**Shi, Q., MacDermid, J. C. Is surgical intervention more effective than non-surgical treatment for carpal tunnel syndrome? A systematic review. J Orthop Surg Res. 2011;6;17.**

PMID: 21477381

Design: meta-analysis of randomized clinical trials

Purpose of study: to compare the outcomes of surgical and non-surgical treatment of CTS

PICOS:

* Patient population: patients with a diagnosis of CTS, regardless of diagnostic criteria used, age, etiology, associated pathology, and gender
* Interventions: surgery for CTS of whatever type
  + Standard open carpal tunnel release (OCTR)
  + Endoscopic carpal tunnel release (ECTR)
  + Open carpal tunnel release with additional procedures such as internal neurolysis, epineurotomy, or tenosynovectomy
  + OCTR using various incision techniques
* Comparisons: a variety of nonoperative interventions
  + Drugs such as oral or local steroids, NSAIDs, diuretics, and pyridoxine
  + Wrist splints
  + Physical therapy, therapeutic exercise, manipulation, ultrasound, laser, yoga, acupuncture
* Outcomes: primary outcome was patient-reported functional and symptom improvement at six months of followup
  + Secondary outcomes were patient-reported functional and symptom improvement at 3 months of followup, patient-reported functional and symptom improvement at 12 months of followup, improvement in neurophysiological parameters, and adverse events/complications/side effects of treatment
* Study types: studies in English which were designated as prospective controlled trials
  + Studies were excluded if they were published before 1970, did not provide data on effectiveness of treatment, or compared the effectiveness of two surgical treatments or of two non-surgical treatments

Study selection:

* Databases were MEDLINE 1980 to June 1010, EMBASE 1980 to June 2010, PEDro searched in 2010, and the Cochrane Central Register 2010
* Study authors independently assessed studies for potential selection, for methodological quality, and data abstraction, using a Cohen’s unweighted kappa to assess interrater agreement on study selection
* Validity assessment was done with the Jadad scale which has three items: randomization, double blinding, and accounting for withdrawals
  + The criterion of randomization is credited with one point if the study is described as randomized, and one additional point is credited if the method to generate the sequence of randomization was described and it was appropriate (table of random numbers, computer generated, etc.)
    - One point is deducted if the method to generate the sequence of randomization was described and it was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc.)
  + The criterion of double blinding is credited if the method of double blinding was described and it was appropriate (identical placebo, active placebo, dummy, etc.)
    - One point is deducted if the study was described as double blind but the method of blinding was inappropriate (e.g., comparison of tablet versus injection with no double dummy)
  + The criterion for dropouts and withdrawals requires that participants who were included in the study but did not complete the observation period or who were not included in the analysis must be described. The number and the reasons for withdrawal in each group must be stated. If there were no withdrawals, it should be stated in the article. If there is no statement on withdrawals, this item must be given no points.
  + Thus, the Jadad scale has a minimum of 0 points and a maximum of 5 points; a score of 3 or greater was designated as high quality
* In addition, an additional 24-item scale for quality evaluation was used, which includes randomization, blinding, and accounting for missing data, and has additional items for description of the primary outcome, description of interventions and participants, and appropriate analysis of results
  + This scale scores 0, 1, or 2 points for each item, making a minimum total score of 0 and a maximum of 48 points
  + Low quality meant the score was 0 to 16
  + Moderate quality meant the score was 17 to 32
  + High quality meant the score was 33 to 48

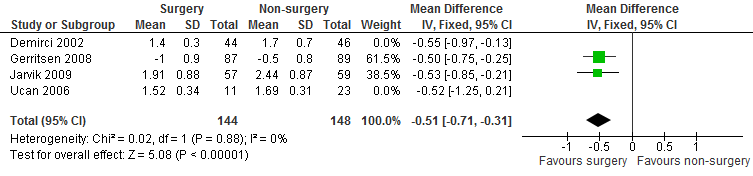
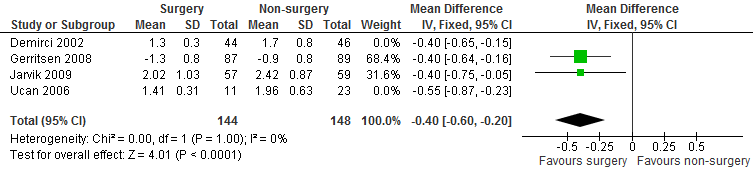
Results:

