

# APPENDIX III-B

## CALCULATING AND INTERPRETING ATTRIBUTABLE RISK AND POPULATION ATTRIBUTABLE RISK

### RISK

Risk as defined for public health planning is the probability of the occurrence of a disease or other health outcome of interest during a specified period, usually one year. Risk is calculated by dividing the number who got the disease during the defined period by the total population of interest during that period. For example if there were 1000 births in a health jurisdiction in one year and 72 of those babies weighed less than 2500 grams, the risk of low birth weight (LBW) in the community would be  $72/1000 = 0.072$  or 7.2%.

However, we seldom rely on a single risk calculation. Most often we are interested in comparing the occurrence or prevalence of a health condition or problem in two groups, one who experiences a risk factor or condition (referred to as the exposed) and one who does not (referred to as the unexposed). This allows us to determine whether a particular exposure does represent a risk, and if so how big a risk. When conducting a risk analysis, epidemiologists generally begin by constructing a 2x2 table as illustrated below where 'a' represents those in the exposed population who experienced the outcome of interest and 'b' those is the exposed population who did not experience that outcome. In this case, the risk of exposure is expressed as 'a/a+b'. Conversely the risk for those not exposed to the risk factor would be 'c/c+d'.

		Disease or Other Health Outcome		
		Yes	No	
Exposure or Person, Place, or Time	Yes	a	b	a + b (n <sub>1</sub> )
	No	c	d	c + d (n <sub>2</sub> )
		a + c (m <sub>1</sub> )	b + d (m <sub>2</sub> )	

The risk can vary from 0 to 1 in the case of a risk factor and from -1 to 0 in the case of a protective factor.

For the rest of this document, we will assume that the data in the 2x2 table represent either the entire population or a random representative sample of the population since those are the types of data used by most public health departments in their surveillance activities. The rest of this discussion will focus on methods for comparing two independent proportions, those exposed to a risk factor or condition and those not exposed. We have used reference material from Handler & Rosenberg, "Analytic Methods in Maternal and Child Health" in developing this resource.<sup>i</sup>

## RISK RATIO OR RELATIVE RISK

Relative risk is the calculated ratio of incidence rates of a health condition or outcome in two groups of people, those exposed to a factor of interest and those not exposed. It is used to determine if exposure to a specific risk factor is associated with an increase, decrease, or no change in the disease or outcome rate when compared to those without the exposure. Relative risk is a statistical measure of the strength of the association between a risk factor and an outcome.

Relative risk can be calculated from a simple 2 X 2 table such as the one above. The formula for calculating relative risk is:

$$\text{Relative risk} = \frac{\frac{a}{(a + b)}}{\frac{c}{(c + d)}}$$

Example: In a particular year in your health jurisdiction, there were 1000 births. In reviewing their birth certificates, you found that 72 had low birthweights (<2500 gms), and 158 had mothers who smoked during pregnancy. Of the mothers who smoked during pregnancy, you found that 19 had low birth weight babies. In order to explore the relationship between maternal smoking and the occurrence of low birth weight, you would construct the following table.

Smoked During Pregnancy	Low Birth Weight		
	Yes	No	
Yes	19	139	158
No	53	789	842
	72	928	1000

$$\text{Relative risk} = \frac{\frac{19}{(19 + 139)}}{\frac{53}{(53 + 789)}} = \frac{\frac{19}{158}}{\frac{53}{842}}$$

$$\text{Relative risk} = \frac{.120}{.063} = 1.9$$

In this example, the risk of LBW in smokers is 12% and in non-smokers 6% and the relative risk of low birthweight associated with cigarette smoking while pregnant is 1.9. In other words, women who smoke while they are pregnant are about twice as likely as those who do not to have a low birthweight infant.

When the relative risk associated with a factor is more than 1, then the factor is called a risk factor. When the relative risk associated with a factor is less than 1, then the factor is called a protective factor.

If the relative risk equals 1, then factor is not associated with the outcome.

Another statistic similar to relative risk and often used by epidemiologists is the odds ratio:

$$\frac{a/b}{c/d}$$

When the difference in incidence or prevalence rate of an outcome is very small, the odds ratio and the relative risk yield similar results, but when the differences are significant the difference is much greater with the odds ratio being higher. For routine public health use, the relative risk and relative prevalence are considered to be preferable to the odds ratio because they are directly related to the probability of developing or having a health outcome.

