

The Hierarchy of Evidence Pyramid



Available from: https://www.researchgate.net/figure/Hierarchy-of-evidence-pyramid-The-pyramidal-shape-qualitatively-integrates-the-amount-of_fig1_311504831 [accessed 12 Dec, 2019]

Hierarchy of Evidence: From Case Reports to Randomized Controlled Trials

*Brian Brighton, MD**; *Mohit Bhandari, MD, MSc***;
Paul Tornetta, III, MD†; and *David T. Felson, MD**

In the hierarchy of research designs, the results of randomized controlled trials are considered the highest level of evidence. Randomization is the only method for controlling for known and unknown prognostic factors between two comparison groups. Lack of randomization predisposes a study to potentially important imbalances in baseline characteristics between two study groups. There is a hierarchy of evidence, with randomized controlled trials at the top, controlled observational studies in the middle, and uncontrolled studies and opinion at the bottom.

This hierarchy has not been supported in two recent publications in the *New England Journal of Medicine* which identified nonsignificant differences in results between randomized, controlled trials, and observational studies. The current authors provide an approach to organizing published research on the basis of study design, a hierarchy of evidence, a set of principles and tools that help clinicians distinguish ignorance of evidence from real scientific uncertainty, distinguish evidence from unsubstantiated opinions, and ultimately provide better patient care.

From the *Boston University School of Medicine and School of Public Health, Boston University, Boston, MA; the **Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada; and the †Department of Orthopaedic Surgery, Boston University, Boston, MA.

Reprint requests to Mohit Bhandari, MD, McMaster University Health Sciences Center, Department of Clinical Epidemiology and Biostatistics, Health Science Centre, Room 2C3, 1200 Main Street West, Hamilton, Ontario, L8N 3Z5 Canada. Phone: 905-525-9140 ext. 22825; Fax: 905-524-3841; E-mail: bhandari@sympatico.ca.
DOI: 10.1097/01.blo.0000079323.41006.12

In the hierarchy of research designs, the results of randomized controlled trials are considered the highest level of evidence.²⁶ Randomization is the only method for controlling for known and unknown prognostic factors between two comparison groups.^{1,26} Lack of randomization predisposes a study to potentially important imbalances in baseline characteristics between two study groups. The role of nonrandomized (observational) studies in evaluating treatments is an area of continued debate: deliberate choice of the treatment for each patient

implies that observed outcomes may be caused by differences among people being given the two treatments, rather than the treatments alone. Unrecognized confounding factors can interfere with attempts to correct for identified differences between groups. There has been considerable debate about whether the results of nonrandomized studies are consistent with the results of randomized controlled trials.^{3, 10, 12–14, 21, 22, 28} Nonrandomized studies, or observational studies, have been reported to overestimate or underestimate treatment effects.^{21, 22}

These considerations have supported a hierarchy of evidence, with randomized controlled trials at the top, controlled observational studies in the middle, and uncontrolled studies and opinion at the bottom. However, these findings have not been supported in two recent publications in the *New England Journal of Medicine* that identified nonsignificant differences in results between randomized, controlled trials and observational studies.^{3, 13}

The current authors provide an approach to organizing published research on the basis of study design, a hierarchy of evidence. The key features and the advantages and disadvantages of specific study designs will be addressed. The concepts presented hopefully will enable clinicians and healthcare personnel to practice in the context of evidence-based orthopaedics.

Types of Study Design

The types of study designs used in clinical research can be classified broadly according to whether the study focuses on describing the distributions or characteristics of a disease or elucidating its determinants.

Descriptive studies are concerned with describing the general characteristics of the distribution of a disease, particularly in relation to person, place, and time. Cross-sectional studies, case reports, and case series represent types of descriptive studies. Information on each can provide clues leading to the generation of a hypothesis that is consistent with existing knowledge of disease occurrence.

Analytic studies focus on determinants of a disease by testing a hypothesis with the ulti-

mate goal of judging whether a particular exposure causes or prevents disease. Analytic design strategies include observational studies, such as case-control and cohort studies, and clinical trials. The difference between the two types of analytic studies is the role that the investigator plays in each of the studies.²⁰ In the observational study, the investigator simply observes the natural course of events. In the randomized controlled trial, the investigator assigns the intervention or treatment. Although an in-depth discussion on assessing the methodologic quality of each particular study is beyond the scope of the current study, the basic strengths and limitations of each research strategy will be addressed.

