CURRENT OPINION

Drug Safety 1999 Aug; 21 (2): 75-79
0114-5916/99/0008-0075/$02.50/0

Biological Interpretation of Relative Risk

Stephan F. Lanes

Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, Connecticut, USA

Abstract

There is widespread interest in assessing the clinical importance of a study result. This goal is impeded, however, by a lack of clarity about the biological interpretability of epidemiological effect measures, such as the relative risk. A relative risk is often interpreted merely as a measure of some vague statistical association, without a view toward a biological effect as an object of measurement. Not infrequently, if it is not statistically significant, the relative risk estimate is ignored completely.

A key to biological interpretation is appreciating the theoretical framework stipulating that outcome rates derived from 2 comparison groups actually represent measures of different effects in the same population. For instance, by using a placebo group to estimate the number of background cases that occurred in the treatment group, an estimate of the number of excess cases that occurred as a result of treatment can be made. This kind of biological entity can be derived from a relative risk, and can be more easily evaluated as to its clinical importance than a statistical association or a statement about statistical significance. Interpretation then becomes a more directed task, with a focus on the validity of certain ancillary hypotheses upon which biological interpretability rests.

The purpose of this paper is to promote the utility of epidemiological effect measures for identifying unintended drug effects. Epidemiological effect measures, such as a relative risk (RR), are sometimes viewed only as measures of statistical association instead of biological measures of causal effect. This distinction, which derives from a theoretical framework for measuring causal effects that has been described in the statistical literature and, more recently, in the epidemiological literature,[1-9] is sometimes overlooked even in formal training in epidemiology. Nevertheless, understanding a causal effect as a conceptual entity distinct from a measure of association can help us sidestep some important pitfalls when study results are interpreted. For instance, by focusing on the magnitude of a causal effect, one can more easily judge that an effect may be clinically important, even though its relative risk estimate may be small (e.g. RR <2) and not statistically significant ($p > 0.05$).

This review is a brief summary of the theoretical framework for causal effect measurement, and its application in one particular setting, namely the estimation of risks (incidence proportions, average risks) and excess case load in a randomised con-

trolled trial. To simplify the presentation further, it has been assumed that preventive effects are absent. For more detailed discussions and generalisations to other measures of disease frequency (rates, time to occurrence) and study designs (non-randomised studies) readers should consult the references provided (Rothman and Greenland[8] provides an excellent starting point).

1. Effect Measurement

When analysing a phase III clinical trial to assess the efficacy and safety of a new drug, there might be interest in identifying any unintended effects that occurred as a result of the new treatment. An effect might be defined, in this context, as any occurrences of an illness caused by the drug.[3,8]

The observed cases of an illness that occur following exposure to a drug may be characterised as belonging to 1 of 2 possible aetiological classes: background cases for whom the drug was irrelevant to the occurrence of the disease, and cases caused by the drug.[3,8] To estimate how many, if any, cases in the treatment group were caused by the drug (excess case load), the number of background cases that occurred in the treatment group must first be estimated. To achieve this goal, it is necessary to seek a comparison group that is identical to the treatment group with regard to the prevalence of each risk factor other than the treatment. In measuring causal effects, then, the comparison group can be viewed as a surrogate or proxy for the treatment group under a particular hypothesised exposure effect (e.g. no treatment effect).

Table I. Data from a hypothetical placebo-controlled trial of a new treatment

Data	Drug	Placebo
Number of cases	15	8
Population at risk	10 000	10 000
Risk	0.0015	0.0008
RD	0.0007	
RR	1.88	
p-value	0.15	
Excess case load	7	

RD = risk difference; **RR** = relative risk.

It is useful at this point to consider the data from a hypothetical study shown in table I. In this study, 10 000 people were randomly assigned to receive the drug and 10 000 people were randomly assigned to receive a placebo. After 1 year of follow-up of these people, 15 cases of a serious illness had occurred in the treatment group and 8 cases had occurred in the placebo group. It is assumed that the study is perfectly valid, in the sense of no measurement error, losses to follow-up, etc., and also that the risk of disease occurrence in the placebo group estimates the background risk of disease occurrence in the treatment group (occurrences unrelated to the treatment). Conditional on these assumptions, the number of background cases that occurred in the treatment group is estimated by multiplying the risk in the placebo group by the denominator in the treatment group. This calculation gives:

(8 background cases/10 000 people untreated) × (10 000 people treated) = 8 background cases in the treatment group

