Risk estimates in medical studies

Epidemiology is a science of making predictions about individual patients by counting clinical events in similar patients, using strong scientific methods for studies of groups of patients to ensure that predictions are accurate. Important approach to obtaining the kind of information clinicians need to make good decisions in the care of their patients, it is evidence based practice. Epidemiology is a tool for the risk assessment.

When reviewing the medical literature, epidemiology studies often provide the most accurate estimates of risks. Epidemiology is truly the scientific basis of public health and a great source of medical evidence, particularly when the focus is on disease risk.

The considerations are:

- Patient's prognosis is expressed as probabilities – estimated by past experience
- Individual clinical observations can be subjective and affected by variables that can cause misleading conclusions
- Clinicians should rely on observations based on investigations using sound scientific principles, including ways to reduce bias

Process by which public health problems are detected, investigated, and analyzed - **Risk estimates.**

Objectives of epidemiology are:

- To determine the rates of disease by person, place and time
  - Absolute risk (incidence, prevalence).

Birth and death rates are also estimates of absolute risk.

- To identify the risk factors for the disease
  - Relative risk (or odds ratio)

Risk factors are identified by determining whether they significantly increase or decrease the risk of developing a disease. The magnitude of increased/decreased risk is expressed as a relative risk or odds ratio

- To develop approaches for disease prevention
  - Attributable risk/fraction

Attributable risk/fraction represents the incidence/proportion of disease risk that is attributable to the risk factor. It represents the amount that could be prevented if the risk factor was eliminated.

Epidemiological studies aim at assessing the relationship between exposures and outcomes. Clinicians are interested in knowing not only whether a link between a given exposure (e.g. smoking) and a certain outcome (e.g. myocardial infarction) is statistically significant, but also the magnitude of this relationship. The 'measures of effect' are indexes that summarize the strength of the link between exposures and outcomes and can help the clinician in taking decisions in every day clinical practice. In epidemiological studies, the effect of exposure can be measured both in relative and absolute terms. The risk ratio, the incidence rate ratio, and the odds ratio are relative measures of effect. Risk difference is an absolute measure of effect and it is calculated by subtracting the risk of the outcome in exposed individuals from that of unexposed.

To determine the rates of disease by person, place, and time:

Absolute risk (incidence, prevalence)

Absolute risk of a disease is your risk of developing the disease over a time period. We all have absolute risks of developing various diseases such as heart disease, cancer, stroke, etc. The same absolute risk can be expressed in different ways. For example, say you have a 1 in 10 risk of developing a certain disease in your life. This can also be said to be a 10% risk, or a 0.1 risk - depending if you use percentages or decimals.

Estimated risk of disease = Number developing disease over study period
Total number in the cohort

- Incidence = number of new cases of a disease occurring in a specified time period divided by the number of individuals at risk of developing the disease during the same time

- Prevalence = total number of affected individuals in a population at a specified time period divided by the number of individuals in the population at the time

- Incidence is most relevant clinically

Example:
In 2006, 4 new cases of an incurable viral disease were diagnosed in a population with 100 individuals. That brought the total number of affected individuals in the population to 16.

The incidence of this disease in this population in 2006 was 4/(100-16). The number at risk was 84 since 16 were already affected.

The prevalence of this disease in this population in 2006 was 16/100.

To identify the risk factors for the disease:

Relative risk (RR), odds ratio (OR)

Relative risk is used to compare the risk in two different groups of people. For example, the groups could be smokers and non-smokers. All sorts of groups are compared to others in medical research to see if belonging to a group increases or decreases your risk of developing certain diseases. For example, research has shown that smokers have a higher risk of developing heart disease compared to (relative to) non-smokers.

