Before we can understand relative risk and odds ratios, we need to understand the difference between **probability** and **odds.**

## Probability and risk

You may be familiar with the idea of probability as:

Consider this example:

There are 5 red tokens and 6 blue tokens in a bag. What is the probability of picking a blue token?

There are 6 blue tokens (ie. 6 “successful” outcomes) and 11 tokens in the bag altogether (ie. 11 possible outcomes). So the probability of choosing a blue token is

In medical contexts, **a risk means a probability**.

For example, if we take a sample of 200 people and 6 of them have a particular disease, this means we can say the probability of someone picked at random from this group having the disease is . If we believe the sample represents the population, we can then say that the **risk** of someone having that disease is 3%.

Probabilities (and risks) are measures of how likely something is to happen.

## Odds

We can also measure how likely something is to happen using the odds.

**Definition:** The odds of an event occurring are:

Sometimes, it can be helpful to think:

The odds compare the “successful” outcomes to the “unsuccessful” outcomes.

Let’s consider our example of token in a bag again:

There are 5 red tokens and 6 blue tokens in a bag. What are the odds of picking a blue token?

In this case the odds of a blue token are

Let's look at one more example to illustrate the difference between probability and odds.

There is a 20% chance that I will catch the bus. What is the probability I will not catch the bus? What are the odds I will catch the bus? What are the odds I will not catch the bus?

The probability I will not catch the bus is 80% (ie 100% - 20%). We can write this as 0.8 or as if we prefer

The odds of me catching the bus are . Again, we can write this number in any way we like, although it is unusual to write odds as percentages.

The odds of me **not** catching the bus are . We can think of this as “I am 4 times more likely **not** to catch the bus than I am to catch the bus”. Note that this is the same as saying that “I am a quarter as likely to catch the bus as I am not to catch it” - this is the same as saying the odds of catching the bus are 0.25 - as we calculated above.

We will now return to our medical example. In our sample of 200 people, 6 of them had the disease. Therefore the odds of having the disease are and the odds of **not** having the disease are . We can interpret this as “we are approximately 32 times as likely **not** to have the disease than we are to have the disease.

## Comparing Odds and Probability

|  |  |
| --- | --- |
| **Probability** | **Odds** |
| Measures how likely something is to happen | Measures how likely something is to happen |
| Impossible things have a value of 0 | Impossible things have a value of 0 |
| Scale 0-1 | Scale 0 to infinity |
| 0.5 - equally likely that an event will happen or not | 1 - equally likely that an event will happen or not |
| >0.5 the event in question is “more likely” to happen  <0.5 the event in question is “less likely” to happen | >1 the event in question is “more likely” to happen  <1 the event in question is “less likely” to happen |

## Relative Risk

Relative risk is a way of comparing probabilities of an event occurring in two different groups.

Since relative risks and odds ratios are most commonly used in medical contexts, we will now stick to examples featuring groups of people either getting a disease or not.

We will imagine that the people in our sample have been “exposed” to a particular set of circumstances.

If this is an exposure to a cause of disease, we would expect more people in the exposed group to get the disease - ie. the risk of disease in the exposed group would be higher.

On the other hand, if the “exposure” is to a treatment, we might find that the risk of getting the disease in the exposed group is lower.

We can display the relevant numbers is a table like so:

#### 

|  |  |  |  |
| --- | --- | --- | --- |
|  | Not exposed | Exposed |  |
| Have the disease | 3 | 40 | 43 |
| Do Not have the disease | 35 | 5 | 40 |
|  | 38 | 45 | 83 |

In this example, 43 people have the disease out of a total of 83. So overall, the risk (probability) of getting the disease is We could write this with mathematical notation like so:

In the exposed group, 40 people out of 45 have the disease. So, in this group, the risk of getting the disease is . Using mathematical notation, this is:

In the group who are not exposed, 3 out of 38 people get the disease. So, in this group, the risk of getting the disease is . Using mathematical notation, this is:

**Relative risk** is used to compare the probability of getting the disease in the exposed group compared to the non-exposed group. We do this by dividing the probabilities like so:

Relative risk of the exposed group compared to the non-exposed group =

We can interpret this as “the probability of getting the disease if you are in the exposed group is 10 times higher than the probability of getting the disease in the non-exposed group”

We could also look at the relative risk the other way round. Is the “risk” of the non-exposed group compared to the exposed group = . ie. th people in the non-exposed group are a tenth as likely to get the disease as the exposed group.

