Bias in Hand surgical randomized controlled trials – systematic review and meta-epidemiological study

Juuso Heikkinen[1], Jarkko Jokihaara[2], Soumen Das De[3], Kati Jaatinen[4], Rachelle Buchbinder[5], Teemu Karjalainen[6]

1. PhD, MD, University of Oulu, Division of Orthopedic and Trauma Surgery, Department of Surgery, Oulu University Hospital, Medical Research Center, Oulu, Finland.
2. Associate Professor; Department of Hand Surgery, Tampere University Hospital, 33520 Tampere, Finland; Faculty of Medicine and Life Sciences, 33014 Tampere University, Finland
3. MBBS, FRCS, MPH, Consultant, Department of Hand & Reconstructive Microsurgery, National University Health System, Singapore
4. PT, University of Jyväskylä, Finland
5. Professor, NHMRC Senior Principal Research Fellow Director, Monash Department of Clinical Epidemiology, Cabrini Institute; Department of Epidemiology and Preventive Medicine, School of Public Health & Preventive Medicine, Monash University; Coordinating Editor, Cochrane Musculoskeletal Group Monash-Warwick Honorary Professor, Clinical Trials Unit, Warwick Medical School, Warwick University
6. Adjunct professor, MD, Unit of Hand Surgery, Department of Surgery, Central Finland Central Hospital, Jyväskylä, Finland

Our study protocol has been published at PROSPERO database (ID: CRD42019122710).

Corresponding author; Juuso Heikkinen, Oulu University Hospital, Kajaanintie 50, 90220 Oulu, phone: +358407031233, email: juuso.heikkinen@fimnet.fi.

We declare no funding or conflict of interest for any of the investigators for this manuscript.

Bias in Hand surgical randomized controlled trials – systematic review and meta-epidemiological study

Abstract

Backround: Inadequately reported or conducted studies do not contribute to valid scientific knowledge and they may decrease the quality of care due to under- or overestimation of the benefits or harms of interventions. Our aim was to evaluate how often hand surgical RCTs use and report adequate methods to ensure internal validity, and if inadequate reporting or methods associate with the magnitude of treatment effect estimates (difference between the groups).

Methods: Data Sources were Cochrane Central Register of Controlled Trials, MEDLINE and Embase databases until November 2020. We included published RCTs investigating the effect of any surgical intervention in hand and wrist region. We assessed internal validity using the Cochrane RoB tool for six domains: selection, performance, detection, attrition, selective reporting and 'other' bias. We extracted the primary outcome and calculated effect size for each study. We used mixed-effect meta-regression to assess if RoB modified magnitude of the effects.

Results:

For 207 assessed trials, risk for bias was unclear or high for 72% in selection, 93% in performance, 88% in detection, 25% in attrition, 22% in selective reporting and 34% in 'other' bias domain. Trials with high of unclear risk of selection bias yielded 0.28 SMD (95% CI 0.02 to 0.55) larger effect sizes compared to studies with low risk. RoB for other domains did not modify the intervention effects. The risk for selection bias declined over time, the OR per additional year for high or unclear risk of bias was 0.90 (95%CI 0.85 to 0.95).

Conclusions and clinical relevance:

The internal validity and credibility of hand surgical RCTs can be improved by using established methods to achieve 1) true randomization, 2) blinding of the participants and study personnel, 3)

publishing the trial protocol and avoiding selective reporting the outcomes, and 4) reporting the trial as recommended in the CONSORT statement.

Registration number: ID: CRD42019122710

Introduction

RCTs are the gold standard for evaluating causal effects of therapy. However, RCTs can yield untrustworthy results due to limitations in their internal and/or external validity [1]. In 2009, Chalmers and colleagues estimated clinical research "waste" to be a staggering 85% of global research investment [2], and one of the reasons for this wastage was methodological flaws in RCTs. Poorly conducted or reported studies are not only a waste of research resources but may cause harm when ineffective or less effective treatment is administered due to lack of precise estimates of the benefits and harms of the procedures.

The utility of an RCT depends on its internal and external validity. External validity relates to the generalizability of the study depending on inclusion criteria and interventions being evaluated. Internal validity relates to how close the results are to the "truth" and can be compromised by flaws in the design, conduct and analysis of the study [3]. Inadequate reporting does not directly cause bias in the study but it leaves uncertainty if the estimates should be trusted.

