Efficacy of Surgical Interventions for Trapeziometacarpal (Thumb Base) Osteoarthritis: A Systematic Review

Tokiko Hamasaki, MSc, 1, 2, 3 Patrick G. Harris, MD, 1, 2, 3 Nathalie J. Bureau, MSc, MD, 4, 5, 6** Nathaly Gaudreault, PhD, 1, 4, 6 Daniela Ziegler, MSI, 1, 6** Manon Choinière, PhD 1, 3, 6, 7

1 Hand Centre of the Centre Hospitalier de l’Université de Montréal, Montreal, Quebec, Canada
2 School of Rehabilitation, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada
3 Centre de Recherche du Centre Hospitalier de l’Université de Montréal, Montreal, Quebec, Canada
4 Department of Surgery, Centre Hospitalier de l’Université de Montréal, Montreal, Quebec, Canada
5 Department of Surgery, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada
6 Department of Radiology, Centre Hospitalier de l’Université de Montréal, Montreal, Quebec, Canada
7 Department of Radiology, Radio-Oncology and Nuclear Medicine, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada
8 School of Rehabilitation, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, Quebec, Canada
9 Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Quebec, Canada
10 Department of Anaesthesiology and Pain Medicine, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada

Abstract

Purpose: This systematic review (SR) aimed to identify the surgical interventions available for trapeziometacarpal osteoarthritis and document their efficacy on pain, physical function, psychological well-being, quality of life, treatment satisfaction, and/or adverse events.

Methods: This PROSPERO-registered SR’s protocol was developed based on the Cochrane review methodology and Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

Results: Among 9049 potential studies identified, 1 SR, 18 randomized controlled trials, and 40 nonrandomized controlled trials were included. We identified 11 categories of surgical techniques: first metacarpal osteotomy, first metacarpal and trapezium partial resection, arthrodesis, trapeziectomy (T), T-ligament reconstruction (LR), T-tendon interposition (TI), T-ligament reconstruction and tendon interposition (LRTI), hematoma distraction arthroplasty (HDA), chondrocostal graft interposition, autologous fat injection, and manufactured implant use. These findings supported by low-quality evidence revealed moderately or largely superior effects of the following interventions: (1) trapeziectomy over T-LRTI using flexor carpi radialis (FCR) and metacarpal tunnel (MT) or using abductor pollicis longus (APL) and FCR for adverse events; (2) trapeziectomy over T-TI with palmaris longus (PL) for pain; (3) T-LR with FCR-MT over T-LRTI with FCR-MT for physical function; (4) trapeziectomy by anterior approach over that by posterior approach for treatment satisfaction and adverse events; (5) T-LRTI using FCR-MT over T-TI with PL for pain; and (6) T-HDA over T-LR using APL-MT-FCR for pain, physical function, and adverse events. GraftJacket (Wright Medical Group, Memphis, TN), Swanson (Wright Medical Group, Letchworth Garden City, UK), and Permacol (Tissue Science Laboratories, Aldershot, UK) implants and hardware (plate/screw) would cause more complications than an autograft. The effect estimates of other surgical procedures were supported by evidence of very low quality.

Conclusions: This SR provided evidence of the efficacy of various surgical interventions for trapeziometacarpal osteoarthritis. Some interventions showed a moderate-to-large superior effect on the studied outcome(s) compared with others. However, these findings must be interpreted with caution because of low-quality evidence. To provide stronger evidence, more randomized controlled trials and methodological uniformization are needed.

Type of study/level of evidence: Therapeutic I.

Declaration of interests: No benefits in any form have been received or will be received related directly or indirectly to the subject of this article.

Corresponding author: Manon Choinière, PhD, CRCHUM, St-Antoine Building, Room 501-126, 850 St-Denis St, Montreal, Quebec, Canada, H2X0A9.
E-mail address: manon.choiniere@umontreal.ca (M. Choinière).

https://doi.org/10.1016/j.jhsg.2021.02.003
2589-5141/ Copyright © 2021, THE AUTHORS. Published by Elsevier Inc. on behalf of The American Society for Surgery of the Hand. This is an open access article under the CC BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/.
Trapeziometacarpal osteoarthritis (TMO) is one of the most prevalent and painful forms of hand osteoarthritis.\(^1\)\(^,\)\(^2\) It not only reduces thumb mobility\(^3\) but also limits hand functions needed for daily activities.\(^4\) The TMO care pathway usually begins with nonsurgical interventions, and when they are unsuccessful, patients might undergo surgery.\(^5\)

Among the numerous surgical techniques available for TMO, trapeziectomy with ligament reconstruction and tendon interposition (T + LRTI) is the most popular among surgeons in the United States\(^5\)\(^–\)\(^8\) despite its higher cost compared to that of other procedures such as simple complete trapeziectomy.\(^9\) Moreover, the choice of a given surgical procedure can affect the length of a patient’s sick leave period.\(^10\) Therefore, surgical procedures must be judiciously chosen, taking into account their benefits, adverse effects, and costs. Unfortunately, evidence of the efficacy of surgical interventions for this pathology is lacking, leaving surgeons and patients unsupported in their decision-making process ("efficacy" here is defined as the performance of a treatment under ideal and controlled circumstances, such as randomized controlled trials [RCTs], as opposed to “effectiveness,” which is defined as the performance of a treatment under usual or “real-world” circumstances\(^11\)). Although some systematic reviews (SRs) examining the efficacy of surgical interventions exist, they are not exhaustive (including solely particular surgical techniques),\(^12\)\(^,\)\(^13\) are methodologically suboptimal (eg, omitting critical appraisal\(^14\)\(^,\)\(^15\)), or are methodologically sound but need to be updated because their included articles were published before 2013.\(^16\)\(^,\)\(^17\) Therefore, we performed a comprehensive SR in terms of surgical procedures using a rigorous methodology. Our explicit questions to be answered by this SR were as follows: (1) What are the surgical interventions available for TMO whose efficacy has been documented?; and (2) What are the effects of these surgical interventions on pain, physical function, psychological well-being, quality of life, treatment satisfaction, and/or adverse events?