So, we remember

**Relative risk** (of the exposed group compared to the non-exposed group) **is given by:**

We sometimes see this written as a formula like so:

#### 

|  |  |  |  |
| --- | --- | --- | --- |
|  | Not exposed | Exposed |  |
| Have the disease | a | b | a+b |
| Do Not have the disease | c | d | c+d |
|  | a+c | b+d | a+b+c+d |

**Relative risk** (of the exposed group compared to the non-exposed group) **is given by:**

## Odds Ratio

Odds ratio, like relative risk, measures how likely the exposed group are to get the disease when compared to the non-exposed group (or vice versa). The difference is that this time, instead of calculating the **risks** and dividing one by the other, we calculate the **odds** and divide one by the other.

So we have

**Odds Ratio** (of the exposed group compared to the non-exposed group) **is given by:**

So in our example:

#### 

|  |  |  |  |
| --- | --- | --- | --- |
|  | Not exposed | Exposed |  |
| Have the disease | 3 | 40 | 43 |
| Do Not have the disease | 35 | 5 | 40 |
|  | 38 | 45 | 83 |

The odds of having the disease in the exposed group are

The odds of having the disease in the non-exposed group are

So the **odds ratio** (of the exposed group compared to the non-exposed group) is

We can interpret this as “the odds of getting the disease for people in the exposed group are nearly 90 times higher than the odds in the non-exposed group”.

Again, we often see the odds ratio written as a formula like so:

#### 

|  |  |  |  |
| --- | --- | --- | --- |
|  | Not exposed | Exposed |  |
| Have the disease | a | b | a+b |
| Do Not have the disease | c | d | c+d |
|  | a+c | b+d | a+b+c+d |

**Odds ratio** (of the exposed group compared to the non-exposed group) **is given by:**

Please note that if the table is drawn the other way round or if the sections of the table are labelled differently, the formula may have the letters in different places.

## Language

We might hear sentences such as:

“The chances of getting the disease in the exposed group are twice as high (as the non-exposed group)”

Or

“The exposed group are 3 times more likely to get the disease”

In these cases we should be careful to specify whether we’re using odds ratio or relative risk - the sentences above **could** mean either. In practice, we find that in informal situations (such as newspaper headlines) the information about how “the chances” are measured is missing. This imprecise language can be misleading.

## When are Odds Ratio and Relative Risk similar?

If the number of people who get the disease in question is small compared to the number of people who don’t get the disease, we find that the odds ratio and the relative risk are similar numbers. This is perhaps the reason we see ambiguous language such as the examples above. In a clinical context, doctors may deliberately choose to use this language because patients are unlikely to be familiar with the differences between odds ratio and relative risk.

To see why this is true, let’s look at this example:

#### 

|  |  |  |  |
| --- | --- | --- | --- |
|  | Not exposed | Exposed |  |
| Have the disease | 1 | 2 | 3 |
| Do Not have the disease | 1000 | 1000 | 2000 |
|  | 1001 | 1002 | 2004 |

The probability of someone in the exposed group having the disease is

The odds of someone in the exposed group having the disease is .

We can see that these are quite similar numbers.

In the same way, the other probabilities and odds will be similar numbers.

In this case, if we carry on the calculations we end up with:

Odds ratio = 2

Relative risk = 1.998004 2

## Comparing Odds Ratio and Relative Risk

|  |  |
| --- | --- |
| **Odds Ratio** | **Relative Risk** |
| <1  means less likely to get the disease  1 means equally likely to get the disease  >1 means more likely to get the disease | <1  means less likely to get the disease  1 means equally likely to get the disease  >1 means more likely to get the disease |
| Compares odds | Compares probabilities |