A systematic error, or *bias*, causes the treatment estimate to consistently deviate from the true value in a particular direction. This must be distinguished from *random error*, which is defined as imprecision related to variations in sampling and measurement. Random error causes the point estimate to deviate in either direction around the "true" outcome and can be decreased by increasing the sample size. But it is impossible to "account for" biases once they have been introduced into a

study. Increasing the sample size or conducting meta-analyses of biased studies can decrease uncertainty around *flawed point estimate* aggravating the problem if the direction of bias is consistent [4]. Numerous meta-epidemiological studies suggest that trials with Risk of Bias (RoB) yield different results than unbiased trials, particularly for studies with subjective outcomes [3].

Much of the evidence in hand surgery is based on biomechanical research, case series and observational studies. High-quality RCTs that should form the basis to the treatment decisions remain infrequent within the specialty[5]. The internal validity of these trials is unknown.

The primary objective of this systematic review and meta-epidemiological study was to assess the extent of risk bias in hand surgical RCTs. We also tested the hypothesis, that studies with inadequate methods or reporting (high or unclear RoB) yield on average larger effect sizes compared with adequately reported and conducted studies (low RoB). We also explored if the risk of bias decreased over time.

Methods

Our study protocol has been published at PROSPERO database (ID: CRD42019122710). We adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

We included all published RCTs with human participants investigating efficacy of surgical intervention in the hand and wrist region without language restrictions. We defined “surgery” as a procedure requiring general, regional or local anaesthesia and a skin incision. We did not consider injections as surgery even though the injection was given with local anaesthetic and/or performed

by a surgeon. All trials investigating surgery were included irrespective of the control arm/s or outcomes being measured.

We excluded studies that included surgical treatment but assessed effects other than the surgery itself, such as pre- or postoperative protocols, anaesthesia, or ex-vivo trials.

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and Embase databases from their inception until 7 November 2020. The search strategies are listed in the eTable 1. Duplicates were removed before screening. Two review authors screened the titles and abstracts for potentially eligible studies. We acquired full texts for eligible trials, and two authors independently read the full texts and identified eligible publications. Any discrepancy between the two assessors was settled by negotiation or by a third arbiter.

Data extraction

Two authors independently extracted data from included studies, and discrepancies were settled through discussion. Extracted data included: Publication year, protocol registration and publication, type of comparison, condition, primary outcome, and effect size for the primary outcome. The hierarchy for the extraction of the primary outcome when authors had not defined is found in the protocol; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=122710

Risk of bias assessment

Two authors independently assessed the RoB in each RCT according to the Cochrane Risk of Bias (RoB) tool (v 1.0). Discrepancies were settled with discussion. This tool divides possible sources of bias into six domains; risk for selection, performance, detection, attrition, selective reporting and 'other' bias [6]. We categorised the RoB in each of the domains as low, unclear, or high RoB

according to the tool, and for the meta-regression we dichotomized the judgement as either 'low risk' or 'high risk' (high or unclear risk)[6, 7].

For detection and performance bias, blinding was deemed impractical when it could not have been achieved with reasonable measures or the act of blinding would obliterate potential benefits of the treatment being evaluated (e.g., external fixation versus cast or percutaneous treatment versus surgery).

Data handling and analyses

All effect sizes were converted to standardized mean differences (SMD). We used standard error of mean (SEM), 95% confidence intervals (95% CI) or p-values to calculate standard deviation (SD) when this was not reported. In studies reporting medians and interquartile ranges (IQR), we calculated SMD by using median value as the approximation of the mean and estimated SD as either IQR/1.35 or range/4. For binary outcomes, we calculated odds ratios (OR) and then converted them to SMD using logit method (SMD= ln OR/1.81) [8, 9]. We excluded the study from the meta-regression if the authors did not report any parameter to calculate or estimate variance.

We performed two meta-regression analyses to assess if the ROB affected the effect size. First, we explored if outcomes were superior in the experimental arm in those studies that expressed a some statement or hypothesis that one treatment (the experimental arm) would have superior outcomes to the control arm/s. We converted the SMD so that a positive SMD signified better outcome in the experimental arm. We then entered the data in a mixed-model (random effect within subgroups and fixed effect between subgroups) meta-regression. Second, we explored if the absolute difference between the groups (regardless of the direction) was larger in studies with high ROB. For this meta-regression, we used data from all studies where we could calculate the effect size with standard

error (SE), and used a mixed-model meta-regression as described. R meta 4.12 package was used for the meta-regression analyses. We used binary logistic regression to assess if the ROB in each of the six domains had changed over time using proportion of studies with low risk as dependent variable and year of publication as a covariate in the model.

Results

We identified 207 trials (Figure 1) with 14539 participants in the analyses. For Lian et al. 2016, we could only extract data from the English abstract [10]. Included study refences are presented in eTable 2.

