The Choice of Controls for Providing Validity and Evidence in Clinical Research

Sunitha Malay, MPH[1] and **Kevin C. Chung, MD, MS**[2]

[1]Research Associate, Section of Plastic Surgery, Department of Surgery, The University of Michigan Health System

[2]Professor of Surgery, Section of Plastic Surgery, Assistant Dean for Faculty Affairs, The University of Michigan Medical School

Abstract

Background—Selection of controls, a group of subjects who are identical to the treatment group in all aspects that affect the outcome except the intervention of interest, is a significant criterion for conducting a study in evidence-based medical research. Few studies emphasize the appropriate selection of control groups in the plastic surgery literature.

Methods—We performed a literature search in *Plastic and Reconstructive Surgery* (*PRS*) from January 1, 2010 through December 31, 2011 for studies in which controls were needed. The number of studies using a control group, control selection criteria and the characteristics of the control populations were evaluated.

Results—327 articles were obtained from our search using key words 'case control studies' and 'retrospective cohort studies.' Among these studies, 121 articles were studies conducted in humans. All these studies based on the study design required a control group, yet only 63 studies (52%) had a comparative control group. Of these studies, we found biases regarding the choice of controls, including selection bias, misclassification bias, and chronology bias.

Conclusion—Our review shows that 48% of the studies published in *PRS* that were required to have a control group failed to incorporate a well-defined control group, which will enhance the validity of a study. Specific details pertaining to the methods used and the success obtained with those methods in recruiting controls needs to be stated explicitly in the article to ensure uniformity, and to support the validity of the research.

Keywords

Controls; selection criteria; bias; case control studies; cohort studies; validity

Introduction

The first use of a distinct control group in a clinical study can be traced back to 1926 when Janet E. Lane-Claypon, an English physician, investigated the causation, prevalence and treatment of breast cancer.[1] In this initial case-control study, cases were women with a history of breast cancer and controls were women without the disease but who were otherwise similar. They found significant associations between the risk of breast cancer and age at menopause, age at pregnancy, number of children and lactation. The introduction of

Corresponding Author: Kevin C. Chung, MD, MS, 1500 E. Medical Center Dr. 2130 Taubman Center, SPC 5340, The University of Michigan Health System, Ann Arbor, MI 48109-0340, Phone: 734-936-5885, Fax: 734-763-5354, kecchung@med.umich.edu.
Conflict of interest - none

randomization to clinical trials by Bradford Hill, an English epidemiologist in the 1950s led researchers to adopt the idea of a randomized control group for appropriate evaluation of a new intervention.[2] A control group is comprised of people similar to the test group in all aspects that affect the outcome except for the treatment/intervention of interest. Controls are selected on the basis of comparability to the target population or the population at risk. This group is essential to demonstrate that a treatment or intervention is superior, less costly, or associated with fewer complications compared to the standard practice. Besides helping in the assessment of safety and efficacy, a control group discriminates outcomes caused by the treatment or intervention of interest from those caused by other factors, such as natural course of disease, patient or observer expectations, or other treatments.

Case-control studies and retrospective cohort studies serve as appropriate study designs that can potentially recruit a control population of an investigator's choice. A case-control study is an observational study in which a group of individuals with disease or condition under investigation (cases) are compared with a group of individuals without that disease (controls). A retrospective cohort study is an observational study in which groups of people are identified and categorized (exposed and non-exposed) based on exposure and followed for a period in the past. Randomized controlled trials (RCTs) and prospective cohort studies are other study designs that include a control group. Due to the random allocation of subjects to treatment or control groups in RCTs, the choice of control group is not at the discretion of the investigator. There is plenty of available literature to inform the control selection process. [3–9] The sources from which controls can be selected is a critical issue because each group offers specific advantages and disadvantages that must be carefully considered in order to closely match the treatment group. A vast diversity of sources from which controls can be used include cancer registries, hospital admission rosters, friends and neighbors of cases, or through random-digit-dialed phone calls. Controls in plastic surgery are incorporated either as concurrent or as historic controls. A concurrent control is a subject enrolled simultaneously with the treatment group from the same source population and followed for the same study period, whereas a historical control is a subject treated in the past with the standard form of care whose outcomes are used to compare with patients receiving the treatment under investigation.