Materials and Methods

The protocol of this PROSPERO-registered SR (PROSPERO CRD42015015623) was developed based on the Cochrane intervention review\(^17\) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines\(^18\) and has been previously published.\(^19\)

Search strategy and criteria

As seen in Figure 1, a literature search of 16 bibliographic databases up to July 4, 2018, was conducted by an experienced medical librarian. The following search terms were used: thumb, trapeziometacarpal (TM) joint, carpometacarpal joint, osteoarthritis, intervention, and pain (an example of our search strategy is presented in Appendix 1 [available on the Journal’s Web site at www.jhsngo.org]). Languages were restricted to English and French. Manual searching of reference lists of included studies was also performed.

Inclusion/exclusion criteria

Two reviewers independently screened the titles/abstracts of studies based on eligibility criteria and then read the full-text copies of potentially relevant studies. Any disagreement was discussed until a consensus was reached. The eligibility criteria were as follows:

- **Study designs:** SR, RCTs, and non-RCTs (NRCTs). The inclusion of SR was considered necessary to avoid redundant work for studies already reviewed. NRCTs were also included in cases where there was no SR/RCT for a given intervention or outcome.
- **Population:** Adults with primary TMO
- **Interventions:** Any type of interventions was first included. Because of the tremendous volume of data, surgical and nonsurgical interventions were separated, and only the surgical ones were included in this SR. An SR of nonsurgical interventions has been previously published.\(^20\)
- **Outcomes:** Pain, physical function, psychological well-being, quality of life, treatment satisfaction, and adverse events were included. These outcomes are considered as core outcomes for chronic pain trials.\(^21\)\(^,\)\(^22\)

Risk-of-bias assessment

Among 9049 potential studies identified after the bibliographic database search, 68 were eligible (Fig. 1), comprising 10 SRs,\(^23\)\(^–\)\(^27\) 18 RCTs,\(^28\)\(^–\)\(^45\) and 40 NRCTs,\(^46\)\(^–\)\(^54\) which represented a total of 3878 thumbs in 3658 participants. The characteristics of the included studies are presented in Appendices 2 (SRs) and 3 (RCTs/NRCTs) (Appendices 2 and 3 are available on the Journal’s Web site at www.jhsngo.org). A total of 276 full-text articles did not meet the selection criteria (the list of the excluded articles is available upon request).

Two reviewers independently assessed the risk of bias (RoB) in the 10 identified SRs using the Assessment of Multiple Systematic Reviews checklist.\(^55\) Any disagreement between the reviewers was discussed until a consensus was reached; otherwise, a third reviewer was consulted. The mean total score of the 10 SRs according to the Assessment of Multiple Systematic Reviews checklist was relatively low at 4.4 ± 2.6 (0 = worst score, 11 = best score: Appendix 2). More than half of the SRs had an unclear or high RoB based on criteria 1, 2, 4, 5, 6, 10, and 11 (Fig. 2A). Because of the suboptimal methodological quality of majority of the SRs, we questioned the validity of these SRs’ findings. Therefore, to integrate only a valid body of evidence from previous SRs into our SR, we decided to apply a minimum set of criteria based on the Agency for Healthcare Research and Quality’s guidance\(^56\) for the 10 identified SRs: (1) multiple data source search; (2) application of pre-defining eligibility criteria for study selection; (3) performing RoB assessment; and (4) considering evidence quality. Only 1 SR\(^57\) met these criteria (Appendix 2).

To assess RoB in RCTs/NRCTs, the Cochrane Effective Practice and Organization of Practice’s risk-of-bias tool\(^58\) was used. The criterion of “similarity of baseline characteristics between experimental and control groups” was assessed by considering the following founders: age, sex, occupation, hand dominance, affected side, radiographic stage, and symptom duration. When >80% of these factors were balanced between groups, RoB was considered low; when 60%–79% were balanced, RoB was considered unclear; when <60% were balanced, RoB was considered high.\(^59\) Blinding was considered for the following 3 parties: participant, performer (surgeon), and assessor. The results of the RoB assessments reported for the included SR\(^60\) were taken into account to ascertain our judgment.\(^57\) Figure 2B shows the results on each item assessed using the risk-of-bias tool: 40 trials suffered from a selection bias due to a nonrandomized design. More than half did not assess the similarity of baseline outcomes and characteristics between groups. When reported, these data were mainly limited to age and sex. Except for 1 RCT,\(^24\) all the trials had unclear or high RoB results regarding blinding participants, either because of the presence of K-wire(s), which had likely informed the participants of the type of surgery,\(^38\)\(^,\)\(^42\)\(^,\)\(^46\) or the retrospective nature of the studies.\(^47\)\(^–\)\(^51\)\(^,\)\(^57\)\(^–\)\(^60\)\(^,\)\(^62\)–\(^65\)\(^,\)\(^67\)–\(^71\)\(^,\)\(^74\)\(^–\)\(^78\)\(^,\)\(^81\)\(^–\)\(^84\) Blinding the
Figure 1. Flow chart of study identification. ACP, American College of Physicians; AHRQ, Agency for Healthcare Research and Quality; CADTH, Canadian Agency for Drugs and Technologies in Health; CINAHL, cumulative index to nursing and allied health literature; EMB, evidence-based medicine; NGC, National Guideline Clearing House; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; OAiSter, Open Archives Initiative; O’Seeker, Occupational Therapy Systematic Evaluation of Evidence; PEDro, Physiotherapy Evidence Database.

