Clin Orthop Relat Res (2019) 477:655-662
DOI 10.1097/CORR.0000000000000612

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Systematic Review

How Often Do Orthopaedic Matched Case-Control Studies Use Matched Methods? A Review of Methodological Quality

Drake G. LeBrun MD, MPH, Tram Tran MPH, David Wypij PhD, Mininder S. Kocher MD, MPH

Received: 28 August 2018 / Accepted: 29 November 2018 / Published online: 21 December 2018

Abstract

Background Case-control studies are a common method of analyzing associations between clinical outcomes and potential risk factors. Matching cases to controls based on known confounding variables can decrease bias and allow investigators to assess the association of interest with increased precision. However, the analysis of matched data generally requires matched statistical methods, and failure to use these methods can lead to imprecise or biased results. The appropriate use of matched statistical methods in orthopaedic case-control studies has not been documented.

Questions/purposes (1) What proportion of matched orthopaedic case-control studies use the appropriate matched statistical analyses? (2) What study factors are associated with the use of appropriate matched statistical tests?

Methods All matched case-control studies published in the top 10 orthopaedic journals according to impact factor from 2007 to 2016 were identified by literature review. Studies using appropriate statistical techniques were identified by two independent evaluators; discrepancies were settled by a third evaluator, all with advanced training in biostatistics. The number of studies using appropriate matched statistical methods was compared with the number of studies reviewed. Logistic regression was used to identify key study factors (including journal, publication year, rank according to impact factor, number of matching factors, number of controls per case, and the inclusion of a biostatistician coauthor) associated with the use of appropriate statistical methods. Three hundred nineteen articles that were initially classified as case-control studies were screened, yielding 83 matched case-control studies. One hundred two of the excluded articles were cohort or cross-sectional studies that were misclassified as case-control studies. The median number of matching factors was 3.0 (range, 1–10) and the median number of controls per case was 1.0 (range, 0.5–6.0). Thirty studies (36%) had a statistician coauthor.

Results Thirty of the 83 included studies (36%) used appropriately matched methods throughout, 11 (13%) used matched methods for multivariable but not univariable analyses, and 42 (51%) used only unmatched methods, which we considered inappropriate. After controlling for the number of controls per case and publication year, we found that the inclusion of a statistician coauthor (70%

Each author certifies that neither he nor she, nor any member of his or her immediate family, has funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

This work was performed at the Harvard T.H. Chan School of Public Health, Boston, MA, USA.

D. G. LeBrun, Hospital for Special Surgery, New York, NY, USA

D. G. LeBrun, T. Tran, D. Wypij , Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

T. Tran, Alpert Medical School of Brown University, Providence, RI, USA

M. S. Kocher, Division of Sports Medicine, Boston Children's Hospital, Boston, MA, USA

M. S. Kocher (✉), Division of Sports Medicine, Boston Children's Hospital, 300 Longwood Avenue, Fegen 2, Boston, MA, 02115, USA, Email: Mininder.Kocher@childrens.harvard.edu

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

versus 38%; odds ratio, 3.6; 95% confidence interval, 1.4–20.3; p = 0.01) and journal were associated with the use of appropriate methods.

Conclusions Although matched case-control studies can be statistically more efficient study designs, in that they are capable of generating more precise effect size estimates than other kinds of retrospective research, most orthopaedic case-control studies use inappropriate statistical methods in their analyses. Additionally, the high degree of study misclassification indicates a need to more rigorously define differences among case-control, cohort, and cross-sectional study designs.

Clinical Relevance Failing to use matched statistical tests may lead to imprecise and/or biased effect estimates, which may lead to a tendency to overestimate or underestimate associations between possible risk factors and clinically relevant outcomes. Orthopaedic researchers should be cognizant of the risks and benefits of matching and should consult individuals with biostatistical expertise as needed to ensure that their statistical methods are appropriate and methodologically rigorous.

