Published in final edited form as:

Am J Ophthalmol. 2009 October ; 148(4): 484–486. doi:10.1016/j.ajo.2009.06.020.

Risk Comparisons

Nicholas P. Jewell

School of Public Health & Department of Statistics, University of California, 107 Haviland Hall, MC 7358, Berkeley, CA 94611

Abstract

We so often use the term ‘risk’ of an event in both common language and technical articles that its meaning in any given context can be vague and misunderstood. The issue becomes even more complicated when we compare risks across groups or contrast different conditions for a single group, in part because there are multiple ways of quantifying the difference—or lack thereof—between two risks.

Definition

Loosely speaking, the risk of a disease is “the probability that an individual without disease will develop disease over a defined age or time interval”¹. Unpacking this definition requires careful thought and precise descriptions including (i) an accurate definition of what constitutes the occurrence of a disease, (ii) the delineation of an appropriate time scale and the window of time when disease is recorded, and (iii) the definition of the population at risk of disease development. This definition of risk is sometimes referred to as *cumulative risk* since it counts all disease occurrences accumulating over the specified time period.

Usually, at least in epidemiological studies, we focus on incident cases of a specified disease as the relevant outcome, thereby limiting the population at risk by requiring that individuals are both disease-free at the beginning of the time interval and at risk of becoming an incident case. This obviously rules out women in studies of prostate cancer incidence, but also needs care in infectious diseases studies, for example, where prior disease experience may confer immunity, determining an individual’s ‘at risk’ status at the beginning of follow-up. Further, inclusion of individuals already suffering from a chronic condition at the beginning of the time interval introduces prevalent cases: most studies of risk comparisons wish to exclude prevalent cases since prevalence is influenced by duration of disease. When incidence is the focus, the term ‘cumulative incidence proportion’ is often used instead of ‘cumulative risk’.

The time scale used may be constant across individuals, e.g. chronological time, or individual-specific, e.g. age, or time since diagnosis. Note that other scales are possible that are only tangentially time-related, e.g. number of unprotected sexual-contacts with an HIV+ partner². Similarly the origin of the time interval may be the same for everyone or specific to individuals e.g. time of diagnosis. The length of the interval is constant across all individuals.

© 2009 Elsevier Inc. All rights reserved.

jewell@berkeley.edu, 510-642-4627.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Sampling Interpretation of Risk

What does the ‘probability’ (in the above definition) refer to when I claim that the risk of an individual becoming legally blind within three years from diagnosis of vision-threatening diabetic retinopathy is 33%?³ The interpretation is not based on any assumption that the outcome in question occurs at random. In fact, disease development may be entirely deterministic albeit according to a mechanism not yet well understood or measurable. What is meant is that the risk is just the probability that a randomly sampled individual will experience the outcome in the appropriate interval according to our precise definitions. Here, the randomness in the probability statement arises from sampling and not the disease mechanism. With this interpretation, it immediately follows that the risk is simply the fraction of the population who experience the event subject to the relevant conditions.

Other Measures of (Cumulative) Risk

Risk, being a proportion, is necessarily quantified by a number lying between 0 and 1 (with both end points possible in extreme circumstances). An alternative quantity measures the risk of disease occurring as compared to the risk of it not occurring, that is, the *odds* of disease development, simply measured by $p/(1 - p)$ if p is the risk. Thus, if the risk is 10%, the odds are automatically 0.11 (1/9, to be precise), reflecting that the outcome is nine times more likely to not occur as it is to occur. The odds of an event must be 0 or greater but has no upper bound. When an event is rare, the risk and odds are almost identical since then $1 - p$ is very close to 1.

Hazard

There are several concerns with the above definition of risk in certain circumstances. For example, suppose a substantial fraction of the population ceases to be at risk during the interval of observation, e.g. because of immunity, or cure. In addition there may be individuals who enter the population during follow-up, technically not counted by the definition since we can't be sure they were at risk at the beginning of the interval. Finally, we may worry that cumulative risk, being a single number, masks dynamic changes over a long time period in that individuals may only experience events at the very beginning—or end—of the interval. We address these issues by breaking the time interval up into smaller windows, measuring risk separately for each period. Continuing this process indefinitely, the resulting plot of risks over consecutive (tiny) periods is known as the *hazard function*. Essentially it can be interpreted as the instantaneous risk at that moment in time (analogous to speed measuring the instantaneous rate of distance covered). The hazard function is the basis of work done in survival analysis where time to event information is exploited⁴, and is closely related to the (average) *Incidence Rate* that assumes a constant hazard over time.