Data collection

Two reviewers independently extracted data from the identified studies using the Cochrane Effective Practice and Organization of Practice’s data abstraction form.37 Missing data were first sought by contacting the authors via e-mail. When medians and (interquartile) ranges were available, means and standard deviations were estimated using an approximation method.92 When graphic data were available, the Plot Digitizer software93 was used to digitize these data.

Meta-analysis methodology

The effect estimates of a given intervention—standardized mean differences for continuous outcomes and risk ratios for dichotomous ones—were first searched for in the included SR.16 If unavailable, we computed them using RCT data. If no RCTs were available, NRCTs were consulted. When a given outcome was measured at different time points, only the effect estimate of the last measured point was included. It was not possible to use a meta-analytic approach because of significant treatment heterogeneity in most of the cases, but when possible, the results were expressed using forest plots (Fig. 3A–C). A publication bias analysis using a funnel plot was not feasible because the number of pooled studies was <10.17

Rating the quality of evidence

The quality of evidence of effect estimates was rated using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system.94 The evidence quality reported for the

Records identified after duplicates removal (n = 9049)
- CINAHL (475), EMB Reviews (Cochrane Database of Systematic Reviews, ACP Journal Club, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, Cochrane Methodology Register, Health Technology Assessment, NHS Economic Evaluation Database) (300), Embase (1701), MEDLINE (1648), O’Seeker (1), PEDro (16), PsychINFO (25), PubMed (4390)
- Grey literature: CADTH (1), Clinical Trials (10), Google Scholar & Google (63), Mednar (337), NGC (4), NICE (0), OAiSter (2), Open Grey (0)
- Hand-searching (76)

Records screened by titles/abstracts for eligibility (n = 4039)

Full-text articles assessed for eligibility (n = 344)

Eligible studies (n = 68)
10 SRs, 18 RCTs, 40 NRCTs

Studies included in quantitative synthesis (n = 59)
1 SR, 18 RCTs, 40 NRCTs

SRs excluded not meeting the minimum AHRQ criteria (n = 9)

Records excluded (n = 3695)
- Not a SR, RCT or NRCT (n = 61)
- No primary TMO adults or no TMO population-specific data (n = 60)
- Individual data of each intervention are not separately reported (n = 14)
- No clear description of the intervention is reported (n = 2)
- Post-surgical rehabilitative intervention (n = 4)
- Non-surgical interventions (n = 57)
- None of the outcomes of interest reported (pain, physical function, psychologic well-being, quality of life, treatment satisfaction, adverse events) or quantitative data are not available (n = 15)
- Outcome data of only the affected thumbs are not available (n = 1)
- Not in English or French (n = 4)
- Articles withdrawn (n = 4)
- Duplicate or same study (n = 9)
- Publication type (e.g., abstract, intervention description, letter) (n = 45)

Full-text articles excluded, with reasons (n = 276)
- No primary TMO adults or no TMO population-specific data (n = 60)
- Individual data of each intervention are not separately reported (n = 14)
- No clear description of the intervention is reported (n = 2)
- Post-surgical rehabilitative intervention (n = 4)
- Non-surgical interventions (n = 57)
- None of the outcomes of interest reported (pain, physical function, psychologic well-being, quality of life, treatment satisfaction, adverse events) or quantitative data are not available (n = 15)
- Outcome data of only the affected thumbs are not available (n = 1)
- Not in English or French (n = 4)
- Articles withdrawn (n = 4)
- Duplicate or same study (n = 9)
- Publication type (e.g., abstract, intervention description, letter) (n = 45)

Two reviewers independently extracted data from the identified studies using the Cochrane Effective Practice and Organization of Practice’s data abstraction form.37 Missing data were first sought by contacting the authors via e-mail. When medians and (interquartile) ranges were available, means and standard deviations were estimated using an approximation method.92 When graphic data were available, the Plot Digitizer software93 was used to digitize these data.

Meta-analysis methodology

The effect estimates of a given intervention—standardized mean differences for continuous outcomes and risk ratios for dichotomous ones—were first searched for in the included SR.16 If unavailable, we computed them using RCT data. If no RCTs were available, NRCTs were consulted. When a given outcome was measured at different time points, only the effect estimate of the last measured point was included. It was not possible to use a meta-analytic approach because of significant treatment heterogeneity in most of the cases, but when possible, the results were expressed using forest plots (Fig. 3A–C). A publication bias analysis using a funnel plot was not feasible because the number of pooled studies was <10.17

Rating the quality of evidence

The quality of evidence of effect estimates was rated using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system.94 The evidence quality reported for the
Figure 2. A Results of each Assessment of Multiple Systematic Reviews (AMSTAR) risk-of-bias item of the included systematic reviews. B Results of each item of the EPOC risk-of-bias tool of the included RCTs and NRCTs. †, low risk of bias; ‡, unclear risk of bias; ‡‡, high risk of bias; AMSTAR, Assessment of Multiple Systematic Reviews; EPOC, Cochrane Effective Practice and Organisation of Care Group.
Results

Question 1: What are the surgical interventions available for TMO whose efficacy has been documented?