Introduction

Case-control studies are a commonly used observational study design in clinical research [16]. They compare subjects with an outcome of interest (cases) and subjects without that outcome (controls) and determine whether various risk factors (exposures) differ between the two groups [4, 5]. Case-control studies have been referred to as "research in reverse" because cases and controls are defined by the presence or absence of an outcome and are analyzed retrospectively for exposure status [18]. Compared with cohort studies that follow subjects with different risk factors over time to determine the risk of developing an outcome, case-control studies are less resource-intensive and can be completed in less time. This makes them particularly useful for studying outcomes that are rare or have long latency periods [18]. Busse et al. [6] described the design, conduct, strengths, and limitations of case-control studies in the context of orthopaedics. Given that many outcomes in orthopaedic surgery are rare (for example, periprosthetic joint infections) or have long latency periods (for example, TKA after ACL reconstruction), case-control studies are helpful to efficiently identify risk factors for these outcomes. It is therefore not surprising that the number of studies with "case-control" and "orthopaedic" in the title or abstract has steadily risen over the past decade.

Matching cases to controls based on known confounders such as age or sex often is performed in case-control studies [3]. Matching reduces but does not eliminate confounding from the matching factors. Importantly, the primary function of matching is to improve the statistical efficiency (that is, precision) of the study by ensuring that cases and controls have similar distributions across confounders, thereby yielding more-precise effect estimates [16]. Most of the time, matched studies require statistical methods (Table 1) that account for statistical dependence across matched cases and controls [3]. The use of unmatched statistical methods in matched studies can lead to biased effect estimates and incorrect conclusions. For example, in a hypothetical study of unilateral TKA in which patients with aseptic loosening are matched individually to patients without aseptic loosening on the basis of age, sex, body mass index (BMI), and implant type, the matched pairs are considered "nonindependent" from each other. The use of unmatched tests would fail to account for the nonindependent subjects and may lead to either (1) a change in the strength of association between the covariates of interest and loosening; or (2) a change in the significance of the findings, especially if there are relatively small numbers of discordant sets (that is, only a few pairs of cases and controls have different exposure statuses). The use of matched tests, on the other hand, would account for nonindependent data and yield correct effect estimates. It is therefore important for orthopaedic researchers to use appropriate statistical techniques when analyzing matched data to minimize bias and strengthen their inferences. To our knowledge, the proportion of orthopaedic case-control studies that use appropriately matched statistical methods is not known, nor are the factors—if any—that might be associated with more-appropriate statistical analysis of these datasets.

Therefore, we asked (1) What proportion of matched orthopaedic case-control studies use the appropriate matched statistical analyses? (2) What study factors are associated with the use of appropriate matched statistical tests?

Materials and Methods

Systematic Review

We performed a systematic review to yield a representative sample of matched orthopaedic case-control studies. The search strategy was constructed with the assistance of a research librarian (CM). We identified studies published in 10 orthopaedic journals over a 10-year period from January 1, 2007 to December 31, 2016. Journals selected were the top 10 orthopaedic journals ranked by impact factor according to the 2017 Journal Citation Reports. We excluded journals focusing on basic science research or physical therapy. The 10 journals were *Acta Orthopaedica*, the *American Journal of Sports Medicine*, *Arthroscopy*, the *Bone & Joint Journal, Clinical Orthopaedics and Related Research®*, the *Journal of Arthroplasty,* the *Journal of Bone and Joint Surgery*, the *Journal of the American Academy of Orthopaedic Surgeons*,

