Published in final edited form as:

Plast Reconstr Surg. 2012 June ; 129(6): 979e–988e. doi:10.1097/PRS.0b013e31824eccb7.

Methodology Rigor in Clinical Research

Lynda J-S. Yang, MD, PhD¹, Kate W-C. Chang, BS, MA¹, and Kevin C. Chung, MD, MS²

¹Department of Neurosurgery, University of Michigan, Ann Arbor, MI, USA

²Section of Plastic Surgery, Department of Surgery, University of Michigan, Ann Arbor, MI, USA

Abstract

Background—Methodology rigor increases the quality of clinical research by encouraging freedom from the biases inherent in clinical studies. As randomized controlled studies (clinical trial design) are rarely applicable to surgical research, we address the commonly used observational study designs and methodologies by presenting guidelines for rigor.

Methods—Review of study designs including cohort, case-control, and cross-sectional studies and case series/reports, as well as biases and confounders of study design.

Results—Details about biases and confounders at each study stage, study characteristics, rigor checklists, and published literature examples for each study design are summarized and presented in this report.

Conclusions—For those surgeons interested in pursuing clinical research, mastery of the principles of methodology rigor is imperative in the context of evidence-based medicine and widespread publication of clinical studies. Knowledge of the study designs, their appropriate application, and strictly adhering to study design methods can provide high-quality evidence to serve as the basis for rational clinical decision-making.

Introduction

Clinical research is “a component of medical and health research intended to produce knowledge valuable for understanding human disease, preventing and treating illness, and promoting health.”¹ The quality of clinical research varies according to the level of freedom from biases that plague research studies—and the rigor by which investigators adhere to the methodology used to reduce bias. Categorization of varying types of clinical evidence provides a basis for evidence-based medicine (EBM)² and subsequent clinical decision-making. The U.S. Preventative Services Task Force created the following system for stratifying quality of evidence derived from clinical studies:³ Level I—evidence obtained from at least 1 properly designed randomized controlled trial; Level II—evidence obtained from well-designed controlled trials without randomization, from well-designed cohort or case-control analytic studies preferably from more than 1 center or research group, and from

Corresponding Author Contact Information Lynda J-S. Yang, MD, PhD Department of Neurosurgery University of Michigan Health System 1500 E. Medical Center Drive, 3552 Taubman Center Ann Arbor, MI, 48109-5338 Phone: 734-936-5017; Fax: 734-936-9294 ljsyang@med.umich.edu.

All authors listed above contributed equally in concept and design of the study, data acquisition, drafting and critical review of the manuscript, and final approval for submission.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

multiple time series with or without the intervention; Level III—opinions of respected authorities based on clinical experience, descriptive studies, or reports of expert committees.

Understanding proper clinical research study design is essential to both the researcher trying to attribute cause-and-effect, and to the clinician trying to interpret whether an outcome is the result of an exposure. Rigor results when strict conduct, accepted specialty standards, and rational interpretation are present. Clinical research studies can be analytic or descriptive. Analytic studies test hypotheses to determine if a causal relationship exists between exposure and outcome,⁴ whereas descriptive studies observe and describe the disease condition relative to person, place, and time.⁵ Analytic studies can be controlled (randomized controlled trial) or observational (cohort or case-control studies), whereas descriptive studies (cross-sectional, case series/reports) can only be observational.⁶ While evidence from properly designed and conducted randomized controlled studies is the ideal goal, they are seldom practical and rarely found in surgical research. Therefore, we address the more commonly applied study designs and methodologies used to achieve rigorous evidence that is not based on clinical trial design.

Bias

When a study fails to measure what it intends to measure, validity of the consequent result declines and evidence provided by the study is suspect. Study results can be erroneous because of bias, confounding, or chance. Methodologically rigorous studies seek to minimize the first two by design and the latter two by analysis. Bias refers to “any trend in the collection, analysis, interpretation, publication, or review of data that can lead to conclusions that are systematically different from the truth.”⁷ Several classifications of bias have been previously proposed.^{8,9} Below, we classify biases according to stage of the study (Figure 1), and suggestions for their amelioration are summarized in Table 1.