The trials were published between years 1982 and 2020 (Figure 2).

Of these, 179 (86%) trials were conducted in a single center and 28 (14%) in two or more centers. Only 38 (18%) of the trials were registered in publicly available trial registries, and a protocol was published for six trials (3%). The conditions evaluated in these trials are summarized in Table 1.

In 171 (83%) studies, two different surgical procedures were compared. Surgery was compared with a non-operative modality in 32 studies (15%) and surgery versus injection or needle fasciotomy in 4 studies (2%). We identified no studies comparing surgery with non-treatment or sham surgery. 142 (69%) trials evaluated treatment of traumatic conditions

Figure 3 illustrates the overview of how studies fared in each bias domain. Main issues were leck of blinding as well as inadequate reporting of random sequence generation or allocation concealment. Only one RCT had low risk of bias in all domains [11]. The assessments of RoB for each included study are presented in the eTable 3.

There was no attempt at blinding in any of the 27 trials that compared surgery with non-operative treatment. Blinding was not deemed feasible in 160 (77%) studies. Only 16 of the 47 (34%) studies where blinding would have been feasible were deemed to use successful blinding of outcome assessment.

The risk for selection bias declined over time, the OR per additional year for high risk of bias (high or unclear risk) was 0.89 (95%CI 0.84 to 0.94). Regarding detection bias (blinding of outcome assessment), we did not find evidence of decline of the ROB over time OR 0.0 (95% CI 0.92 to 1.07).

172 studies (83%) reported sufficient data to calculate SMD to assess if the risk of bias modified the size of the treatment effect. Of these, we estimated SD based on IQR or range for 22 studies and converted OR to SMD for 41 studies. We could differentiate between the experimental and control arms in 97 (47%) studies, and these studies contributed to the meta-regression assessing if the risk of bias modified the treatment effect size and direction (i.e. benefit or harm)

Trials with a high risk of selection bias yielded, on average, 0.28 SMD larger benefits for experimental arm compared to with studies with low risk (Table 2). The risk of bias in the other domains did not significantly modify the effect sizes but the confidence intervals were wide regarding detection and performance bias due to low number of studies with adequate blinding.

Discussion

Poor reporting or inadequate methods can bias the results, and even when they don't, they cause uncertainty if the evidence reflects the true efficacy of interventions. Our systematic review assessing the methods and reporting in hand surgical RCTs reveal that important methodological processes to safeguard against biases are frequently ignored or not reported. We observed small improvement in reporting or methods along with the time. Simple measures such as true randomization, blinding of the participants and personnel as well as transparent reporting would improve the reliability results and may improve clinical practice.

The RCT is the gold standard to determine *causal* effects of interventions. The key principle of an RCT is clinical equipoise, i.e. the experimental and control groups have, on average, similar expected future outcome distribution. If the equipoise is maintained throughout the study, the observed difference in the outcome (beyond random error) is caused by the assigned treatment. [11, 12]. Flaws in the study design that introduce inherent systematic errors – or bias – at any point of the study will cause the treatment effect to deviate from the true value, often in unknown ways. Although it is unclear how these flaws affect the results of one study, they affect the degree of trust we can place on the results [13]

Although there was improved reporting with regard to selection bias over time, only 65% of studies had a low risk of selection bias between 2015 and 2020, several years after the CONSORT statement was published (year 1996), and the journals integrated them into their guidelines. Adequate randomization is the fundamental first step for rigorous RCT, and it is based on the premises that there is true random sequence generation with adequate allocation concealment. Prior knowledge of the manner of allocation may cause conscious or unconscious selection of participants based on their baseline characteristics. For example, investigators may exclude a

potential participant with a more comminute distal radius fracture if they were aware of the coming allocation to the experimental implants.

In this study, risk of selection bias overestimated benefits in the experimental arm by 0.28 SMD. This corresponds to a small- to medium effect size. Our findings corroborate the findings of previous meta-epidemiological studies. A sample of studies from the Cochrane pregnancy and childbirth database found that studies with inadequate allocation concealment resulted in 40% higher OR compared to studies with adequately concealed allocation [14]. Another large meta-epidemiological study found that studies with inadequate random sequence generation and allocation concealment exaggerated effect estimates by 10% [15].