Table 1. Some common unmatched and matched statistical analyses

Dependent variable (outcome)	Independent variable (predictor)	Unmatched analysis	Matched analysis
Binary	Binary	Pearson's chi-square test	McNemar's test
		Fisher's exact test	Mantel-Haenszel test
		Unconditional logistic regression	Conditional logistic regression
	Categorical	Pearson's chi-square test	Conditional logistic regression
		Fisher's exact test	
		Unconditional logistic regression	
	Continuous	Unconditional logistic regression	Conditional logistic regression
Continuous	Binary		Paired t-test (assuming normality of paired differences)
		Two-sample independent t-test (assuming normality of outcomes)	Wilcoxon signed rank test (nonparametric)
		Wilcoxon rank-sum test (nonparametric)	Linear regression (including fixed or random effects for matching factors)
	Categorical		One-way repeated measures ANOVA (assuming normality of outcomes)
		One-way ANOVA (assuming normality of outcomes)	Friedman's two-way ANOVA (nonparametric)
		Kruskal-Wallis test (nonparametric)	Linear regression (including fixed or random effects for matching factors)
	Continuous	Pearson's correlation	
		Linear regression	Linear regression (including fixed or random effects for matching factors)
		Spearman's rank correlation (nonparametric)	

Knee Surgery-Sports Traumatology-Arthroscopy, and *The Spine Journal*. We conducted a search with the following terms and Boolean operators: "case-control studies" (MeSH: NoExp) OR case control* AND match*.

One author (DGL) performed the initial search. All retrieved studies underwent an abstract review by two independent reviewers (DGL, TT) with the assistance of Rayyan systematic review software (Qatar Computing Research Institute, Doha, Qatar) [14]. Studies were excluded hierarchically if they met any of the following criteria: (1) not a clinical study (defined as a study reporting patient data and clinical outcomes); (2) not a case-control study; (3) not matched; or (4) used propensity score-matching. These criteria were chosen to be consistent with a prior similar study that was not limited to orthopaedics [13]. Case-control studies were defined according to Breslow and Day [4] as: "…an investigation into the extent to which persons selected because they have a specific disease (the cases) and comparable persons who do not have the disease (the controls) have been exposed to the disease's possible risk factors to evaluate the hypothesis that one or more of these is a cause of the disease." This definition is consistent with the classic epidemiologic understanding of case-control studies and similar evaluative studies in other specialties [4, 11, 12, 18, 19]. Only individually or frequency-matched studies were used, because other matching strategies (for example, propensity score-matching) require different statistical methods [1].

To ensure that we were capturing all matched case-control studies published in the journals chosen, we performed a hand-search of two journals (the *Journal of Bone and Joint Surgery* and the *American Journal of Sports Medicine*), each over a 1-year period (2016), to identify any studies that may have been missed by our original electronic search strategy. Of note, no additional studies were identified via the hand-search, suggesting that our electronic search strategy was rigorous enough to capture most, if not all, of matched case-controls studies in our target journals.

Data Extraction

All studies passing abstract review subsequently underwent a full-text review by two independent reviewers (DGL, TT). Studies were excluded if they were determined to have met any of the aforementioned criteria on full-text review. Extracted data for all studies included journal, publication year, body part involved (for example, hip, knee, shoulder,

etc), number of controls per case, number of matching factors, journal impact factor, and the inclusion of a coauthor with statistical or epidemiologic expertise. The latter was a binary (yes/no) variable determined by having either (1) an author affiliation with a department of biostatistics or epidemiology or a school of public health; or (2) an author with an advanced degree in biostatistics or epidemiology (for example, MPH, MS, or PhD). Authors with PhDs and nonMPH master's degrees who were not affiliated with a department of biostatistics or epidemiology or a school of public health were searched online to identify the subject of their advanced degree. These authors fell into three categories: (1) authors with confirmed biostatistics or epidemiology degrees (n = 2), (2) authors with advanced degrees who were confirmed to be something not related to biostatistics or epidemiology (n = 8), and (3) studies with MD/PhDs whose PhD topics were not confirmed but presumed to be in something other than biostatistics or epidemiology (n = 9). These criteria are consistent with other orthopaedic studies that have assessed statistical expertise [2, 10].