It can never be certain, of course, that the comparison group estimates the background risk in the treatment group; as in all science, inference is conditional on hypotheses that may be criticised and tested, but which can never proven.[10] By subtracting the risk in the placebo group (8 background cases) from the risk in the treatment group (15 total cases), an excess case load of 7 cases caused by the drug is calculated. These 7 cases constitute a causal effect of the drug that can be measured either by a risk ratio (RR) or a risk difference (RD). The risk difference (attributable risk) measures directly the excess case load in relation to the underlying population that received the treatment:

RD = (15 cases/10 000 people treated) – (8 cases/10 000 people untreated) = (7 cases due to treatment/10 000 people treated)

A more common measure is the risk ratio, or relative risk, which is described as the ratio of the risk in the treatment group divided by the risk in the comparison (placebo) group. Because in estimating causal effects the comparison group repre-

sents the treatment group under a hypothetical treatment effect (e.g. no treatment effect), the 2 components of effect measurement share a common denominator (i.e. the number of people receiving the drug). When we take the ratio of these risks, therefore, the denominators of the risks cancel, so it is simpler to refer to the numerators (i.e. cases): the total number of cases in the treatment group divided by the number of background cases, or 15/8 = 1.88. For this reason, the relative risk also is sometimes referred to as the total number of cases *observed* divided by the number of cases *expected*, where the expectation refers to the null hypothesis of no treatment effect (i.e. background cases). While any ratio of 2 risks is a risk ratio and a measure of association, not every risk ratio also has a biological interpretation as a measure of causal effect.

In our example, the excess case load can be derived from the relative risk estimate (RR) and the total number of exposed cases (a), as follows:

Total cases – background cases = a – (a/RR) = 15 – (15/1.88) = 15 – 8 = 7 cases caused by the drug

The cases caused by the drug are not observable, in that there is no empirical way to tell which 7 of the 15 treated cases were caused by the drug. We can estimate, however, that the probability that a treated case was caused by the drug is 7 out of 15, i.e. ≈47%. The ratio of the number of cases caused by the drug divided by the total number of exposed cases may be referred to as the excess fraction.[8]

2. Implications

2.1 Both Components of a Causal Effect Measure Refer to the Same Population

The idea that both components of a causal effect measure refer to the same population may seem odd to anyone who has been taught that the relative risk is simply the risk in an exposed group divided by the risk in an unexposed group. These apparently conflicting notions are reconciled once it is recognised that traditional teaching merely describes the mechanical procedure for computing a relative risk as a measure of association. Nevertheless, causal inference requires that the 2 components of effect measurement (e.g. the overall risk and the risk in the absence of a treatment effect) refer to the same population.[8] Indeed, when the comparison group is dissimilar to the treatment group with regard to extraneous risk factors, the procedure of dividing the risk in the treatment group by the risk in the untreated group usually will not produce a valid measure of causal effect. In this situation, some form of statistical adjustment, such as standardisation, is needed in order to estimate the background risk in the target population (e.g. the treated group) using the risk in the comparison (e.g. untreated) group.

2.2 Magnitude of the Relative Risk does not Indicate the Magnitude of the Effect

It is common to mistake the size of the relative risk with the size of the effect.[11] The effect (e.g. excess case load) is measured directly and in relation to the underlying population at risk by the risk difference. The relative risk measures the total number of cases of an illness among people exposed to a drug divided by the number of background cases in the same population. In the example, the relative risk of 1.88 indicates that the risk following treatment is about twice the background risk or, equivalently, that the excess case load due to the treatment is similar in size to the background risk.

This background risk, however, could be either large or small. Small relative risks, therefore, may correspond to large effects (and vice-versa). Consider that the effect (in terms of excess case load) of cigarette smoking on the occurrence of heart disease is generally much larger than the effect of smoking on the occurrence of lung cancer in the same population, even though relative risk estimates for lung cancer (e.g. RR = 15) typically are much larger than relative risk estimates for heart disease (e.g. RR = 1.5). The reason, of course, is that the background risk of heart disease is usually much larger than the background risk of lung cancer.

3. Magnitude of the p-Value Does Not Indicate the Magnitude of the Effect

The misconception that the p-value is 'the probability that the finding is due to chance'[11] induces investigators to focus unduly on findings with very small p-values (e.g. results that are 'statistically significant'). The inevitable consequence is that we sometimes overlook clinically important findings merely because they are not statistically significant.[12-21] This is especially true in clinical trials that lack statistical power to detect rare unintended treatment effects. The sizable literature on this topic attests to the need for a reminder of the correct interpretation of a p-value.