Risk Comparisons

For simplicity, consider two subgroups of the population with the risk in each of these two groups, p_1 and p_0 , defined identically. There are several ways to quantify a difference in these risks. The *Risk Difference*, or *Excess Risk*, is merely the absolute difference $p_1 - p_0$. The *Relative Risk*, p_1/p_0 , compares the two risks multiplicatively. Similarly, the *Odds Ratio* is the relative

change in the odds, defined by $\frac{p_1/(1 - p_1)}{p_0/(1 - p_0)}$. In survival analysis, the ratio of the hazard functions is the *Relative Hazard* and will generally vary in time. However, it is often assumed that Relative Hazards are constant over time, the so-called proportional hazards assumption, in which case the Relative Hazard is a single quantity applying to the whole interval.

When the disease is rare in both subgroups, the Relative Risk is very similar to the Odds Ratio. In such circumstances, if hazard functions are proportional, the Relative Hazard is itself close to both the Relative Risk and Odds Ratio. The *Incidence Rate Ratio* is a version of the Relative Hazard when hazard functions are assumed to be constant.

When choosing a suitable measure to compare risks, the level of the underlying risks are often crucial to interpretation, and only Excess Risk captures this characteristic. For example, a statin drug may reduce the risk of stroke by an apparently impressive 20% over 6 years, and yet the Excess Risk be as low as 0.6% or 0.006, reflecting a reduction of 6 cases per thousand over 6 years. Then, in one year, treating an entire at risk population of 1,000 individuals would eliminate only one stroke victim, potentially at very high cost⁵.

On the other hand, as Excess Risk naturally varies with the size of the risks, metaanalyses may be better applied to summarizing the Relative Risk—potentially less variable across studies—with the goal that the resulting average Relative Risk be more precisely estimated.

Estimation of Risk Comparisons

Usually, we do not have available data from the entire population, but rather only information from samples. With random samples from appropriate risk groups, as in a cohort study, it is straightforward to estimate the Excess Risk, Relative Risk and Odds Ratio. However, depending on the nature of the disease, cohort studies can be extremely expensive in resources and time and case-control studies⁶ become an attractive alternative. Case-control sampling yields separate random samples of individuals who experience the event (cases) during the relevant time interval and those who do not (controls). In such studies, it is not possible to estimate the Excess or Relative Risks absent any external information since the apparent disease frequency in the sample is entirely dependent on the size of the case and control samples, manipulated at will by the investigator. However, case-control studies do yield estimates of the Odds Ratio. Variants of the case-control design can also be used to estimate the Relative Hazard under the proportional hazards assumption, or the Relative Risk⁷. Case-control studies are more precise than cohort studies with rare outcomes, although it is also important to consider bias issues when choosing a design.

All study designs are subject to selection biases, particularly when choosing controls in either cohort or case-control investigations. Errors in measurement of exposure variables is an important issue for case-control studies, particularly when data is collected retrospectively on exposures that occurred in the far past. This point is of special concern when the measurement error is differential between cases and controls. Misclassification of disease status must also be considered, again most importantly, when the likelihood of such errors changes depending on exposure levels. It is important to note, however, that measurement errors distort estimation of risk comparisons even when the errors occur at random and are thus not differential across groups.

While it is easy to focus on incidence in cohort studies by eliminating prevalent cases at the beginning of the time interval of interest, this is sometimes not implemented with some cross-sectional studies where prevalent cases are compared with non-diseased ‘controls’. Prevalent cases are also used in some case-control studies in a similar way. In both of these situations, investigators must be cautious in their interpretation of observed risk comparisons since, for example, prevalence Odds Ratios are not the same as the (incidence) Odds Ratios defined above. For example, apparent exposure effects, captured by a prevalence Odds Ratio, may be due to their influence on duration of disease rather than on incidence⁸.

Finally, in many studies, using a consistent time interval for all study participants, as assumed in the definition of cumulative risk, may be difficult in practice. For example, in cohort studies,

follow-up information on some individuals may cease at various times throughout the study for many reasons. Accommodating differential length of follow-up in estimation of the various risk comparisons is one of the primary motivations to use survival analysis techniques that focus on hazard functions.