- RR = ratio of incidence of disease in exposed individuals to the incidence of disease in non-exposed individuals (from a cohort/prospective study)
    - If RR > 1, there is a positive association
    - If RR < 1, there is a negative association
    - If RR =1, there is no association
- OR = ratio of the odds that cases were exposed to the odds that the controls were exposed (from a case control/retrospective study) – is an estimate of the RR
    - Interpretation is the same as the RR

Relative risks are estimated from prospective studies. Individuals who are exposed (and not exposed) to a risk factor are identified at baseline. Overtime, those who develop the disease of interest are identified. The incidence of the disease is determined, separately, for the groups who were and were not exposed to the risk factor. The ratio of these rates is generated. If the ratio is statistically significantly different from 1.0, then the risk factor is said to be associated with the disease.

Example:
The incidence of lung cancer among smokers is 0.96/1000/yr. The incidence of lung cancer among non-smokers is 0.07/1000/yr. The relative risk associated with smoking in this population is 0.96/100/yr divided by 0/.07/1000/yr = 13.7. Therefore, smokers are ~14 times more likely to develop lung cancer than non-smokers.

In statistics, an **odds** of an event is the ratio of:

− The probability that the event WILL occur to the probability that the event will NOT occur

Example,
in 100 births, the probability of a delivery being a boy is 51% and being a girl is 49%
The odds of a delivery being a boy is 51/49 = 1.04

In simpler term, an odds of an event can be calculated as:
− Number of events divided by number of non-events

Let's borrow the concept of odds and apply it to disease and non-disease. So, the odds of having the disease is the ratio of the probability that the disease will occur to the probability that the disease will not occur. Or, the odds of having the disease can be calculated as the number of people with the disease divided by the number of people without the disease.
[Note: in the exposure-disease 2x2 table, the odds of having a disease in the exposed group is the same as the odds that an exposed person develops the disease].

Odds ratio is the ratio of the odds of disease in the exposed to the odds of disease in the non-exposed (in a *cohort study*).

Odds ratio is the ratio of the odds of disease in the exposed to the odds of disease in the non-exposed (in a *case-control study*).

Odds ratio can be calculated in a cohort study and in a case-control study
− The exposure odds ratio is equal to the disease odds ratio
Relative risk can only be calculated in a cohort study.

***Odds are another way of presenting a risk or probability (P). It is a ratio of probabilities -- the probability that an event will occur divided by the probability that the event will not occur. Mathematically, odds = P / 1-P. For most people, odds are less intuitive than percentages.

What are the odds of MI with experimental drug A?

$$\text{Odds of MI} = \frac{\text{Probability of MI}}{\text{Probability of no MI}} = \frac{0.10}{0.90} = 0.11$$

The odds of MI are 0.11 or "0.11 to 1" or "11:100" or "11 to 100." For people taking drug A, there are 11 people with MI for every 100 people without MI. Note that this calculation focuses on drug A alone and does not account for placebo.***

To develop approaches for disease prevention:

Attributable risk (AR)/fraction (AF) is a measure of excess risk that is attributed to the exposure.

- AR = the amount of disease incidence that can be attributed to a specific exposure
  - Difference in incidence of disease between exposed and non-exposed individuals
  - Incidence in non-exposed = background risk
  - Amount of risk that can be prevented

- AF = the proportion of disease incidence that can be attributed to a specific exposure (among those who were exposed)
  - AR divided by incidence in the exposed X 100%

In general, all individuals, whether they have or have not been exposed to a risk factor, have some chance of developing a disease if no prevention measures have been taken. The AR/AF estimates the risk above and beyond this baseline risk that all people have.
Example:
The attributable risk associated with smoking is 0.96/1000/yr – 0.7/1000/yr = 0.89/1000/yr. The incidence of lung cancer attributed to smoking is 0.89/1000/yr.

The proportion of lung cancer cases that could be prevented in this population of smokers if they would quit smoking is 0.89/1000/yr divided by 0.96/1000/yr = 92.7%.

Number needed to treat or ***Number needed to harm*** are ways of expressing the effectiveness and safety of an intervention in a way that is clinically meaningful. In general, NNT is always computed with respect to two treatments A and B, with A typically a drug and B a placebo. A defined endpoint has to be specified If the probabilities pA and pB of this endpoint under treatments A and B, respectively, are known, then the NNT is computed as 1/(pB-pA).