* The literature search identified 1333 articles, but after the abstracts review, only 10 articles potentially met the inclusion criteria, and 7 of these (5 RCTs and 2 comparative cohort studies) were included in the primary review
* A total of 4 studies were considered “high quality” on the Jadad scale, but no study had a Jadad score higher than 3 because all studies had 0 points for the double-blinding criterion
* 5 studies used patient self-administered function and symptom questionnaires, and 4 of these, which used disease-specific scales for hand function, could be pooled in a meta-analysis
  + One of these 4 pooled studies compared surgery with multi-modality treatment, one compared surgery with splinting plus local steroid injection, on compered surgery with two-dose steroid injection, and one compared surgery with splinting
  + The hand-specific scale which was used, the Boston Questionnaire, has 11 items for symptom severity and 8 items for functional difficulties
  + Each item in the questionnaire is scored from 1 to 5, with higher scores indicating worse symptomatology or function
  + The hand score entered into the analysis was the mean of all 11 symptom items and all 8 function items, such that the score was a number between 1 and 5
  + For patient-reported functional improvement at 6 months, surgery was better than non-surgery; the difference was that the functional impairment was a mean of 0.36 points lower than non-surgery (95% confidence interval from 0.22 to 0.47)
    - In spite of the diversity of non-operative comparisons, the heterogeneity, as estimated by the I2 statistic, was 40%, which is not generally considered to be heterogeneous, justifying the fixed-effect meta-analysis which was used by the authors
  + For self-reported symptom improvement at 6 months, surgery was also better than non-surgery; the difference was 0.43 points in favor of surgery (95% CI 0.29 to 0.57); the I2 statistic was 0%, meaning no heterogeneity
* Five studies evaluated the electrophysiologic improvement at 6 months followup; two high-quality studies measured median motor nerve distal latency and three moderate quality studies assessed the number of normal nerve conduction tests at 6 months
  + Surgery was found to be superior to non-surgery in the 2 studies which measured median motor nerve distal latency (0.5 ms, 95% CI 0.16 to 0.95), with borderline heterogeneity with I2  of 50% for the pooled studies
* Complications and side effects were more common in surgically than in non-surgically treated patients
  + There was large variation across studies in complication rates, due to some authors reporting all complications regardless of severity with other authors declaring only clinically important complications
  + For surgery, a number of minor adverse effects were seen, such as wound hematoma, skin irritation, painful scar or stiff wrist which resolved over a few weeks
  + For non-surgery, splinting sometimes led to swelling of the wrist, finger, or hand

Authors conclusions:

* Both surgical and nonsurgical conservative interventions are beneficial in the management of CTS
* Surgery provides better outcomes in terms of symptoms and return of nerve conduction, but there are more complications with surgery than with conservative treatment
* The size of the incremental benefit of surgery over non-surgery is small to moderate, and conservative treatment is effective can circumvent the need for surgery in a certain proportion of patients, making it a justified first line treatment

Comments:

* The authors acknowledge the limitations of the Jadad scale (attached) as a quality measure for assessing clinical trials, and supplement it with a 24-item Structured Effectiveness Quality Evaluation Scale (SEQES) which is attached
* Neither the Jadad nor the SEQES has a specific question relating to concealment of allocation, which is an important criterion for avoiding bias in randomized trials
  + However, the authors awarded only one point for randomization to the study by Ucan 2006
  + Ucan 2006 randomized patients with “previously prepared, randomly enumerated closed envelopes” without specifying that the envelopes were opaque, and it is possible that this led to the authors not awarding two points to the study for adequate randomization, giving it a Jadad score of 2, while the studies by Gerritsen 2002 and Jarvik 2009 had Jadad scores of 3 for making the method of allocation concealment more explicit
* The principal outcome of symptomatic and functional improvement at 6 months was estimated by pooling 4 studies which used similar questionnaires for symptoms and for function
  + Of these 4 studies, one was not randomized (Jadad score of 0), and the other was Ucan with a Jadad score of 2; the other two studies in the pooled meta-analysis were Gerritsen and Jarvik
  + Even though there were differences in the non-surgical intervention (Jarvik using multi-modality, Ucan using 3 months of splinting plus a steroid injection, Gerritsen using 6 weeks of splinting, Demirci using 2 steroid injections), the I2 heterogeneity measure is 40%, which does not exceed the 50% generally considered to indicate considerable heterogeneity
  + Ucan had two conservative treatment groups, one using only splinting (n=23) and one using splinting plus one steroid injection (n=23); the latter group was selected by the authors for comparison with surgery
  + Jarvik treated the conservative group with NSIDs and 6 sessions of individualized therapy with a hand therapist, which included instruction on splint use
* Figure 1 pools functional improvement data at 6 months from these 4 studies, while figure 2 pools symptom improvement at 6 months
  + The functional improvement from the 4 studies favored surgery, with the pooled estimate in Figure 1 being equal to 0.35 on a scale from 1 to 5
  + When the two weaker studies (Ucan 2006 and Demirci 2002) are removed from the analysis in Figure 1, and only the two higher quality studies are retained, the effect size in favor of surgery is larger, with a mean difference of 0.51 points: 
  + This is of interest because it is generally suspected that the effect of an intervention will appear greater in studies with poor control of bias than in studies with better control of bias
  + When Figure 2, symptom improvement at 6 months, is analyzed with only the two higher-quality studies, the estimate of the effect size for surgery is essentially the same, 0.40 in favor of surgery rather than 0.43: 
* The authors state that the treatment advantage of surgery is relatively small, but for the Boston questionnaire which was used to measure symptom and functional improvement, there is not an agreed upon minimum clinically important difference (MCID) against which the effect size can be compared
* However, it is possible to pool the effect sizes using standardized mean differences (SMD), which measure the differences in outcome in terms of standard deviations, and there is a commonly agreed upon scale for SMD
  + Less than 0.2 SD is no difference
  + 0.2 SD is “small”
  + 0.5 SD is “moderate”
  + 0.8 SD or greater is “large”
  + For Figure 1, the treatment advantage of surgery from the four studies is 0.58, which would qualify as a “moderate” advantage
* For the two higher-quality studies, patients qualified for inclusion without having had a prolonged period of CTS symptoms: Jarvik required 2 weeks and Gerritsen did not have a minimal duration of symptoms, while both Ucan and Demirci required 6 months of symptoms for entry into the study
* The authors conclude that further randomized trials comparing surgery versus non-surgery would be less helpful than good prognostic studies to identify which CTS patients require surgery after a fair trial of conservative treatment
  + This is a reasonable opinion, since the meta-analysis of randomized trials can estimate the average treatment advantage of surgery, but do not identify which patients required eventual surgery
* There is strong evidence from randomized trials that the benefits of steroid injection are transient in nature, being most effective in the very short term, but showing waning benefits in the intermediate and longer term

Assessment: High-quality systematic review and meta-analysis which supports strong evidence that in patients with CTS which has not become chronic, carpal tunnel release leads to a moderate treatment advantage with respect to functional improvement 6 months after surgery. Because the benefits of steroid injection are transient, the initial nonoperative treatment plan should not expect these injections to resolve the problem. However, there is considerable benefit to conservative treatment such as splinting and individualized hand therapy, which are appropriate for first-line treatment, since there is insufficient evidence to identify which patients are likely not to benefit from conservative treatment sufficiently to avoid surgery.

References:

Demirci S, Kutluhan S,et al: Comparison of open carpal tunnel release and local steroid treatment outcomes in idiopathic carpal tunnel syndrome. Rheumatol Int 2002, 22(1):33-37.