We defined an appropriate statistical analysis as one that used any matched statistical test to evaluate exposure-outcome relationships, assuming that the study matched at or near the individual level. The reported statistical methods of all included studies were independently evaluated by two reviewers (DGL, TT). Disagreements on the appropriateness of the methods were resolved by a senior biostatistician (DW). Studies were categorized according to the following system: (1) matched methods used for all analyses; (2) matched methods used for multivariable but not univariable analyses; and (3) only unmatched methods used. The first two categories were deemed to have used appropriate statistical methods, although studies in the second category used unmatched methods for univariable analyses. This decision was made to provide a more-lenient appraisal of the included studies. Additionally, it is unlikely, though not impossible, for studies to find different results if they use matched as opposed to unmatched univariable methods, assuming that their multivariable methods are matched regardless of their univariable methods.

The initial search strategy returned 319 articles. Two hundred thirty-six of these were subsequently excluded (Fig. 1), yielding 83 matched case-control studies. No additional studies were identified from a hand-search of two journals' studies published in 2016. Of the 236 excluded articles, 231 were excluded because they were not case-control studies. Of these, 102 reported their study design in the abstract or full text as a case-control study. However, on review of their methods, these studies were deemed to be more consistent with cohort or cross-sectional studies. Thus, the estimated proportion of study misclassification based on this sample is 55% (95% CI, 48%–62%).

Among the 83 retained studies (Table 2), the median number of matching factors was 3.0 (range, 1–10). The most common matching factors used were age (n = 67; 81%), sex (n = 60; 72%), BMI (n = 15; 18%), and surgery date (n = 13; 16%). One study did not mention the number or type of matching factors used. The median number of controls per case was 1.0 (range, 0.5–6.0). The median total sample size was 118 (range, 20–40,329). Thirty of 83 studies (36%) had a statistician coauthor. Twenty-eight studies (34%) were published from 2007 to 2011 and 55 studies (66%) were published from 2012 to 2016. The *American Journal of Sports Medicine* and *Journal of Bone & Joint Surgery* published the most matched case-control studies with 23 and 15 over 10 years, respectively.

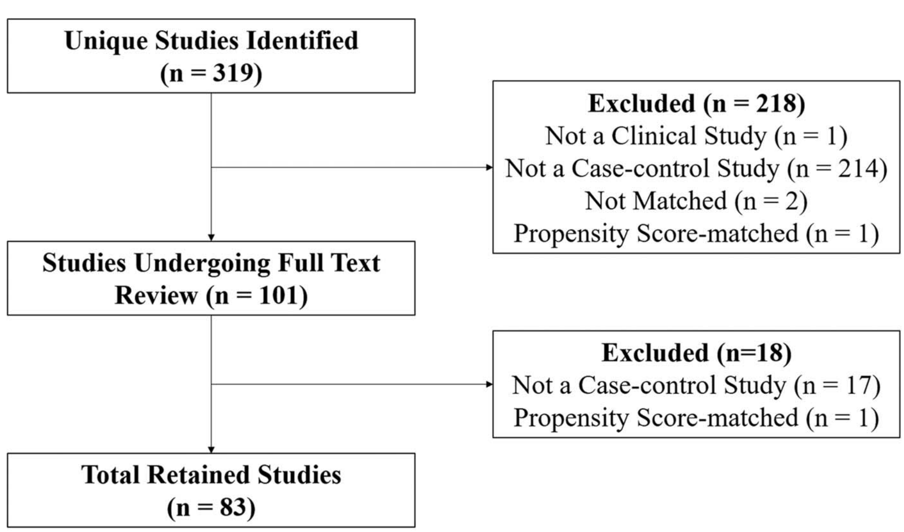

Fig. 1 This is a flow diagram of our search demonstrating the total number of studies reviewed and reasons for exclusion.

