An overview of clinical research: the lay of the land

David A Grimes, Kenneth F Schulz

Many clinicians report that they cannot read the medical literature critically. To address this difficulty, we provide a primer of clinical research for clinicians and researchers alike. Clinical research falls into two general categories: experimental and observational, based on whether the investigator assigns the exposures or not. Experimental trials can also be subdivided into two: randomised and non-randomised. Observational studies can be either analytical or descriptive. Analytical studies feature a comparison (control) group, whereas descriptive studies do not. Within analytical studies, cohort studies track people forward in time from exposure to outcome. By contrast, case-control studies work in reverse, tracing back from outcome to exposure. Cross-sectional studies are like a snapshot, which measures both exposure and outcome at one time point. Descriptive studies, such as case-series reports, do not have a comparison group. Thus, in this type of study, investigators cannot examine associations, a fact often forgotten or ignored. Measures of association, such as relative risk or odds ratio, are the preferred way of expressing results of dichotomous outcomes—e.g., sick versus healthy. Confidence intervals around these measures indicate the precision of these results. Measures of association with confidence intervals reveal the strength, direction, and a plausible range of an effect as well as the likelihood of chance occurrence. By contrast, p values address only chance. Testing null hypotheses at a p value of 0·05 has no basis in medicine and should be discouraged.

Clinicians today are in a bind. Increasing demands on their time are squeezing out opportunities to stay abreast of the literature, much less read it critically. Results of several studies indicate an inverse relation between knowledge of contemporary care and time since graduation from medical school.1 In many jurisdictions, attendance at a specified number of hours of continuing medical education courses is mandatory to maintain a licence to practise. However, the failure of these courses to improve patient care2 emphasises the importance of self-directed learning through reading. Many clinicians in practice, though, report that they feel unqualified to read the medical literature critically.3 Scientific illiteracy is a major failing of medical education.4

We have written this series of short essays on research methods for busy clinicians and active researchers. The needs of clinicians predominate; hopefully, this primer will produce more critical and thoughtful consumers of research, and thus better practitioners. The needs of clinicians overlap with those of researchers throughout the essays, but that overlap becomes most pronounced in the discussion of randomised controlled trials. For readers to assess randomised trials accurately, they should understand the relevant guidelines on the conduct of trials, emerging from methodological research. In presenting those discussions to clinicians, our essays will hopefully help researchers who do randomised trials as well.

We will cover descriptive studies, cohort studies, case-control studies, bias, and screening tests in separate articles, but will devote five articles to randomised controlled trials. This disproportion is intentional; randomised controlled trials are the gold standard in clinical research, and The Lancet publishes large numbers of them. Randomised controlled trials help to eliminate bias, and research has identified the important methodological elements of trials that minimise bias.5

Finally, because trials are so important, clinicians might be more likely to act on their results than on those of observational studies; hence, investigators need to ensure that trials are done and reported well. Here, we provide a brief overview of research designs and discuss some of the common measures used.

A taxonomy of clinical research

Analogous to biological taxonomy, a simple hierarchy can be used to categorise most studies (panel).6 To do so, however, the study design must be known. As in biology, anatomy dictates physiology. The anatomy of a study determines what it can and cannot do. A difficulty that readers encounter is that authors sometimes do not report

Rating clinical evidence

Assessment system of the US Preventive Services Task Force

Quality of evidence

I Evidence from at least one properly designed randomised controlled trial.
II-1 Evidence obtained from well-designed controlled trials without randomisation.
II-2 Evidence from well-designed cohort or case-control studies, preferably from more than one centre or research group.
II-3 Evidence from multiple time series with or without the intervention. Important results in uncontrolled experiments (such as the introduction of penicillin treatment in the 1940s) could also be considered as this type of evidence.
III Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.

Strength of recommendations

A Good evidence to support the intervention.
B Fair evidence to support the intervention.
C Insufficient evidence to recommend for or against the intervention, but recommendation might be made on other grounds.
D Fair evidence against the intervention.
E Good evidence against the intervention.

Lancet 2002; 359: 57-61

Family Health International, PO Box 13950, Research Triangle Park, NC 27709, USA (D A Grimes ☏, K F Schulz ☏)

Correspondence to: Dr David A Grimes (e-mail: dgrimes@fhi.org)
the study type or provide sufficient detail to figure it out. A related problem is that authors sometimes incorrectly label the type of research done. Examples include calling non-randomised controlled trials randomised, and labelling non-concurrent cohort studies case-control studies. The adjective case-controlled is also sometimes (inappropriately) applied to any study with a comparison group.