Despite the available guidance and sources to choose controls, we hypothesize that this comparative group, which is essential to validate the study results, is infrequently used in plastic surgery studies. The aim of our paper is to assess the use of control groups in plastic surgery trials as reported in *Plastic and Reconstructive Surgery* (*PRS*) and to recommend the inclusion of a valid control group, where necessary, in all future manuscripts. We also intend to inform our readers about the qualities that constitute a good or a poor control group.

Methods

We performed a literature search in *PRS* from January 2010 to December 2011 inclusive, using the key words 'case control studies' and 'retrospective cohort studies.' Our selected search terms reflect the fact that in observational studies such as case control and retrospective cohort studies, an investigator has the potential ability to choose the controls.

Initial title and abstract search was done for each article to identify if the study required and/ or included a control group. For the studies including a control group, we assessed the source population, type of control (concurrent or historic), and number of subjects enrolled in both the control group and treatment group. Lastly, we assessed the methods of control selection, such as matching based on the variables of outcome in the study. We included only studies conducted with humans in our analysis and excluded studies conducted in non-humans, systematic reviews, chart reviews, case series, discussions, editorials, and replies.

Results

Our literature search found 327 articles. (Figure 1) Among those, 140 articles were studies conducted in human beings. After the exclusion of 19 duplicate articles, 121 articles remained that required a control group to comparatively evaluate the results observed in the intervention group. Only 63 studies (52%) contained a control group, leaving 58 studies (48%) that required, but did not include a control group to validate their results.

We found that nearly all of the studies used concurrent controls (61/63 studies) rather than historical controls (2/63 studies). Of the concurrent controls, a control with different treatment was the most commonly used (41/61), followed by a no treatment control (9/61). Few studies used more than one control (8/61) and a very few studies enrolled patients to act as their own controls (3/61). (Table 1) The proportion of the number of controls to those in the treatment group varied; however, controls were equal or more in number in 46/61 of the studies. The method of selection of controls was not indicated in the articles for 58/61 of the studies conducted. However, in 3/61 studies, controls were selected based on matching variables of interest such as age, sex, body mass index, comorbidities, and wound characteristics.[10,11,12] In 3/61 studies subjects were self-controls, in which they received treatment on one side of the body and other side acted as a control, which is not a good control because the effect on one body part will influence the outcomes for another body part in the same subject. Additionally, the source of controls was from patients with the same illness but undergoing different treatment in 58/61 studies, from hospital patients with other illnesses in 2/61 studies and from general population samples in 1/61 studies.

We observed the studies that included a control group failed to mention the specific principles applied in the selection of controls. Details like how controls were selected and why specific controls were pursued were not explicitly stated, suggesting a lack of scientific approach in the choice of controls by the investigators. A criterion for the choice of control was often the receipt of a different type of treatment or not receiving any treatment, especially in retrospective cohort studies. The relevance and generalizability of the inferences drawn from clinical studies are contingent on the strict criteria regarding the selection of suitable controls that are established prior to the beginning of a study. Studies invariably had specific inclusion and exclusion criteria for the selection of cases but this rigorousness was not evident with regards to control selection. Incomparability of cases and controls will result in selection bias that influences the outcomes and study conclusions. Selection bias occurs due to unaccounted differences in the characteristics between cases and controls. This is seen more often in case control and retrospective cohort studies in which exposure and outcomes have already occurred prior to subjects' selection into the study.[13] Table 2 discusses the possible biases in control selection and measures to overcome those biases. Therefore, non-inclusion or inclusion of an inappropriate control renders the quality of study results to be weaker.

