Fundamental Epidemiology Terminology and Measures: It Really Is All in the Name

Thomas R. Vetter, MD, MPH,* and Christine A. Jesser, ScD†

Epidemiology is the study of how disease is distributed in populations and the factors that influence or determine this distribution. Clinical epidemiology denotes the application of epidemiologic methods to questions relevant to patient care and provides a highly useful set of principles and methods for the design and conduct of quantitative clinical research. Validly analyzing, correctly reporting, and successfully interpreting the findings of a clinical research study often require an understanding of the epidemiologic terms and measures that describe the patterns of association between the exposure of interest (treatment or intervention) and a health outcome (disease). This statistical tutorial thus discusses selected fundamental epidemiologic concepts and terminology that are applicable to clinical research. Incidence is the occurrence of a health outcome during a specific time period. Prevalence is the existence of a health outcome during a specific time period. The relative risk can be defined as the probability of the outcome of interest (eg, developing the disease) among exposed individuals compared to the probability of the same event in nonexposed individuals. The odds ratio is a measure of risk that compares the frequency of exposure to a putative causal factor in the individuals with the health outcome (cases) versus those individuals without the health outcome (controls). Factors that are associated with both the exposure and the outcome of interest need to be considered to avoid bias in your estimate of risk. Because it takes into consideration the contribution of extraneous variables (confounders), the adjusted odds ratio provides a more valid estimation of the association between the exposure and the health outcome and thus is the preferably reported measure. The odds ratio closely approximates the risk ratio in a cohort study or a randomized controlled trial when the outcome of interest does not occur frequently (<10%). The editors, reviewers, authors, and readers of journal articles should be aware of and make the key distinction between the absolute risk reduction and the relative risk reduction. In assessing the findings of a clinical study, the investigators, reviewers, and readers must determine if the findings are not only statistically significant, but also clinically meaningful. Furthermore, in deciding on the merits of a new medication or other therapeutic intervention, the clinician must balance the benefits versus the adverse effects in individual patients. The number needed to treat and the number needed to harm can provide this needed additional insight and perspective. (Anesth Analg 2017;125:2146–51)

Oh, you want a definition. I hate definitions...

Benjamin Disraeli (1804–1881),
British politician, novelist, and essayist

Use the word ‘zeitgeist’ as often as possible. Ideally, you want to find words that sound familiar but people don’t really know their definitions: ‘zeitgeist,’ ‘bildungsroman,’ ‘doppelganger’ — better yet, anything Latin.

Stephen Colbert (b. 1964),
American comedian, television host, actor, and writer

Epidemiology is “the study of how disease is distributed in populations and the factors that influence or determine this distribution.”1 Clinical epidemiology “denotes the application of epidemiologic methods to questions relevant to patient care” and “provides a highly useful set of principles and methods for the design and conduct of quantitative clinical research.”2

The reader is referred to the 2 previous basic statistical tutorials in this series that discuss the practical application—including the advantages and limitations—of the array of available experimental, quasi-experimental, and descriptive and analytic observational study designs.3,4

Clinical epidemiologists and clinical researchers alike are often interested in examining the relationship between an exposure or therapeutic intervention and a specific health outcome. A health outcome can be broadly defined as a disease, medical condition, change in health status, or other event (including death).5

Validly analyzing, correctly reporting, and successfully interpreting the findings of a clinical research study often require an understanding of the epidemiologic terms and measures that describe the patterns of association between an exposure (eg, a treatment or intervention) and a health outcome (eg, a “disease”).

This basic statistical tutorial thus discusses the following selected fundamental epidemiologic concepts and terminology that are applicable to clinical research:
MEASURES OF FREQUENCY: INCIDENCE VERSUS PREVALENCE

Incidence is the occurrence of a health outcome during a specific time period (eg, an acutely elevated postoperative serum troponin level after noncardiac surgery).\(^6\) Incident cases are individuals who change in status from 1 state of the health outcome to another (eg, from a nondisease state to a disease state) over a specific period of time—for example, during the duration of the study-related observation and data collection\(^5,6\) or relative to a reference event such as birth, for example, incidence of breast cancer among women before 40 years of age.

The incidence is defined as the number of new events or cases of the health outcome occurring during a specified time period in a population of interest that is at risk for experiencing the health outcome\(^5,6\):

\[
\text{Incidence} = \frac{\text{incident cases}}{\text{total population at risk}}
\]

The incidence is synonymous with the event rate.\(^9\) Because incidence is a nondimensional proportion, it has no unit of measure and ranges from 0 to 1.0 (or 0% to 100%).\(^6,10\) In a randomized controlled trial (RCT), researchers often report and compare the respective incidence or event rate observed in each study group during the course of the study (eg, postoperative nausea and vomiting [PONV] during the first 24 hours after surgery).