We identified the following 11 TMO surgical techniques: first metacarpal osteotomy,\(^\text{18}\) first metacarpal and trapezium partial resection,\(^\text{52,71}\) arthrodesis,\(^\text{72}\) trapeziectomy,\(^\text{58,69}\) trapeziectomy and ligament reconstruction (T+LR),\(^\text{32,37,38,53}\) trapeziectomy and tendon interposition (T+TI),\(^\text{32,63,71,83,84}\) trapeziectomy with ligament reconstruction and tendon interposition (T+LRTI),\(^\text{32,63,71,83,84}\) trapeziectomy with hematoma distraction arthroplasty (HDA),\(^\text{52,69}\) trapeziectomy with hematoma distraction arthroplasty (HDA),\(^\text{52,69}\) trapeziectomy with hematoma distraction arthroplasty (HDA),\(^\text{52,69}\) trapeziectomy with hematoma distraction arthroplasty (HDA),\(^\text{52,69}\) trapeziectomy with hematoma distraction arthroplasty (HDA),\(^\text{52,69}\) trapeziectomy with hematoma distraction arthroplasty (HDA),\(^\text{52,69}\) trapeziectomy with hemato-

Question 2: What is the efficacy of the identified surgical interventions?

The effect estimates of each TMO surgical procedures are presented in Table 1 for studies with evidence of low quality (due to a RoB and imprecision) that are still the best studies available for their subject matter. We can summarize the efficacy of TMO surgeries whose effect estimates are moderate or large as follows.

T+LRTI

T+LRTI using ½ flexor carpi radialis (FCR)-MT was moderately inferior to trapeziectomy by posterior approach for reducing the number of adverse events (Table 1\(^\text{31,32,38}\) and Fig. 3B), largely inferior to T+LR with ½ FCR-MT for physical function improvement (Table 1\(^\text{36}\) and moderately superior to T+TI using PL for pain reduction (Table 1\(^\text{32}\) ). T+LRTI (½ FCR-APL-½ FCR) was largely superior to arthrodesis by plate-screw fixation for reducing pain improving physical function, and reducing the number of adverse events (Table 1\(^\text{40,44}\) and moderately superior to T+GraftJacket allograft for reducing the number of adverse events (Table 1\(^\text{42}\) ). T+LRTI using APL-FCR-APL was moderately inferior to trapeziectomy by posterior approach (Table 1\(^\text{42}\) ) and largely superior to T+Swanson implant for reducing the number of adverse events (Table 1\(^\text{35}\) ).

T+LR

T+LR using ½ FCR-MT was largely superior to T+LRTI with ½ FCR-MT for physical function improvement (Table 1\(^\text{36}\) ). T+LR using APL-MT-FCR was moderately-to-largely inferior to HDA for reducing pain improving physical function, and reducing the number of adverse events (Table 1\(^\text{42}\) ).
Table 1
Summary of Findings Supported by Evidence of Low Quality