Discussion

Only 52% of studies published in *PRS* in 2010 and 2011 validated their results with a control group. Forty-eight percent of the research studies conducted in plastic surgery failed to incorporate a control group, which undermined the studies' conclusions.

Studies mostly enrolled a concurrent control rather than a historic control (61 vs. 2 studies). Among the studies with concurrent controls, a control group with a different treatment (41/61 studies) was the most frequently used type, where two or more different treatments or a new treatment and standard therapy were compared. Most of the retrospective cohort studies used this type of control. Each patient group served as a control to the other group

thereby comparatively evaluating the safety and effectiveness of various treatments. For example, Colwell et al. compared breast reconstruction performed with single stage acellular dermal matrix (experimental group) with those of two stage reconstruction without acellular matrix (control group).[14] Use of such a control group, demonstrated that single stage implant reconstruction was cost-effective and had lower complication rate compared to two stage reconstruction. However, there is a possibility of channeling bias with this type of control. This bias occurs when an investigator assign treatments to subjects based on their comorbidities and prognostic factors.[13]

No treatment control, in which a control group is enrolled and observed without any intervention done, was another type of control that was used in plastic surgery (9/61 studies). Because the control group does not undergo any treatment, the outcomes observed in the intervention group can be attributed to the treatment alone, thereby establishing the effectiveness of the treatment. Case-control studies mainly employed this type of control. Such a control can be used only when not providing treatment does not deprive the control group of any benefits that would have been otherwise possible. For example, Caviggioli et al. compared the outcomes of fat tissue grafting in patients suffering from post-mastectomy pain syndrome with patients who did not receive fat grafting after mastectomy but suffered from pain.[15] This control group established the fat grafting procedure to be a safe, noninvasive, and effective treatment of post-mastectomy pain.

Occasionally, more than one control group is used in plastic surgery research (8/61 studies). Having more than one control group increases the power of a study and reduces the bias. Increase in power is achieved because of greater comparability as more information is available regarding potential differences in response between cases and controls. This is useful especially in a retrospective study of a rare disease with small number of available cases. However, the increase in power will not be observed after a case to control ratio of 1:4 because this ratio achieves maximum power. [16] Enrollment of two or more types of controls allows the researchers to reduce any possible bias due to the choice of a single control group.[9] For example, in a study of pelvic cancer defects reconstruction, the authors compared the outcomes of vertical rectus abdominis flap (VRAM) reconstruction performed with six technical modification groups.[17] Fascia sparing, component separation, mesh reinforcement, deepithelized skin paddle, extended VRAM flap and omental flap were the techniques used. Patients in each group served as a control to other groups. Use of more than one control group helped to establish the effectiveness of total VRAM flap deepithelization technique in minimizing donor and recipient-site complications.

Patients are infrequently used as self-controls in plastic surgery research (3/61 controls). Although, some investigators claim that it minimizes within subject bias, it is not ideal to use patients themselves who are undergoing the treatment as controls. For example, in a study of patients with hemifacial microsomia, each affected side of the face was treated as a case, with the unaffected side as the control.[18] The assessment and the effect of surgery on one side of the face will be influenced by the other, which cannot definitively establish the effectiveness of a treatment. Another example of this bias is to perform two different face-lift procedures in a patient and attempts to compare outcomes of the two techniques. The changes in one side of the face will certainly affect the other side, which makes comparison imprecise.