Levels of Evidence

Investigators have attempted to minimize potentially harming patients by basing clinical decisions of the sorts of evidence that are least likely to be wrong. Two studies defined what was thought to be the evidence providing the least biased estimate of the effect of an intervention: a systematic review documenting homogeneity in results of a large number of high-quality randomized controlled trials (randomized with concealment, blinded, complete followup, and intention-to-treat analysis).^{8, 25} This was termed Level 1 Evidence. These investigators additionally categorized studies of an intervention based on an increasing degree of potential bias: systematic reviews with randomized controlled trials that reveal differences in treatment effect (heterogeneity); individual high-quality randomized controlled trials (Level IB evidence); less rigorous randomized controlled trials; cohort or observational studies (Level 2 evidence); case-control studies (Level 3 evidence); case series (Level 4 evidence), and expert opinion (Level 5 evidence).²³ Based on the various levels of evidence of a particular treatment, grades of recommendation can be determined.^{25, 26} For example, the following grades of recommendations have been proposed: (1) Grade A, consistent Level 1 studies; (2) Grade B, consistent Level 2 or Level 3 studies; (3)

Grade C, Level 4 studies; and (4) Grade D, Level 5 studies.^{25,26}

Case Report and Case Series

Case reports are an uncontrolled, descriptive study design involving an intervention and outcome with a detailed profile of one patient. An early example from the orthopaedic literature of a case report is Birkett's 1869 description of a fracture-dislocation of the hip.⁷ Expansion of the individual case report to include multiple patients with an outcome of interest is a case series. In 1981, a famous case series involving five homosexual men in Los Angeles, CA with *Pneumocystis carinii* between 1980 and 1981 marked the beginning of the AIDS epidemic in the United States.⁹ Although descriptive studies are limited in their design to make causal inferences about the relationship between risk factors and an outcome of interest, they are helpful in developing a hypothesis that can be tested using an analytic study design.

Case-Control Study

One type of observational study is the case-control study that starts with the identification of individuals who already have the outcome of interest, cases, and are compared with a suitable control group without the outcome event. The relationship between a particular intervention or prognostic factor and the outcome of interest is examined by comparing the number of individuals with each intervention or prognostic factor in the cases and controls. Case-control studies can be used to study prognostic factors. One example of a case-control study to investigate prognostic factors would be to identify patients with nonhealing fractures and a similar group of patients with well-healed fractures to see whether the patients with nonhealing fractures were more likely to smoke.

Case-control studies are described as retrospective because they are done looking back in time at information collected about past exposure to possible attributable factors. From this information, an odds ratio can be calculated to describe the odds of a particular factor in individuals with the outcome of interest compared

with those without the outcome. Although the odds ratio can be used to estimate the relative risk, it cannot provide information about excess risk. The case-control study can be useful in studying rare outcomes, outcomes with multiple potential etiologic factors, or looking at outcomes that take a considerable length of time to develop. Additionally, they can be done in a short time with small sample sizes, and for less money than other types of studies. However, because the information usually is collected from patients or their hospital records, data may be inaccurate because of the effect of recall bias and measurement bias.

Cohort Study

The term cohort comes from the Roman word for a group of soldiers that marched into battle together.¹⁵ In the cohort study design, the cohort represents a group of people followed up with time to see whether an outcome of interest develops. Ideally this group meets a level of certain predetermined criteria representative of a population of interest and is followed with well-defined outcome variables. The Framingham Heart Study² is an example of a large cohort study involving residents of a Massachusetts community with identifiable cardiovascular risk factors being followed up for cardiovascular events. Usually this group is matched with a control population selected on the presence or absence of exposure to a factor of interest. The purpose of this type of study is to describe the occurrence of certain outcomes with time and to analyze associations between prognostic factors and those outcomes.

Cohort studies can be prospective in nature meaning that they begin at a specified point and are followed forward in time to evaluate the influence of certain prognostic factors or interventions on the desired outcomes. Examples include prospective cohort studies such as one evaluating refractures in patients initially treated for a fracture.²⁴ The strengths of a prospective cohort study are the ability for the investigator to study several outcomes with time, and ensure that the data collected are relevant and accurate. The drawbacks are the ex-

pense of involving a large number of subjects and requirement of a long study period.

A retrospective cohort or historic cohort involves identifying patients from past records and following this group backward in time from the present to the past records. Retrospective cohort studies have the advantage of being shorter in duration compared with prospective studies but they lack the ability to control the selection of subjects and lack the control over outcome measurements. Cohort studies are observational in nature and subject to systematic bias. Confounding variables may be introduced because random allocation is not used thereby affecting the outcome rather than the factors being examined.

Randomized Controlled Trials

Randomized controlled trials classically are held as the standard to which all other designs should be measured.^{1,26} In a randomized controlled trial, subjects are assigned to a treatment group or a control group. The control group usually receives an accepted treatment or no treatment at all, whereas the treatment group is assigned the intervention of interest. Randomized controlled trials are thought to represent the highest quality of evidence based on their methodologic strengths of randomization of patient assignment and blinding of intervention and outcome. Studies are randomized to eliminate selection bias and to balance confounding factors between both groups. Blinding of the subjects, investigators, or both (double-blinded) involves concealing patient assignment so as not to influence the outcome. Controlled trials without randomization occasionally are done but represent a class of evidence with less internal validity and are subject to selection bias. The importance of randomization comes from the neurosurgical literature.^{1, 6} During the 1970s and early 1980s, surgeons frequently did extracranial to intracranial bypass (anastomosis of a branch of the external carotid artery, the superficial temporal, to a branch of the internal carotid artery, the middle cerebral). They thought it prevented strokes in patients whose sympto-

matic cerebrovascular disease otherwise was surgically inaccessible. Comparisons of outcomes among nonrandomized cohorts of patients who, for various reasons, did or did not have this operation, fueled their conviction. These studies suggested that patients who had surgery seemed to fare much better.^{1,6} However, to the surgeons' surprise, a large multicenter trial in which patients were allocated to surgical or medical treatment using a process analogous to flipping a coin (a randomized control trial), showed that the only effect of surgery was to increase adverse outcomes in the immediate postsurgical period.¹⁹

The advantages of a randomized controlled trial are the quality of the study associated with its inherent internal validity because potential confounding variables can be controlled for, thereby potentially providing strong evidence for cause and effect relationships. Randomized controlled trials may not always be suitable for answering some research questions for technical or ethical reasons.

