Title: Research in Hand Surgery: Types of Study Design

ABSTRACT

Modern advances in surgery are most robustly achieved through empirical observation, hypothesis generation, methodical data collection and analysis and, often, a willingness to challenge the status quo of current practices. Clinical research takes many forms, and it is important to understand the nuances of different study designs and their indications. We have illustrated this through selection of ten example research articles and we have suggested several tools to assist with critical appraisal of study quality within each category. We hope the reader will find this a useful resource for future reference when interpreting research studies or indeed designing their own.

INTRODUCTION

A scientific study is built on first and foremost on a research question and then the appropriate study design to answer the question. One of the most important considerations when planning the study should therefore be the type of study design. The correct study design depends entirely on the type of question and should generate the most robust data possible for analysis and interpretation, while minimising risk of bias. This article outlines some of the considerations in study design for the hand surgeon.

Study designs can be divided into two categories: experimental and observational. In observational studies, the researcher(s) describe characteristics of participants/patients and perform an analysis of this data, such as the impact of risk factors on disease occurrence or treatment outcomes. In experimental studies, the researcher(s) manipulate one or more groups through some form of intervention (e.g., pharmacological, surgical, physiotherapeutic,

socio/psychological), and analyze the outcome of this intervention. In addition to the above primary research, systematic reviews (with or without meta-analyses), provide a synthesis of two or more of the above study designs. Here we describe the most commonly used study designs, providing examples of each and tools that readers can use to determine the quality and reliability of the studies they encounter.

OBSERVATIONAL (DESCRIPTIVE) STUDY DESIGNS

Case Reports

Single case reports constitute a type of observational study. The researcher has identified a clinical case encountered in routine practice, and recognized the unique or nuanced element in the presentation, progression, or management of the clinical problem. The aim of the report is to raise awareness or inform the scientific and clinical community of uncommon variants of existing conditions, or indeed to identify novel conditions and prompt further discussion, hypothesis generation, and literature review. As these reports are anecdotal in nature and only describe the experience of an individual patient, they are highly susceptible to bias. However, they are usually of interest to the hand surgery community and add to the existing literature. The authors of the following case report used this format to discuss mass-related complications of the wrist post-arthroscopy and highlight inconsistency of terminology and classification in current literature (Chen et al., 2022). We suggest authors should refer to the CARE guidelines (for **CA**se **RE**ports), https://www.care-statement.org/) when preparing a case report.

Case Series

In a case series, the researcher has identified a group of participants with a known shared exposure (for example, to a particular disease or treatment) and should provide a simple, descriptive report on the experience of these participants after the exposure. In this example, sensory, motor, and functional sequelae were described and compared in patients who either underwent distal long finger amputation through the distal interphalangeal joint or the diaphysis of the middle phalanx (Haddad et al., 2022). Case series are commonly utilised to aggregate information regarding the natural history of diseases, and do not provide analytical insight to causative relationships. Consecutive case series are of higher quality than non-consecutive as this reduces the risk of bias. They are typically retrospective in nature but can be prospective. The main weakness of a case series is the lack of a pre-determined protocol or control group and small numbers. We suggest the reader refers to the Quality Assessment Tool for Case Series Studies (US National Heart, Lung and Blood Institute – https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools).

OBSERVATIONAL (ANALYTICAL) STUDY DESIGNS

Cohort Studies

A cohort study involves longitudinal analysis of one or more cohorts of participants. Cohort studies may be single-arm or comparative, the latter usually involving two cohorts who are either exposed or unexposed to a specified risk factor or intervention. For example this propensity score-matched cohort study looked retrospectively at the data from well-validated Swedish nationwide health registries to investigate the association of bariatric surgery and Dupuytren's disease (Burkard et al., 2022). The researcher would have recorded the incidence of a particular outcome in the exposed and unexposed groups. Comparing outcomes between groups may be used to infer associations between risk factors and disease