### ***Confidence Intervals Around Relative Risk***

To calculate the 95% confidence intervals for relative risk, we use the following formula:

$$CI = (r_1/r_2) \text{ plus or minus } 1.96 \times \text{square root of } \{(1/a \times b/n_1) + (1/c \times d/n_2)\}$$

$$\text{Where } r_1 = a/(a+b) \quad \text{and} \quad r_2 = c/(c+d)$$

$n_1$  = total number of births in group 1, those with the risk factor. In this example, it is the total number of births to smokers, which is 158.

$n_2$  = total number of births in group 2, those without the risk factor. In this example, it is the total number of non-smokers who gave birth, which is 842.

Using the example above:

$$r_1 = a/(a+b) = 19 / (19 + 139) = .120$$

$$r_2 = c/(c+d) = 53 / (53 + 789) = .063$$

$$\begin{aligned} r_1 / r_2 &= .120 / .063 \\ &= 1.9 \end{aligned}$$

$$\begin{aligned} (1/a \times b/n_1) &= 1/19 \times 139/158 \\ &= .052 \times .880 \\ &= .819 \end{aligned}$$

$$(1/c \times d/n_2) = 1/52 \times 789/842$$

$$= .019 \times .937$$

$$= .018$$

$$\text{square root of } (1/a \times b/n_1) + (1/c \times d//n_2)$$

$$= \text{square root of } .819 + .018$$

$$= \text{square root of } .837$$

$$= .915$$

$$\text{CI} = (r_1/r_2) \text{ plus or minus } 1.96 \times \text{square root of } (1/a \times b/n_1) + (1/c \times d//n_2)$$

$$= 1.9 \text{ +/- } 1.96 \times .915$$

$$= 1.9 \text{ +/- } 1.79$$

$$\text{Relative Risk} = 1.9 \text{ (95\% CI .11, 3.69)}$$

## ATTRIBUTABLE RISK AND POPULATION ATTRIBUTABLE RISK

As discussed above, relative risk helps you determine whether and how strongly a precursor is associated with a particular outcome. Attributable risk helps you determine how much of an outcome may be attributable to a particular risk factor (i.e. an estimate of the excess risk) in a population exposed to that factor. This is a valuable measure, since it provides estimates of the relative impact of the poor outcome that could be achieved if the risk factor were reduced or eliminated. On the other hand, one might want to know the proportion of all cases of an outcome in the total population that could be attributed to the exposure to the risk factor. This is called the population attributable risk and when expressed as a percent, the population attributable risk percent. Calculating the population attributable risk percent allows you to determine what percent of an outcome could possibly be prevented if a risk factor were to be removed from the population.

### *Attributable Risk and Attributable Risk Percent*

To calculate the attributable risk, one simply subtracts the risk for the non-exposed group from the risk for the exposed group. Thus, attributable risk is sometimes called the Risk Difference, or Excess Risk. The excess risk is “attributed” to the exposure.

$$\text{Attributable risk (AR)} = p_1 - p_2$$

Where:

$p_1$  =  $a/a+b$ , the proportion of smokers with LBW infants and

$p_2$  =  $c/c+d$ , the proportion on nonsmokers with LBW

$$p_1 = a/a+b = 19 / (19 + 139) = .120$$

$$p_2 = c/(c+d) = 53 / (53 + 789) = .063$$

$$\text{AR} = .120 - .063$$

$$= .057$$

### *Calculating Confidence Intervals for Attributable Risk*

To calculate the 95% confidence intervals for attributable risk, we use the following formula:

$$\text{Confidence Interval} = \text{Attributable risk} \pm 1.96 \times \text{Square Root of } [p \times q (1/n_1 + 1/n_2)]$$

$p$  = the risk, which is the number of adverse outcomes divided by the total number of events. In the example on low birth weight, low birth weight is the adverse outcome and the total number of events is the total number of births. From the 2x2 tables, we see that there were 72 low birth weight births out of a total of 1000 birth, so the risk is 72/1000 or .072.