| Surgical Interventions Compared | Outcome(s): Pain, PF, Psychological Well-Being, QoL, TS, AE | Study Design, Reference, Year of Publication, (Mean Follow-Up Time) | Effect Estimate*: SMD or Risk Ratio (RR) [95% CI] (No. of Thumbs Included) | Magnitude of Effect Estimate | Favored Intervention |
|---------------------------------|-------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------|-----------------------------|----------------------|
| **1. T versus T + LRTI (FCR-MT)** | Pain (VAS) | RCT, De Smet et al.10 2004, (T group 34 mo; T + LRTI group 26 mo) | SMD 0.39 [-0.16, 0.94] (55) SMD 0.23 [-0.31, 0.78] (55) | Small Small | T + LRTI (FCR-MT) T + LRTI (FCR-MT) |
|                                | PF (DASH) |                                                                                      | RR 1.07 [0.85, 1.35] (72)                                      | Trivial Trivial | T + LRTI (FCR-MT) T + LRTI (FCR-MT) |
|                                | Pain (PRWHE pain) | RCT, Spekrejse et al.25 2015 (mean 5.3 y) | SMD 0.05 [-0.41, 0.52] (72) SMD 0.05 [-0.52, 0.41] (72) | Trivial Trivial | T + LRTI (FCR-MT) T + LRTI (FCR-MT) |
|                                | PF (DASH) |                                                                                      | SMD −0.41 [−0.88, 0.05] (72) | Small Small | T + LRTI (FCR-MT) T + LRTI (FCR-MT) |
|                                | TS (no. of patients with good/excellent satisfaction) | RCT, Spekrejse et al.25 2015 (1−5 y) | RR 1.18 [0.75, 1.85] (72) | Trivial Trivial | T + LRTI (FCR-MT) T + LRTI (FCR-MT) |
|                                | AE (scar tenderness, sensory changes, infection, tendonitis, neurona, carpal tunnel syndrome, and worsening scaphotrapezial osteoarthritis) |                                                                 |                                    |                               |                     |
| **3. T by posterior approach versus T + LRTI (%FCR-MT)** | Pain (VAS, 0–6, the patient evaluating measure pain score) | 3 RCTs, Field and Buchanan,23 2007 (12 mo) Gangopadhyay et al.,29 2012 (median 6 y) Salem and Davis,30 2012 (6 y) | SMD −0.09 [−0.32, 0.15] (142) (Fig. 3A) | Trivial T by posterior approach |                     |
|                                | PF (DASH) |                                                                                      | RR 0.71 [0.14, 3.68] (31) | Trivial Trivial | T by posterior approach T by posterior approach |
|                                | AE (nerve dysfunction, tendon pulling sensation, tender scar, CRPS, and De Quervain’s disease) | 3 RCTs, Field and Buchanan,23 2007 (12 mo) Gangopadhyay et al.,29 2012 (median 6 y) Salem and Davis,30 2012 (6 y) | RR 0.55 [0.35, 0.89] (286) (Fig. 3B) | Moderate Moderate |                     |
| **4. T + LR (%FCR-MT) versus T + LRTI (%FCR-MT)** | Pain (no. of patients with occasional to constant pain) | RCT, Kriegs-Au et al.,36 2004 (48.2 mo) | RR 0.85 [0.47, 1.57] (31) | Trivial T + LR (%FCR-MT) |                     |
|                                | PF (no. of patients with mild-to-severe difficulty) | RCT, Kriegs-Au et al.,36 2004 (6 mo) | RR 0.30 [0.07, 1.24] (31) | Large Large | T + LR (%FCR-MT) T + LR (%FCR-MT) |
|                                | TS (no. of satisfied patients) | RCT, Salem and Davis,36 2012 (6 y) | SMD 0.14 [−0.23, 0.51] (114) | Large Large | T + LR (%FCR-MT) T + LR (%FCR-MT) |
|                                | AE (radial nerve irritation and CRPS) | SR, Wajon et al.,37 2015, and RCT, Kriegs-Au et al.,36 2004 (48.2 mo) | RR 1.32 [0.98, 1.78] (31) | Moderate Moderate | T + LR (%FCR-MT) T + LR (%FCR-MT) |
| **5. T + LRTI (%FCR-MT) versus T + TI (PL)** | Pain (0–6) | RCT, Gangopadhyay et al.,32 2012 (median 6 y) | SMD −0.56 [−0.97, −0.16] (100) | Moderate Moderate | T + LRTI (%FCR-MT) T + LRTI (%FCR-MT) |
|                                | AE (nerve dysfunction, tendon pulling sensation, tender scar, and CRPS) | RCT, Gangopadhyay et al.,32 2012 (median 6 y) | RR 0.90 [0.54, 1.50] (100) | Trivial Trivial | T + LRTI (%FCR-MT) T + LRTI (%FCR-MT) |
| **6. Arthrodesis (K-wire) versus T + LRTI (%FCR-MT - K-wire)** | Pain (no. of patients with occasional to constant pain) | RCT, Kriegs-Au et al.,36 2006 (early after surgery) | RR 1.00 [0.16, 6.42] (40) | No effect difference T + LR (%FCR-MT) T + LR (%FCR-MT - K-wire) |                     |
|                                | PF (no. of patients with mild-to-severe difficulty) | RCT, Gervin et al.,33 1997 (23 mo) | RR 1.00 [0.83, 1.20] (20) | No effect difference T + LR (%FCR-MT - K-wire) T + LR (%FCR-MT) |                     |
|                                | TS (no. of satisfied patients) | RCT, Kriegs-Au et al.,36 2006 (early after surgery) | RR 0.92 [0.51, 1.65] (20) | Trivial Trivial | T + LR (%FCR-MT - K-wire) T + LR (%FCR-MT) |
| **7. T + LR (%FCR-MT-Minimitek) versus T + LRTI (%FCR-MT-Minimitek)** | Pain (PRWHE pain) | RCT, Spekrejse et al.,34 2016 (5 y) | SMD 0.85 [0.18, 1.52] (38) | Large Large | T + LR (%FCR-MT-Minimitek) T + LR (%FCR-MT-Minimitek) |
|                                | PF (DASH) |                                                                                      | SMD 1.46 [0.73, 2.18] (38) | Large Large | T + LR (%FCR-MT-Minimitek) T + LR (%FCR-MT-Minimitek) |
|                                | TS (no. of satisfied patients) | RCT, Vermeulen et al.,40 2014 (12 mo) | SMD 0.62 [0.38, 1.00] (38) | Large Large | T + LR (%FCR-MT-Minimitek) T + LR (%FCR-MT-Minimitek) |
|                                | AE (tendinitis, neurona, nonunion requiring surgery, and CRPS) | RCT, Spekrejse et al.,34 2016 (5 y) | RR 6.18 [0.80, 47.96] (38) | Large Large | T + LR (%FCR-MT-Minimitek) T + LR (%FCR-MT-Minimitek) |
| **8. Arthrodesis (plate/screws) versus T + LRTI (%FCR-APL-%FCR)** | Pain (Michigan Hand Questionnaire pain) | RCT, Spekrejse et al.,34 2016 (5 y) | SMD 0.33 [−0.85, 0.19] (58) | Small Small | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
|                                | PF (DASH) |                                                                                      | SMD 0.26 [−0.78, 0.26] (58) | Small Small | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
|                                | QoL (SF-12 physical component) | RCT, Spekrejse et al.,34 2016 (5 y) | SMD 0.11 [−0.41, 0.62] (58) | Trivial Trivial | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
|                                | QoL (SF-12 mental component) | RCT, Spekrejse et al.,34 2016 (5 y) | RR 0.50 [0.19, 1.28] (60) | Moderate Moderate | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
| **9. T + LRTI (%FCR-APL-%FCR) versus T + Minijacket allograft** | Pain (Michigan Hand Questionnaire pain) | RCT, Spekrejse et al.,34 2016 (5 y) | SMD 0.33 [−0.85, 0.19] (58) | Small Small | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
|                                | PF (DASH) |                                                                                      | SMD 0.26 [−0.78, 0.26] (58) | Small Small | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
|                                | QoL (SF-12 physical component) | RCT, Spekrejse et al.,34 2016 (5 y) | SMD 0.11 [−0.41, 0.62] (58) | Trivial Trivial | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
|                                | QoL (SF-12 mental component) | RCT, Spekrejse et al.,34 2016 (5 y) | RR 0.50 [0.19, 1.28] (60) | Moderate Moderate | T + LRTI (%FCR-APL-%FCR) T + LRTI (%FCR-APL-%FCR) |
### Table 1 (continued)

