Clin Orthop Relat Res (2011) 469:909–913
DOI 10.1007/s11999-010-1504-4

IN BRIEF

In Brief

Statistics in Brief

Study Designs in Orthopaedic Clinical Research

Brigitte M. Jolles MD, MSc, Estelle Martin PhD

Published online: 5 August 2010

Question

What is the best study design for your clinical research question and why?

Discussion

The validity of a clinical study depends on its ability to generate unbiased data, thereby improving the current scientific knowledge, and to generate results that would be applicable to the majority of patients. To achieve these goals, rigorous scientific methods should be used in clinical research.

When designing clinical research to address a specific question, the first step is to establish a detailed study design. The best and most practical study design should be used in all cases and should minimize bias. Levels of evidence describe the relative risks of bias for each design [2–5], with Level I having the least risk and Level V having the greatest risk. However, these levels are frequently low in orthopaedic clinical research because of specific difficulties related to this specialty and surgery in general. We describe the most commonly used study designs in orthopaedic surgery clinical research and show the best ways to use them by providing practical examples.

B. M. Jolles
Department of Orthopaedic Surgery and Traumatology, Centre Hospitalier Universitaire Vaudois, University of Lausanne, Lausanne, Switzerland

E. Martin
Department of Geriatric Medicine, Centre Hospitalier Universitaire Vaudois, University of Lausanne, Lausanne, Switzerland

B. M. Jolles (✉)
CHUV Site Hôpital Orthopédique, 4, Avenue Pierre Decker, 1011 Lausanne, Switzerland
e-mail: brigitte.jolles-haeberli@chuv.ch

Study Designs

When designing clinical research, the first step is to define whether one wishes to describe or observe events or to study a treatment or diagnostic or prognostic tool and examine the events afterward (Table 1). This decision divides study designs into two distinct categories: descriptive studies and analytic studies (Fig. 1).

Descriptive Studies

Descriptive studies are those in which the researcher merely describes a situation or some events. They offer no explanations about the events or the type of links (ie, causal or noncausal) between those events and potential risk factors. However, they can give rise to hypotheses that could be confirmed or refuted through additional studies. Descriptive studies include cross-sectional studies, correlational studies, case series, and case reports. A cross-sectional study shows the incidence or prevalence of an event in a specified population. An example would be the study of the incidence of residual pain after ankle arthroplasty. A correlational study examines potential relations between two variables. An example would be the study of osteoporosis prevalence and wrist fracture in men older than 65 years. Case series typically provide a detailed description of patients, typically more than 10. When

Table 1. Advantages and disadvantages of descriptive and analytic studies

General purpose of study	Types of study designs	Specific purpose of study	Advantages	Disadvantages
Observe events: retrospective descriptive studies	Cross-sectional study	Determine the prevalence of an event	Low cost and time	Cannot explain the event
	Correlational study	Determine potential relations between two variables	Low cost and time	Cannot explain the type of relation between the two variables (time relationship, presence of potential confounding factors)
	Case series	Describe rare events (diseases, complications); generate new hypotheses (diagnostic methods, treatments, association of diseases)	Low cost and time	No control group Selection bias No idea of the cases frequency
Determine effects of treatment or diagnostic or prognostic tools: analytic studies	Retrospective case-control study (control group)		Takes into account known patient confounding factors if matched-paired or crossover Low cost and time (rare or long latency diseases) No risk to subjects	Selection bias
	Prospective (longitudinal) cohort study		Outcomes measured after exposures True incidence rates Avoids recall bias	No control group Selection bias Cost and time (rare or long latency diseases)
	Prospective nonrandomized controlled trial		Control group Avoids experimental bias and placebo effect	Selection bias Cost and time Not always possible in surgery for ethical reasons
	Prospective randomized controlled trial		Takes into account known and unknown patient confounding factors Avoids treatment, attrition, assessment bias if blinded	Cost and time Not always possible in surgery for ethical reasons

describing rare diseases or rare secondary effects or complications of treatments, case reports might be used; these reports typically offer detailed descriptions of fewer than 10 patients. Findings from these latter two designs can help generate new questions or hypotheses related to diagnostic methods, potential risks, treatments, or associations of diseases. They then would help the researcher justify a future investigation with another study design and more patients.

In case-series studies one cannot determine whether the cases are rare or frequent (ie, their incidence). The group of cases usually is selected and subject to bias as even rare events may appear simultaneously without necessarily being connected. To explore the incidence and the cause-effect relationship while limiting as much bias as possible, one needs analytic studies.