$q = 1 - p$ . In the birth weight example,  $1 - .072 = .928$

$n_1$  = total number of births in group 1, those with the risk factor. In this example, it is the total number of births to smokers, which is 158.

$n_2$  = total number of births in group 2, those without the risk factor. In this example, it is the total number of non-smokers who gave birth, which is 842

$$\begin{aligned} (1/n_1 + 1/n_2) &= 1/158 + 1/842 \\ &= .006 + .001 \\ &= .007 \end{aligned}$$

$$\begin{aligned} [p \times q (1/n_1 + 1/n_2)] &= 0.072 \times 0.928 \times .007 \\ &= .00047, \text{ which we can round to } .0005 \end{aligned}$$

$$\begin{aligned} \text{Square Root of } [p \times q (1/n_1 + 1/n_2)] &= \text{square root of } .0005 \\ &= .022 \end{aligned}$$

$$\begin{aligned} \text{Confidence Interval} &= \text{Attributable risk} \pm 1.96 \times \text{Square Root of } [p \times q (1/n_1 + 1/n_2)] \\ &= .057 \pm 1.96 \times .022 \\ &= .057 \pm .04 \end{aligned}$$

$$\text{Attributable risk} = .057 \text{ (95\% CI = .014, .100)}$$

Attributable risk is often measured as a percent. The formula for attributable risk percent (ARP) is:

$$\text{ARP} = (p_1 - p_2) / p_1 \times 100$$

Where:

$$\begin{aligned} p_1 &= a/a+b, \text{ the proportion of smokers with LBW infants and} \\ p_2 &= c/c+d, \text{ the proportion on nonsmokers with LBW} \end{aligned}$$

Refer to the data presented in the 2 X 2 table above for smoking during pregnancy and low birth weight.

Assuming that we are measuring incidence as a percentage, one can calculate the incidence of low birth weight attributable to smoking during pregnancy:

$$\begin{aligned} \text{AR\%} &= (.120 - .063) / .120 \times 100 \\ &= .475 \times 100 \\ &= 47.5\% \end{aligned}$$

This means that, among newborns with mothers who smoke and had low birth weight babies, almost half (48%) of those low birth weights are attributable to smoking. Keep in mind that 6.3% of mothers who don't smoke will have low birth weight babies, while 12% of mothers who smoke will have low birth weight babies, indicating that smoking almost doubles the risk of low birth weight babies.

A fairly simple calculation gives us the number of low birthweight babies where the condition is attributable to smoking. We know that 158 babies had mothers who smoked during pregnancy and that 5.7% of those had low birthweights as a consequence of smoking. Thus, we can estimate that about nine babies had low birthweights as a consequence of maternal smoking.

### *Calculating Confidence Intervals for Attributable Risk Percent*

The formula for calculating the 95% confidence intervals for the attributable risk percent is:

$$95\% \text{ CI} = \text{ARP} \pm \text{ARP} \times (\text{C.I. range from the attributable risk} / \text{the attributable risk})$$

Using the numbers from the examples presented so far, we have the following:

$$\begin{aligned} \text{ARP} &= 47.5\% \\ \text{C.I. range from the attributable risk} &= .04 \\ \text{Attributable risk} &= .057 \end{aligned}$$

Thus, the formula becomes:

$$\begin{aligned} 95\% \text{ C.I.} &= 47.5\% \pm 47.5\% \times (.04 / .057) \\ &= 47.5\% \pm 47.5\% \times .7 \\ &= 47.5\% \pm 33.3 \end{aligned}$$

$$\text{Attributable Risk Percent} = 47.5\% \text{ (95\% CI} = 14.2\%, 80.8\%)$$

## **POPULATION ATTRIBUTABLE RISK AND POPULATION ATTRIBUTABLE RISK PERCENT**

While attributable risk helps us estimate the excess risk among the exposed that can be attributed to the risk factor, from a public health perspective it is often more useful to re-define the attributable risk in terms of the whole population, and thus to know the proportion of cases in the total population that can be attributed to the risk factor. For this calculation, we use the

population attributable risk (PAR). Population attributable risk depends not only on the excess risk imposed by the exposure, but also on the share of the total population that is exposed. Two formulas can be used to calculate the PAR:

$$\text{PAR} = p_0 - p_2 \quad \text{or an alternate formula,} \quad \text{PAR} = (p_1 - p_2) \times n_1/N$$