Literature Search

A literature search cannot access all related studies and cannot reveal existing systematic differences within those unaccessed citations. English-language studies with positive results and recently published articles are most likely to be found and quoted. The effect of unaccessed reports that may lead to wrong conclusions remains unknown. Researchers must recognize this bias and strive to identify relevant studies from all available resources. If publication bias is not recognized until final stages of the study (e.g., data analysis), a funnel plot can estimate each study's effect, with the x-axis representing estimated effect (odds ratio or mean difference), and y-axis representing inverse of error or sample size.¹⁰ When no publication bias is present, the plot is a roughly symmetrical inverse funnel showing large studies with smaller variation and more precision of estimated effect, and small studies with larger variation and less precision. More information on funnel plots can be found on Cochrane Collaboration website and elsewhere.¹¹

Study Design and Execution

Selection bias occurs in the design and execution stage if selected characteristics are systematically dissimilar from the study target.¹² Reduction of this bias is achieved by accurate selection among the comparison group and adherence to predetermined criteria. Prevalence-incidence bias occurs if the time gap between exposure and selection of study participants is enough to distort frequency of exposure and affect measurement among groups. Detection bias refers to the situation in which innocent exposure might be mistakenly considered as cause of the disease, or cause signs or symptoms that initiate a search for the disease condition.⁹ Bias of the volunteer effect is also known as self-selection

bias—subjects who volunteer to participate in a study may systematically differ from those who are qualified to participate but do not volunteer.

Study participants lost to follow-up create a significant bias. Participants without proper follow-up will have exposure or outcome differences from those who remain in the study.¹³ Drop-out refers to participant withdrawal from a study, necessitating that distribution of outcome effect must be analyzed on an intent-to-treat basis.¹⁴

Data Collection

Bias results from inaccurately measuring or classifying exposures/outcomes among subject groups. Inaccurate data influences analysis result, invalidates study effect, and misleads the investigation. Other common data collection biases are summarized in Table 1.

Confounders

Confounders confuse the results of a study--the apparent effect of the exposure of interest is distorted because of the mistaken effect of an extraneous factor (confounder) when analyzing or interpreting study results.

Data Analysis and Interpretation

Confounders are extraneous risk factors associated with both exposure (causal factor) and disease (outcome) that can affect data analysis. Although the measurement might be accurate, misinterpretation of data will lead to confusion. When the confounder is independently associated with outcome, the distorted result could be an overestimation or underestimation of the effect.^{15, 16} Randomization, matching, and restriction are commonly used to prevent confounding. Confounders should be considered in the design stage and during data analysis, using techniques to offset inadequacies of study design and reduce data misinterpretation. For example, stratification divides the study population into subgroups to compensate for confounding effects and increase participant representation for single confounders; multivariable analysis is useful for controlling multiple potential confounders simultaneously.

Cohort Studies

Of the observational analytic studies, the cohort study is most reliable for providing evidence to support potential cause-and-effect (intervention and outcome) relationships with fewest systematic errors. The integral design feature is the comparison of 2 groups (cohorts), of which 1 has been exposed to treatment or disease (study group) and the other has not (unexposed group). The investigator does not control the intervention, but only records the natural progression after intervention has occurred. Unlike the randomized study design, the cohort study does not interfere with clinical practice and is more palatable to most clinicians and patients and is more feasible and less demanding to perform (Table 2).

Cohort studies can be prospective or retrospective. In prospective studies, the intervention is recorded at initiation of the study (Figure 2A); in retrospective studies, data is collected after intervention has already been initiated (Figure 2B). In both, participants progress naturally, and prevalence of the outcome of interest is determined afterward. The prospective study has significantly fewer systematic errors and bias than the retrospective study but costs more and takes longer to complete. For both investigator and clinician-reader, knowledge of the appropriate research method can help determine the study's adherence to the method, and therefore support or refute validity of the results (Table 3).