Gerritsen AAM, de Vet HCW, et al. Splinting vs. Surgery in the Treatment of Carpal Tunnel Syndrome. JAMA 2002;288:1245-1251.

Jarvik JG, Comstock BA, et al. Surgery versus non-surgical therapy for carpal tunnel syndrome: a randomised parallel-group trial. Lancet 2009, 374(9695):1074-81.

Ucan H, Yagci I, Yilmaz L, Yagmurlu F, Keskin D, Bodur H: Comparison of splinting, splinting plus local steroid injection and open carpal tunnel release outcomes in idiopathic carpal tunnel syndrome. Rheumatol Int. 2006, 27(1):45-51.

##### Additional file 2

Jadad et al. Scale\*

Please read the article and try to answer the following questions (see attached instructions):

1. Was the study described as randomized (this includes the use of words such as randomly, random, and randomization)?

2. Was the study described as double blind?

3. Was there a description of withdrawals and dropouts?

Scoring the items:

Either give a score of 1 point for each “yes” or 0 points for each “no.” There are no in-between marks.

Give 1 additional point if: For question 1, the method to generate the sequence of randomization was described and it was appropriate (table of random numbers, computer generated, etc.)

And/or: If for question 2 the method of double blinding was described and it was appropriate (identical placebo, active placebo, dummy, etc.)

Deduct 1 point if: For question 1, the method to generate the sequence of randomization was described and it was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc.)

And/or For question 2, the study was described as double blind but the method of blinding was inappropriate (e.g., comparison of tablet versus injection with no double dummy)

Guidelines for assessment

1. Randomization

A method to generate the sequence of randomization will be regarded as appropriate if it allowed each study participant to have the same chance of receiving each intervention and the investigators could not predict which treatment was next. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should be not regarded as appropriate.

2. Double blinding

A study must be regarded as double blind if the phrase “double blind” is used. The method will be regarded as appropriate if it is stated that neither the person doing the assessments nor the study participant could identify the intervention being assessed, or if in the absence of such a statement the use of active placebos, identical placebos, or dummies is mentioned.

3. Withdrawals and dropouts

Participants who were included in the study but did not complete the observation period or who were not included in the analysis must be described. The number and the reasons for withdrawal in each group must be stated. If there were no withdrawals, it should be stated in the article. If there is no statement on withdrawals, this item must be given no points.

**\* From Jadad, A. R., Moore, A., Carroll, D., et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control. Clin. Trials* 17: 1, 1996.**

##### Additional file 3

Structured Effectiveness Quality Evaluation Scale (SEQES)

# Evaluation Criteria Score Study question

**Study question**

1. Was the relevant background work cited to establish a foundation for the research question?

**Study design**

2. Was a comparison group used?

3. Was patient status at more than one time point considered?

4. Was data collection performed prospectively?

5. Were patients randomized to groups?

6. Were patients blinded to the extent possible?

7. Were treatment providers blinded to the extent possible?

8. Was an independent evaluator used to administer outcome measures?

**Subjects**

9. Did sampling procedures minimize sample/selection biases?

10. Were inclusion/exclusion criteria defined?

11. Was an appropriate enrollment contained?

12. Was appropriate retention/follow-up obtained?

**Intervention**

13. Was the intervention applied according to established principles?

14. Were biases due to the treatment provider minimized (ie attention, training)?

15. Was the intervention compared to appropriate comparator?

**Outcomes**

16. Was an appropriate primary outcome defined?

17. Were appropriate secondary outcomes considered?

18. Was an appropriate follow-up period incorporated?

**Analysis**

19. Was an appropriate statistical test(s) performed to indicate differences related to the intervention?

20. Was it established that the study had significant power to identify treatment effects?

21. Was the size and significance of the effects reported?

22. Were missing data accounted for and considered in analyses?

23. Were clinical and practical significance considered in interpreting results?

**Recommendations**

24. Were the conclusions/clinical recommendations supported by the study objectives, analysis, and results?

**Total quality score (sum of above)=**