incidence or to identify differences in the temporal history of a disease process relating to the risk factor exposure status. However, cohort studies are susceptible to misinterpretation of inter-variable relationships if confounding variables are not accounted for. These studies can also be costly to run as it may take many years for some outcomes to occur; as such they are best suited for studying risk factors relating to common outcomes or diseases. To reduce cost, multiple outcomes may be studied for a single risk factor – either from the inception of the study or, alternatively, secondary outcomes may be added retrospectively to ongoing cohort studies. With long follow-up periods, high attrition rates may become a source of bias if group matching is affected by the end of the study. Cohort studies differ importantly from case series in the method of participant entry into the study. For example, a case series would comprise a report on patients undergoing a specific type of operation with outcomes of that operation described descriptively. In contrast, a cohort study would consist of a group of participants/patients who have a certain condition that may have undergone a type of intervention, different types of intervention or no intervention, with outcomes reported for the whole group or groups. When appraising cohort studies, researcher should use the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (US National Heart, Lung and Blood Institute – https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools).

Case-Control Studies

A case-control study is designed inversely to cohort studies, in that the researcher will identify and prospectively recruit participants who already have a particular outcome or disease (cases) and matched controls without the same condition. The researcher will then look for evidence of prior exposure to specified risk factors in all participants. For example,

this study (Titchener et al., 2013) employed a case-control design to explore and identify the poorly understood risk factors for lateral epicondylitis (tennis elbow). Note that although this is a relatively common condition, the case-control design is an appropriate choice as it would be impractical to recruit patients to a cohort study based on exposure to unknown risk factors and, secondly, onset of lateral epicondylitis may require long follow-up for detection, increasing attrition rates. The rate of exposure in cases versus controls may be used to calculate an odds ratio and infer association or causality between risk factors and disease outcomes. However, poor recall of prior events, symptoms, or treatments – or poor historical data entry – can significantly affect analysis and interpretation of the results. Case-control studies are often used to investigate risk factors for rare diseases as there is no requirement for longitudinal follow-up whilst waiting for rare events to occur. If the risk factor itself is rare, a cohort study is better suited to longitudinally compare patients with known exposure to matched, unexposed controls. The researcher is referred to the Quality Assessment Tool for Quality Assessment of Case-Control Studies (US National Heart, Lung and Blood Institute – https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools).

Cross-Sectional Studies

A cross-sectional study may be used to ascertain information about the prevalence of characteristics, risk factors, or diseases affecting a population at one point in time. Consequently, it is impossible to discern causal relationships from data collected as there is no longitudinal follow-up, but associations between risk factors and outcomes may be inferred. This recent example looked at the ulnar variance and triangular fibrocartilage thickness in adolescents, by using single data points from magnetic resonance imaging (MRI) scans of healthy participants (van der Post et al., 2022). Due to single time point data

collection, these studies are relatively inexpensive and may be used to sample large numbers of individuals. Cross-sectional studies may still be affected by selection bias if samples are not representative of the wider population being studied. We suggest the researcher should refer to the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (US National Heart, Lung and Blood Institute – https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools).

EXPERIMENTAL STUDY DESIGNS

Non-randomized Experimental Studies

Experimental studies are designed to identify and analyse causal relationships between interventions and outcomes. In uncontrolled experimental studies, all participants recruited are treated with the same intervention and outcomes are measured without comparison to other groups. These differ from case series in that they participants are actively recruited in advance according to eligibility criteria and receive more standardised interventions and prospective follow-up determined *a priori*. Case series are usually retrospective, descriptive analyses of patients who have had more *ad hoc* intervention and assessment. In an uncontrolled trial, analytical (rather than descriptive) statistics would be applied to test a pre-defined hypothesis. The following example of an uncontrolled experimental study examined effects of partial wrist denervation in wrist osteoarthritis through patient-reported outcomes and objective function (Swärd et al., 2022). The authors themselves highlight key limitations of this study design, namely the lack of control group or blinding leading to high risk of bias and confounding variables (mitigated somewhat by use of statistical methods). Despite this, in some cases where the benefit of surgery is self-evident and outcomes far exceed the

minimum clinically important difference, uncontrolled trials may be appropriate and sufficient.

