29. Relative Risk, Odds Ratio, and Hazard Ratio

François Noël, June 2022
(revised by Álvaro Vigo, UFRGS)

During the COVID-19 pandemic, how many times did we have to examine studies assessing the **efficacy or effectiveness** of an antiviral drug or a vaccine? How were such evaluations performed? Which measures were used, and what do they mean? At the outset, we can point out that there is considerable confusion in the use (or interpretation) of these efficacy measures in the international literature (Castro et al., 2019; George et al., 2021; Spruance et al., 2004), and the situation becomes even worse when technical terms are translated into Portuguese. Some Brazilian authors find it preferable to use the English terms to avoid confusion, while others recommend using the Portuguese terminology followed by the English acronym in parentheses, which seems to us the most appropriate approach.

1. Relative Risk (Risk Ratio: RR)

Many studies assessed the Relative Risk (RR), which is nothing more than a ratio of probabilities, or incidences, as shown below.

- Considering the table below, in which the numbers (A and C) of events (for example, deaths due to COVID-19) are observed in the treated group (of size $n_1 = A+B$) and in the control group ($n_2 = C+D$), the relative risk (RR) is calculated as follows:

$$\mathbf{RR = [A/(A+B)] / [C/(C+D)]} \qquad \text{Equation 1}$$

Group	Event	Without event
Treatment	A	B
Control	C	D

Thus, the relative risk tells us the ratio between the probability of an event occurring in the treatment group [A/(A+B)] and in the control group [C/(C+D)]. Relative risk is a powerful and relatively easy-to-understand statistical tool because it directly shows whether a treatment has an effect (George et al., 2020).

– How should relative risk be interpreted? Because it is a ratio expressing how many times more (or less) likely the outcome is in the treated group, an RR of 0.7, for example, means that *"Those who were treated had 0.7 times the risk of experiencing the outcome compared with those who were not treated"*. In this case, it is also correct (and more intuitive) to say that the **relative efficacy** of the intervention, [(1 − RR) × 100], was 30%,

since *"Those who were treated had 30% less risk of experiencing the outcome compared with those who were not treated."*

Limitations in the use and/or interpretation of relative risk

- RR can only be estimated when information about outcome status is obtained using epidemiological study designs that involve following individuals over time, assessing the presence or absence of disease at the end of the follow-up period, as in cohort studies, case-control studies, and randomized controlled clinical trials (RCTs). Unlike RCTs, considered the gold standard for estimating treatment effects, other designs are susceptible to confounding, and therefore adjusted (more sophisticated) analyses must be used to estimate relative risk in such cases.

- When follow-up times differ among individuals, regression methods must be used, such as Cox regression, which estimates a measure of association (or effect, in the case of RCTs) known as the hazard ratio (HR: see item 3).

- Reporting RR alone can exaggerate or minimize the clinical relevance of the treatment.

> For example, compare a situation where an RR of 0.5 arises from very low risks (e.g., 0.05% in the treated group vs. 0.1% in the control group) with another where the same RR is obtained despite much higher risks (e.g., 45% vs. 90%). The impact of the treatment on public health will be very different, even though the treatment has the same *relative* efficacy.

Thus, it is important to emphasize that relative efficacy does not reflect the magnitude of the intervention's impact. An alternative measure to assess the magnitude of impact and intervention costs is the **Number Needed to Treat (NNT)**, which represents the number of individuals who must receive the intervention to prevent one case of disease or death (Szklo and Nieto, 2019). NNT is defined as the inverse of the Absolute Risk Reduction (ARR), i.e.,

$$\mathbf{NNT} = \frac{\mathbf{1}}{\mathbf{ARR}} = \frac{\mathbf{1}}{\mathbf{R_{Control} - R_{Treated}}}. \qquad \text{Equation 2}$$

Thus, many authors emphasize the importance of reporting the **absolute risk** in both groups, in addition to the RR, in order to better contextualize the impact of the disease and of the treatment.

2. Odds Ratio (OR)

On the other hand, many clinical studies use the **Odds Ratio (OR)**, another measure of association widely employed in epidemiology (Schmidt and Kohlmann, 2008), although it has a much less intuitive interpretation than relative efficacy, as shown below:

Considering the case described in the table above, we must first define an **odds** as the probability of occurrence of the event (or exposure) divided by the probability of non-occurrence of that same event:

Odds in the treatment group = [**A/(A+B)] / [B/(A+B)]** Equation 3
Odds in the control group = [**C/(C+D)] / [D/(C+D)]** Equation 4

Thus, the **Odds Ratio (OR)** is calculated by dividing (3) by (4), which gives:

OR = (A/B) / (C/D) Equation 5

ORs are frequently interpreted as if they were equivalent to RRs, although they always overestimate RRs (Schmidt and Kohlmann, 2008).