Blinding of participants and investigators is essential throughout the trial after the randomization. The nature of the interventions in surgery makes blinding unfeasible in many instances, such as comparing external fixation versus cast immobilization of distal radius fractures. Knowledge of the received treatment may cause both the study subjects and investigators to systematically deviate from the study protocol (performance bias). Examples of this include differential utilization of post-operative therapy and medications guided by pre-conceived perceptions. Awareness of the received intervention may also affect the reporting or recording of outcomes, referred as detection bias. For instance, a surgeon who prefers plates over K-wires for phalangeal fractures may be more likely to report complete union on radiographs when plates were used.

We found that blinding was adequately achieved and reported in less than 11% of all the studies, and even when we deemed it was feasible, only 34% of trialists had an attempt at doing so. We were unable to detect any effect of blinding on the effect size. However, the limited number of studies with adequate blinding diminished the power to detect small- to medium effects. There is a

paucity of meta-epidemiological studies evaluating blinding specifically in surgery, but several recent systematic reviews that have investigated this issue have demonstrated that blinding may affect the outcomes, effect sizes varying from no effect to 36% exaggeration of effect [16-24].

Unlike trials involving medications, blinding in surgical trials can be extremely challenging. Placebo (or sham) surgery may be the only method to reliably achieve blinding in trials comparing surgery versus a non-operative treatment, but ethical and practical issues inevitably arise. The investigators may consider blinded evaluation of outcome or an outcome that is not prone to bias. The problem with objective outcomes is that patients may not consider differences meaningful in their daily lives.

Attrition bias was low in most (75%) studies. Our study likely underestimates the true rate of attrition as many of the studies did not specifically report details or the reasons for participant attrition. Missing follow-up data does not automatically bias the results if the "missingness" is random. However, it is likely that systematic reasons for loss of follow-up, such as a good outcome or the occurrence of adverse events, result in differential attrition. Thus, it is important to report the reasons for missing data along with the numbers. The existing evidence regarding the effect of attrition bias is generally inconsistent and estimates are imprecise, especially in surgical trials [3, 19].

Trials with high risk of selective reporting did not impact between-group differences in this study. However, we were unable to asses properly this RoB for most of the trials, since only three trial protocols were published and only 13% of the trials were registered. Trials with selective reporting are more likely to show a significant treatment effect [25-27]. It is also important that the statistical methods used for comparisons are reported *a priori* as the methods may affect the conclusions [28]. Thus, to achieve better transparency, we recommend that hand surgical trials are preregistered.

Finally, "other" biases relate to any important concerns not addressed in the other domains of the Cochrane RoB tool. These include problems with inclusion and exclusion criteria, usage of co-interventions, unplanned deviations from the protocol, differential diagnostic activity, potential conflict of interest and selective reporting of subgroups [6]. Decrease in high risk of other bias over years may reflect a general progress in the conduct and reporting of surgical trials.

There are several limitations in this study. First, the quality of the data in any meta-analysis is only as good as what is reported. Many studies lacked methodological details, and this limited our ability to assess true RoB. Second, very few studies were blinded, and this decreased the power to detect any effects of performance and detection bias. Third, it is possible that we may have missed some hand surgery RCTs despite performing a search. Fourth, we estimated the bias in an indirect way as we cannot know the "true" underlying effect to which effect sizes should be compared. Also, the confidence intervals were wide and suggested that the effect of inadequate randomization may be also close to zero. Imprecision of the estimates from meta-regression is related to the small sample sizes of included trials and could not be controlled in any way. Finally, the assessment of bias itself may introduce a degree of subjectivity.

In conclusion, we observed limitations in the internal validity of hand surgical RCTs arising mainly from poor conduct and/or reporting of randomisation process and inadequate blinding. The flaws in randomization or failure to report it adequately seemed to associate with larger treatment effects. Complying with the established guidelines in planning and conducting the trials (SPIRIT and CONSORT) would greatly improve credibility of the results, ultimately resulting in better clinical practices. Investigators, peer reviewers and journal editors are in key position to address these problems. [29]