Distinct from concurrent controls are historic controls, occasionally employed in plastic surgery research (2/63 studies), in which a set of subjects outside the study population who were treated in the past. Use of historical control tends to overestimate the effect of the more current treatment. Differences in diagnostic methodology and outcome assessment methodology used for the historic and intervention groups may result in bias.[19] Chronology

bias occurs when historical controls are used. It is the effect of different time and place between the two comparative groups, which can confound the results.[13] The improvement in technology and expertise over time will tend to favor the more current treatment. Therefore, historical control use should be restricted to those a study in which enrollment of concurrent controls is not possible. For example, Pannucci et al. examined the effect of post-operative enoxaparin (chemoprophylactic agent) to prevent venous thromboembolism in high-risk plastic surgery patients operated in 2009 by comparing it with a historical control group comprised of patients operated in 2006–2008.[20] The controls did not receive chemoprophylaxis after surgery because venous thromboembolism was not identified as a major patient safety issue before 2008. With similar eligibility criteria for historical controls as that of intervention patients, the authors ensured comparability and thus validity of the study.

We found several studies that did not include a control group despite having the possibility to do so. To illustrate we will discuss a few articles. Sainsbury et al. did not use a control group in the study of innovative intralesional Bleomycin injection treatment for vascular birthmarks.[21] This vital study for a common yet significant condition in infants could have used a control group treated with other therapies to assess Bleomycin's relative efficacy or ability to minimize complications. Chang et al. did not enroll a control group in their study about the treatment of breast cancer–related lymphedema with lymphaticovenular bypass.[22] Women using standard methods such as exercise, bandages, manual drainage, or pumps and drugs for reducing post-cancer lymphedema could have been incorporated as controls to determine if the bypass method was better than standard methods. Arneja et al. in their study of resections performed for pericoular hemangioma in children failed to incorporate a control group with standard corticosteroid therapy. [23] The study found that the new treatment also reduces astigmatism and prevents amblyopia but the lack of a control group made it difficult for comparative evaluation to determine if the standard therapy also produced the same outcomes. Roostaeian et al. in their retrospective study on immediate implant placement in post-mastectomy women found that it is a safe and viable option that can provide very good aesthetic results in appropriately selected candidates.[24] A control group of women who underwent delayed implant placement could have been incorporated and outcomes compared.

In the studies reviewed, the investigator determined the type of control groups used, if they were used at all. Often controls were a convenience sample, one that can be easily obtained but not necessarily in accordance with the principles of control selection. Such a method would lead to selection bias. Additionally, incorrect ascertainment of the disease status or exposure status by the investigator may result in improper categorization of cases and controls leading to misclassification bias. In the event of selecting controls without well-defined criteria, the investigator is prone to generate conditions leading to channeling bias.[13] In this bias, subjects without existing comorbidities and better prognostic factors tend to be allocated more aggressive treatments when compared to subjects with comorbidities who will be assigned treatments with less risk. To overcome these biases and other issues encountered in recruiting controls, a standard approach needs to be followed. Selected controls should be comparable to the cases and represent the variability of the source population. Lack of such comparability may result in an association where one does not truly exist or may fail to reveal true associations. For example, for each case of rectus femoris flap reconstruction performed for groin wounds, an age and sex matched control was selected from patients presenting to a physiotherapy department for other problems.[10] Careful matching based on age and sex and obtaining from the same hospital ensured the validity of the control. Table 3 provides a quick view of the qualities that render a control to be good. Additionally, published articles need to clearly illustrate the details pertaining to control selection such as source population, sampling periods, recruitment, participation

rates, and matching criteria. When controls are selected based on matching variables from a source population, care must be taken to not overmatch. If the matched variables can affect the outcome, overmatching occurs and leads to biased results. For example, in a case-control study evaluating the role of alcohol consumption and benign proliferative epithelial disorders of breast in women, there were two control groups.[25] One group comprised of biopsy negative women whose biopsies were performed at the same laboratory as the cases were and another group was age-matched controls from community. Overmatching could have arisen from the biopsy controls as well as community controls because area of residence may serve as a social proxy for alcohol consumption.

Our study has a limitation that our review is restricted to one journal and to a time period of two years. Our inferences may not reflect all the studies published in other plastic surgery journals. However, *Plastic and Reconstructive Surgery* publishes plastic surgery studies with the greatest impact and papers presented in this journal should reflect the most influential publications in this specialty.