Initiation of the cohort study includes firm definitions of the intervention and outcome of interest (that has not been obtained previously). Exclusion and inclusion criteria must be set and strictly applied to each potential subject. If significant differences exist between patients who receive the intervention and those who do not, erroneous attribution of cause-and-effect is the result. Likewise, comparisons to unexposed/untreated groups that are historical to the treatment group are compromised by changes in standard treatment practices over time. Study drop-out and loss to follow-up should be reported accordingly.

Other confounding factors are variables associated with intervention (risk factors for the outcome of interest). Confounders should be considered to avoid stating an association of cause-and-effect when such an association does not really exist. Strict exclusion and/or matching of individual patients according to the confounding variable are acceptable options for reducing erroneous study results. Other interventions that are not strictly part of the study can affect the outcome of interest and confound results of a cohort study; these co-interventions must be specified, and the study analysis adjusted.

Finally, the cohort study design must be appropriate for the research question—one that purports to establish incidence of the outcome of interest in the exposed study group versus the unexposed group. Vogelin and colleagues¹⁷ tested the hypothesis that patients who underwent carpal tunnel decompression had greater decrease in the cross-sectional area (CSA) of the median nerve than those who had conservative management. The research question and selection of the appropriate comparison group were well-defined and concurrent. All participants were subjected to predetermined inclusion and exclusion criteria. The measurement of CSA was clearly delineated, and measurement procedures were comparable between surgical and nonsurgical groups. Confounders were identified, and stratification was appropriately used to reduce the effect. Study limitations were recognized and considered to reach the conclusion that surgical decompression was associated with a greater decrease in median nerve CSA than nonsurgical treatment.

Case-Control Studies

Another observational analytic study design used to test potential cause-and-effect is the case-control study. When the outcome of interest is rare or lengthy, the cohort study cannot be reasonably employed; the case-control study method is best in these circumstances. The integral design feature is initial identification of the outcome of interest and the patients who have it (cases), followed by comparison to patients without the outcome of interest (controls) to determine if a specific exposure (intervention) is associated with the outcome. Surgical case-control studies must be performed retrospectively, thereby suffering from systematic errors inherent to retrospective studies (Table 4).

Because case-control studies are performed after reaching the outcome, they are relatively time- and cost-efficient to conduct. This design does not allow for calculation of prevalence of the outcome, because both case and control patients are distinctly chosen. However, this design is well-suited when exposure is more prevalent than the outcome, and when the outcome is rare or occurs over a long period of time. Because this design begins with the outcome, determination of the exposure often relies on patient recall (potential recall bias) when medical records are incomplete and verification of exposure is difficult. Selection of appropriate control patients is critical, and false associations of exposure to outcome can result from comparisons to inappropriate control patients.

Outcome of interest is the core of case-control studies (Figure 3). An unambiguous definition of the outcome and the criteria to determine presence or absence of the outcome must be explicitly stated. To avoid biases incurred by time (e.g., changes in treatment practices), case patients should be drawn from a discrete and reasonably short time interval.

In these patients, important characteristics that potentially contribute to the outcome are determined and explicitly stated in the inclusion and exclusion criteria.

Control patients possess similar patient determinants as case patients, but control patients do not have the outcome of interest. Control patients represent the population at risk of potentially experiencing the outcome and are selected without regard to the specific exposure (intervention). Investigators must (1) specify the exposure then determine patient exposure status, then (2) compare frequency of a particular exposure in case versus control patients to measure likelihood of a true association between exposure and outcome. Unlike the cohort study, the case-control method allows for analysis of multiple exposures that may be associated with the outcome of interest (Table 5).

Investigators employing the case-control method should report the likelihood that exposure is related to outcome as an odds ratio. The odds ratio measures strength of the relationship between exposure and outcome; it only describes the strength of association or non-independence between exposure and outcome. It does not generally describe risk of developing the outcome relative to exposure: only when outcome is less than 5% can the odds ratio approximate the relative risk (probability of the event occurring in case versus control patients).

