In this lecture, I am going to describe an important epidemiological measure, Attributable Risk, as well as review the calculation of the relative risk and odds ratio.

Let's start with the attributable risk.

The reason for the Shakespeare reference is that it seems in epidemiology we use multiple names for the same thing. By the way, trivia quiz 4 asks what play this quote was taken from. Attributable risk has also been called risk difference and population attributable risk has been called population risk difference, with proportion used instead of percent. I apologize for the confusion this may cause. I have chosen to teach the concepts using the simpler names. However, in this slide I have put the names used in the book so when you read any of these you can follow what they are talking about. Honestly, the names I used are ones that are more commonly used in epidemiology, and I think they will serve you better. In quizzes and tests, I will try to remember to use these names and book the book ones in parentheses but I may forget at times. So you have this handy slide to refer to. There are a number of measures of attributable risk among studies.

The four we want to discuss here are:
<table>
<thead>
<tr>
<th>Concept</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributable Risk</td>
<td>AKA Rate or risk difference</td>
</tr>
<tr>
<td>Attributable Risk Percent</td>
<td>AKA Attributable proportion among the exposed</td>
</tr>
<tr>
<td>Population Attributable risk</td>
<td>AKA Population rate difference</td>
</tr>
<tr>
<td>Population Attributable Risk Percent</td>
<td>AKA Attributable proportion among the total population</td>
</tr>
</tbody>
</table>

The reason for the Shakespeare reference is that it seems in epidemiology we use multiple names for the same thing. By the way, trivia quiz 4 asks what play this quote was taken from. Attributable risk has also been called risk difference and population attributable risk has been called population risk difference, with proportion used instead of percent. I apologize for the confusion this may cause. I have chosen to teach the concepts using the simpler names. However, in this slide I have put the names used in the book so when you read any of these you can follow what they are talking about. Honestly, the names I used are ones that are more commonly used in epidemiology, and I think they will serve you better. In quizzes and tests, I will try to remember to use these names and book the book ones in parentheses but I may forget at times. So you have this handy slide to refer to.
Let’s start with a discussion of attributable risk among the exposed. Attributable risk is also called the risk difference, and you will see why shortly. It is a measure of the public health impact of an exposure on a given disease. It identifies how much of the disease among the exposed could be eliminated if the exposure were eliminated. The good news is that the attributable risk in the exposed is very easy to calculate. You simply subtract the incidence of disease in the unexposed from the incidence of disease in the exposed.

For most diseases, causation results from a number of risk factors. Let’s consider heart disease. We know risk factors include smoking, diet, exercise, and genetics. What the attributable risk among the exposed seeks to do is to determine how much of the disease is due to the exposure. So a question might be stated as “How much heart disease is caused by smoking? The reason for that question is if we calculate the attributable risk due to smoking, we can determine how much of heart disease could be eliminated if we eliminated smoking. Look at the graph on the right. The group on the left are those who smoke and the group on the right are those who do not smoke. The blue section indicates the amount of heart disease in both groups. We know that both people who smoke and those who do not have heart disease. So we subtract the “base” level of heart disease from the smokers to identify the amount of heart disease resulted from smoking.

This number is usually expressed as a certain number per 100, 1,000 or whatever makes sense.
This calculation expresses the attributable risk among the exposed as a proportion, which we call attributable risk percent. Think about the public health meaning of this. If we know the proportion of disease among the exposed that is caused by the exposure, we can make decisions as to what intervention program is best. Attributable risk has important public health implications. When we look at it as a percent, it is a bit easier to understand the significance of the exposure. The limitation of this measure though is that we are looking only at the exposed group, and not the general population. Thus, we may have a high attributable risk among an exposed group but if the exposure is rare, it may not have much effect.

To accomplish this, all you need to do is divide the attributable risk in the exposed by the incidence in the exposed. Again, mathematically, it is quite simple.

Let’s calculate these measures using real numbers as I think it will make it clearer. This example is looking at the association between smoking and coronary heart disease from a hypothetical cohort study of 3,000 cigarette smokers and 5,000 nonsmokers. This is a traditional 2 by 2 table but I added a total column and the calculation of the incidence rates among the exposed and non-exposed. If you want to see how the numbers were calculated, click on the markers.