References

1. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ. 2001 Jul 7;323(7303):42-6. doi: 10.1136/bmj.323.7303.42.
2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet. 2009 Jul 4;374(9683):86-9. doi: 10.1016/S0140-6736(09)60329-9. Epub 2009 Jun 12.
3. Berkman ND, Santaguida PL, Viswanathan M, Morton SC. 2014 Sep.
4. Boutron I PM, Higgins JPT, Altman DG, Lundh A, Hróbjartsson A. . Chapter 7: Considering bias and conflicts of interest among the included studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). 2019.
5. Sugrue CM, Joyce CW, Sugrue RM, Carroll SM. Trends in the Level of Evidence in Clinical Hand Surgery Research. Hand (N Y). 2016 Jun;11(2):211-5. doi: 10.1177/1558944715627619. Epub 2016 Feb 26.
6. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011 Oct 18;343:d5928.(doi):10.1136/bmj.d5928.
7. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019 Aug 28;366:l4898.(doi):10.1136/bmj.l4898.
8. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook. 2019.
9. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med. 2000 Nov 30;19(22):3127-31. Epub 2000/12/13.
10. Lian ZM YJ, Zhang TL, Ma C, Liu Q, Yang GZ. . Bridging extema/ fixation combined with Kirschner-wire fixation versus volar locked plate fixation for unstable fractures of the distal radius. Zhongguo Zuzhi Gongcheng Yanjiu. 2016;2016;20(44)·6590-6598.
11. Freedman B. Equipoise and the ethics of clinical research. N Engl J Med. 1987 Jul 16;317(3):141-5. doi: 10.1056/NEJM198707163170304.
12. AJ L. Clinical equipoise: foundational requirement or fundamental error? In: Steinbock B, ed The Oxford Handbook of Bioethics Oxford, UK: Oxford University Press; 2007:571-595. 2007:571-95.
13. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924-6. doi: 10.1136/bmj.39489.470347.AD.
14. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995 Feb 1;273(5):408-12. doi: 10.1001/jama.273.5.408.
15. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012 Sep 18;157(6):429-38. doi: 10.7326/0003-4819-157-6-201209180-00537.

16. Saltaji H, Armijo-Olivo S, Cummings GG, Amin M, da Costa BR, Flores-Mir C. Influence of blinding on treatment effect size estimate in randomized controlled trials of oral health interventions. BMC Med Res Methodol. 2018 May 18;18(1):42. doi: 10.1186/s12874-018-0491-0.
17. Saltaji H, Armijo-Olivo S, Cummings GG, Amin M, da Costa BR, Flores-Mir C. Impact of Selection Bias on Treatment Effect Size Estimates in Randomized Trials of Oral Health Interventions: A Meta-epidemiological Study. J Dent Res. 2018 Jan;97(1):5-13. doi: 0.1177/0022034517725049. Epub 2017 Aug 16.
18. Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in Physical Therapy Trials and Its Association with Treatment Effects: A Meta-epidemiological Study. Am J Phys Med Rehabil. 2017 Jan;96(1):34-44. Epub 2016/05/06.
19. Moustgaard H, Clayton GL, Jones HE, Boutron I, Jorgensen L, Laursen DRT, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. BMJ. 2020 Jan 21;368:l6802.(doi):10.1136/bmj.l6802.
20. Moustgaard H, Bello S, Miller FG, Hrobjartsson A. Subjective and objective outcomes in randomized clinical trials: definitions differed in methods publications and were often absent from trial reports. J Clin Epidemiol. 2014 Dec;67(12):1327-34. Epub 2014/09/30.
21. Hrobjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Rasmussen JV, Hilden J, et al. Observer bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. Int J Epidemiol. 2014 Jun;43(3):937-48. Epub 2014/01/23.
22. Hrobjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. Cmaj. 2013 Mar 5;185(4):E201-11. Epub 2013/01/30.
23. Hrobjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. Bmj. 2012 Feb 27;344:e1119. Epub 2012/03/01.
24. Dechartres A, Trinquart L, Faber T, Ravaud P. Empirical evaluation of which trial characteristics are associated with treatment effect estimates. J Clin Epidemiol. 2016 Sep;77:24-37.(doi):10.1016/j.jclinepi.2016.04.005. Epub Apr 29.
25. Fergusson D, Laupacis A, Salmi LR, McAlister FA, Huet C. What should be included in meta-analyses? An exploration of methodological issues using the ISPOT meta-analyses. Int J Technol Assess Health Care. 2000 Autumn;16(4):1109-19. doi: 10.017/s0266462300103150.
26. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ. 2004 Sep 28;171(7):735-40. doi: 10.1503/cmaj.1041086.
27. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004 May 26;291(20):2457-65. doi: 10.1001/jama.291.20.2457.
28. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database--update and key issues. N Engl J Med. 2011 Mar 3;364(9):852-60. Epub 2011/03/04.
29. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010 Jun 1;152(11):726-32. doi: 10.7326/0003-4819-152-11-201006010-00232. Epub 2010 Mar 24.

Figure legends

Figure 1. PRISMA flow diagram.

Figure 2. Included hand surgical trials, trial number for year.

Figure 3. Bias domains: Selection bias (sequence generation and allocation concealment), Performance bias (blinding of participants and personnel), Detection bias (blinding of outcome assessors), Attrition bias (incomplete outcome data), Selective reporting and “Other” bias.

