be 100, and it cannot be concluded with sufficient confidence that the observed number of cases is not the result of chance and reflects a real cancer increase or decrease. Statistical significance is not assessed when fewer than five cases are observed.

In addition to the range of the estimates contained in the confidence interval, the width of the confidence interval also reflects the stability of the SIR estimate. For example, a narrow confidence interval (e.g., 103--115) allows a fair level of certainty that the calculated SIR is close to the true SIR for the population. A wide interval (e.g., 85--450) leaves considerable doubt about the true SIR, which could be much lower than or much higher than the calculated SIR. This would indicate an unstable statistic.