From the data in table I, it is possible to compute a 2-tailed p-value of 0.15. This p-value indicates that, if there were no effect of the drug (i.e. if all of the cases in the treatment group were unrelated to the treatment), and if each of these patients (and his or her outcome status) were repeatedly reassigned at random (with equal probability) to the treatment group or the placebo group, for large samples, a test statistic at least as large as was observed would be obtained in one of the groups in 15% of these random allocations. The p-value, therefore, tells us about the probabilistic performance of the randomisation procedure; it does not tell us the magnitude of the drug effect. Indeed, computation of the p-value under the null hypothesis proceeds on the assumption that there is no effect of the drug. Because the p-value is a function of both the size of the difference in the study and the size of the study, a small p-value does not indicate a large effect, nor does a large p-value indicate a small effect. The magnitude of a causal effect should be interpreted separately from the magnitude of its p-value.

4. Conclusion

Viewing epidemiological effect measures as measures of causal effects rather than mere measures of association enhances their utility in interpreting study results. Consider, for instance, that a conventional approach to interpretation of a relative risk from the hypothetical study shown in table I could perfunctorily dismiss the result as a weak association that is not statistically significant. By way of contrast, a causal interpretation of the relative risk of 1.88 reveals that among the 15 cases of a serious illness that occurred among 10 000 patients treated for one year, factors besides the treatment were responsible for 8 cases and the drug caused an excess case load of 7 cases. The importance of this estimate can then be evaluated critically, with a full evaluation considering, among other things, the likely impact of the allocation procedure on the result (as indicated, for instance, by the p-value), as well as the benefits of the treatment and the risks and benefits of alternative treatments available.

No single piece of information can provide us with everything we would like to know about a study result. A causal interpretation of epidemiological effect measures, however, can be a useful tool in identifying clinically important effects.

Acknowledgements

The author would like to thank Sander Greenland and Ken Rothman for their helpful comments regarding content and citations.

References

1. Rubin DB. Neyman (1923) and causal inference in experiments and observational studies [comment]. Stat Sci 1990; 5: 472-80
2. Lewis D. Causation. J Philos 1973; 70: 566-7
3. Rothman KJ. Causes. Am J Epidemiol 1976; 104: 587-92
4. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Stat 1978; 6: 34-58
5. Holland PW. Statistics and causal inference. Am Stat Assoc 1986; 81: 945-60
6. Greenland S, Robins JM. Identifiability, exchangeability, and epidemiologic confounding. Int J Epidemiol 1986; 15: 412-8
7. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol 1974; 66: 688-701
8. Rothman KJ, Greenland S. Mod Epidemiol, Philadelphia: Lippincott-Raven, 1998: 58-62
9. Rubin DB. Practical implication of modes of statistical inference for causal effect and the critical role of the assignment mechanism. Biometrics 1991; 47: 1213-34
10. Popper KR. The logic of scientific discovery. New York: Harper & Roe, 1968. [Originally published as Logik der Forschung. Vienna: Springer, 1934]
11. Angell M. The interpretation of epidemiologic studies. N Engl J Med 1990; 323: 823-5
12. Morrison DE, Henkel RE, editors. The significance test controversy. Chicago: Aldine Publishing Company, 1970

13. Rothman KJ. A show of confidence. N Engl J Med 1978; 299: 1362-3
14. Salsburg DS. The religion of statistics as practiced in medical journals. Am Stat 1985; 39: 220-3
15. Rothman KJ. Significance questing. Ann Int Med 1986; 105: 445-7
16. Simon R. Confidence intervals for reporting results of clinical trials. Ann Intern Med 1986; 105: 429-35
17. Gardner NJ, Altman DG. Confidence intervals rather than p-values: estimation rather than hypothesis testing. BMJ 1986; 292: 746-50
18. Poole C. Beyond the confidence interval. Am J Pub Health 1987; 77: 195-9
19. Savitz D. Is statistical significance testing useful in interpreting data? Reprod Toxicol 1993; 7: 95-100
20. Cohen J. The earth is round (p<.05). Am Psychol 1994; 47: 997-1003
21. Borenstein M. The case for confidence intervals in controlled clinical trials. Control Clin Trials 1994; 15: 411-28

Correspondence and reprints: Dr *Stephan F. Lanes*, Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Road, P.O. Box 368, Ridgefield, CT 06877-0368, USA.
E-mail: slanes@rdg.boehringer-ingelheim.com