An example of a case-control study is one that identified risk factors for ulnar nerve compression at the elbow.¹⁸ The outcome of interest was identified as ulnar nerve compression. The “cases” were patients with ulnar nerve compression who underwent surgical decompression, and the “controls” were patients admitted to the same hospital who were treated for herniated spinal discs. The investigators clearly defined target outcome, inclusion and exclusion criteria, method for selecting “case” and “control” patients, and the treating institution. Appropriate statistical methods were used to determine possible risk factors and odds ratio. Discussion of the limitations and possible biases demonstrated that methodological rigor appears to have been sustained.

Cross-sectional Studies

Both cohort and case-control studies test hypotheses concerned with cause-and-effect (treatment-outcome). When the primary question concerns prevalence of a condition and association of an exposure or treatment to the condition, the cross-sectional study design is ideal. It allows examination of a sample population at a single point in time (Figure 4). The sample population should be appropriately selected to reflect composition of the population. Pre-defined measurements and assessments are carefully established and often use questionnaires or surveys. The goal of cross-sectional design is to (1) estimate prevalence of the condition of interest, and (2) establish an association between a potential “treatment” or “exposure” and an “outcome.”

Despite the inability of this design to establish a causal relationship, the cross-sectional study has the distinct advantage of time: no follow-up is required. This study design can accommodate secondary data analysis, using data possibly collected by others for another purpose. There is close control over the measurement/assessment process, and key data usage can be maximized.

Conversely, the disadvantages of cross-sectional design are derived from these same concepts. If secondary data analysis is used for efficiency, there is no control over purpose or method used to collect the data. Because the study concerns only a single time-point, no causal relationships can be postulated. If a rare outcome is of interest, the sample size required will be very large, and generalizability of the results will be limited by the sample population (Table 6).

This method requires strict attention to the following steps (Table 7). Identification and definition of the population of interest must adhere to pre-determined criteria. Outcomes must be defined with precision. Exposure/treatment criteria should be set. Data collection instruments must be created with proper standardization, and administered with attention to protocol adherence by trained staff. Analysis of collected data requires consideration of sampling design, and adjustment for confounders is usually necessary.

An example of a cross-sectional study illustrating these principles investigated the prevalence of peripheral nerve conduction abnormalities in a population exposed to arsenic, and the potential association of these nerve conduction abnormalities with arsenic exposure.¹⁹ Both the population of interest and outcome were well-defined. Collection of data took place at a single point in time. Parameters for measuring nerve conduction involved a strict protocol for the neurophysiologists. Arsenic levels were measured in well water and in participants' urine; tests were augmented with standard surveys and interviews. Limitations in the context of potential bias and sample size were discussed. Despite limitations, methodology rigor supported findings suggesting that chronic arsenic poisoning from drinking water is unlikely to affect nerve conduction velocity within the range examined.

Case Series/Reports

Case series/reports are descriptive observational studies and are mostly retrospective. Case series/reports do not use a comparison group to analyze treatment effect on outcome: they cannot attribute cause-and-effect because they do not test hypotheses of treatment efficacy.²⁰ A case series only describes the progress of patients with similar diagnoses or treatments over time (Figure 5). However, well-structured case series are useful for generating hypotheses about potential treatment/outcome causal relationships that can be tested using more analytic study methods:⁶ e.g., the presentation of an innovative procedure can be useful clinically and lead to further study.

Case series/reports are inexpensive and quick to conduct. They can describe a new disease entity, trend, or rare phenomenon without the constraints of randomization or controlled studies. A strong argument has been made for the relevance of publishing case series/reports with regard to clinical relevance (external validity).²¹

Advantages of case series/reports are offset by the limitations, because these descriptive studies cannot establish a cause-and-effect relationship. They describe interventions or treatments from a single institution or practitioner. Patients are likely chosen for the study depending on outcome (selection bias). No control group exists, so it is impossible to determine if outcomes are related to treatment (Table 8).