To obtain the attributable risk you simply take the incidence among the exposed (smokers) and subtract the incidence in the non-exposed (non-smokers). The incidence in the non-smokers serve as the background level of disease. In this case the attributable risk equals 28.0-17.4 which equals 10.6. Since the original numbers were calculated in 1,000 per year, we indicate the AR as 10.6 per thousand.

### Attributable risk percent among the exposed

#### Answers the question

**What proportion of the risk in exposed people is due to the exposure?**

#### Formula

\[
\frac{I_{\text{exposed}} - I_{\text{unexposed}}}{I_{\text{exposed}}} = \text{Attributable risk in exposed}
\]
If we want to express the attributable risk as a proportion, then we take the attributable risk and divide it by the incidence in the exposed. This gives us the proportion of disease among the exposed which could be eliminated if we eliminate the exposure. The calculation for this is to take 28.0 - 17.4 which equals 10.6. Then divide this number by the incidence in the exposed which is 28. The result is 0.379 which we can express as 37.9%.

Now you try some calculation. I have set up two calculations, the first is a review of the relative risk and the second is the attributable risk. They can be filled in by dragging and dropping the numbers directly out of the 2 by 2 table which also has the incidence among exposed and non-exposed as I did earlier. The final answers for the relative risk and attributable risk need to be selected from the number choices below the calculations. The relative risk calculation is set up with the blue diagonal line standing for divided by as is the blue line between both rows in the calculations. It is a bit awkward but I think it is doable. First calculate the relative risk which you learned when we studies cohort studies. If you have forgotten that formula look back over your notes or in the book.

Then calculate the attributable risk among the exposed using the formula I just showed you. Good luck. You can do this multiple times if you get it wrong and if you can’t figure it out, let us know.
The population attributable risk addresses a different issue, the amount of disease among the total population which would be eliminated if the exposure were eliminated. Remember previously, we were only considering disease among the exposed. The population attributable risk takes into account two things: Incidence in the total population and incidence in the unexposed. Mathematically, this formula is easy to calculate except we often do not have the incidence in the total populations. We can, however, calculate this number if we know the prevalence of exposure in the entire community. We can use the prevalence of exposure to calculate a weighted average of attributable risk among the total population.

Has two components:
- Attributable risk due to smoking in the exposed
- Prevalence of exposure in the entire community

The effect runs from the maximum if everyone in the community is exposed to nothing if no one in the community is exposed.

In the same way, we could express the attributable risk among the exposed as a percent, we can express the population attributable risk as a percent. It is the population attributable risk divided by the incidence in the total populations.
Here is our example from before looking at the relationship between smoking and cardiovascular disease. While the attributable risk in the exposed answers an important question, the Population Attributable Risk can be much more important as we want to understand the impact an exposure would have in an entire community. But as I said we may not know the incidence in the total population but we can calculate it. We need to know the following:

1. Incidence among smokers
2. Incidence among non-smokers
3. The proportion of the total population that smokes

Let’s say we know the proportion of smokers in the population is 44% which means the proportion of non-smokers is 56%. 100-44=56.

We obtain a weighted average

\[
\text{Population Attributable Risk} = \frac{\text{Incidence in smokers} \times \% \text{ smokers}}{\text{Total population}} + \frac{\text{incidence in nonsmokers} \times \% \text{ nonsmokers}}{\text{Total population}}
\]

If you click on the i, you can see how this calculation is done.

\[
\begin{align*}
28 & \quad 17.4 \\
1000 & \quad 1000
\end{align*}
\]

So with this figure, we can calculate the next two measures we need the Population Attributable risk.

Once we know the incidence in the total population as I said the population attributable risk is simple to calculate. To calculate the PAR, take the Incidence in the total population which we calculated previously and subtract the incidence in the unexposed. In this example, it is

\[
\frac{22.1 - 17.4}{1000} = \frac{4.7}{1000}
\]

When all parts of a formula are divided by the same number, 1,000, you can remove it.

To determine the population attributable risk percent, which is the population attributable risk expressed as a proportion,
you divide by population attributable risk by the incidence in the total population.

\[
\frac{22.1 - 17.4}{22.1} = 21.3\%
\]

This number tells us what percent of heart disease in the entire population is attributed to smoking, and therefore what percent of heart disease we could eliminate if we eliminated smoking in the population. It is clearly an important number to know.