Figure 1

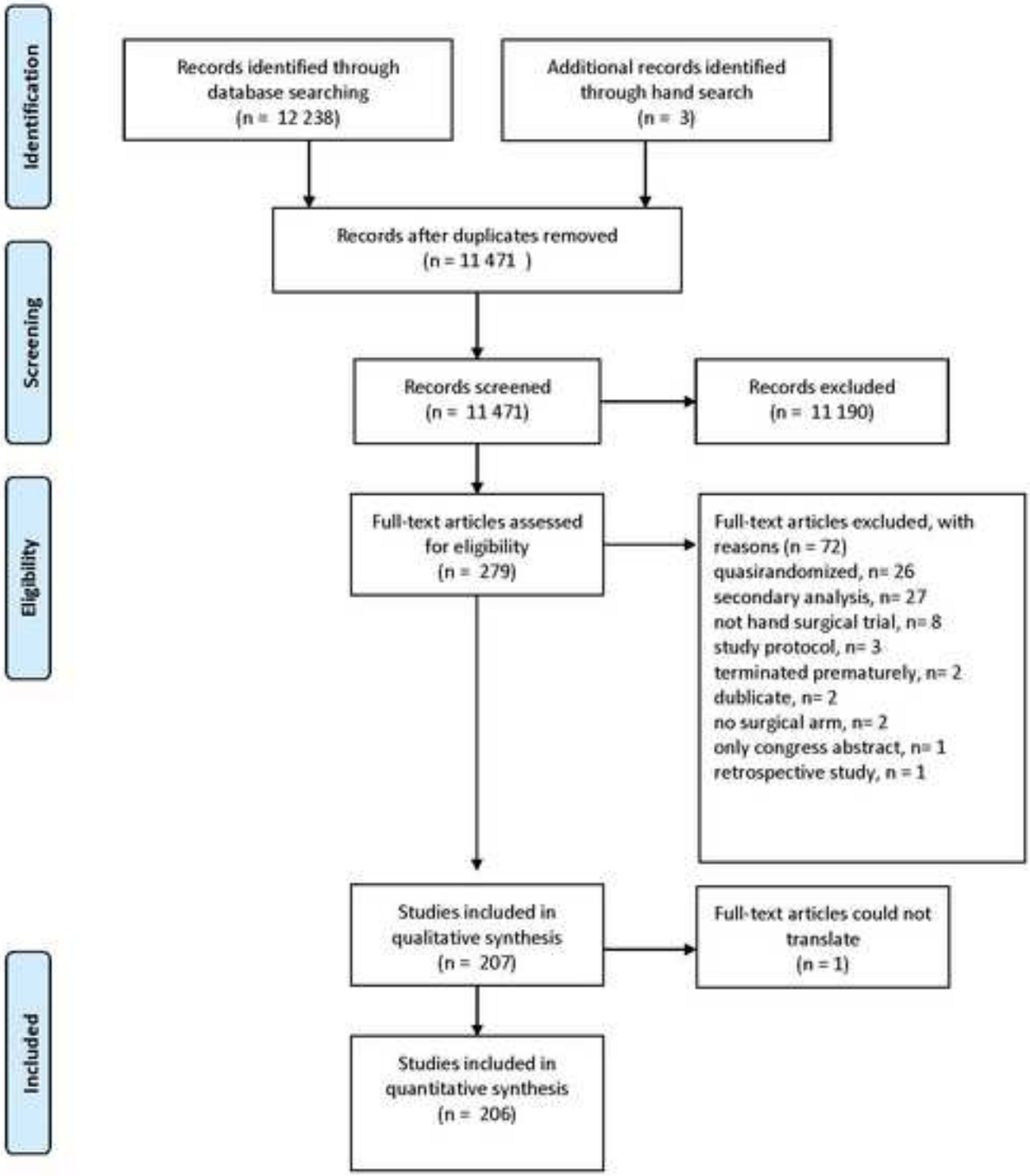

Figure 2

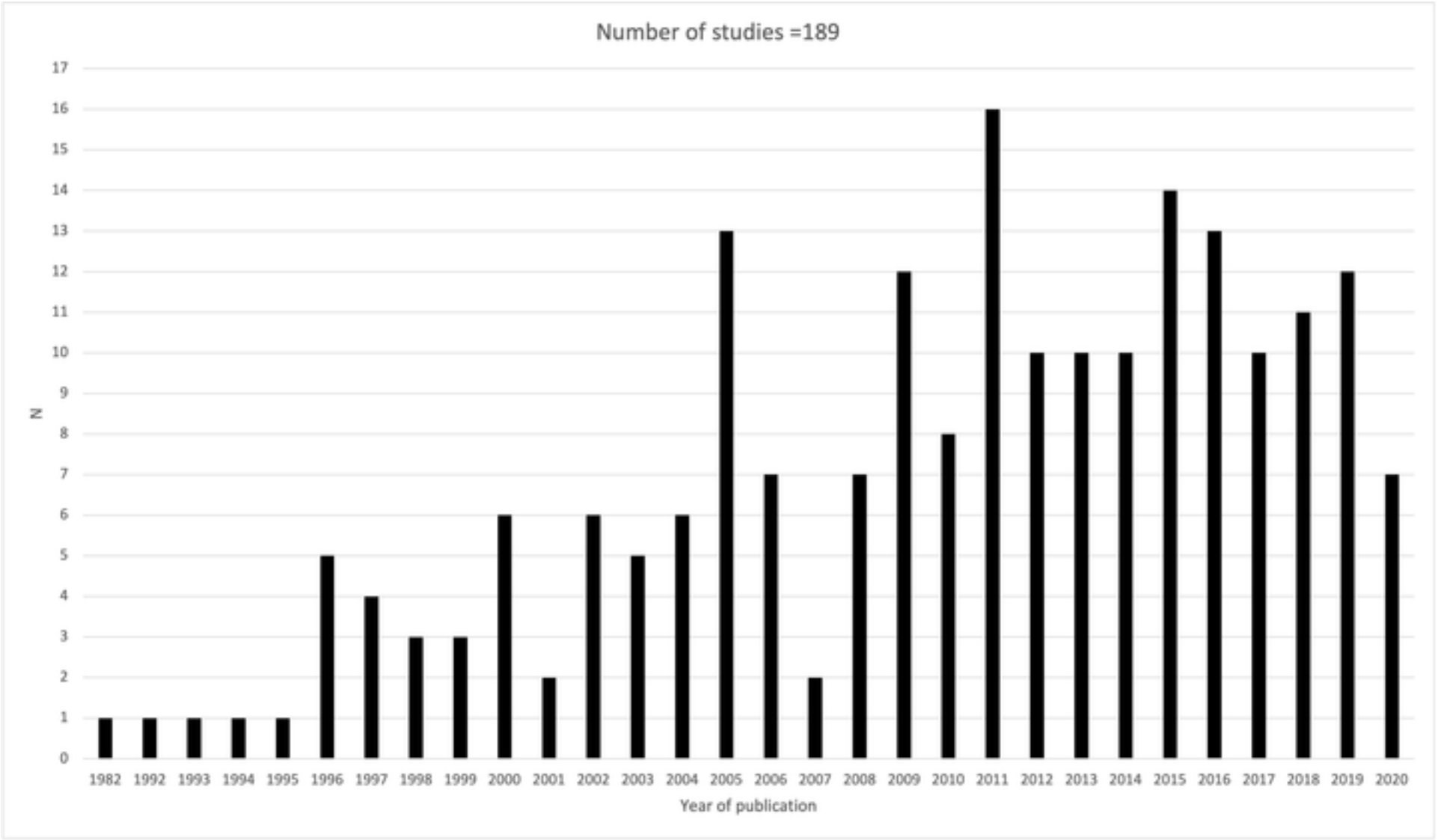

Figure 3

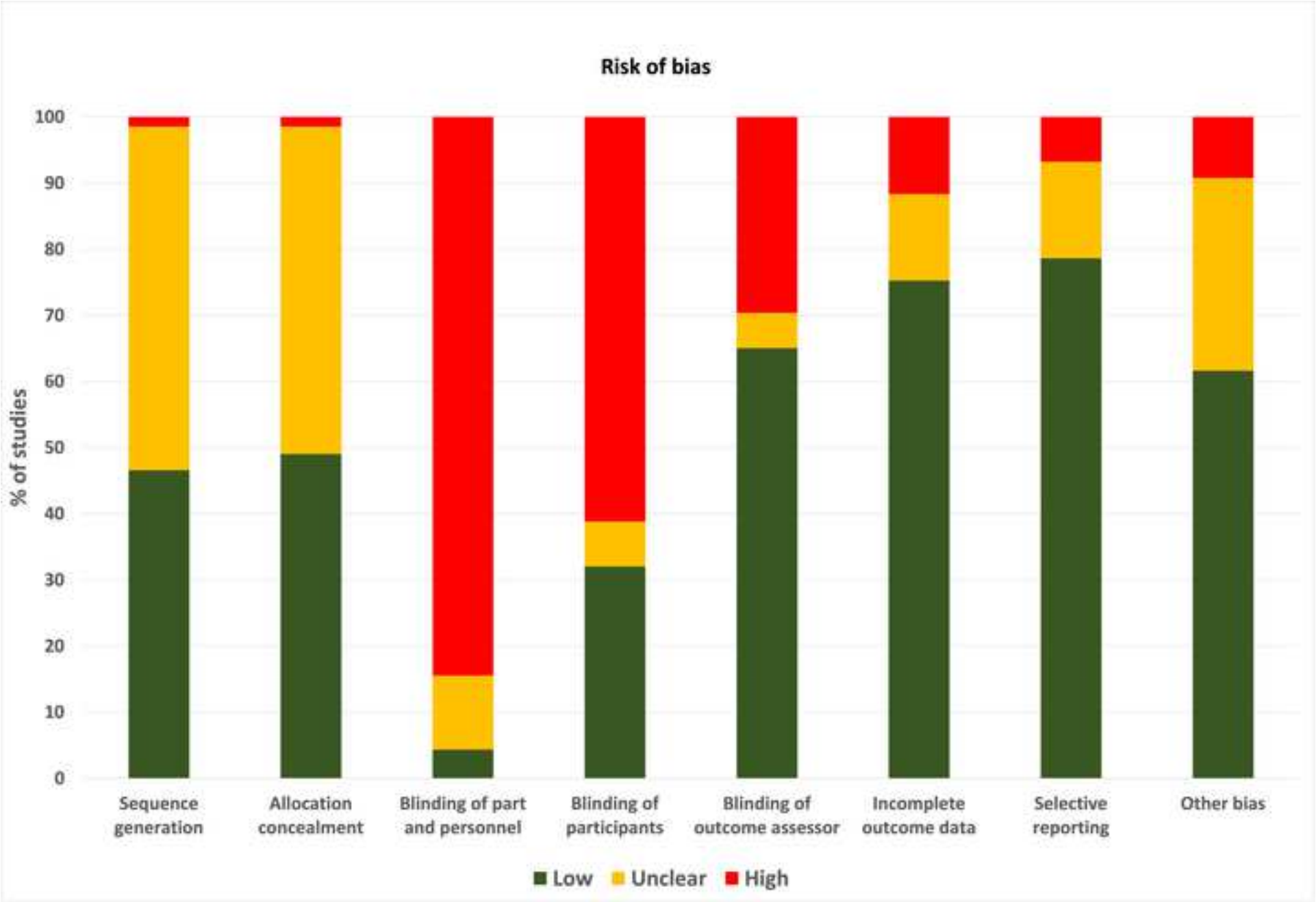

Table 1

Condition	N	%
Radius fractures	101	49.0
Carpal tunnel syndrome	20	9.7
Trapeziometacarpal osteoarthrosis	18	8.7
Non-traumatic tendon conditions e.g. tendinopathy	10	4.8
Tendon injuries	9	4.4
Carpal bone fractures	7	3.4
Metacarpal fractures	7	3.4
Rheumatoid arthritis	6	2.9
Miscellanous	6	2.9
Nerve injuries	4	1.9