Good case studies must adhere to the study design (Table 9). Because no hypothesis is tested, the study question should be focused on defining the study sample, diagnosis or treatment of interest, and outcome. Pre-defined inclusion and exclusion criteria should be strictly applied. A literature review should be conducted to justify the study, and the case series/report should provide new information of clinical significance. Outcome measures should be standardized to reduce measurement bias, and only descriptive statistics should be used (no P-values).

The strength of case series/reports is in details of the study. The case description should provide the reader with adequate data, acknowledging that chance and bias are present. Completeness in recording patient characteristics and length of follow-up is critical. When conducted properly, formulation of a reasonable hypothesis for further study can occur.

For example, Teboul et al.²² reported a series of patients with brachial plexus injury who underwent a nerve transfer procedure. The sample is well-defined and treatment is novel. Outcome is documented in the context of neurologic recovery. Cases are presented with adequate detail to assess validity of the study, and duration and follow-up are reasonable. Descriptive statistics were appropriately applied. The methods justify the conclusion that nerve transfer is a potentially effective treatment for brachial plexus injury which generated a novel hypothesis, prompting questions for multiple other analytic studies. The unique features and clinical significance of the study have withstood the test of time, as the reported transfer has become standard practice.

Conclusions

Mastering the principles of methodology rigor by those interested in pursuing clinical research is imperative in this age of EBM and widespread publication of clinical studies. Attention to reducing bias at each stage of a study increases validity of study results and can be used to support EBM and clinical decision-making.

Acknowledgments

Financial Disclosure and Products

This work is supported in part by grants from the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the National Institute on Aging (R01 AR062066), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (2R01 AR047328-06), and a Midcareer Investigator Award in Patient-Oriented Research (K24 AR053120) (to Dr. Kevin C. Chung).

References

1. [July 1, 2011] Clinical Research: A National Call to Action (American Association of Medical Colleges and the American Medical Association). Anonymous Available at: <http://www.umassmed.edu/uploadedfiles/NationalCalltoAction.pdf>.
2. Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. *Health Aff (Millwood)*. 2005; 24:18–28. [PubMed: 15647212]
3. [July 1, 2011] Guide to Clinical Preventive Services (U.S. Preventive Services Task Force). Anonymous Available at: <http://odphp.osophs.dhhs.gov/pubs/guidecps/pcpstoc.htm>.
4. Brighton B, Bhandari M, Tornetta P 3rd, Felson DT. Hierarchy of evidence: from case reports to randomized controlled trials. *Clin Orthop Relat Res*. 2003; (413):19–24. [PubMed: 12897592]
5. Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. *Lancet*. 2002; 359:145–149. [PubMed: 11809274]
6. Kooistra B, Dijkman B, Einhorn TA, Bhandari M. How to design a good case series. *J Bone Joint Surg Am*. 2009; 91(Suppl 3):21–26. [PubMed: 19411496]
7. Last, JM.; Spasoff, RA.; Harris, SS.; Thuriaux, MC. *A Dictionary of Epidemiology*. 4 Ed.. Oxford University Press; Oxford: 2001.
8. Miettinen OS, Cook EF. Confounding: essence and detection. *Am J Epidemiol*. 1981; 114:593–603. [PubMed: 7304589]
9. Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979; 32:51–63. [PubMed: 447779]
10. Sutton, AJ.; Abrams, KR.; Jones, DR.; Sheldon, TA.; Song, F. *Methods for Meta-Analysis in Medical Research*. John Wiley; Chichester: 2000.
11. Chung KC, Burns PB, Kim HM. A practical guide to meta-analysis. *J Hand Surg Am*. 2006; 31:1671–1678. [PubMed: 17145390]
12. Kleinbaum, DG.; Kupper, LL.; Morgenstern, H. *Epidemiologic Research: Principles and Quantitative Methods*. Lifetime Learning Publications; Belmont, CA: 1982.
13. Hennekens, CH.; Buring, JE. *Epidemiology in Medicine*. 1 Ed.. Little, Brown and Co.; Boston: 1987.