Analytic Studies

Analytic studies are designed to answer a scientific hypothesis while minimizing bias as much as possible. Bias can be minimized by using a control group or/and prospective enrollment or/and randomization of patients or/and blinding. Analytic studies include retrospective case-control studies where a control-group is added and prospective studies. Prospective studies may include no control group (ie, longitudinal cohort studies) or one or more groups (controlled trials) (Fig. 1).

A case-control study is a retrospective case series to which a control group is added. As the number of patients is increased there is a lesser chance of observing a random result and therefore is prone to less bias than a study without such a control group. For example, a group of patients who experienced one episode of dislocation of their THA prosthesis could be compared with a group of patients in whom the hip prosthesis did not dislocate and who had their surgery performed during the same period to determine factors related to the presence of dislocation. Without controls, the real effect of potential factors could not be identified conclusively because they could be present in subjects who did not experience dislocation.

A matched-paired case-control study further reduces the risk of bias of a retrospective case-control study design by matching each patient with a control patient by presumed confounding factors. Confounding factors, such as age, gender, and comorbidities, have the potential to influence outcome if they do not occur in the same proportions in both groups. The best control for the patient would be him- or herself. If the patient can be studied with two distinct treatments at different times, the study then is called a crossover case-control study. This design would be possible only if the baseline status of the patients can be measured before observing them in a second situation. However, even if some bias can be avoided, case-control studies remain retrospective designs and, as such, may give rise to time-related biases. For example, (1) with time the patient could have an imperfect memory of the symptoms

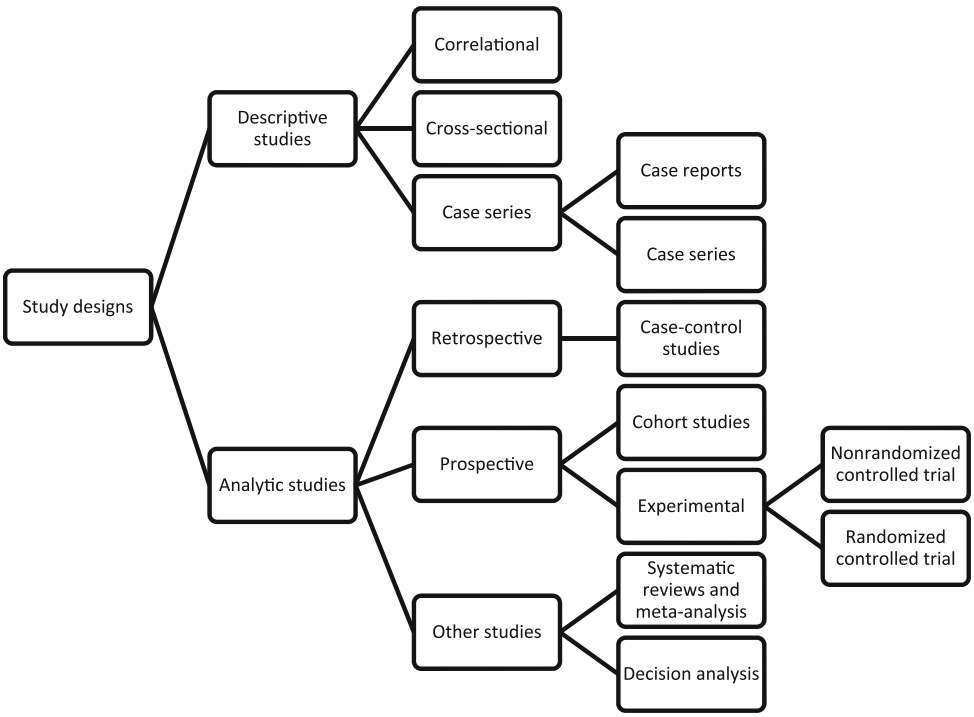

Fig. 1 The available study designs and categories are shown.

before and after the event or could interpret the event with time therefore giving more or less emphasis on one or the other symptoms he/she had before or after the event; (2) with time the treatments and/or assessments and/or record keeping might have been more or less subtly changed.

A longitudinal cohort study is a prospective study in which one group of patients is followed longitudinally while the baseline parameters and their evolution are recorded. The measurement tools are chosen before the patients are included in the study. For example, one could study quality of life and ROM of the knee in patients before and after TKA using the WOMAC questionnaire and an ambulatory gait analysis device. However, in this prospective design, there is still a potential selection bias: investigators can choose which patients to enroll unless all patients with a given diagnosis are included (ie, a consecutive series). Even if all patients are included, there still can be other forms of bias such as referral bias (in which the patients are limited to those seeking care at a given institution or institutions) or diagnostic bias (in which specific criteria are required for diagnosis but potentially excluded other patients).