The appropriate conduct of a study and its results are not only dependent upon the intervention of interest but also on the thoughtful selection of controls. It is therefore essential that investigators employ control groups, when possible, to test the validity of their treatment effect. Similarly, standardized reporting of control selection protocol in peer-reviewed journals will allow researchers to evaluate appropriate methods to recruit potential controls. This effort, which is currently lacking in the studies conducted in plastic surgery, will enhance the validity of specific studies and thus contribute to evidence-based medical practice.

Acknowledgments

References

1. Paneth N, Susser E, Susser M. Origins and early development of the case-control study: Part 2, The case-control study from Lane-Claypon to 1950. Soz Praventivmed. 2002; 47(6):359–365. [PubMed: 12642996]

2. Stolberg HO, Norman G, Trop I. Randomized controlled trials. AJR Am J Roentgenol. Dec; 2004 183(6):1539–1544. [PubMed: 15547188]

3. Miettinen OS. The "case-control" study: valid selection of subjects. J Chronic Dis. 1985; 38(7):543–548. [PubMed: 4008595]

4. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. Am J Epidemiol. May 1; 1992 135(9):1019–1028. [PubMed: 1595688]

5. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. II. Types of controls. Am J Epidemiol. May 1; 1992 135(9):1029–1041. [PubMed: 1595689]

6. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. III. Design options. Am J Epidemiol. May 1; 1992 135(9):1042–1050. [PubMed: 1595690]

7. Miller AB. Hospital or population controls? It depends on the question. Prev Med. May; 1994 23(3): 263–266. [PubMed: 8078845]

8. Feinstein AR, Horwitz RI. On choosing the control group in case-control studies. J Chronic Dis. 1983; 36(4):311–313. [PubMed: 6833449]

9. Lasky T, Stolley PD. Selection of cases and controls. Epidemiol Rev. 1994; 16(1):6–17. [PubMed: 7925728]

10. Sbitany H, Koltz PF, Girotto JA, Vega SJ, Langstein HN. Assessment of donor-site morbidity following rectus femoris harvest for infrainguinal reconstruction. Plast Reconstr Surg. Sep; 2010 126(3):933–940. [PubMed: 20811226]

11. Chen CL, Shore AD, Johns R, Clark JM, Manahan M, Makary MA. The impact of obesity on breast surgery complications. Plast Reconstr Surg. Nov; 2011 128(5):395e–402e. [PubMed: 21788831]

12. Lerman B, Oldenbrook L, Eichstadt SL, Ryu J, Fong KD, Schubart PJ. Evaluation of chronic wound treatment with the SNaP wound care system versus modern dressing protocols. Plast Reconstr Surg. Oct; 2010 126(4):1253–1261. [PubMed: 20885246]

13. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. Plast Reconstr Surg. Aug; 2010 126(2):619–625. [PubMed: 20679844]

14. Colwell AS, Damjanovic B, Zahedi B, Medford-Davis L, Hertl C, Austen WGJ. Retrospective Review of 331 Consecutive Immediate Single-Stage Implant Reconstructions with Acellular Dermal Matrix: Indications, Complications, Trends, and Costs. Plastic and Reconstructive Surgery. 2011; 128(6):1170–1178. [PubMed: 22094736]

15. Caviggioli F, Maione L, Forcellini D, Klinger F, Klinger M. Autologous fat graft in postmastectomy pain syndrome. Plast Reconstr Surg. Aug; 2011 128(2):349–352. [PubMed: 21788826]

16. Gail M, Williams R, Byar DP, Brown C. How many controls? J Chronic Dis. Nov; 1976 29(11): 723–731. [PubMed: 791962]

17. Campbell CA, Butler CE. Use of adjuvant techniques improves surgical outcomes of complex vertical rectus abdominis myocutaneous flap reconstructions of pelvic cancer defects. Plast Reconstr Surg. Aug; 2011 128(2):447–458. [PubMed: 21788836]