- If this measure is not intuitive, why is it so widely used in clinical research? There are three main reasons:

1. It does not suffer from the limitations associated with estimating relative risk, such as in case-control studies; in other words, it can be estimated from data generated by *any* epidemiological design (George et al., 2020);
2. Many statistical software packages historically offered only ORs as the measure of association (Schmidt and Kohlmann, 2008). With the extensive development of computational methods in recent years, this justification is no longer valid;
3. In observational studies, adjusted analyses are required to obtain unbiased association estimates, which is generally performed using logistic regression, even though this model estimates odds ratios rather than relative risks.

3. Hazard Ratio (HR)

The instantaneous risk (hazard) is a speed-based risk measure at a given point on the time curve (hence, the adjective “instantaneous” which unfortunately is seldom used in practice). Risk at time *t* is defined as the limit of the following expression when the time interval Δt approaches zero; thus, it provides information about the rate at which the event occurs within a given time interval:

Probability of an event in [t, t+Δt] / Δt Equation 6

Thus, the hazard ratio (HR) describes the ratio of the instantaneous hazard rates of the event (outcome) at a given moment in time, defined for each of the groups being compared (treatment vs. control). The HR can be estimated using the Cox regression model, which has the advantage of using all available information, including data from patients who do not complete the study (Spruance et al., 2004).

HR is widely used in survival curves, which show the temporal progression of an event within a group, such as death when evaluating treatment efficacy, or infection when evaluating vaccine efficacy. HR is also used in therapeutic trials where the question is how much a treatment can shorten the duration of illness (or hospitalization due to COVID-19, for example) (Spruance et al., 2004). It is important to note that HR has become a standard measure of efficacy/effectiveness and is routinely interpreted similarly

to relative risk, despite their intrinsic differences. In fact, HR can be misleading when used to evaluate the magnitude of treatment benefit, since a relatively large HR may correspond to only a small treatment effect (Spruance et al., 2004).

Conclusion: Use and Differences Between RR, OR, and HR

In conclusion, it is important to stress that the metric to be used depends on the study design and the statistical analysis employed (Castro et al., 2019). The hazard ratio (HR) differs from the relative risk (RR) and the odds ratio (OR) in that RR and OR are cumulative over the entire duration of the study, whereas HR pertains to the rate of change (George et al., 2021).

The table below provides a didactic summary of some characteristics and differences among the three metrics discussed in this section (George et al., 2021).

	RR	OR	HR
Use	Indicates how an intervention changes a risk (causality or association)	States whether there is an association between an intervention and an outcome (association)	Indicates how an intervention changes the rate of occurrence of an event
Limitations	The study design must accurately represent the population – It cannot be used in case-control studies or in cross-sectional studies	It can be applied in all cases, but it is not always a useful statistic by itself. – It overestimates relative risks	To be useful, the rate of change in both groups must be relatively consistent
Timeline	Static: does not take rates into account – Summarizes an overall study	Static: does not consider rates – Summarizes the study overall	Rate-based – Provides information on how a study progresses over time

References

Castro S.M.J., Mancuso A.C.B., Leotti V.B., Hirakata V.N., Camey S.A. Bioestatística e epidemiologia: perguntas que você sempre quis fazer, mas nunca teve coragem. *Clin. Biomed. Res.* 39(3):258-265, 2019.

George A., Stead T.S., Ganti L. What's the Risk: Differentiating Risk Ratios, Odds Ratios, and Hazard Ratios? *Cureus* 12(8):e10047, 2020.

Schmidt C.O., Kohlmann T. When to use the odds ratio or the relative risk? *Int. J. Public Health* 53:165-167, 2008.

Spruance S.L., Reid J.E., Grace M., Samore M. Hazard Ratio in Clinical Trials. *Antimicrob. Agents Chemother.* 48(8):2787-2792, 2004.

Szklo M., Nieto J. Epidemiology Beyond the Basics. Fourth Edition. Jones & Bartlett Learning, Burlington, MA, 2019.