VI. Clinical Relevance

Juan O. Talavera, Rodolfo Rivas-Ruiz, Marcela Pérez-Rodríguez

In clinical practice, the maneuver that is usually selected is the one that achieves an outcome with at least 10 % of direct superiority or when the number needed to treat is ≈10. Although these parameters serve for estimating the magnitude of an association, we are forced to differenciate the measures of impact (attributable risk, preventable fraction), association (relative risk, odds ratio, risk ratio) and frequency (incidence and prevalence), which are applicable when the outcome is nominal. We also have to identify the way for measuring the strength of association and the magnitude of association when the outcome variable is quantitative. Not unfrequently, association measures are interpreted as if they were impact measures, v.gr., for a relative risk of 0.68, a 32 % of outcome reduction is assumed without considering that this is a relative reduction that can be generated by a ratio of 0.4/0.6, 0.04/0.06 or 0.00004/0.00006 as well; however, the direct reduction is 20 % (60-40 %), 2 % and 2 per 100 000, respectively. Therefore, in order to estimate the impact of a maneuver, it is important that the direct difference or the number needed to treat is available.

Key words
association measures
exposure
risk or outcome
relative risk
number needed to treat

This article was originally published in Rev Med Inst Med Seguro Soc 2011; 49 (6): 631-635 and it has been reviewed for this issue.

Introduction

Even with a well-designed trial, with an adequate statistical analysis and sample size, in which statistical significance in the association between a maneuver and an outcome is shown (whether it is the association between a risk factor or preventive maneuver and the occurrence of a disease, or between a prognostic factor or therapeutic maneuver and the course of the disease), the clinician needs to identify the magnitude of this association —impact of the maneuver— in order to consider its usefulness in common clinical practice, in which most of the time, the benefit of a therapeutic maneuver is considered and it is usually selected that which achieves a favorable outcome with at least 10 % of direct superiority over others. This means that, for example, if the outcome is survival and the selected maneuver is *A*, it is expected for it to be 10 % superior than standard maneuver *B* (70 % two-year survival for maneuver *A* versus 60 % for maneuver *B*), or if the outcome is the level of glucose, then a reduction of at least 10 % is expected (from 140 to 126 mg/dL). And if the outcome is heart failure, a reduction of at least 10 % is expected in the degree of heart failure (overall, at least 10 % more of patients improving their heart failure grade). It should be noted that the substraction of a proportion from another was made directly, whereas for quantitative data, 10 % is estimated based on the reference value.

In public health or preventive medicine, direct differences lower than 10 %, and even as low as 4 to 7 %, are highly relevant, since susceptible populations may include millions of subjects. The same happens in clinical care, where the rate of unwanted outcomes is around 10 %, for which any expected reduction will be lower than this and its relevance will depend on the severity and cost of the disruption. On the other hand, in case of adverse events, differences even lower than 10 % are significant, especially depending on the severity of the event. Nevertheless, in most clinical situations, a minimum gain of 10 % is considered desirable.

While for clinicians it is common and understandable a percentage difference to estimate the impact of an association, in literature there is a series of calculations known as *impact measures* that, in spite of being discretely more elaborated, turn out to be an association between proportions. In the process of obtaining the impact measures, *association measures* are estimated (indicators that assess the strength at which a variable or feature is associated with another), which would be meaningless if they would not be accompanied by the certainty that such association is real and not due to chance, and for this purpose, *statistical significance* is estimated (an association is real when the

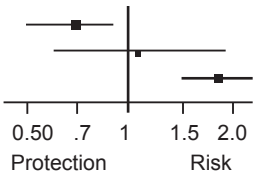
Table I Double input table for measures of relative frequency (example), association and impact