In their analysis of data from the National Anesthesia Clinical Outcomes Registry for the period from 2010 to 2013, Nunnally et al\(^1\) reported that the incidence of cardiac arrest during intraprocedure and postanesthesia care was 5.6 per 10,000 cases or 0.056% (95% confidence interval [CI], 0.03%–0.13%), with an associated mortality rate from these cardiac arrests of 58.4% (95% CI, 54.6%–60.9%).

Prevalence is the existence of a health outcome during a specific time period (eg, preoperative pregnancy in pediatric patients presenting for outpatient surgery). Prevalent cases are all individuals with the health outcome of interest within a specified time frame, regardless of when they developed or were diagnosed with the health outcome.\(^5\)

The prevalence is the proportion or percentage of individuals in a population with a specific health outcome:\(^5,6\)

\[
\text{Prevalence} = \frac{\text{prevalent cases}}{\text{total population}}
\]

Because prevalence is a nondimensional proportion, it has no unit of measure and ranges from 0 to 1.0 (or 0% to 100%).\(^5,6\) In a RCT, researchers often report and compare the respective prevalence of a pertinent comorbidity observed in each study group before the study intervention (eg, preoperative coronary artery disease, congestive heart failure, diabetes mellitus, etc).

There are 2 types of prevalence. The point prevalence is the proportion or percentage of individuals in a population with an existing health outcome at a specific point in time (eg, coronary artery disease in noncardiac surgery patients as diagnosed preoperatively with cardiac stress testing and/or cardiac catheterization). The period prevalence represents how many individuals experienced the health outcome at any point during a time period or over usually an extended range of time (eg, any episode of acute low back pain during the previous 10 years).\(^5,7,10\)

In their cross-sectional study of anesthesiology residents, de Oliveira et al\(^12\) reported the point prevalence of high burnout to be 41% (95% CI, 38%–43%) and the point prevalence of depression to be 23% screened (95% CI, 19%–24%). The prevalence of suicide ideation among those residents who screened positive for depression was 23% (95% CI, 18%–28%).

MEASURES OF PROBABILITY: RISK AND THE RISK RATIO OR RR

Risk can be generally defined as the probability or likelihood of an untoward event.\(^8,14\) Risk factors are characteristics that are associated with an increased likelihood of experiencing a specific adverse health outcome. There is usually not a direct, one-to-one (A → B) relationship between a putative risk factor and a particular adverse health outcome (eg, disease).\(^11\) Epidemiologists instead classically refer to a “web of causation”—in which a disease often has multiple and relationally complex causes, some identified and others less or even not apparent.\(^14,15\)

Therefore, it should be noted that simply because a risk factor appears to predict an adverse health outcome, it does not assuredly follow that it causes this adverse outcome—hence, the adage or cliche of “association does not imply causation.”\(^14–17\) Authors should thus generally avoid the common temptation to refer in their articles to an observed statistically significant association as validly identifying a risk factor, but instead state the factor as being associated or correlated with the outcome.\(^16,18\)

The incidence of an adverse health outcome represents its absolute risk.\(^13,19\) While this measure of effect has utility, epidemiologists and clinical researchers are often more interested in comparative risk. Namely, “how many times more likely are exposed persons to become diseased, relative to nonexposed persons?”\(^19\) Conversely, “how many times less likely are treated patients to become diseased, relative to nontreated patients?”

The RR can be defined as the probability of the outcome occurring in exposed individuals compared to the probability of the same event in nonexposed individuals.\(^13,19,20\)

\[
\text{RR} = \frac{\text{risk in the exposed (treated)}}{\text{risk in the nonexposed (untreated)}}
\]

Furthermore, in a cohort study or RCT, the RR can be calculated directly from the data in a conventional 2 × 2 contingency table (Figure 1)\(^13,21\):
This is because at the start of a cohort study or RCT, all of the enrolled study subjects have yet to be treated (exposed) or to experience the health outcome and the study group sizes (number treated versus untreated) is known.

Interpreting the RR is straightforward but can be a bit counterintuitive:13

• If the RR = 1 → the risk in the exposed individuals equals the risk in the nonexposed individuals.

• In an RCT: If the RR = 1 → the risk in the treated patients equals the risk in the untreated patients.

• If the RR > 1 → the risk in the exposed individuals is greater than the risk in the nonexposed individuals, which represents a positive association between the exposure of interest and outcome of interest.

• In an RCT: If the RR > 1 → the risk of adverse outcome in the treated patients is greater than the risk of adverse outcome in the untreated patients, which represents a positive but clinically detrimental association.

• If the RR < 1 → the risk of adverse outcome in the exposed individuals is less than this risk in the nonexposed individuals, which represents a negative association between the exposure and outcome, possibly indicating a protective effect of the exposure.