| Surgical Interventions Compared | Outcome(s): | Study Design, Reference, Year of Publication, (Mean Follow-Up Time) | Effect Estimate*: SMD or Risk Ratio (RR) [95% CI] (No. of Thumbs Included) | Magnitude of Effect Estimate | Favored Intervention |
|--------------------------------|-------------|------------------------------------------------------------------|-----------------------------------------------------------------------|-----------------------------|---------------------|
| **10. T by posterior approach versus T+LRTI (APL-FCR-APL)** | Pain (VAS) | RCT, Belcher and Nicholl, 2000 (14 mo) | SMD –0.11 [−0.71, 0.50] (42) | Small | T by posterior approach |
| | PF (VAS) | | SMD –0.41 [−1.03, 0.20] (42) | | T by posterior approach |
| | AE (recurrent pain, instability, neuroma, sensory loss, and FCR rupture) | | RR 0.40 [0.09, 1.77] (42) | Moderate | T by posterior approach |
| **11. T-LRTI (APL-FCR-APL) versus T+Swanson silastic implant** | Pain (VAS) | RCT, Tagli and Kopylov, 2002 (2–4 y) | SMD 0.20 [−0.47, 1.07] (26) | Small | T-Swanson silastic implant |
| | TS (no. of satisfied patients) | AE (implant dislocation) | RR 0.85 [0.65, 1.11] (26) | Large | T-LRTI (APL-FCR-APL) |
| **12. T+HDA versus T+LR (APL-MT-FCR)** | Pain (VAS) | RCT, Corain et al, 2016 (T+HDA group 6.6 y; T+LR group 7 y) | SMD –3.06 [−3.60, −2.53] (120) | Large | T+HDA |
| | PF (DASH) | | SMD –0.76 [−1.13, −0.39] (120) | Moderate | T+HDA |
| **13. T by posterior approach versus T+TI (PL)** | Pain (0–6) | RCT, Gangopadhyay et al, 2012 (median 6 y) | RR 0.56 [−0.97, −0.16] (99) | Small | T by posterior approach |
| | AE (nerve dysfunction, tendon pulling sensation, tender scar, and CRPS) | | RR 0.58 [0.31, 1.07] (99) | Small | T by posterior approach |
| **14. T by anterior approach versus T by posterior approach** | Pain (VAS) | RCT, Ritchie and Becher, 2008 (33 mo) | SMD –0.28 [−0.90, 0.34] (40) | Trivial | T by posterior approach |
| | PF (VAS) | | SMD –0.11 [−0.73, 0.51] (40) | Large | T by anterior approach |
| | TS (VAS) | | SMD –0.94 [−1.60, −0.29] (40) | Large | T by anterior approach |
| | AE (sensory alteration around scar) | | RR 0.33 [0.08, 1.46] (40) | Large | T by anterior approach |
| | AE (numbness around scar) | | RR 0.25 [0.03, 2.05] (40) | Large | T by anterior approach |
| | AE (tenderness in scar) | | RR 0.33 [0.01, 7.72] (40) | Large | T by anterior approach |
| | AE (infection) | | RR 1.00 [0.07, 14.90] (40) | Large | T by anterior approach |
| **15. T by posterior approach versus T+Permacol porcine xenograft** | Pain (VAS) | RCT, Belcher and Zic, 2001 (median 6 mo) | SMD –0.94 [−1.75, −0.12] (26) | Moderate | T by posterior approach |
| | PF (VAS) | | SMD –0.51 [−1.30, 0.27] (26) | Moderate | T by posterior approach |
| | TS (VAS) | | SMD –1.27 [−2.12, −0.41] (26) | Moderate | T by posterior approach |
| | AE (neuroma, sensory change, erythema, pain, and instability) | | RR 0.38 [0.13, 1.11] (26) | Moderate | T by posterior approach |
| **16. Elektra uncemented cup versus Elektra cemented cup** | AE (cementing failure, trapezium fracture, cup loosening, and migration) | RCT, Hansen and Stilling, 2013 (2 y) | RR 1.08 [0.18, 6.57] (28) | Trivial | Elektra uncemented cup |

ADL, activities of daily living; AE, adverse event; APL, abductor pollicis longus; CI, confidence interval; CRPS, complex regional pain syndrome; DASH, disabilities of the arm, shoulder and hand; FCR, flexor carpi radialis; HDA, hematoma distraction arthroplasty; LR, ligament reconstruction; LRTI, ligament reconstruction and tendon interposition; MT, metacarpal tunnel; PF, physical function; PL, palmaris longus; RCT, randomized controlled trial; RR, ratio of continuous outcomes and RRs for dichotomous outcomes were used to establish the relative effect of each intervention. A relative effect was considered trivial when SMD $<0.2$, or RR $<1$, or CRPS. For dichotomous outcomes, when the outcome was non desirable (eg, number of patients in pain, adverse events), RR $<1$ indicates the superiority of the first interventions; RR $>1$ indicates the superiority of the second intervention. When the outcome was desirable (eg, number of satisfied patients), RR $>1$ indicates the superiority of the first interventions and RR $<1$ indicates the superiority of the second intervention.

* Interpreation of effect size.

**T+TI**

T+TI using the PL tendon was moderately inferior to T+LRTI with $\frac{1}{2}$FCR-MT and trapeziectomy by posterior approach for reducing pain (Table 125). **Trapeziectomy**

Trapeziectomy by posterior (dorsolateral) approach was largely inferior to trapeziectomy by anterior approach (Table 127) for treatment satisfaction and reducing the number of adverse events. However this technique was moderately-to-largely superior to T+LRTI with $\frac{1}{2}$FCR-MT and T+LRTI with APL strip-FCR-APL strip in terms of reducing the number of adverse events (Table 128 and Fig. 3B); T+TI using PL for relieving pain (Table 132); and T+Permacol xenograft for reducing pain improving physical function, ensuring treatment satisfaction, and preventing complications (Table 130). **Arthrodesis**

Arthrodesis by plate-screw fixation was largely inferior to T-LRTI with $\frac{1}{2}$FCR-APL-$\frac{1}{2}$FCR for relieving pain, improving physical function, and reducing the number of adverse events (Table 140). **HDA**

HDA demonstrated moderate-to-large superiority compared to T-LR performed with APL-MT-FCR for pain, physical function, and adverse events (Table 142).
Manufactured implants

The GraftJacket allograft, Permacol xenograft, and Swanson implant were moderately-to-largely inferior to their comparators (T+LRTI with ½FCR-APL-½FCR, Table 13; T+LRTI with APL-FCR-APL Table 13); and trapeziectomy by posterior approach Table 128 respectively) for relieving pain improving physical function, ensuring treatment satisfaction, and/or reducing the number of adverse events.