Biology has animal and plant kingdoms. Similarly, clinical research has two large kingdoms: experimental and observational research. Figure 1 shows that one can quickly decide the research kingdom by noting whether the investigators assigned the exposures—eg, treatments—or whether they observed usual clinical practice. For experimental studies, one needs to distinguish whether the exposures were assigned by a truly random technique (with concealment of the upcoming assignment from those involved) or whether some other allocation scheme was used, such as alternate assignment. An example of the latter would be a trial alternating months of liberal versus restricted access to electronic fetal monitoring for women in labour.

With observational studies, which dominate the literature, the next step is to ascertain whether the study has a comparison or control group. If so, the study is termed analytical. If not, it is a descriptive study (figure 1). If the study is analytical, the temporal direction of the trial needs to be identified. If the study determines both exposures and outcomes at one time point, it is termed cross-sectional. An example would be measurement of serum cholesterol of men admitted to a hospital with myocardial infarction versus that of their next-door neighbour. This type of study provides a snapshot of the population of sick and well at one time point.

If the study begins with an exposure—eg, oral contraceptive use—and follows women for a few years to measure outcomes—eg, ovarian cancer—then it is deemed a cohort study. Cohort studies can be either concurrent or non-concurrent. By contrast, if the analytical study begins with an outcome—eg, ovarian cancer—and looks back in time for an exposure, such as use of oral contraceptives, then the study is a case-control study.

Studies without comparison groups are called descriptive studies. At the bottom of the research hierarchy is the case report. When more than one patient is described, it becomes a case-series report.

**What studies can and cannot do**

Is the study design appropriate for the question?

Starting at the bottom of the research hierarchy, descriptive studies are often the first foray into a new area of medicine. Investigators do descriptive studies to describe the frequency, natural history, and possible determinants of a condition. The results of these studies show how many people develop a disease or condition over time, describe the characteristics of the disease and those affected, and generate hypotheses about the cause of the disease. These hypotheses can be assessed through more rigorous research, such as analytical studies or randomised controlled trials. An example of a descriptive study would be the early reports of Legionnaire's disease and toxic-shock syndrome. An important caveat (often forgotten or intentionally ignored) is that descriptive studies, which do not have a comparison group, do not allow assessment of associations. Only comparative studies (both analytical and experimental) enable assessment of possible causal associations.

**Cross-sectional study: a snapshot in time**

Sometimes termed a frequency survey or a prevalence study, cross-sectional studies are done to examine the presence or absence of disease and the presence or absence of an exposure at a particular time. Thus, prevalence, not incidence, is the focus. Since both outcome and exposure are ascertained at the same time (figure 2), the temporal relation between the two might be unclear. For example, assume that a cross-sectional study finds obesity to be more common among women with than without arthritis. Did the extra weight load on joints lead to arthritis, or did women with arthritis become involuntarily inactive and then obese? This type of question is unanswerable in a cross-sectional study.

**Cohort study: looking forward in time**

Cohort studies proceed in a logical sequence: from exposure to outcome (figure 2). Hence, this type of research is easier to understand than case-control studies. Investigators identify a group with an exposure of interest and another group or groups without the exposure. The investigators then follow the exposed and unexposed groups forward in time to determine outcomes. If the exposed group develops a higher incidence of the outcome than the unexposed, then the exposure is associated with an increased risk of the outcome.

The cohort study has important strengths and weaknesses. Because exposure is identified at the outset, one can assume that the exposure preceded the outcome. Recall bias is less of a concern than in the case-control study. The cohort study enables calculation of true incidence rates, relative risks, and attributable risks. However, for the study of rare events or events that take years to develop, this type of research design can be slow to yield results and thus prohibitively expensive. Nonetheless, several famous, large cohort studies continue to provide important information.

---

**Figure 1: Algorithm for classification of types of clinical research**

Did investigator assign exposures?  

- Yes: Experimental study  
  - Random allocation?  
    - Yes: Randomised controlled trial  
    - No: Non-randomised controlled trial  

- No: Observational study  
  - Comparison group?  
    - Yes: Analytical study  
    - No: Descriptive study  

Exposure→Outcome  
Exposure and outcome at the same time  
Exposure→Outcome

---

**Figure 2: Algorithm for classification of types of clinical research**

Did investigator assign exposures?  

- Yes: Experimental study  
  - Random allocation?  
    - Yes: Randomised controlled trial  
    - No: Non-randomised controlled trial  

- No: Observational study  
  - Comparison group?  
    - Yes: Analytical study  
    - No: Descriptive study  

Exposure→Outcome  
Exposure and outcome at the same time  
Exposure→Outcome
Outbreaks of food-borne diseases are a prototype for study designs. The randomised controlled trial is the only known way to avoid selection and confounding biases in clinical research. This design approximates the controlled experiment of basic science. It resembles the cohort study in several respects, with the important exception of randomisation of participants to exposures (figure 2).