• In an RCT: If the RR < 1 → the risk of adverse outcome in the treated patients is less than this risk in the untreated patients, which represents a negative statistical association but a clinically beneficial association.

A RR of 1 is referred to as the null value. A CI can and typically should be generated and reported for a sample-based point estimate of an RR.23

Last, the risk ratio should be distinguished from the rate ratio, which is (unfortunately) also abbreviated as the RR. The rate ratio is more accurately referred to as the incidence rate ratio because it is a measure of incidence rate in the exposed group versus the nonexposed group.10

In their Cochrane review, Guay et al24 observed that an epidural containing a local anesthetic, with or without the addition of an opioid, accelerates the return of the gastrointestinal transit and that an epidural containing a local anesthetic with an opioid decreases pain after an abdominal surgery. However, they did not find a statistically significant difference in the incidence of vomiting within 24 hours, with the RR of 0.84 (95% CI, 0.57–1.23), and did not find a statistically significant difference in the incidence of gastrointestinal anastomotic leak, with the RR of 0.74 (95% CI, 0.41–1.32). In the case of both of these estimates of the risk ratio, the 95% CI spans the null value of 1, indicating statistical nonsignificance with an α of .05.

MEASURES OF PROBABILITY: ODDS AND THE ODDS RATIO OR RELATIVE ODDS

Odds is defined as the ratio of 2 probabilities, specifically13,20,25,26:

\[
\text{Odds} = \frac{\text{probability of the event}}{1 - \text{probability of the event}}
\]

In a case-control study, investigators identify a subset of individuals with the outcome (eg, a disease) of interest—the cases, and a subset of individuals without the outcome of interest—the controls.4,27 Because these 2 groups were selected by the investigators based on certain specific criteria, there is no way to determine the incidence rate in the 2 groups vis-à-vis the underlying populations.21,27

In a case-control study, one instead can compare the frequency of exposure to a putative causal factor in the cases versus in the controls. This generates the OR, a measure of risk that is conceptually and mathematically similar to the RR (Figure 1).21,26,27 The OR can be simply calculated as the cross-products ratio of the data in a conventional 2 × 2 contingency table (Figure 1).13,21,22

Interpreting the OR is essentially the same as the RR13,22:
• If the OR = 1, the exposure is not related to the health outcome (disease).
• If the OR > 1, the exposure is positively related to the health outcome (disease).
• If the OR < 1, the exposure is negatively related to the health outcome (disease).

An OR of 1 is referred to as the null value. A CI can and typically should be generated and reported for a given sample-based point estimate of an OR.23

CRUDE (UNADJUSTED) OR VERSUS aOR

Once again, there is seldom a direct, one-to-one (A → B) relationship between a putative risk factor and a particular adverse health outcome.14 Instead the health outcome often has multiple and relationally complex causes, some identified and others less or even not apparent.14,15 These other contributing factors are referred to as “extraneous variables” or “covariates.”19 As discussed in a previous statistical tutorial in this series, these other contributing factors can function as “confounders,” which can compromise the validity of an observational study.28

The OR that is directly calculated as the cross-products ratio in a 2 × 2 contingency table (Figure 1) is referred to as the crude or unadjusted OR.21 Mathematically accounting for or adjusting for extraneous variables (covariates) generates an aOR.21 This adjustment is typically accomplished with multivariable analysis (eg, logistic regression for a dichotomous outcome or a Cox proportional hazards model for time to an event).19 These techniques will be the subject of future statistical tutorials.

Because the aOR takes into consideration the contribution of extraneous variables (covariates), it is innately a more valid estimation of the relationship (association) between the exposure and the adverse health outcome (disease) and thus is usually the preferably reported measure.

In their case-control study, Dunham et al29 reported that odds of stroke after cardiac surgery are increased in patients with a low minimum Pao2 within 24 hours of surgery, with an aOR of 1.23 (95% CI, 1.07–1.41) per 10 mm Hg lower nadir Pao2.

OR VERSUS RISK RATIO (RR)

In a case-control study, only the OR can be validly calculated and reported as a measure of association. In contrast, in a cohort study or an RCT, either the risk ratio or the OR can be validly calculated and reported as a measure of association.13

A common and important question is “when is the OR a good estimate of the RR?” Or alternatively stated “when does the OR closely approximate the RR?” Essentially, the OR closely approximates the RR in a cohort study or an RCT when the health outcome (disease) does not occur frequently (<10%).13,26,28 However, if a study outcome is common (>10%), the OR will be further from the null value of 1 than the corresponding RR—thus likely exaggerating the association and effect.31,32 As the event becomes more common, this exaggeration grows, and the OR is no longer a valid proxy for the RR.26

In their retrospective cohort study, Ramachandran et al33 reported that odds of postoperative intubation are increased in patients with a high Perioperative Sleep Apnea Prediction score, with an aOR of 2.3 (95% CI, 1.5–3.7). Their reporting of an OR rather than a risk ratio is valid given the observed, very low 0.3% overall incidence of postoperative respiratory complications.