Controlled experimental studies compare outcomes in groups with matched characteristics (e.g., demography and disease status) but differing exposure to interventions applied by the researcher. Provided that groups are well-matched, controlled studies allow the researcher to isolate the impact of the chosen intervention while minimizing the influence of other variables on the outcome. Here (Shibata et al., 2023), the authors have identified patients >70 years old who underwent anterior locking plate fixation for distal radius fracture plus either a) K-wire fixation, b) locking plate fixation, or c) Darrach procedure for distal ulna fracture. This allowed comparisons between these three surgical techniques in terms of functional outcomes and complications. We direct the reader to our suggested tool for quality assessment of non-randomized studies: Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool (*https://methods.cochrane.org/methods-cochrane/robins-i-tool*).

Randomized-Controlled Trials (RCTs)

This is the gold standard study design for testing effectiveness of interventions, generating the most robust and reliable data. Unintended bias from group allocation may be further mitigated through randomization – wherein the researcher(s) do not influence the allocation of participants to experimental or control groups, but this process is determined by software or some other random process. It is worth noting that even with appropriate randomization there may be statistical differences between group characteristics due to random chance, especially with low numbers of participants. The decision to account for poorly matched groups by adjusting outcomes for these variables should be influenced by their clinical

relevance to the study outcome; adjustment based purely on statistical significance can lead to erroneous results, as explored previously (Broekstra et al., 2022). Risk of bias may be reduced further by blinding participants, researchers, or both ("double blind") to group allocation for the duration of the study. In surgical trials, this may be both ethically and practically challenging as it may be impossible to conceal the intervention from patients and practitioners or researchers. However, "sham" surgeries – in which control patients typically undergo the same process of admission to hospital; anaesthesia; surgical approach, exploration, and closure; without the therapeutic intervention in question – are increasingly used as controls in the evaluation of surgical interventions. An example RCT is the DRAFFT2 trial, which compares patient reported outcome measures (PROMs) after manipulation of dorsally displaced distal radius fractures plus either moulded cast application or K-wire fixation (Costa et al., 2022). This was a randomized superiority trial conducted across 36 UK (NHS) centres providing a pragmatic assessment of real-world outcomes. We suggest the reader refers to the Cochrane risk-of-bias tool for randomized trials (RoB 2) for critical appraisal of such studies (https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0).

EVIDENCE SYNTHESES

Systematic Review

Systematic reviews (SRs) apply a rigorous, systematic approach to identify, evaluate, and assimilate all relevant published studies on a particular topic in order to present an up-to-date and comprehensive account of the current state of knowledge in that field, to answer a specific question. They can provide a narrative summary of the available evidence on a subject, and if data is available, a quantitative meta-analysis that provides inferential statistics

based on synthesis of primary data. Note that systematic reviews differ from literature reviews by employing methodological rigour to avoid selection bias of specific studies, and therefore the results from a systematic review should be reproducible. This occurs via a thorough literature search, selection of studies based on strict pre-defined inclusion and exclusion criteria, followed by critical assessment of the quality and validity of the findings, and synthesis of a concluding narrative or discussion. This recent example (Deshmukh et al., 2021) examined the heterogeneity of reported outcomes from 160 randomized, quasi-randomized, or large prospective observational studies on hand fractures and joint injuries, to demonstrate the need for a new core outcome set and promote consistency and reproducibility of future studies in this field. As they involve careful and systematic assessment of potentially large numbers of studies, SRs are generally considered a high level of evidence but this is dependent on the quality of the individual studies included. Systematic reviews often encounter a high rejection rate not only because of a poor research question, but also the small number of selected studies and poor quality of evidence. Often, no conclusion can be made from the review. SRs are often produced in combination with meta-analysis and are frequently used to inform health policies and clinical guidelines. We refer readers to the following quality assessment tool: A Measurement Tool to Assess Systematic Reviews 2 (AMSTAR 2 – https://amstar.ca/Amstar-2.php).