The hallmark of randomised controlled trials is assignment of participants to exposures purely by the play of chance. Randomised controlled trials reduce the likelihood of bias in determination of outcomes. When properly implemented, random allocation precludes selection bias. Trials feature uniform diagnostic criteria for outcomes and, often, blinding of those involved as to the exposure each participant is receiving, therefore reducing information bias. A unique strength of this study design is that it eliminates confounding bias, both known and unknown. Furthermore, the trial tends to be statistically efficient. If properly designed and done, a randomised controlled trial is likely to be free of bias and is thus especially useful for examination of small or moderate effects. In observational studies, bias might easily account for small to moderate differences.13

Randomised controlled trials have drawbacks too. External validity is one. Whereas the randomised controlled trial, if properly done, has internal validity—ie, it measures what it sets out to measure—it might not have external validity. This term indicates the extent to which results can be generalised to the broader community. Unlike the observational study, the randomised controlled trial includes only volunteers who pass through a screening process before inclusion. Those who volunteer for trials tend to differ from those who do not; for example, their health might be better.14 Another limitation is that a randomised controlled trial cannot be used in some instances, since intentional exposure to harmful substances—eg, toxins, bacteria, or other noxious exposures—would be unethical. As with cohort studies, the randomised controlled trial can be prohibitively expensive. Indeed, the cost of large trials runs into the tens of millions of US dollars.

**Measurement of outcomes**

**Confusing fractions**

Identification and quantification of outcomes is the business of research. However, slippery terminology often complicates matters for investigators and readers alike. For example, the term rate (as in maternal mortality rate) has been misused in textbooks and journal articles for decades. Additionally, rate is often used interchangeably with proportions and ratios.14 Figure 3 presents a simple approach to classification of these common terms.
Measures

Relative risk (also termed the risk ratio) is another venerable misnomer is actually a ratio, not a rate, a fact

relatedness of numerator and denominator—divides ratios into two groups: those in which the numerator is included in the denominator—eg, rate and proportion—and those in which it is not.

A rate measures the frequency of an event in a population. As shown in figure 3, the numerator (those with the outcome) of a rate must be contained in the denominator (those at risk of the outcome). Although all ratios feature a numerator and denominator, rates have two distinguishing characteristics: time and a multiplier.

Proportion is often used synonymously with rate, but the former does not have a time component. Like a rate, a proportion must have the numerator contained in the denominator. Since the numerator and denominator have the same units, these divide out, leaving a dimensionless quantity; a number without units. An example of a proportion is prevalence—eg, 27 of 100 at risk have hay fever. This number indicates how many of a population at risk have a condition at a particular time—eg, 11 cases of tuberculosis per 100 000 persons per year.

Proportion is often used synonymously with rate, but the former does not have a time component. Like a rate, a proportion must have the numerator contained in the denominator. Since the numerator and denominator have the same units, these divide out, leaving a dimensionless quantity; a number without units. An example of a proportion is prevalence—eg, 27 of 100 at risk have hay fever. This number indicates how many of a population at risk have a condition at a particular time—eg, 11 cases of tuberculosis per 100 000 persons per year.

Proportion is often used synonymously with rate, but the former does not have a time component. Like a rate, a proportion must have the numerator contained in the denominator. Since the numerator and denominator have the same units, these divide out, leaving a dimensionless quantity; a number without units. An example of a proportion is prevalence—eg, 27 of 100 at risk have hay fever. This number indicates how many of a population at risk have a condition at a particular time—eg, 11 cases of tuberculosis per 100 000 persons per year.

Proportion is often used synonymously with rate, but the former does not have a time component. Like a rate, a proportion must have the numerator contained in the denominator. Since the numerator and denominator have the same units, these divide out, leaving a dimensionless quantity; a number without units. An example of a proportion is prevalence—eg, 27 of 100 at risk have hay fever. This number indicates how many of a population at risk have a condition at a particular time—eg, 11 cases of tuberculosis per 100 000 persons per year.

Proportion is often used synonymously with rate, but the former does not have a time component. Like a rate, a proportion must have the numerator contained in the denominator. Since the numerator and denominator have the same units, these divide out, leaving a dimensionless quantity; a number without units. An example of a proportion is prevalence—eg, 27 of 100 at risk have hay fever. This number indicates how many of a population at risk have a condition at a particular time—eg, 11 cases of tuberculosis per 100 000 persons per year.


21 Fetal EF, Rosenbush EJ, Lee MJ, Del Priore G. Distribution of study designs in four major US journals of obstetrics and gynecology. Gynecol Obstet Invest 2001; 51: 8–11.