Now I want you to calculate the population attributable risk and population attributable risk percent. To be nice to you, I am providing the incidence per 1,000 in the total population. So you will not have to calculate the weighted average to obtain this figure. First calculate the population attributable risk and then calculate the population attributable risk percent. Be sure to match the colors in your answers or you will be marked wrong even if you get it right.

This slide contains a pie chart of the different population Attributable Risk factors for Liver cancer. As you can see the in the column on the right, the odds ratios for liver cancer due to Hepatitis B and Hepatitis C are the same. But the population Attributable risk Percent is much higher for Hepatitis B. This is because the rate of HBV infection is much higher than the rate of HCV infection in Gambia. Approximately 20% of the people have HBV infection while HCV prevalence is described as low, although clearly not that low since it accounts for 20% of the cases. If one were to develop a liver cancer prevention program, it would make sense to target hepatitis B first. Since there is a vaccine for hepatitis B and not hepatitis C, there are other reasons why this might be a good choice.
This slide also illustrates the difference in meaning between an Odds Ratio and the Attributable Risk. The odds ratio and the relative risk are simply measures of association. An odds ratio of 16.7 means that people with HBV are 17 times more likely to have liver cancer than those without HBV. The same is true for HCV. In contrast the attributable risk takes into account the association but also the prevalence of the exposure. More common exposures will show a higher risk for disease than less common exposures even when they have the same odds ratios.

This slide shows a comparison of the relative risk and attributable risk for the relationship between smoking and lung cancer as well as smoking and heart disease. It demonstrates a similar concept as shown in the prior slide, but this time the relative risks are quite different. The relative risk for smoking and lung cancer is 14.0 while the relative risk for smoking and heart disease is 1.6. In this example, we are looking at the same exposure but the difference in the Attributable Risk is due to the fact that heart disease is a much more common illness than lung cancer. Thus, even though the risk for lung cancer from cigarettes is higher than for heart disease, the prevalence of heart disease is much higher in the population.

**Thinking about the meaning of relative risk and attributable risk**

**Conclusion:** Cigarette smoking is a much stronger risk factor for lung cancer but (assuming smoking is causally related to both diseases) the elimination of cigarettes would prevent far more deaths from coronary heart disease. Why is this so?

<table>
<thead>
<tr>
<th></th>
<th>Lung Cancer</th>
<th>Coronary Heart Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cigarette Smoker</td>
<td>140</td>
<td>669</td>
</tr>
<tr>
<td>Non Smoker</td>
<td>10</td>
<td>413</td>
</tr>
<tr>
<td>RR</td>
<td>14.0</td>
<td>1.6</td>
</tr>
<tr>
<td>At</td>
<td>130/100,000/YR</td>
<td>25/100,000/YR</td>
</tr>
</tbody>
</table>

Please match the measure to its definition

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Risk</td>
<td>Measure of association in cohort study</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>Measure of association in case-control study</td>
</tr>
<tr>
<td>Attributable risk (Risk difference)</td>
<td>Measure of increased risk in exposed</td>
</tr>
<tr>
<td>Population attributable risk</td>
<td>Measure of increased risk in population</td>
</tr>
<tr>
<td>Attributable risk percent</td>
<td>Proportion of increased risk in exposed</td>
</tr>
<tr>
<td>Population attributable risk percent</td>
<td>Proportion of increased risk in populations</td>
</tr>
</tbody>
</table>
Which factors are needed to calculate the population attributable risk when you do not have the incidence in the total population?

- Incidence in exposed
- Incidence in non-exposed
- Relative risk
- Odds ratio
- Prevalence of exposure
- Prevalence of disease

Which measure best shows the strength of association between an exposure and an outcome in a case-control study?

- Relative risk
- Attributable risk percent
- Odds ratio
- Population attributable risk percent
- Incidence in the exposed

You are the director of the Pasco County Health Department and you have been given a grant to improve the health of diabetics in your community. You know the relative risk for lack of exercise for diabetes is 4.2 and the population attributable risk percent is 10.4%. The relative risk of diet for diabetes is 2.4 and the population attributable risk for diet is 12.8%. You only have enough money to do one of these programs. Which one do you select?