The effect estimates of other surgical interventions supported by evidence of very low quality are presented in Appendix 4 (Appendix 4 is available on the Journal's Web site at www.jhsogo.org).

Discussion

TMO is one of the most prevalent and painful forms of hand osteoarthritis.1–3 The TMO care regimen usually starts with nonsurgical interventions; when they are unsuccessful, surgery may be chosen.3 Unfortunately, evidence of the efficacy of comprehensive surgical interventions for TMO was lacking. Therefore, we conducted an SR to identify all the surgical interventions for TMO and review their efficacy.

We identified 12 categories of surgical techniques: first metacarpal osteotomy, first metacarpal and trapezium partial resection, trapeziectomy, trapeziectomy and chondrocostal graft interposition, arthrodesis, and the use of manufactured implants.

Pertaining to the efficacy of these surgical techniques, the evidence quality was low at best. Although some readers may view evidence of such quality as unreliable, one must consider that generating robust evidence in the field of surgery is nearly impossible because of numerous obstacles in conducting high-quality RCTs (eg, impossibility to blind surgeons, urgent situations, learning curves, or patients' reluctance for randomization or participation in a trial).95–98

If we systematically consider low-quality evidence unreliable, this research field will never generate new evidence through an SR (unless cumulative homogeneous RCTs reach a pooled sample size of 400), and all the efforts put in by the trialists of surgery research will be completely in vain. Therefore, we suggest considering low-quality evidence in such a context as "the best available" evidence. Indeed, a trivial or small effect estimate supported by low-quality evidence would not be indicative of anything, but a moderate-to-large effect estimate can at least suggest some superiority of 1 intervention over another. From this perspective, we can conclude the findings of the present review as follows: T+LRTI using ½FCR-APL should not be recommended when compared to trapeziectomy, which has a simpler procedure, because of higher complication rates in T+LRTI groups.25,31,32,36

Choosing T+LRTI with ½FCR-MC is also not justified because of its largely inferior effect on physical function compared with the simpler procedure T+LR using ½FCR-MC.36 T+TI using PL is also not recommended because of its moderate inferiority compared with trapeziectomy and T+LRTI with ½FCR-MC for relieving pain. HDA would be an interesting choice for surgeons because of its simpler procedure and capacity for maintaining “a metacarpal height to the scaphoid, an important aspect for restoring force” because it was moderately-to-largely superior to T+LR (APL-MTR-FCR) for 3 outcomes (pain, physical function, and adverse events) in a relatively large sample size (n = 120).27

Trapeziectomy by anterior approach—less practiced than that by posterior approach—has shown large effects in terms of adverse events and higher treatment satisfaction compared with the latter,27 probably because “removal of the trapezium via the anterior approach is easier and, perhaps, can be achieved less traumatically.”27 The use of the Swanson silastic prosthesis, Permacol xenograft, or GraftJacket allograft and hardware (screws/plate) increases the complication rate when compared to an autograft.30,36,39,40,44

Pertaining to first metacarpal osteotomy, first metacarpal and trapezium partial resection, chondrocostal graft interposition, autologous fat injection, and the use of manufactured implants (Arex, ARPE, Artelon, De la Caffinière, Guepar, Ivory, Ledoux, Maia, P2, PyroDisk, Pyrocardan), the evidence quality was low (Appendix 4); thus, we have little confidence in their effect estimates.

Our results diverged from a previous SR conducted by Wajon et al,16 who concluded that “of the surgical options included, this systematic review has failed to identify any additional benefit in terms of pain, physical function, patient global assessment, strength, adverse events of any procedure over another.” The discrepancies between their SR and ours are due to the diverged interpretations of what is “not statistically significant,” which is commonly interpreted as “no effect difference” based on an arbitrary P value.30 Yet, it should be interpreted as “no strong evidence that the intervention effects are different.”99 Secondly, our SR included scientific papers published before July 4, 2018, whereas the SR conducted by Wajon et al included publications from before August 8, 2013. Consequently, our SR added knowledge generated from 6 new RCTs,48–54 which is presented in Table 12.

Despite its rigorous methodological quality based on the Cochrane and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses recommendations, the limitations of this SR need to be mentioned. First, the evidence provided by this SR was, at best, of low quality, thereby limiting our confidence in the effect estimates. Second, the literature search was limited to English and French. Yet, there is some evidence of the fact that the effect estimates of English language-restricted meta-analyses do not greatly differ from those that include other languages.100 Thus, the bias related to languages in our findings may be negligible. Third, certain trials31,36,41–43,47,49,51,59,65,67,71,72,76,81,82 included patients with scaphotrapezial osteoarthritis, a common comorbidity of TMO. Other trials30,33,37–39,50,52,54–57,61–64,69,70,78 did not report whether they included patients with this type of osteoarthritis or not. This may have compromised the validity of our findings. However, according to a cross-sectional study,101 the presence of scaphotrapeziotrapezoid osteoarthritis accounted solely for 1% of the variance of pain intensity in patients with TMO. We, therefore, believe that the inclusion of patients with this type of osteoarthritis in our review did not threaten the validity of its findings.