Where:

- $p_0$  = the proportion of ALL births with the outcome of interest  $(a + c / a + b + c + d)$
- $p_1$  = the proportion of births with the outcome of interest to mothers WITH the risk factor  $(a / a + b)$
- $p_2$  = the proportion of births with the outcome of interest to mothers WITHOUT the risk factor  $(c / c + d)$
- $N$  = the total number of births  $(a + b + c + d)$
- $n_1$  = the number of births exposed to the risk factor  $(a + b)$

Using the same data,

$$\begin{aligned} \text{PAR} &= p_0 - p_2 \\ &= 72/1000 - 53/842 \\ &= .072 - .063 \\ &= .009 \end{aligned}$$

or using the alternate formula, we have:

$$\begin{aligned} \text{PAR} &= (p_1 - p_2) \times n_1/N \\ &= (.120 - .063) \times 158/1000 \\ &= .057 \times .158 \\ &= .009 \end{aligned}$$

This means that the overall risk of low birth weight for the total population imposed by smoking is about nine per 1,000 births. This is a fairly small added risk, mainly because smokers are a relatively small share of the study overall population.

Another way of looking at the numbers is population attributable risk percent (PAR%), which tells us the percent of cases in the total population that can be attributed to smoking.

$$\begin{aligned} \text{PAR} &= p_0 - p_2 / p_0 \times 100 \\ &= .072 - .063 / .072 \times 100 \\ &= 12.5\% \end{aligned}$$

This means that out of 72 low birthweight cases, including those born to both smoking and non-smoking mothers, 9 cases or 12.5% can be attributed to smoking. The calculation helps estimate the percent of cases in the total population that might be prevented by removing the exposure .

## SUMMARY

To summarize, both relative risk and attributable risk are valuable tools for determining the contribution of risk factors to an adverse outcome. Relative risk (RR) is a measure of the strength of the association or causal link between a risk factor and an outcome. Attributable risk (AR) helps measure the excess risk associated with the risk factor. Population attributable risk (PAR) gives the added risk in relation to the total population. Population attributable risk percent (PAR%), gives the percent of cases in the total population that can be attributed to the risk factor.

The PAR% is an especially useful, and underutilized tool in program planning. It can be used to predict the impact of public health interventions on adverse outcomes, since it considers both the excess risk associated with the exposure and the proportion of the population that is exposed. A risk factor with a large excess risk and widespread exposure poses the most severe public health risk. One with a relatively small excess risk and relatively rare exposure poses the lowest public health risk. Factors with small excess risk and wide exposure, or large excess risk and relatively rare exposure form an intermediate group of public health risks.

The PAR% quantifies the contribution of the risk factor to the outcome and can thus help direct interventions. The higher the PAR%, the greater the proportion of the outcome that is attributable to the risk factor. One can compare the values of population attributable risk percents for selected risk factors to identify those risk factors that are most important for planning interventions.

Most of the time when we examine risk factors, we look at behaviors, medical conditions, and environmental factors. It makes sense to talk of "eliminating" these kinds of risk factors. There are other risk factors, such as race or age that cannot be changed. Instead, identifying people at risk for an adverse outcome by race or age groups provides populations for targeting interventions.

It is important to note, however, that population attributable risk percents calculated from a 2 x 2 table are crude measures of attributable risk. Because the outcome is compared to only one risk factor at a time, there is no way to know if other risk factors may underlie or explain the associations found in a 2 x 2 table. More advanced statistical methods, such as multivariate analysis, can be used to calculate attributable risks for individual risk factors that adjust for the influence of other potential risk factors. Another option is to review the literature for studies on the outcomes of interest that use multivariate analyses to assess the impact of the risk factors of interest.

---

<sup>1</sup>Rosenberg, D. & Handler A., Measures of Association and Hypotheses Testing, In: *Analytic Methods in Maternal and Child Health*, Handler A., Rosenberg D, Monahan C., & Kennelly J. (eds), Maternal and Child Health Bureau, HRSA, 1998.