Table 2. Percent of included studies appropriately analyzing matched data by study characteristic (n = 83)

Study characteristic	Studies appropriately analyzing matched data
Journal	
Journal of Bone and Joint Surgery	12/15
American Journal of Sports Medicine	13/22
Arthroscopy	1/10
Clinical Orthopaedics and Related Research®	2/8
Knee Surgery, Sports Traumatology, Arthroscopy	5/10
Bone & Joint Journal	2/5
The Spine Journal	2/3
Acta Orthopaedica	2/2
Journal of Arthroplasty	2/8
Body part involved	
Knee	24/44
Hip	5/11
Knee and hip	2/7
Upper extremity	3/10
Spine	3/5
Foot/ankle	2/4
Multiple	2/2
Year of publication	
2007-2008	4/9
2009-2010	8/15
2011-2012	7/14
2013-2014	13/19
2015-2016	9/26
Statistician coauthor	
No	20/53
Yes	21/30
Number of controls per case	
≤ 1	22/50
> 1 and ≤ 2	7/18
> 2	12/15
Number of matching factors	
Not reported	0/1
1	5/10
2	8/26
3	12/22
≥ 4	16/24

Data are presented as # studies appropriately analyzing dependent observations/total # studies in subgroup. Journals are listed in order according to decreasing impact factor.

Statistical Methods

We used univariable logistic regression to determine whether journal, body part, year of publication, inclusion of a biostatistician coauthor, rank according to impact factor, number of controls per case, and number of matching factors were associated with the use of any appropriate matched methods. These factors were chosen to assess whether any associations in the use of appropriate matched methods emerged across the included studies and to be consistent with a prior similar study in an epidemiology journal [13]. Journals were compared in reference to the *Journal of Bone and Joint Surgery*. We analyzed year of publication and rank according to impact factor as ordinal variables. Number of controls per case and number of matching factors were analyzed as continuous variables. Subsequently, we used multivariable logistic regression to identify factors associated with the use of any appropriate matched methods. Any covariate with a p value of less than 0.2 on univariable analysis, or any covariate that we believed might be associated with the use of matched methods (publication year, journal, rank according to impact factor, and inclusion of a coauthor with statistical expertise), was included in the multivariable model. Missing data was not a significant issue in this study, as all articles were accessible. Only one study did not provide the number of matching factors. All other covariates were fully reported.

A two-sided type I error rate of 0.05 was used to indicate statistical significance. All calculations were performed using STATA 15.0 (Stata Corp, College Station, TX, USA).

Results

Proportion of Studies Using Matched Methods

Of the 83 case-control studies, 30 (36%) used appropriate analyses throughout, 11 (13%) used matched analyses for multivariable but not univariable analyses, and 42 (51%) used only unmatched analyses, which we considered inappropriate. For the purpose of our analysis, these studies were dichotomized into those that used any matched analysis (n = 41 [49%]) and those that did not (n = 42 [51%]). One study justified the use of unmatched analyses as a result of minimal dependence within matched groups [9]. This study matched patients based on surgeon and surgery date only, leading to reportedly highly variable demographic characteristics within matched groups and therefore minimal statistical dependence. Such justification is helpful in allowing the reader to determine whether data dependence was considered; however, no other studies that used unmatched methods provided a similar justification.

Study Factors Associated with the Use of Matched Methods

After controlling for the number of controls per case and publication year, we found that the inclusion of a statistician coauthor (OR, 3.6; 95% CI, 1.4–20.3; p = 0.01) and journal remained associated with the use of appropriate methods (Table 3). Rank according to impact factor was not included in the multivariable analysis due to collinearity with journal; of note, rank according to impact factor was not associated with the use of matched methods on univariable analysis.

Discussion

The use of appropriate statistical methods is critical to ensuring that study results are precise and valid. Most matched orthopaedic case-control studies rely on pair matching or small matched sets, therefore necessitating the use of matched statistical tests to control for data dependence between matched cases and controls. However, to our knowledge, the proportion of orthopaedic case-control studies using matched statistical tests has not previously been documented, despite the utility of this study design in answering clinical research questions that may not be easily answered in other ways. In this review of 83 orthopaedic case-control studies published over the past 10 years, we found that most matched case-control studies used unmatched statistical methods. Moreover, we found that the use of matched methods was associated with the inclusion of an author with statistical expertise, which points to an intervention that clinician-scientists can use to ensure the appropriateness of their statistical analysis.