18. MacQuillan A, Biarda FU, Grobbelaar A. The incidence of anterior belly of digastric agenesis in patients with hemifacial microsomia. Plast Reconstr Surg. Oct; 2010 126(4):1285–1290. [PubMed: 20885248]

19. Pandis N. Use of controls in clinical trials. Am J Orthod Dentofacial Orthop. Feb; 2012 141(2): 250–251. [PubMed: 22284295]

20. Pannucci CJ, Dreszer G, Wachtman CF, et al. Postoperative enoxaparin prevents symptomatic venous thromboembolism in high-risk plastic surgery patients. Plast Reconstr Surg. Nov; 2011 128(5):1093–1103. [PubMed: 22030491]

21. Sainsbury DCG, Kessell G, Fall AJ, Hampton FJ, Guhan A, Muir T. Intralesional Bleomycin Injection Treatment for Vascular Birthmarks: A 5-Year Experience at a Single United Kingdom Unit. Plastic and Reconstructive Surgery. 2011; 127(5):2031–2044. [PubMed: 21532430]

22. Chang DW. Lymphaticovenular bypass for lymphedema management in breast cancer patients: a prospective study. Plast Reconstr Surg. Sep; 2010 126(3):752–758. [PubMed: 20811210]

23. Arneja JS, Mulliken JB. Resection of amblyogenic periocular hemangiomas: indications and outcomes. Plast Reconstr Surg. Jan; 2010 125(1):274–281. [PubMed: 20048618]

24. Roostaeian J, Pavone L, Da Lio A, Lipa J, Festekjian J, Crisera C. Immediate placement of implants in breast reconstruction: patient selection and outcomes. Plast Reconstr Surg. Apr; 2011 127(4):1407–1416. [PubMed: 21460648]

25. Rohan TE, McMichael AJ. Alcohol consumption and risk of breast cancer. Int J Cancer. May 15; 1988 41(5):695–699. [PubMed: 3366491]
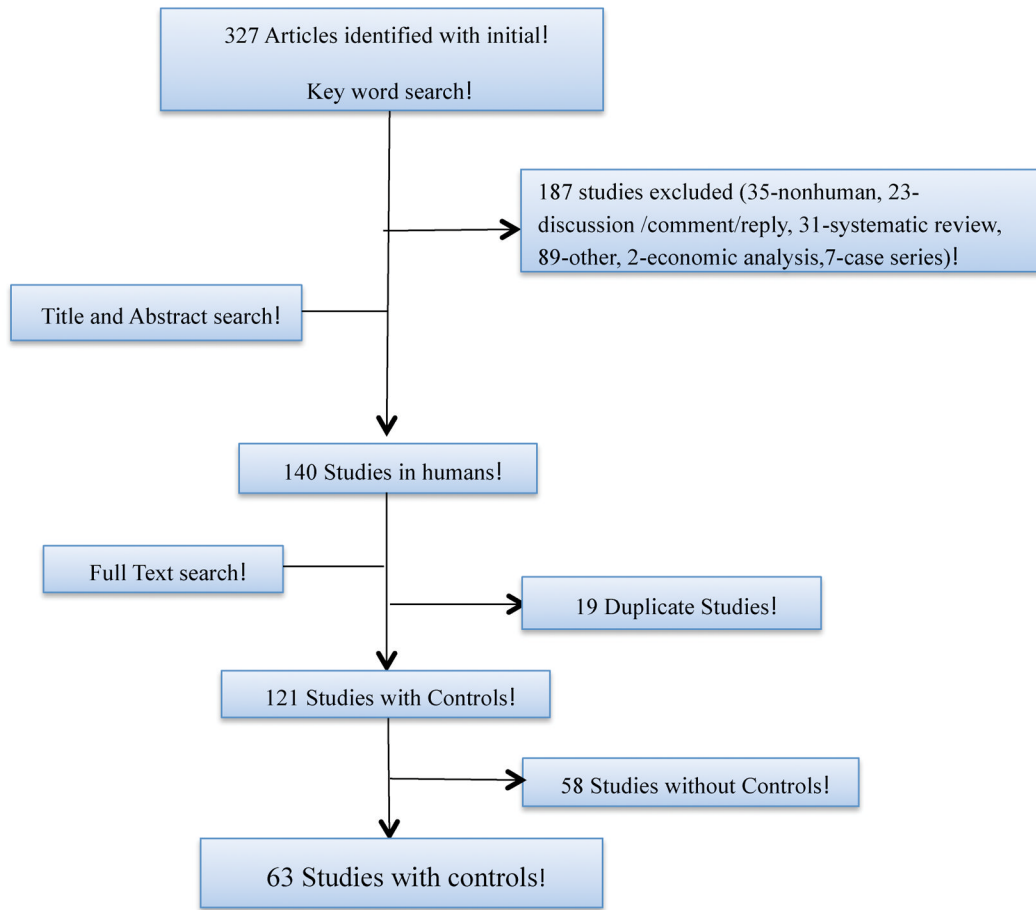
327 Articles identified with initial!