	Outcome +	Outcome –	Total
Exposed (treated)	a 5	b 95	a + b = 100
Non-exposed (placebo)	c 15	d 85	c + d = 100
Total	a + c = 20	b + d = 180	
Clinical trial and cohort	**Formula**	**Example**	**Interpretation**
Exposed incidence (Ei)	Ie = a/a + b	5/100 = 0.05	5 new cases in 100 subjects or 5 %
Incidence of observed or non-exposed (Io)	Io = c/c + d	15/100 = 0.15	15 new cases in 100 subjects or 15 %
Relative risk (RR)	RR = Io – Ie	0.05/0.15 = 0.33	A protection exists. Relative or risk reduction. The risk is below the unit
Absolute risk reduction (ARR) (attributable risk [AR])	RR = Io – Ie	0.15 – 0.05 = 0.1	The direct reduction of risk attributed to treatment is 10 %
Number needed to treat (NNT)	NNT = 1/RAR	NNT = 1/0.1 = 10	10 people have to be exposed to observe the beneficial effect in one
Attributable fraction (AF) (for RR > 1)	Ie – Io/Ie	Since in this example RR is > 1, AF is not calculated	Interpreted as the proportion of cases exposed due to the risk factor
Relative risk reduction (RRR) (for RR < 1, preventable fraction)	RRR = 1 – RR x 100	1 – 0.33 x 100 = 67 %	67 % of cases were prevented due to the exposition factor
Case-controls, and cross-sectional survey			
Prevalence of exposed (Pe) (only in cross-sectional survey)	Pe = a/a + b		Number of events in the exposed group (used in cross-sectional studies)
Prevalence of non-exposed (Po) (only in cross-sectional survey)	Po = c/c + d		Number of events in non-exposed group or control (used in cross-sectional studies)
Exposition factor prevalence in cases	PfrCa = a/a + c	5/20 = 0.25	25 % of cases were exposed to exposition factor
Exposition factor prevalence in controls	PfrCo = b/b + d	95/180 = 0.527	52.7 % of controls were exposed to exposition factor
Odds ratio (OR)	a x d/b x c	RM = 5 x 85/15 x 95 RM = 425/1.425 RM = 0.29	The exposed group is protected. The risk is below the unit

Incidence and prevalence are frequency measures; relative risk and odds ratio are considered association measures; and absolute risk reduction and relative risk reduction are impact measures. Another association measure is the risk ratio, obtained in the Cox proportional hazards survival analysis (Hazard risk ratio, HRR). Attributable risk and preventable fraction can also be estimated based on the OR (instead of using Ei using Pe and instead of Io, Po)

p-value is < 0.05). Before these two types of measures, during the process of data management, we have to make use of what is known as *frequency measures*, which estimate the absolute number of events. It should be emphasized that, in most cases, what we observe in articles are *relative frequency measures*, in which the number of events is related with the total number of individuals in the population or sample under study, so that comparisons can be made at a later stage between groups with different *n* (Table I).

Table II Examples of RR and 95 % confidence intervals

	A		B			
Study examples	**Events**	**Total**	**Events**	**Total**	**RR (CI 95 %)**	**RR (CI 95 %)**
Aspirin (A) versus placebo	65	5000	95	5000	0.68 (0.50, 0.94)	
Coffee consumption (A) versus placebo (B)	25	5003	24	5000	1.04 (0.60, 1.82)	
With dyslipidemia (A) versus healthy (B)	205	5000	115	5000	1.78 (1.42, 2.23)	

RR = relative risk; 95 % CI = 95 % confidence interval; RRR = relative risk reduction
Aspirins have a statistically significant RRR of 32 %; dyslipidemia has a statistically significant RR increase of 78 %. Coffee consumption has a non-statistically significant relative increase of 4 %.

In clinical practice, measurements of the association between two variables (maneuver and outcome) by means of relative risk (RR), odds ratio (OR) and hazard ratio (Hazard risk ratio, HR) are common and are interpreted similarly; variables with a value below 1 are considered protective, whereas those with values above 1 are considered risk variables. This way, we have that common risk for the population or sample of suffering or having the event of interest without identifying any factor, either protective or of risk is 1 (which corresponds to the incidence or prevalence of the event in the entire sample or popu-

Table III Association measures and equivalents for quantitative variables

Qualitative dependent variable (nominal)			**Quantitative dependent variable**	
Frequency measures	Association measures	Impact measures	Power of association	Magnitude of association
Incidence • Incidence rate • Cumulative incidence	RR (cumulative incidence ratio)	Attributable risk (etiologic fraction, ARR and NNT)	r^2	% of difference of the means *b* coefficient
	HR *(Hazzard risk ratio)*	RRR, AF (attributable fraction)	R^2	% of difference of the means through the regression equation *(ŷ= a + b1X1)*
Prevalence • Point prevalence • Period prevalence	OR (prevalence odds ratio or crossover products)			
		r *b* coefficient R^2	% of proportion differences through the probability equation $\hat{y} = 1/1 + e^{-(a + b1X1...)}$	

$\hat{y} = 1/1 + e^{-(a + b1X1...)}$ *= probability of the event*
RRR = relative risk reduction
The NNT (number needed to treat) is a relatively new way for estimating the magnitude of association

lation under study). But if we identify a risk factor, we observe that the incidence in this subgroup increases and that in those without this risk factor, it decreases in relationship with the risk of the entire population or sample. For example, if we consider the use of aspirin to prevent myocardial infarction in a population where the one-year incidence is 1.6 %, the incidence in the aspirin-exposed group will be 1.3 %, while in the control group it will be 1.9 % with a relative risk of 0.68 (0.013/0.019), which means that there is a relative risk reduction of 32 %. So far, there seems to be an association between the use of aspirin and the reduction of infarction, but the confidence interval of 95 % for such relative risk will have to be examined: if the interval within its limits (lower and upper) is below the unit, it is considered to be statistically significant, but if the upper value exceeds the unit (1), then it is not statistically significant and, therefore, the possibility that the observed point value of 0.68 is due to chance can not be ruled out. Similarly, when we talk about a risk factor, the lower limit of the 95 % confidence interval is expected to be above the unit (1) in order for it to be statistically significant (Table II).