Table 3. Multivariable analysis of associations between characteristics of included studies and the appropriate analysis of matched data (n = 83)

Study characteristic	OR (95% CI)	p value
Journal		0.01
Journal of Bone and Joint Surgery	ref	
American Journal of Sports Medicine	0.22 (0.04-1.23)	
Arthroscopy	0.04 (0.003-0.46)	
Clinical Orthopaedics and Related Research®	0.05 (0.004-0.46)	
Knee Surgery, Sports Traumatology, Arthroscopy	0.46 (0.07-2.97)	
Bone & Joint Journal	0.19 (0.02-1.95)	
The Spine Journal	0.86 (0.05-13.67)	
Acta Orthopaedica	n/a	
Journal of Arthroplasty	0.06 (0.006-0.56)	
Statistician coauthor	3.6 (1.4-20.3)	0.01
Year of publication	1.0 (0.8-1.2)	0.89
Number of controls per case	1.2 (0.7-2.1)	0.56

Odds ratios, confidence intervals, and p values are derived from multivariable logistic regression analysis; for ordinal predictor variables, such as year of publication and number of controls per case), odds ratios are reported per unit increase; a multiple degrees-of-freedom test was used to determine the multivariable p value associated with journal across its categorical variables; *The Journal of Bone and Joint Surgery* was used as a reference for other odds ratios; both studies from *Acta Orthopaedica* correctly used matched methods; therefore, no odds ratio could be calculated. CI = confidence interval; OR= odds ratio; ref = reference group.

Our study has limitations. First, it is important to acknowledge that in matched case-control studies in which there are many subjects per matched stratum or minimal dependence among matched subjects, it may not be necessary to use matched analyses and it may, in fact, be preferable to use an ordinary unmatched stratified analysis [15]. Although this may hold true in select circumstances, most studies in orthopaedics are matched at or near the individual level with sparse data within strata and therefore should be analyzed using matched methods. This was true in our study sample in which most studies matched at the individual level, necessitating a matched analysis. Additionally, our methodology of classifying authors with biostatistical expertise is imperfect. Most studies had coauthors who clearly did or did not have credentials consistent with biostatistical expertise. Specifically, 30 studies had authors with either an advanced degree in statistics or an affiliation with a department of biostatistics or epidemiology. Similarly, 44 studies had authors with no advanced degrees in statistics and no affiliations with a department of biostatistics or epidemiology. A subset of studies (n = 9) had authors with MD/PhDs whose PhDs were not clearly identifiable. It is possible that some of these authors, as well as authors without other advanced graduate degrees or biostatistical affiliations, may indeed have relevant biostatistical expertise. Although this is an important limitation of our study, we do not feel that this detracts from the overarching finding that including individuals with advanced biostatistical training can help researchers ensure that their methods are statistically rigorous and correct for the type of analysis being performed.

Additionally, we relied on a sample of matched case-control studies from 10 orthopaedic journals over a 10-year period. Although it is possible that including additional studies from other orthopaedic journals or using a longer time period would change our findings, we believe that our sample is sufficiently large to provide an appropriate snapshot of orthopaedic research over the past decade. Lastly, we did not assess other study designs with matched data that may require matched statistical methods. For

example, matched cohort studies match exposed and unexposed subjects based on potential confounding factors rather than on their outcome status [8]. Although cohort studies are conceptually different from case-control studies, the analysis still requires matched statistical methods. Given the relative efficiency of matched case-control studies in orthopaedics, we limited this review to case-control studies, but it is possible that similar findings might be seen in matched cohort studies.