14. Stewart LA, Parmar MK. Bias in the analysis and reporting of randomized controlled trials. *Int J Technol Assess Health Care*. 1996; 12:264–275. [PubMed: 8707499]
15. Ahlbom, A.; Norell, S. *Introduction to Modern Epidemiology*. 2 Ed.. Epidemiology Resources; Chestnut Hill, MA: 1990.
16. Rothman, KJ.; Greenland, S. *Modern Epidemiology*. 2 Ed.. Lippincott-Raven; Philadelphia: 1998.
17. Vogelien E, Nuesch E, Juni P, Reichenbach S, Eser P, Ziswiler HR. Sonographic follow-up of patients with carpal tunnel syndrome undergoing surgical or nonsurgical treatment: prospective cohort study. *J Hand Surg Am*. 2010; 35:1401–1409. [PubMed: 20807617]
18. Bartels RH, Verbeek AL. Risk factors for ulnar nerve compression at the elbow: a case control study. *Acta Neurochir (Wien)*. 2007; 149:669–674. [PubMed: 17502986]
19. Fujino Y, Guo X, Shirane K, et al. Arsenic in drinking water and peripheral nerve conduction velocity among residents of a chronically arsenic-affected area in Inner Mongolia. *J Epidemiol*. 2006; 16:207–213. [PubMed: 16951540]
20. Carey TS, Boden SD. A critical guide to case series reports. *Spine (Phila Pa 1976)*. 2003; 28:1631–1634. [PubMed: 12897483]
21. Audige L, Hanson B, Kopjar B. Issues in the planning and conduct of non-randomised studies. *Injury*. 2006; 37:340–348. [PubMed: 16483576]
22. Teboul F, Kakkar R, Ameer N, Beaulieu JY, Oberlin C. Transfer of fascicles from the ulnar nerve to the nerve to the biceps in the treatment of upper brachial plexus palsy. *J Bone Joint Surg Am*. 2004; 86-A:1485–1490. [PubMed: 15252097]

Figure 1.
Biases or confounding are encountered at all stages of study.

Figure 2.

(A) Prospective cohort studies begin by recording the study population intervention, whereas (B) retrospective cohort studies require that data be collected after the intervention has been initiated.

Figure 3.
Case-control studies begin with identification of the outcome of interest and patients who have the outcome of interest.

Figure 4.
Cross-sectional studies examine a sample population at a single point in time.

Figure 5.
Case series/reports describe the progress of patient interventions over time.

Table 1**Biases and Their Prevention at Various Stages of Study**

Study Stage	Biases	Study Design Affected	Preventions
Literature search	Publication bias	Meta-analysis, systematic review	Funnel plot to examine effect and sample size
Design and execution	Selection (sampling) bias		
	Admission (Berkson's bias)	Case-control, cross-sectional, cohort (hospital-based)	Use population-based cases and controls; avoid using hospital samples
	Prevalence-incidence (Neyman's bias)	Cross-sectional, case-control, cohort (prospective)	Also include non-exposure subjects and incident cases
	Detection bias	Case-control, cohort	Subjects undergo identical detection methods
	Healthy worker bias	Cohort (retrospective)	Multiple comparison cohorts
	Volunteer effect (self-select, response bias)	Observational	Achieve response rate of at least 80% and compare responders and non-responders
Data collection	Loss of follow-up	Cohort	Regular follow-up in order to maintain follow-up rate
	Information bias		
	Misclassification	Experimental and observational	Information should be gathered in the same way; verify exposure with record, if possible
	Measurement	Experimental and observational	Use standard measurement with objective criteria for exposures and outcomes
	Interviewer (intra/inter)	Survey, experimental, and observational	Well-trained and experienced interviewer blinded to study hypothesis
	Questionnaire	Survey, experimental, and observational	Careful wording without misleading and obvious demonstration of study hypothesis
	Recall	Observational	Difficult to prevent, might be helpful to double-check with record
	Diagnostic suspicion	Case-control, cohort	Use identical diagnostic examination among groups
	Exposure suspicion	Case-control, cross-sectional	Use comparable methods to measure/observe study groups
	Missing data	Experimental and observational	Analyze missing data on intent-to-treat basis
Analysis	Confounding bias	Experimental and observational	Randomization
			Stratification (frequency, severity, any confounders)
			Restriction
			Matching