The evidence quality supporting the efficacy of surgical interventions included in this SR was found to be low or very low. Indeed, it is highly challenging to conduct an RCT with minimal RoB in surgical research because of the aforementioned obstacles.95–98

To be able to generate more robust evidence (at least of moderate-level quality), more RCTs are needed to obtain a pooled sample size of 400 (200 per group), as required for the GRADE precision criteria for continuous variables.102 Performing a network meta-analysis comparing directly and indirectly different surgical interventions for TMO could also be useful to summarize their relative efficacy. However, this type of analysis requires clinical and methodological homogeneity among trials. Thus, concerted efforts toward harmonization of trial methodology are desired (eg, using the same outcome measures and assessment time points).

Acknowledgments

We thank Mr. Sylvain Laprise, MSc, for his precious involvement in critical appraisal and data extraction; Dr. Hervé Zamhouhoun, PhD, (Quebec SPOR-SUPPORT Unit, Canada) for methodological advice;
and Mr. Denis Bélanger for the English review of the manuscript. Thanks are due to Dr. Gordon H. Guyatt for his valuable methodological advice, as well as Drs. Isam Atroshi, Simonetta Odella, David Gallinet, Arlindo G. Pardini Jr, and Anne Wajon for their providing us with further information by personal communication. T.H. was supported by a doctoral training award of the Fonds de Recherche du Québec—Sante (grant number 30823) and by a doctoral scholarship from P.G.H.’s funds of the Centre Hospitalier de l’Université de Montréal (CHUM) Foundation and M.C.’s funds of the Centre de Recherche du Centre Hospitalier de l’Université de Montréal (CRCHUM). This systematic review was supported by discretionary funds from CRCHUM to M.C. and by the Multidisciplinary Councillor of CHUM.

References

1. Haugen IK, Englund M, Alabadi P, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis. 2011;70(9):1581–1586.
2. Marshall M, van der Windt D, Nicholls E, et al. Radiographic thumb osteoarthritis: frequency, patterns and associations with pain and clinical assessment findings in a community-dwelling population. Rheumatology (Oxford). 2011;50(4):735–739.
3. Bijsterbosch J, Visser W, Kroon HM, et al. Thumb base involvement in symptomatic hand osteoarthritis is associated with more pain and functional disability. Ann Rheum Dis. 2010;69(3):585–587.
4. Gehrmann SV, Tang J, Li ZM, et al. Motion de cit of the thumb in CMC joint arthritis. J Hand Surg Am. 2010;35(9):1449–1453.
5. Kloppenburg M, Kroon FP, Blanco Fj, et al. 2018 update of the EULAR recommendations for the management of hand osteoarthritis. Ann Rheum Dis. 2019;78(1):16–24.
6. Deutsch Z, Niedermeier SR, Aran HM. Surgeon preference, influence, and treatment of thumb carpometacarpal arthritis. Hand (NY). 2017;13(4):403–411.
7. Yuan F, Aliu O, Chung KC, et al. Evidence-based practice in the surgical treatment of thumb carpometacarpal osteoarthritis. J Hand Surg Am. 2017;42(2):104–112.e101.
8. Wolf JM, Delaronde S. Current trends in nonoperative and operative treatment of trapeziometacarpal osteoarthritis: a survey of US hand surgeons. J Hand Surg Am. 2012;37(7):77–82.
9. Mahmoudi E, Yuan F, Lark ME, et al. Medicare spending and evidence-based approach in surgical treatment of thumb carpometacarpal joint arthritis: 2001 to 2010. Plast Reconstr Surg. 2016;137(6):980e–989e.
10. Wolf JM, Atroshi I, Zhou C, et al. Sick leave after surgery for thumb carpometacarpal arthritis. Ann Intern Med. 2014;160(9):434–438.
11. Reviscki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. PharmacoEconomics. 1999;15(5):423–434.
12. Li YK, White C, Ignacy TA, Thoma A. Comparison of trapeziectomy and trapezium arthroplasty with or without tendon interposition: a systematic literature review. Plast Reconstr Surg. 2011;128(1):199–207.
13. Huang K, Hollevoet N, Siddins G. Thumb carpometacarpal joint total arthroplasty: a systematic review. J Hand Surg Eur. 2015;40(4):338–350.
14. Adams R. Does arthroscopic debridement with or without interposition material address carpometacarpal arthritis? Clin Orthop Relat Res. 2014;472(2):1166–1172.
15. Lin JD, Karl JW, Strauch RJ. Trapeziometacarpal joint stability: the evolving importance of the dorsal ligaments. Clin Orthop Relat Res. 2014;472(4):1138–1145.
16. Wajon A, Vyn_IRQI)T C, Carr E, et al. Surgery for thumb (trapeziometacarpal joint) osteoarthritis. Cochrane Database Syst Rev. 2015;2015(2):CD004631.
17. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions version 5.1.0. 2011. The Cochrane Collaboration, Available at: https://training.cochrane.org/handbook. Accessed June 15, 2014.
18. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264–269.
19. Hamasaki T, Lalonde L, Harris P, et al. Efficacy of treatments and pain and function outcomes for trapeziometacarpal (thumb base) osteoarthritis: protocol for a systematic review. BMJ Open. 2015;5,5.
20. Hamasaki T, Laprise S, Harris PG, et al. Efficacy of non-surgical interventions for trapeziometacarpal arthritis (thumb base) osteoarthritis: A systematic review. Arthritis Care Res (Hoboken). 2020;72(10):1720–1735.
21. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMPACT recommendations. Pain. 2005;113(1-2):9–19.
22. Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMPACT recommendations. Pain. 2010;156(3):337–349.
23. Martus G, Veltri K, Thoma A. Surgical treatment of osteoarthritis of the carpometacarpal joint of the thumb: a systematic review. Plast Reconstr Surg. 2004;114(2):421–432.