Our finding that the majority of orthopaedic case-control studies incorrectly used unmatched tests is consistent with two other studies that evaluated the appropriateness of reported statistical methodologies in "core clinical journals" across various medical specialties [13] and nursing [7] (Table 4). Unlike our study, Niven et al. [13] found that a higher impact factor and having multiple controls per case were both associated with the use of appropriate methods. We did not identify differences across rank according to impact factor in our study; however, Niven et al. [13] used a set of 100 journals with a wider range of impact factors, which may explain why they identified this predictor of appropriate methodology. Similarly, we did not identify an association between the ratio of controls per case and the use of appropriate methods. Although having more controls per case may lead to a more precise effect estimate, there is no clear relationship between having more controls per case and correctly using matched methods. They did not assess the inclusion of a statistician coauthor, a finding that was associated with the use of appropriate methods in our study. This finding is not surprising, because investigators with a formal background in biostatistics or epidemiology may be more cognizant of the analytical differences between matched and unmatched studies and in the appropriate analysis of matched data. To this end, we recommend that clinician-scientists conducting any type of analytical observational study (including case-control, cohort, cross-sectional, and case series) seek statistical expertise to ensure the soundness and appropriateness of their analysis. Given the substantial amount of time and effort that goes into conducting analytical observational studies, it seems reasonable to include authors with statistical expertise to ensure that their methods are statistically rigorous and correct for the type of analysis being performed.

Additionally, orthopaedic clinician-researchers and reviewers should be cognizant of a few key limitations of matching. First, matching does not completely eliminate confounding by the matching factor if there is a remaining association between the matching factor and the exposure of interest. Second, matching can introduce additional bias by making the controls and cases more similar to each other with respect to exposure. Matching can also reduce sample size by excluding unmatched cases and controls within a sample, which may lead to selection bias in the final sample. Third, matching in a case-control study precludes estimating the effect of the matching variable on the outcome because the association is artificially changed through matching. As mentioned, the major reason for matching is not to control for confounding, but rather to produce a more statistically efficient (precise) estimates. Choosing the appropriate matching factors to achieve this goal is complex. For example, matching on a factor that is associated with the exposure only (that is, overmatching) may decrease precision. Conversely, matching on a factor that is associated with the outcome only (that is, unnecessary matching) may be uneconomical without increasing statistical precision. Lastly, matching on a factor that is associated with an intermediate variable between exposure and outcome may introduce an irreparable bias in both a crude and matched analysis.

An unanticipated but important finding from this study was a high degree of study misclassification. More than 100 studies were self-identified as case-control studies but were determined to be more consistent with cross-sectional or cohort studies. This is similar to a study of six high-impact surgical journals, which found that only 35% of self-identified "case-control" studies were truly case-control studies [11]. Many of the misreported studies in our sample retrospectively identified patients with different exposures (for example, different interventions or different comorbidities) and assessed outcomes between exposed and unexposed groups. This study design is more consistent with a retrospective cohort study than a case-control study. The high degree of study misclassification in our sample suggests a need to more clearly outline differences between observational study designs in orthopaedic research. Study misclassification threatens the clarity of a study by misrepresenting its primary aim. For example, the

Table 4. Comparison of our study results and other similar studies in the medical literature

Characteristic Subject area	Current study Orthopaedics	Niven et al. [14] Core clinical specialties	Conway et al. [7] Nursing
Number of studies reviewed	83	37*	41
Number (%) of studies using unmatched methods	42 (51%)	21 (57%)	31 (76%)

*Only one of 37 studies reviewed was published in an orthopaedic journal.

overarching objective of case-control studies is to identify exposures (risk factors) associated with specific outcomes. Assessing the incidence of different outcomes across exposure groups is not possible with case-control studies, as these studies, by definition, start by identifying groups of individuals with and without a given outcome of interest. Portraying a retrospective cohort study as a case-control study suggests that a study's primary aim is to evaluate possible risk factors for a specific outcome, even though the authors may have actually intended to assess differential risk of outcomes across groups with different risk factors. Additional studies should be performed to more thoroughly characterize the degree of study misclassification in orthopaedics and the extent to which that misclassification may affect readers' interpretations of the data. Furthermore, matching represents one of several potential pitfalls associated with case-control studies. Sackett et al. [17] described 35 biases arising from case-control analyses. A followup study evaluating the degree to which orthopaedic case-control studies acknowledge and account for these various sources of bias would be both helpful in contextualizing past research as well as ensuring higher quality analyses in future orthopaedic work.