Meta-analysis

Meta-analyses use a similar approach to systematic reviews in order to identify published studies in a particular field. Unlike SRs, meta-analyses use data extraction and statistical approaches to combine results from multiple studies and calculate an overall effect size for a given intervention. These typically include analysis of randomized controlled trials and

observational studies; however, care must be taken to ensure both that studies are directly comparable and that correct methodological and statistical approaches are applied to the data. Also, the aim is not to answer a specific question, unlike SR, but to collate a number of findings from the evidence. The following example (Wade et al., 2018) analyzed data from five randomized trials to assess the effects of using absorbable versus non-absorbable sutures for skin closure following elective carpal tunnel decompression, quantitatively comparing the impact on a number of different outcomes including postoperative pain, function, scar satisfaction and adverse events. Readers may use the same quality assessment tool as for SRs in appraising meta-analyses (AMSTAR 2 – https://amstar.ca/Amstar-2.php). Researchers are also encouraged to apply the GRADE approach (The Grading of Recommendations Assessment, Development and Evaluation) to determine the certainty of evidence and strength of recommendations in SRs or meta-analyses. More information can be found at the GRADE working group (https://www.gradeworkinggroup.org) or Cochrane organization (https://training.cochrane.org/grade-approach).

SUMMARY

The choice of study design in surgical research may depend on several factors relating both to the inherent nature of the question to be investigated and practical elements such as cost, duration, technical and/or human resources and expertise. This article has sought to define and rationalize common study designs as well as key considerations for critical appraisal of studies in each category. All studies contain sources of bias and studies with similar design may differ significantly in quality dependent on the individual methodology; hence, we additionally provide a selection of tools to assist in assessing quality of evidence for each

238 study design discussed. In **Table 1** we have summarised the information above for quick

239 reference.

240 **REFERENCES**

241

242 Broekstra DC, de Boer MR, Stunt JJ. Statistics in publishing: the (mis)use of the p-value
243 (part 1). J Hand Surg Eur Vol. 2022, 47: 677–80.

244 Burkard T, Lane JCE, Holmberg D, Thorell A, Burden AM, Furniss D. The association of
245 bariatric surgery and Dupuytren's disease: a propensity score-matched cohort study. J Hand
246 Surg Eur Vol. 2022, 47: 288–95.

247 Chen JS, Shaughnessy PJ, Catalano LW. Synovial herniation following wrist arthroscopy: a
248 case report. J Hand Surg Eur Vol. 2022, 47: 327–8.

249 Costa ML, Achten J, Ooms A, et al. Surgical fixation with K-wires versus casting in adults
250 with fracture of distal radius: DRAFFT2 multicentre randomised clinical trial. BMJ. 2022,
251 376: 1–7.

252 Deshmukh SR, Mousoulis C, Marson BA, et al. Developing a core outcome set for hand
253 fractures and joint injuries in adults: a systematic review. J Hand Surg Eur Vol. 2021, 46:
254 488–95.

255 Haddad B, Benhamou S, David E, Klein C. The value of preserving the full middle phalanx
256 in distal long finger amputations: A series of 59 cases. J Hand Surg Eur Vol. 2022, 47: 773–
257 5.

258 van der Post AS, Jens S, Jacobs K, Smithuis FF, Obdeijn MC, Maas M. Ulnar variance and
259 triangular fibrocartilage thickness in adolescents: a cross-sectional MRI study of healthy
260 participants. J Hand Surg Eur Vol. 2022, 47: 722–7.

261 Shibata R, Tokutake K, Takegami Y, Natsume T, Matsubara Y, Imagama S. Comparison of

surgical treatments for distal ulna fracture when combined with anterior locking plate fixation of distal radius in the over 70 age group. J Hand Surg Eur Vol. 2023: 175319342211504.

Swärd EM, Franko MA, Wilcke MK. The effects of partial wrist denervation in wrist osteoarthritis: patient-reported outcomes and objective function. J Hand Surg Eur Vol. 2022, 47: 798–804.

Titchener AG, Fakis A, Tambe AA, Smith C, Hubbard RB, Clark DI. Risk factors in lateral epicondylitis (tennis elbow): A case-control study. J Hand Surg Eur Vol. 2013, 38: 159–64.

Wade RG, Wormald JCR, Figus A. Absorbable versus non-absorbable sutures for skin closure after carpal tunnel decompression surgery. Cochrane Database Syst Rev. 2018, 2018.