Key word search!

187 studies excluded (35-nonhuman, 23-discussion /comment/reply, 31-systematic review, 89-other, 2-economic analysis,7-case series)!

Title and Abstract search!

140 Studies in humans!

Full Text search!

19 Duplicate Studies!

121 Studies with Controls!

58 Studies without Controls!

63 Studies with controls!

Figure 1.
Trial flow diagram for studies obtained with PRS search

Table 1

Results of controls included in studies in PRS (years: 2010–2011)

Type Control Used	n
Historical control	2
Concurrent control	61
Different treatment control	40
No treatment control	9
Subjects as self-control	3
More than one control group	9

n- Number of studies

Table 2

Possible Bias with Controls and Preventive Measures

Type of control	Type of bias	Measures to prevent bias
Different treatment control	Channeling bias[1]: Patient comorbidities and prognostic factors compel the investigator to decide which treatment arm they will be placed, a more aggressive one versus one with fewer risks.	Randomization
No treatment control	Misclassification bias[2]: Improper assignment of the exposure or disease status. The bias may overestimate or underestimate the treatment effect.	Clear definitions, blinding, standard procedures, good measuring methods
Hospital controls	Do not represent the source populations from which cases arise.	Use population based controls, controls with disease not related to exposure. Select only for hospital cases
Friend/Spouse/Sibling control	Selection bias[3]: Due to overlapping nature of these controls. The bias may overestimate or underestimate the treatment effect.	Same selection criteria for cases and controls, high rate of follow-up and high response rate
Historic control	Chronology bias[1]: Due to differences in time and place of controls and treatment groups. Bias also possible due to differences in diagnostics, treatments and outcome assessment methods between the two comparative groups. The bias overestimates the treatment effect.	Use historic controls from very recent past or conduct a prospective cohort study or a randomized controlled trial

[1]Pannucci CJ, Wilkins EG. Identifying and Avoiding Bias in Research. *Plastic and Reconstructive Surgery*. 2010;126(2):619–625

[2]Jennings JM, Sibinga E. Research and statistics: understanding and identifying bias in research studies. *Pediatr Rev.* Apr 2010;31(4):161–162

[3]Austin H, Flanders WD, Rothman KJ. Bias arising in case-control studies from selection of controls from overlapping groups. *Int J Epidemiol.* Sep 1989;18(3):713–716

Table 3

Qualities of a Good Control

Good Control
Comparable to the cases because they are selected using a similar criteria as cases
Represent the population base from which cases arise
Identified during the same time period as the cases are sampled
Participate in only one clinical trial at a time
Free of exposure/intervention under study