The present analysis demonstrated that most matched case-control studies in high-impact orthopaedic journals used inappropriate statistical methods. Importantly, given that problems with matched methods were seen across the full span of the study period and in all the journals we evaluated, orthopaedic investigators, reviewers, and journal editors should be aware of statistical tests for matched data and, when necessary, consult individuals with expertise in biostatistics to obtain more statistically efficient and potentially less-biased findings.

Acknowledgments We thank Carol Mita, a research librarian at Countway Library of Medicine, Harvard Medical School, for her assistance formulating and revising the search strategy.

References

1. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res.* 2011;46:399–424.
2. Bhandari M, Morrow F, Kulkarni A, Tornetta P. Meta-analyses in orthopaedic surgery. A systematic review of their methodologies. *J Bone Joint Surg*. 2001;83-A:15–24.
3. Bland J, Altman D. Statistics notes: matching. *BMJ*. 1994;309:1128.
4. Breslow N, Day N. Statistical methods in cancer research. Volume I - The analysis of case-control studies. *IARC Sci Publ*. 1980;32:5–338.
5. Breslow NE. Statistics in epidemiology: the case-control study. *J Am Stat Assoc.* 1996;91:14–28.
6. Busse JW, Obremskey WT. Principles of designing an orthopaedic case-control study. *J Bone Joint Surg.* 2009;91:15–20.
7. Conway A, Rolley JX, Fulbrook P, Page K, Thompson DR. Improving statistical analysis of matched case-control studies. *Res Nurs Health*. 2013;36:320–324.
8. Cummings P, McKnight B, Greenland S. Matched cohort methods for injury research. *Epidemiol Rev.* 2003;25:43–50.
9. Krych AJ, Sousa PL, Morgan JA, Levy BA, Stuart MJ, Dahm DL. Incidence and risk factor analysis of symptomatic venous thromboembolism after knee arthroscopy. *Arthroscopy*. 2015;31:2112–2118.
10. Leopold SS, Warme WJ, Fritz Braunlich E, Shott S. Association between funding source and study outcome in orthopaedic research. *Clin Orthop Relat Res.* 2003;415:293–301.
11. Mihailovic A, Bell CM, Urbach DR. Users' guide to the surgical literature: Case-control studies in surgical journals. *Can J Surg.* 2005;48:148–151.
12. Nesvick C, Thompson C, Boop F, Klimo P. Case-control studies in neurosurgery: A review. *J Neurosurg.* 2014;121:285–296.
13. Niven DJ, Berthiaume LR, Fick GH, Laupland KB. Matched case-control studies: a review of reported statistical methodology. *Clin Epidemiol*. 2012;4:99–110.
14. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A, Chalmers T, Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A, Uman L, Moher D, Liberati A, Tetzlaff J, Altman D, Wallace B, Trikalinos T, Lau J, Brodley C, Schmid C, Tsafnat G, Glasziou P, Choong M, Dunn A, Galgani F, Coiera E, O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S, Cortes C, Vapnik V, Cohen A, Hersh W, Peterson K, Yen P. Rayyan—a web and mobile app for systematic reviews. *Syst Rev.* 2016;5:210.
15. Pearce N. Analysis of matched case-control studies. *BMJ*. 2016:i969.
16. Rose S, van der Laan M. Why match? Investigating matched case-control study designs with causal effect estimation. *Int J Biostat.* 2009;5:1–24.
17. Sackett DL. Bias in analytic research. *J Chronic Dis.* 1979;32:51–63.
18. Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet*. 2002;359:431–434.
19. Song J, Chung K. Observational studies: cohort and case-control studies. *Plast Reconstr Surg*. 2010;126:2234–2242.