An NNT of 1 is the most effective and means each patient treated responds. An NNT of 2 or 3 indicates that a treatment is quite effective (with one patient in 2 or 3 responding to the treatment). An NNT of 20 to 40 can still be considered clinically effective.

The evidence-based journals (Evidence Based Medicine and ACP Journal Club) have achieved consensus on some terms they use to describe both the good and bad effects of therapy. We will bring them to life with a synthesis of three randomized trials in diabetes which individually showed that several years of intensive insulin therapy reduced the proportion of patients with worsening retinopathy to 13% from 38%, raised the proportion of patients with satisfactory hemoglobin A1c levels to 60% from about 30%, and increased the proportion of patients with at least one episode of symptomatic hypoglycemia to 57% from 23%. Note that in each case the first number constitutes the ***"experimental event rate" (EER)*** and the second number the ***"control event rate" (CER).*** We will use the following terms and calculations to describe these effects of treatment: when the experimental treatment reduces the probability of a bad outcome (worsening diabetic retinopathy).

RRR (relative risk reduction) = (EER – CER)/CER - the proportional reduction in rates of bad outcomes between experimental and control participants in a trial, accompanied by a 95% confidence interval (CI). In the case of worsening diabetic retinopathy, (EER – CER)/CER = (13% - 38%)/38% = 66%.

ARR (absolute risk reduction) = (EER – CER) - the absolute arithmetic difference in rates of bad outcomes between experimental and control participants in a trial, accompanied by a 95% CI. In this case, (EER – CER) = (13% - 38%) = 25%. (This is sometimes called the risk difference)

NNT (number needed to treat) = 1/ARR - the number of patients who need to be treated to achieve one additional favorable outcome, accompanied by a 95% CI. In this case, 1/ARR = 1/25% = 4.

When the experimental treatment increases the probability of a good outcome (satisfactory hemoglobin A1c levels):

RBI (relative benefit increase) = (EER – CER)/CER - the proportional increase in rates of good outcomes between experimental and control patients in a trial, accompanied by a 95% confidence interval (CI). In the case of satisfactory hemoglobin A1c levels, (EER – CER)/CER = (60% - 30%)/30% = 100%.

ABI (absolute benefit increase) = (EER – CER) **-** the absolute arithmetic difference in rates of good outcomes between experimental and control patients in a trial, accompanied by a

95% confidence interval (CI). In the case of satisfactory hemoglobin A1c levels, (EER – CER) = (60% - 30%) =30%

NNT (number needed to treat) = 1/ARR - the number of patients who need to be treated to achieve one additional good outcome, calculated as and accompanied by a 95% CI. In this case, 1/ARR = 1/30% = 3.

When the experimental treatment increase the probability of a bad outcome (episodes of hypoglycemia):

RRI (relative risk increase) = (EER – CER)/CER - the proportional increase in rates of bad outcomes between experimental and control patients in a trial, accompanied by a 95% confidence interval (CI). In the case of hypoglycemic episodes, (EER – CER)/CER = (57% - 23%)/23% = 148%. (RRI is also used in assessing the impact of "risk factors" for disease.)

ARI (absolute risk increase) = (EER – CER) - the absolute arithmetic difference in rates of bad outcomes between experimental and control patients in a trial, accompanied by a 95% confidence interval (CI). In the case of hypoglycemic episodes, (EER – CER) = |57% - 23%| = 34%. (ARI is also used in assessing the impact of "risk factors" for disease.)

NNH (number needed to harm) = 1/ARI

The number of patients who, if they received the experimental treatment, would result in one additional patient being harmed, compared with patients who received the control treatment, accompanied by a 95% CI. In this case, 1/ARI = 1/34% = 3.

References:

Hennekens CH, Buring JE. Epidemiology in Medicine. 1st ed. (Mayrent SL, ed.). Little, Brown and Co. 1987.

Dawson B, Trapp R. Basic & Clinical Biostatistics. 4th ed. McGraw-Hill Medical; 2004.