Other fracture	4	1.9
Other osteoarthrosis	3	1.4
Amputation	3	1.4
Burns	3	1.4
Dupuytren’s contracture	3	1.4
Phalanx fractures	3	1.4
Total	207	100.0

Table 2. Standardized Mean Differences (SMD) from the meta-regression model

All studies with sufficient data to calculate SMD, n=172 studies				
Domain	Low risk of bias	Unclear or high risk	95% CI	p-value
Selection bias	Ref.	0.12	-0.07 to 0.31	0.23
Performance bias	Ref.	0.35	-0.22 to 0.92	0.23
Detection bias	Ref.	0.08	-0.30 to 0.45	0.68
Attrition bias	Ref.	-0.11	-0.32 to 0.09	0.29
Selective reporting	Ref.	0.11	-0.12 to 0.34	0.35
Other bias	Ref.	0.04	-0.17 to 0.25	0.71
Studies with distinguishable experimental and control arms, n=97 studies				
Selection bias	Ref.	**0.28**	**0.02 to 0.55**	**0.03**
Performance bias	Ref	0.48	-0.07 to 1.0	0.09
Detection bias	Ref	0.23	-0.16 to 0.62	0.25
Attrition bias	Ref	-0.23	-0.54 to 0.08	0.16
Selective reporting	Ref	-0.18	-0.48 to 0.13	0.25
Other bias	Ref	-0.0	-0.29 to 0.29	0.99