A triple randomized clinical trial (RCT) of sufficient sample size, where the randomization assignment has been followed, is considered the gold standard design for clinical research because randomization remains the only way to minimize selection bias: the control of all potential confounding factors is enhanced and, given adequate power, even unknown factors will be distributed more equally between the two groups of patients under study. For example, this avoids selection of younger and healthier patients for a new type of implant, or more compliant patients for the treatment with possible increased side effects. The randomization should remain coded and not determined by a simple method such as an alternative process (ie, ABABA). Randomization may be performed by blocks and stratification to improve the allocation procedure: it then achieves an approximate balance of important characteristics especially in small studies (age distribution for example) [1].

To avoid systematic subjectivity and treatment bias, blinding to knowledge of the treatment should be added when possible. If a patient, surgeon, or observer knows that a specific treatment is available, he or she probably will act differently. Patients might have better care from the surgeons' preferred treatment modality. Surgeons could make different decisions with respect to whether to stop the study for a patient because of his or her generally poor medical condition when knowing the treatment group (attrition bias). Researchers might evaluate the patients differently (radiographic measurement, side effects, etc) if they think one treatment is not as effective as another (assessment bias). Blinding therefore is particularly important if possible, and would best apply to patients, surgeons, observers, and/or statisticians. A study can be single-, double-, or triple-blind depending on the people involved. A single-blind study would have only the patient, or the surgeon, or the researcher blinded; that is, one of these three categories of people would be blinded, while the other two would not be, often because it simply is not possible or ethical. If two of these categories could be blinded, then the study would be a double-blind study. If all these three categories could be blinded, then the study would be triple-blind. Blinding should be maintained until the end of the study (after statistical analysis) to achieve the best level of confidence in the data recorded. Studies without randomization or low blinding have been described as having more therapeutic effects than double-blind randomized controlled clinical trials [6]. Triple-blind randomized controlled trials therefore provide the highest level of reliable evidence. However, additional analysis of residual confounding should be performed during data analysis to ensure the full validity of the study findings [7].

Systematic reviews, meta-analysis studies, and decision analysis studies are based on previous publications of a specific topic and allow more global views if properly conducted. The conclusions of these studies rely on the quality and availability of data, more than the design.

Myths and Misconceptions

A RCT is not always better than a retrospective study: if a RCT is not blinded, has a small sample size, or has many protocol violations, the quality of the study will be low. Matching of patients in retrospective case-control studies does not always reduce bias if the control group is not properly selected and matched to the appropriate characteristics of the populations. Repeated measurements on subjects do not increase the number of independent observations.

Conclusion

After formulating an addressable question, the first step in designing a study is to determine how one might minimize the various forms of bias. The randomized, controlled, triple-blind study generally is identified as the gold standard. If this design is the most reliable and objective method to eliminate bias and produce solid evidence, it also is far from always being practical in orthopaedic clinical research. Even though levels of evidence reflect potential for bias, these levels per se should not be at the forefront of one's mind when designing a study. Rather, one should think about minimizing bias: (1) whether there

are controls, (2) whether there is blinding of the treatment to patients, and/or observers, and (3) the randomization process. Ethical considerations and the individual skill and technique of the surgeons often dictate the choices that are made. The case report that informs surgeons about a rare but devastating complication is still essential. All study designs are useful. The best study design therefore is the one that could provide the best evidence to answer the research question.

References

1. Altman DG, Bland JM. How to randomise. *BMJ*. 1999;319:703–704.
2. American Society of Plastic Surgeons. Scales For Rating Levels Of Evidence. Available at: http://www.plasticsurgery.org/Medical_Professionals/Health_Policy_and_Advocacy/Health_Policy_Resources/Evidence-based_GuidelinesPractice_Parameters/Description_and_Development_of_Evidence-based_Practice_Guidelines/ASPS_Evidence_Rating_Scales.html. Accessed May 28, 2010.
3. Centre for Evidence-based Medicine. Levels of evidence. Available at: http://www.cebm.net/index.aspx?o=1025. Accessed June 11, 2010.
4. Clinical Orthopaedics and Related Research. Level of Evidence Classification Scheme. Available at: http://www.editorialmanager.com/corr/account/LOE.doc. Accessed May 28, 2010.
5. Gibbons RJ, Smith S, Antman E. American College of Cardiology/American Heart Association clinical practice guidelines: Part I: where do they come from? *Circulation.* 2003;107:2979–2986.
6. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273:408–412.
7. Vavken P, Culen G, Dorotka R. Management of confounding in controlled orthopaedic trials: a cross-sectional study. *Clin Orthop Relat Res*. 2008;466:985–989.