References

1. American Medical Association: User's Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. In Guyatt GH, Rennie D (eds). Ed 2. Chicago, American Medical Association Press 2001.
2. Anderson KM, Castelli WP, Levy D: Cholesterol and mortality: 30 years of follow-up from the Framingham Study. *JAMA* 257:2176–2180, 1987.
3. Benson K, Hartz AJ: A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 342:1878–1886, 2000.
4. Bhandari M, Guyatt GH, Montori V, Devereaux PJ, Swiontkowski MF: User's guide to the orthopaedic literature: How to use a systematic literature review. *J Bone Joint Surg* 84A:1672–1682, 2002.
5. Bhandari M, Guyatt GH, Swiontkowski MF: User's guide to the orthopaedic literature: How to use an article about a prognosis. *J Bone Joint Surg* 83A:1555–1564, 2001.
6. Bhandari M, Guyatt GH, Swiontkowski MF: User's guide to the orthopaedic literature: How to use an article about a surgical therapy. *J Bone Joint Surg* 83A:916–926, 2001.
7. Birkett J: The Classic: Description of a dislocation of the head of the femur, complicated with its fracture. *Clin Orthop* 377:4–6, 2000. (Reprinted from Birkett J: Description of a dislocation of the head of the femur, complicated with its fracture. *Med Circ Trans* 52:133–138, 1869.)
8. Canadian Task Force on the Periodic Health Examination: The periodic health examination. *CMAJ* 121:1193–1254, 1979.
9. Centers for Disease Control: Pneumocystis pneumonia: Los Angeles. *MMWR Morbid Mortal Weekly Rep* 30:250–252, 1981.
10. Chalmers TC, Celano P, Sacks HS, Smith Jr H: Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 309:1358–1361, 1983.
11. Cochrane Collaboration: <http://www.cochrane.org>.
12. Colditz GA, Miller JN, Mosteller F: How study design affects outcomes in comparisons of therapy. *Med Stat Med* 8:441–454, 1989.
13. Concato J, Shah N, Horwitz RI: Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 342:1887–1894, 2000.
14. Emerson JD, Burdick E, Hoaglin DC, et al: An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Controlled Clin Trials* 11:339–352, 1990.
15. Guralnik DB (ed): Webster's New World Dictionary of the American Language. New York, Prentice Hall Press 1986.
16. Guyatt GH: Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA* 268:2420–2425, 1992.
17. Guyatt GH, Haynes RB, Jaeschke R: User's guide to the medical literature: XXV: Evidence-based medicine: Principles for applying the user's guide to patient care. *JAMA* 284:1290–1296, 2000.
18. Guyatt GH, Sackett DL, Sinclair JC: User's guide to the medical literature: IX: A method for grading health care recommendations. *JAMA* 274:1800–1804, 1995.
19. Haynes RB, Mukherjee J, Sackett DL, et al: Functional status changes following medical or surgical treatment of cerebral ischemia: Results in the EC/IC Bypass Study. *JAMA* 257:2043–2046, 1987.
20. Hennekens CH, Buring JE: *Epidemiology in Medicine*. Ed 1. Boston, Little, Brown, and Company 1987.
21. Ioannidis JP, Haidich AB, Pappa M, et al: Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 286:821–830, 2001.
22. Kunz R, Oxman AD: The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 317:1185–1190, 1998.
23. Oxford Guidelines: <http://minerva.minervation.com/cebmdocs/levels.html>.
24. Robinson CM, Royds M, Abraham A, et al: Refractures in patients at least 45 years old: A prospective analysis of twenty-two thousand and sixty patients. *J Bone Joint Surg* 84A:1528–1533, 2002.
25. Sackett DL: Rules of evidence and clinical recommendations on use of antithrombotic agents. *Chest* 89 (2 Suppl):2S-3S, 1986.

26. Sackett DL, Haynes RB, Guyatt GH, Tugwell P: *Clinical Epidemiology: A Basic Science for Clinical Medicine*. Ed 2. Boston, Little, Brown, and Company 1991.
27. Sackett DL, Rosenberg WM, Gray JA: Evidence-based medicine: What it is and what it isn't. *BMJ* 312:71-72, 1996.
28. Sacks HS, Chalmers TC, Smith Jr H: Sensitivity and specificity of clinical trials: Randomized versus historical controls. *Arch Intern Med* 143:753-755, 1983.