Table 2**Study Characteristics, Cohort Studies****Requirement**

- Disease-free (at-risk) population
- From at-risk population, cohort groups identified as having been exposed or unexposed to treatment/complication/factor
- Cases within each cohort are identified as diseased or non-diseased
- Same source population for exposed and unexposed groups
- Data can be collected prospectively or retrospectively

Advantages

Able to:

- Examine multiple outcomes for a given exposure
- Investigate rare exposures
- Identify causality by following subjects from exposure to outcome
- Perform relatively inexpensively and quickly when conducting retrospectively
- Design with specific target when conducting prospectively
- Measure rates of disease in exposed and unexposed over time (longitudinal)

Pitfalls

- Need large subject numbers in order to assess rare outcomes
- Cohort groups could be misclassified if exposure or disease definition is unclear
- Long follow-up period and expensive for prospective design
- Difficult to maintain subject follow-up or withdrawals
- Less control for variables when conducted retrospectively
- Recall bias could cause invalid data

Table 3**Rigor Checklist for Study Design, Cohort Studies**

Criteria	Checklist
1. Population	
Disease-free (at risk)	Who are the targeted subjects? What are the inclusion and exclusion criteria?
Exposure	Define exposure status clearly. When exposed, and how long?
Outcome/disease	What is the outcome? Define the result clearly.
2. Design	
Suitability	Is it appropriate based on the research question?
Importance	Is it able to provide quality evidence?
3. Data and Sampling	
Cohort groups	Are the populations classified with rationale?
Sources	Are the samples representative?
Sample size	Are samples large enough to assess relative risk/incidence?
Duration and follow-up	Determine the period of follow-up necessary.
Drop-out and missing	How much missing data? What's the distribution in missing data?
Exposure timing	Ensure that cases are exposed during a similar time period.
4. Measure Instruments	
Outcomes status	Detailed definition of outcomes; are they standardized? How should outcomes between 2 groups be compared?
Tools	What are the measuring instruments? Are they valid and reliable?
5. Analysis	
Statistics	What are the tests used? Relative risk or disease incidence?
Confounders	Are the confounders identified?
6. Interpretation	
Discussion	Are the results related to the research question?
Limitations	Any further applications?

Table 4**Study Characteristics, Case-Control Studies****Requirement**

- Cases with disease/treatment/complication/outcome
- Controls without disease/treatment/complication/outcome
- Same source population for cases and controls
- Data collected retrospectively
- Identify exposure rates to risk factor(s)

Advantages

- Able to examine rare outcomes
- Relatively inexpensive
- Relatively quick to conduct
- Able to identify multiple risk factors

Pitfalls

- Only one outcome of interest at a time
- Inappropriate group selection
- Imbalance of confounders in case and control groups because of lack of randomization
- Recall bias could cause invalid data
- Unable to define exposed and unexposed rates when exposure is rare
- Unable to provide prevalence of outcome and causality

Table 5**Rigor Checklist for Study Design, Case-Control Studies**

Criteria	Checklist
1. Population	
Case definitions	Who are the targeted subjects? What are the inclusion and exclusion criteria?
Control definitions	Do the controls have similar characteristics as the cases?
Exposure	What are possible risk factors? Are controls and cases treated differently?
2. Design	
Suitability	Is it appropriate based on the research question?
Importance	Is it able to provide quality evidence?
3. Data and Sampling	
Sample selection	Are the population sources for cases and controls similar?
Sources	Are the samples representative?
Sample size	How many samples were in the study?
Duration and follow-up	How long?
Drop-out and missing	How much missing data? What's the distribution in missing data?
4. Measure Instruments	
Outcomes status	Detailed definition of outcomes; are they standardized?
Tools	What are the measuring instruments? Are they validated and reliable?
5. Analysis	
Statistics	What are the tests used? Odds ratio?
Confounders	Are the confounders identified?
6. Interpretation	
Discussion	Are the results related to the research question?
Limitations	Any further application?