Further Issues

Many other issues must be considered to properly interpret estimates of risk comparisons. First, these quantities reflect association and not necessarily causation. Assessment of the role of other factors in distorting crude associations, as described above, is crucial in observational studies where the groups being compared are not determined by the investigator. This is known as *confounding* in the literature⁷. Randomization of risk groups as in therapeutic clinical trials is extremely valuable since it effectively removes the potential for confounding. In examining the role of other factors, we must also consider *effect modification* (or *interaction* to use statistical terminology) where, for example, the Relative Risk depends on the level of another factor. For example, the Singapore Malay Eye Study determined that Malay adults with lower educational levels had significantly higher prevalence of age-related macular degeneration, finding moreover that the relationship was stronger in never-smokers (based on examining Odds Ratios)⁹. In assessing statistical interaction, one must consider carefully whether the Excess or Relative Risk (or Odds Ratio) is used as the comparative measure, as the interpretation of effect modification is quite different depending on which measure is employed⁷.

More complex methods are often necessary in allowing for the role of multiple risk factors simultaneously. Methods to address confounding then usually require the use of statistical models such as logistic regression¹⁰ or Cox regression⁴ where it is important to assess the validity of model assumptions. Additional complications arise when the outcomes for individual units under study are not independent. This is common in ophthalmology when each individual under study provides information on two eyes¹¹, or in cases where units under study are retinal zones of a single eye that are being examined for evidence of neuropathy¹².

Acknowledgments

Funding support through NIAID R01-AI070043. No financial conflict of interest. The author has provided expert witness testimony in the previous two years on various cases involving Cox-2 inhibitors, other pain relievers, and defibrillators.

References

1. Gail, MH. Risk. In: Armitage, P.; Colton, T., editors. Encyclopedia of Biostatistics. John Wiley & Sons Ltd; Chichester, England: 1999.
2. Jewell NP, Shiboski S. Statistical analysis of HIV infectivity based on partner studies. *Biometrics* 1990;46:1133–1150. [PubMed: 2085629]Medline. doi:10.2307/2532454
3. The Eye Diseases Prevalence Research Group. The prevalence of diabetic retinopathy among adults in the United States. *Arch Ophthalmol* 2004;122:552–563. [PubMed: 15078674]Medline. doi: 10.1001/archophth.122.4.552
4. Lemeshow S, Hosmer DW. Survival analysis: Applications to ophthalmic research. *Am J Ophthalmol* 2009;147:957–958. [PubMed: 19463538]Medline. doi:10.1016/j.ajo.2008.07.042
5. Abramson, J. Overdosed America: The broken promise of American medicine. HarperCollins; New York: 2004.
6. Breslow, NE.; Day, NE. Statistical Methods in Cancer Research, Volume 1 -- The Analysis of Case-Control Studies. International Agency for Research on Cancer; Lyon, France: 1980.
7. Jewell, NP. Statistics for Epidemiology. Chapman & Hall/CRC; Boca Raton: 2004.

8. Lang, JM. Case-control study, prevalent. In: Armitage, P.; Colton, T., editors. Encyclopedia of Biostatistics. John Wiley & Sons Ltd; Chichester, England: 1999.
9. Cackett P, Tay WT, Aung T, et al. Education, socio-economic status and age-related macular degeneration in Asians: the Singapore Malay Eye Study. *Br J Ophthalmol* 2008;92:1312–1315. [PubMed: 18664503]Medline. doi:10.1136/bjo.2007.136077
10. Lemeshow S, Hosmer DW. Logistic regression analysis: Applications to ophthalmic research. *Am J Ophthalmol* 2009;147:766–767. [PubMed: 19376329]Medline. doi:10.1016/j.ajo.2008.07.042
11. Cumberland WG. Analysis of correlated data. *Am J Ophthalmol*. forthcoming
12. Han Y, Schneck ME, Bearse MA Jr, et al. Formulation and evaluation of a predictive model to identify the sites of future diabetic retinopathy. *Invest Ophthalmol Vis Sci* 2004;45:4106–4112. [PubMed: 15505062]Medline. doi:10.1167/iovs.04-0405