ABSOLUTE RISK REDUCTION

As noted above, the incidence of an adverse health outcome (disease) represents its absolute risk.13,19 The absolute risk reduction is the simple arithmetic difference between the 2 event rates observed in the clinical study groups.9,20 Conversely, if applicable, the absolute risk increase can be calculated for the 2 event rates. As the observed event rates decrease, the absolute risk reduction also decreases (Figure 2).9

RR REDUCTION

The RR reduction is the difference in the event rates between 2 study groups, expressed as a proportion of the event rate in the untreated group.9,20 Conversely, if applicable, the RR increase can be calculated for the 2 event rates.

![Figure 2](image-url). Results of 4 hypothetical perioperative studies, each which report the ARR and/or the RRR. The NNT equals 100/ARR (as a percentage) or 1/ARR (as a proportion). ARR indicates absolute risk reduction; NNT, number needed to treat; RRR, relative risk reduction.)
The RR reduction represents the “efficacy” of a medication or other intervention. As the observed event rates decrease, the RR reduction generally remains constant as the efficacy of the intervention often remains constant.

**RR Reduction Versus Absolute Risk Reduction**

“RR reduction is often more impressive than absolute risk reduction. Furthermore, the lower the event rate in the control group, the larger the difference between RR reduction and absolute risk reduction.” The editors, reviewers, authors, and readers of journal articles should be aware of and make this key distinction.

**NUMBER NEEDED TO TREAT AND NUMBER NEEDED TO HARM**

In assessing the findings of a clinical study, the investigators, reviewers, and readers must determine if the findings are not only statistically, but also clinically significant. Furthermore, in deciding on the merits of a new medication or other therapeutic intervention, the clinician must balance the benefits versus the adverse effects in individual patients. The NNT and the NNH can provide this needed additional insight and perspective.

The NNT is calculated as 100 divided by the absolute risk reduction expressed as a percentage or 1 divided by the absolute risk reduction expressed as a proportion. The NNT represents the number of patients needed to be treated to prevent 1 more adverse event.

In their updated systematic review and meta-analysis of RCTs of prevention of PONV, De Oliveira et al reported that a 4–5-mg dose of dexamethasone reduced 24-hour PONV with an NNT of 3.7 (95% CI, 3.0–4.7) and an 8–10-mg dose of dexamethasone dose reduced 24-hour PONV with an NNT of 3.8 (95% CI, 3.0–4.3). Given that the NNT is the same, the more conservative dose of dexamethasone appears indicated.

The NNH is calculated as 100 divided by the absolute risk increase expressed as a percentage or 1 divided by the absolute risk increase expressed as a proportion. The NNH represents the number of patients needed to be treated to harm one more of them because of a medication-related side effect or another therapeutic complication.

In their systematic review and meta-analysis of outcomes of RCTs, Brandt et al reported that the incidence of cardiac arrest was lower in the prophylactic treatment group compared to the control group with a RR reduction of 0.41 (95% CI, 0.31–0.55) and an NNH of 128 (95% CI, 90–184). The NNH here provides further context in “weighing the value” of continuous lumbar plexus blockade. These authors specifically generated a Peto OR for such binary studies with rare events.

While seemingly straightforward and innately attractive—and thus increasingly reported for clinical studies—the NNT and NNH have potential limitations in certain settings (eg, with a time-to-event study design), which should be consistently acknowledged and addressed. Calculating a CI for a point estimate of the NNT or NNH can also be fraught with problems.

**CONCLUSIONS**

Clinical epidemiologists and clinical researchers share a typical interest in the relationship between an exposure or therapeutic intervention and a specific health outcome—whether it is a disease, medical condition, change in health status, or other event (including death). Research reports need to consistently use the correct terminology and measures of effect and association. Validly analyzing, correctly reporting, and successfully interpreting the findings of a clinical research study thus requires an understanding of the epidemiologic terms and measures that describe the patterns of association between a treatment or intervention (an “exposure”) and a health outcome (a “disease”).

**DISCLOSURES**

**REFERENCES**

4. Vetter TR. Magic mirror, on the wall—which is the right study design of them all? Part II. Anesth Analg. 2017;125:328–332.

Copyright © 2017 International Anesthesia Research Society. Unauthorized reproduction of this article is prohibited.
20. Scheckman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat—which of these should we use? Value Health. 2002;5:431–436.