Table 6**Study Characteristics, Cross-sectional Studies**

Requirement
<ul style="list-style-type: none"> • Disease-free population • Subjects are classified by well-defined, health-related characteristics and then all are sorted simultaneously by the presence/absence of disease • Exposure and outcome (disease) are determined simultaneously • Same source population for exposed and unexposed groups
Advantages
<ul style="list-style-type: none"> • Relatively inexpensive and quick to conduct • Flexible design for different research purposes (risk, prevalence, test new intervention, etc.) • Able to examine association between exposure and outcome
Pitfalls
<ul style="list-style-type: none"> • Both groups need to be similar in size in order to assess the relationship • Cohort groups could be misclassified if the exposure or disease definition is unclear • Confounders may distributed unequally • Recall bias could cause invalid data

Table 7**Rigor Checklist for Study Design, Cross-sectional Studies**

Criteria	Checklist
1. Population	
Disease-free (at risk)	Who are the targeted subjects? What are inclusion and exclusion criteria?
Exposure	Define the expose status clearly. When exposed, and how long?
Outcome/disease	What is the outcome? Define the result clearly.
2. Design	
Suitability	Is it appropriate based on the research question?
Importance	Is it able to provide quality evidence?
3. Data and Sampling	
Cohort groups	Are the populations classified with rationale?
Sources	Are the samples representative?
Sample size	Are the samples large enough to assess the relative risk/incidence?
Drop-out and missing	How much missing data? What's the distribution in missing data?
Exposed timing	Ensure cases have been exposed during a similar time period.
4. Measure Instruments	
Outcomes status	Detailed definition of outcomes; are they standardized?
Tools	What are the measuring instruments? Are they validated and reliable?
5. Analysis	
Statistics	What are the tests used? Relative risk or disease incidence?
Confounders	Are the confounders identified?
6. Interpretation	
Discussion	Are the results related to the research question?
Limitations	Any further application?

Table 8

Study Characteristics, Case Series/Reports

Requirement
<ul style="list-style-type: none">• Cases that report new information or unexpected variation in diseases or syndromes• Cases inconsistent with known diseases or syndromes• Patient characteristics and information should be described in detail (demographics, medical record, past/current interventions, and results)• A potential hypothesis should be presented and should be testable by another study type
Advantages
<ul style="list-style-type: none">• Quick and inexpensive to conduct• May identify new disease, trend, or rare phenomena• May detect new drug side-effects or potential uses• Shared educational experience
Pitfalls
<ul style="list-style-type: none">• Cases may not reflect general situation• Intervention or treatment are based on single-institution and single-physician experience• Causes or associations may have other explanations• Lack of comparison group for evidence support• Low evidence-base; the study might report insufficient information• Might over-emphasize misleading events• Selection or measurement bias

Table 9**Rigor Checklist for Study Design, Case Series/Reports**

Criteria	Checklist
1. Introduction	
Population	Who are the targeted subjects? What is their background and unique features? What are the inclusion and exclusion criteria?
Literature review	Do results of related literature review justify the study?
Treatment/intervention	Do treatment/intervention options differentiate subjects?
Outcome/disease	What is the outcome? Define the result clearly. Is it different from the regular one? If so, at what level?
Objective	Should provide evidence with clinical significance or teaching value
2. Case presentation	
Case history	What are the case histories, demographics, and situations presented?
Sources	Single or multiple institutions?
Sample size	Report single case or multiple cases
Outcomes status	Detailed definition of outcomes; are they standardized?
Duration and follow-up	Determine length of follow-up interval
Statistics	Descriptive statistics
Detail	Study should provide enough detail for audience to judge validity of the study
3. Discussion	
	Unique features of current case(s)
	Potential clinical implications
	Teaching value
	Potential hypotheses for further study
Limitations	Any further application?