Frequency, association and impact measures are based on the presence or not of an event or outcome and, therefore, these are nominal variables, but, in clinical practice, there are numerous outcome variables that are measured through the change in the value of a quantitative variable, in which there is equal interest in knowing the strength and magnitude of the association, and thus, it is important to have an equivalent.

Table III shows the relative frequency, association and impact measures in a global context, basically described for a nominal dependent variable. Other measures also applicable that can define the power of association are added —association measures—:

- The determination ratio r^2, which measures the percentage of explanation of one variable based on the other and which is the square of the *r* obtained in a correlation, in this case the phi coefficient.
- The beta coefficient, which is the value obtained in a regression model (in this case logistic), which corresponds to the odds ratio logarithm.
- The R^2 similar to r^2, whose result is obtained from the regression model.

As for the magnitude of association, the estimated probability of a phenomenon occurrence can be obtained from the result of a regression model ($y = 1/1 + e^{-(a + b1X1...)}$), which in the basis of the equation for its calculation adds the beta coefficients of the different variables, and finally, calculates its global OR. With this equation, if two treatments are compared, the difference of such probability (difference of proportions) can be estimated, even if adjusted for multiple variables of interest; similarly, the different probabilities for a phenomenon to occur by exposure to different values of a quantitative variable can be compared.

The same table III shows when the dependent variable is quantitative: the units to measure the strength of association are limited to Pearson's r^2, coefficient *b* and R^2, the latter two as a result of the linear regression model.

Finally, to assess the association magnitude of a quantitative variable, the mean differences are used, more specifically the mean difference ratio, either directly estimated or as a result of the regression equation (in the linear regression, the value of the dependent variable is obtained directly).

A measure for the association magnitude that has become widely accepted is the *number needed to treat* (NNT = 1/RAR), which refers to the number of subjects that have to be treated in order to obtain the benefit in one when compared with placebo; when this number is negative, it is known as *number needed to harm*. Therefore, to define if a maneuver is clinically significant, a direct difference of 10 % can still be used or the number needed to treat (NNT), in which although there is no pre-established parameter, a value around 10 is considered ideal, which would represent treating 10 subjects to obtain the desired benefit in one (equivalent to 10 %). It is worth mentioning that, generally, placebo is rarely used as the comparative group in clinical trials; therefore, this number may be underestimated when comparing it with other active maneuver.

Comments

Proper use of measures of frequency, association or impact and their equivalents is essential to avoid common errors committed in clinical practice. It is not uncommon to interpret association measures as if they were impact measures; for example, if the OR, RR or HR of a maneuver is 0.68, a 32 % reduction of the outcome is assumed. However, it should be considered that this is a relative reduction that the same can be generated by a 0.4/0.6 ratio than from a 0.04/0.06 or 0.00004/0.00006 ratio (RR = 0.66); nevertheless, in the first case, the NNT is 5, in the second 50, and in the third 50 000. Therefore, for estimating the impact of a maneuver, it is important that the direct difference or NNT (RAR) is available.

Bibliography

1. Cordell WH. Number needed to treat (NNT). Ann Emerg Med 1999;33:433-6.
2. Feinstein AR. Principles of medical statistics. New York, NY: Chapman and Hall/CRC; 2002.
3. Guyatt GH, Sackett DL, Cook DJ. Users guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help in caring for my patients? Evidence Based Medicine Working Group. JAMA 1994;271:59-63.

For online calculation

4. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ. 1995; 310:452-4.
5. KT Clearing House. [Website]. Odds ratio to NNT converter. Disponible en http://ktclearinghouse.ca/cebm/practise/ca/calculators/ortonnt
6. Sociedad Española de Hipertensión/Liga Española para la Lucha contra la Hipertensión Arterial. [Sitio web]. Odds ratio, riesgo relativo y número necesario a tratar. Available at http://www.seh-lelha.org/oddsratio.htm
7. University of British Columbia. [Sitio web]. UBC clinical significance calculator. Available at http://spph.ubc.ca/sites/healthcare/files/calc/clinsig.html