

$3^{\text {rd }}$ lecture in Epidemiology: Odds Ratio

The previous lecture was about: Relative Risk (RR), Attributable Risk Percentage (AR \%), and Population Attributable Risk Percentage (POPAR \%).

RR: it is used in cohort studies where you can calculate [scenario 1: incidence rate of the disease (example: Bronchitis) for those who are exposed to the risk factor (example: Smoking) ], and [ scenario 2: the incidence rate of the disease for those who are not exposed to the risk factor].

You divide the incidence rate of the first scenario over the incidence rate of the second scenario:


RR could be:

- $\quad R R=1$ : this means there is no association between the two variables whatsoever, because the incidence rate of bronchitis for smokers (exposed) equals incident rate of bronchitis for non smokers (nonexposed).
- RR> 1: this indicates a positive association.
- $R R<1$ : this indicates a protective association; the factor is protective against occurrence of the disease.

The principle of cohort studies is basically (regarding IR, AR\%, and POPAR\%) is to start by risk factor, a group of people are exposed to this risk factor (like smokers) and another group are non exposed (non smokers). [Note: all participants in this study, at the beginning of the study should be free from the disease to be investigated (free from bronchitis) because we are looking for new cases of that disease. So if we take 500 participants for this study, all those that already have bronchitis must be excluded because we are only

interested in new cases that would occur over -for example- the next two months, let's say 50 of them already have it therefore they are excluded from the study and we remain with 450 participants, now these 450 participants are divided according to their exposure to the risk factor | (smokers and non-smokers).] | $\begin{array}{l}\text { Follow up with the participants for a period of } \\ \text { two months, let's say the study started by the } \\ \text { beginning of December and until January, by } \\ \text { the end of January, you will be able to } \\ \text { identify those who have bronchitis (new } \\ \text { cases of bronchitis) }\end{array}$ |
| :--- | :--- |



By this way you get the IR of bronchitis in smokers and non smokers.

Odds Ratio: it measures the degree of association between risk factor and the disease occurrence in a case-control study.

Let's take this table:

|  | Disease +ve | Disease -ve |
| :--- | :---: | :---: |
| Risk Factor <br> +ve | A | B |
| Risk Factor <br> -ve | C | D |

Fig. 2

> Disease: Bronchitis
> Risk Factor: Smoking

Done by: Mona Makhamreh

A: has bronchitis and smokes. (Has the risk factor and the disease)

B: no bronchitis, and smokes. (Has the risk factor and not the disease)

C: has bronchitis, and don't smoke. (No risk factor, has the disease)

D: no bronchitis, and don't smoke. (No risk factor, no disease)

Step 1: start by recognizing whether the disease is present or absent.

Step 2: take 100 cases of bronchitis, then take an equal number of cases (100) for those who don't have bronchitis. [In such case studies conducted, \# of cases is usually equal to \# of controls]

Those who have bronchitis are referred to as "case group" ( $\mathrm{A}+\mathrm{C}$ in the table) while those who don't have bronchitis are referred to as "control group" (B+D in the table), we will be able to compare cases with control groups later on.

Step 3: start filling the table by taking the Case Group, which are a total of a 100, now divide them according to those who have the risk factor before the onset of the disease (group A in the table) (simply by asking were you smoking before the signs and symptoms?) and those who don't have the risk factor (group C in the table), This is to achieve temporality (the effect occurs after the cause).

It turned out that $\mathrm{A}=40 \mathrm{C}=60$.

Step 4: take the control group (don't have bronchitis), which are also a 100 participants, and when asked about smoking habits; smokers $(B=20)$ and nonsmokers ( $D=80$ )

Now that we have all the variables, the study is almost done, such studies are simple, cheap and easy to be done, plus the sample size needed is quite small.

So what is the degree of association of bronchitis and smoking in this study?
To answer that, we need to calculate the Odds Ratio, also known as the cross product ratio.

Step 5: Calculate OR by cross multiplying the variables in the table
Odds Ratio $=$ A*D/ B*C, example (40*80)/ (20*60) $=$ almost 3

Always determine the first variable which is cases with the risk factor (here abbreviated by A) and from the table you can directly know the cross product ratio. If a table was like the one in fig. 3 then " $U$ " would be cases with risk factor (equivalent to A in fig.2) which leads us to multiply it by $P$ divided by ( $\left.C^{*} J\right)$ to get $O R$
(The cases are not always on the left side of the table and the controls aren't always on the right, different studies have different tables, therefore it is important that you are able to identify which number in the table corresponds to the equivalent letter in the equation. The doctor said he might bring a question resembling this one in the exam. He could also just give us the data needed to calculate $O R$ in the question and we'll be required to draw the graph ourselves and conclude the answer)

| Risk Factor |
| :--- | :---: | :--- |
| +ve |$\quad$ Control $\quad$ cases

Now assume that the $P$ value is calculated by the computer/statistical package used and turned out to be 0.02 , significance level is $5 \%$ : (<0.05)

- From these values you get that there is a significant relationship between bronchitis and smoking ( $\mathrm{OR}=3, \mathrm{P}=0.02$ )
- The principle of OR is similar to RR;
- $\quad O R=1$ : this means there is no association between smoking and bronchitis.
- OR> 1: this indicates a positive association.
- $\quad \mathrm{OR}<1$ : this indicates a protective association; the factor is protective against occurrence of the disease.

So what does the value of $\mathrm{OR}=3$ really mean?

Notice in the design of the study, you start by cases and controls, then end by the outcome, hence OR interpretation must be opposite to RR.

So the \#3 means: Cases with bronchitis are 3 times more likely to be smokers when compared to controls (completely opposite to RR due to the design of the study).

The $p$ value was significant (0.02), what does this mean? (Almost the same explanation as the one for the RR) it means that an OR of 3 is significantly different from an OR of 1 (and 1 means that there's no association between variables).
-Null Hypothesis: you assume there is no relationship between smoking and bronchitis, so you test this hypothesis by getting the $P$ value, in the before mentioned example $\mathrm{P}=0.02$, which is significant, and the null hypothesis in this case is rejected.
> * If a question was given about this type of study and you were asked: what is the prevalence rate as well or incidence rate of this disease from the given information? The answer would be can't be determined as this type of study is specific to OR.

## Types of Epidemiological Studies

We have two types: Observational studies and Experimental studies

- Observational studies are divided into two main types:

1. Descriptive Studies: have four subtypes, we're only interested in cross sectional studies
2. Analytic Studies: classified into two subtypes:
a. Case Control Studies
b. Cohort Studies

- Experimental studies are carried out on humans to determine for example the efficacy or the efficiency of a new drug, and they are also divided into two types:

1. Clinical trials
2. Community Trials
