Int J Public Health 53 (2008) 118–119
1661-8556/08/020118-2
DOI 10.1007/s00038-008-7120-3

Risk quantification in epidemiologic studies

Carsten Oliver Schmidt, Thomas Kohlmann

Institute for Community Medicine, Methods in Community Medicine, University of Greifswald

Submitted: 12 December 2007; Revised: 16 January 2008; Accepted: 21 January 2008

Background and objectives

Epidemiologists commonly use statistical methods to quantify the occurrence of health related events. Measures of disease frequency are used to describe the incidence and prevalence of diseases (e. g. cumulative incidence, hazard rate, point prevalence, period prevalence). Risk measures are required to relate predictors to disease frequency (e. g. risk difference, relative risk, odds ratio). This contribution briefly reviews measures of risk.

One frequent underlying question is whether or not a group with exposure to a risk factor shows a higher risk of developing a disease or a symptom compared to a group without exposure to this risk-factor. A prominent example is the effect of smoking on cancer. For example, an American cohort study among men revealed a 25-year incidence rate of 0.6 % for lung cancer among non-smokers whereas the rate among current smokers was 6.3 % [1]. The incidence rate for all cancer types was 13.4 % among non-smokers vs. 22.6 % among smokers. As can easily be calculated, the absolute difference between incidence rates among exposed and unexposed is much larger with regard to overall cancer as compared to lung cancer (approx. 10 % vs. 5 % points). However, in relative terms, the incidence rate for lung cancer is about ten-fold the incidence rate for non-smokers while the risk for developing any kind of cancer is not even twice as high among smokers. This example illustrates that different effect measures may yield different interpretations even though they are based on the same data. As several risk measures are available for the quantification of such effects we aim at giving a brief overview.

Overview on risk measures

Risk measures may roughly be classified into two groups: measures of absolute difference, and measures of relative difference. Formulas for all subsequently introduced measures are provided in Table 1. They are discussed in more detail in common epidemiology text books (e. g. [2]).

Measures of absolute difference

Probably the most common absolute measure is the *risk difference* (*RD*, also denoted as *absolute risk reduction (ARR), excess risk* or *attributable risk*). It is computed as the differ-

Table 1 Calculation of basic measures of risk

Formal display

		Disease	
		yes	no
Exposed	yes	a	b
	no	c	d

Incidence among exposed (I^+)	$I_1 = \frac{a}{a+b}$
Incidence among unexposed (I^-)	$I_0 = \frac{c}{c+d}$
Measures of absolute difference	
Risk Difference (RD)	$RD = I_1 - I_0 = \frac{a}{a+b} - \frac{c}{c+d}$
Numbers needed to expose (NNE)	$NNE = \frac{1}{RD}$
Measure of relative difference	
Relative Risk (RR)	$RR = \frac{I_1}{I_0} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$
Odds exposed (O_1)	$O_1 = \frac{a}{b}$
Odds unexposed (O_0)	$O_0 = \frac{c}{d}$
Odds Ratio (OR)	$OR = \frac{O_1}{O_0} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \times d}{b \times c}$

Int J Public Health 53 (2008) 118–119

ence between the incidence of an event among the exposed and unexposed. Values of a risk difference range between –1 and 1. A value of 0 denotes no difference among the exposed and unexposed.

While the risk difference is straightforward to interpret, researchers should be aware of the different practical importance a risk difference may have, depending on whether the incidence of a disease is low or high. An incidence change from 1 % to 2 % for a harmful side effect is potentially much more relevant compared to a change from 50 % to 51 % for a "nuisance" symptom. Yet, both render the same risk difference of 1 %point.

A second important measure, the *numbers needed to expose* (NNE, also called "exposure impact number" - IEN) is derived from the risk difference by taking its reciprocal. The NNE tells, how many people need to be exposed or treated to observe one additional outcome of interest. This and other measures of potential impact are described in more detail in an upcoming Hints & Kinks.

Measures of relative difference

Contrary to the measures of absolute difference, measures of relative difference only have a meaning on a proportional scale. One of the most important measures of this type is the *relative risk* (*RR*). It is computed as the ratio between the incidence of an outcome of interest among exposed and the unexposed cases. It ranges between zero and infinity. In case of equal incidences among the exposed and unexposed, the relative risk takes on the value 1. Values higher than 1 signify a greater risk among the exposed, values lower than 1 denote a lower risk.

Measures of relative difference complement absolute risk measures in important ways: In the above example an incidence change from 1 % to 2 % leads to a RR of 2. A change from 50 % to 51 % corresponds to a RR of only 1.02. However the risk difference is the same in both cases (1 %point). Vice-versa, the same RR may correspond to highly different risk differences: a RR of 2 as in case of an incidence change from 1 % to 2 % may also be caused by an incidence change from 40 % to 80 %. While the first corresponds to a RD of 1.0 %point the latter corresponds to a RD of 40 %points. This example illustrates that the incidence among the unexposed should always be taken into account when interpreting relative risks or risk differences.

A widely used measure of relative difference is the *odds ratio* (OR). The popularity of the OR is due to the fact that it is directly computable from multivariate logistic regression models in virtually all major statistical software packages. It is computed as the ratio between the odds of observing the outcome of interest among exposed and the unexposed cases. Ranges of OR are the same as for RR. While the OR is often used as a proxy for the RR, the former is always biased away from the null, thus an OR overestimates a risk or a protective factor if interpreted as a RR [3]. How to adequately estimate the RR directly from multivariable analysis or how to transform both measures into each other will be discussed in an upcoming Hints & Kinks contribution.

Remarks on software

The introduced risk measures may easily be computed by hand. However major statistical software packages allow for their computation together with their confidence intervals (for example in SAS, use the TABLES statement and the "risk" subcommand, in SPSS, use CROSSTABS and the STATISTICS "risk" subcommand, STATA offers highly versatile commands, check "epitab", to get an overview).

References

1. Chyou PH, Nomura AM, Stemmermann GN. A prospective study of the attributable risk of cancer due to cigarette smoking. Am J Public Health 1992;82(1):37–40.

2. Gordis L. Epidemiology. Philadelphia: Saunders Company 2004.

3. Davies HT, Crombie IK, Tavakoli M.When can odds ratios mislead? BMJ 1998;316(7136): 989–91.

Address for correspondence

Carsten Oliver Schmidt
Institute for Community Medicine
Methods in Community Medicine
University of Greifswald
Walther-Rathenau-Str. 48
D-17487 Greifswald
Tel.: +49 3834 867760
Fax: +49 3834 867766
e-mail: carsten.schmidt@uni-greifswald.de