| CMAJ Open 2014-0130 | |
|---|---|
| Title | Operative versus nonoperative interventions for common fractures of the clavicle: a meta-analysis of randomized controlled trials |
| Authors | Tahira Devji BSc, Ydo Kleinlugtenbelt MD, Nathan Evaniew MD, Bill Ristevski MD MSc, Shoghag Khoudigian BSc, Mohit Bhandari MD PhD |
| Reviewer 1 | |
| Name | Waddell, James P. |
| Institution | St. Michael's Hospital, Orthopaedic Surgery |
| General comments | Page 6<br>How did you define "routine hardware removal"? I'm not sure how this would be included in the database since I'm sure not all studies would have defined it as such.<br>Page 8<br>Your decision to treat those that were lost to follow-up as not having an event is contrary to the generally accepted practice of assuming they did have an event. I'm not sure why you would assume that people that didn't come back did not have an event – they may well have gone to another doctor because they were unhappy about the event occurring in the hands of their primary orthopaedic surgeon. I will need a little explanation of why you did that.<br>Page 9/10<br>I'm not sure why you included trials that compared one operative implant to another if the purpose of your study was to determine the value of operative treatment versus non-operative treatment. I don't see that including those trials of one device versus another would add anything to that debate.<br>Page 11<br>You state that long term function favoured operatively treated patients which seems clear enough (my bias personally is towards non-operative care) but then you state the pooled estimate did not exceed the threshold of + 1.33 SD for the MID. You will have to explain this a little more clearly for the statistically challenged such as myself – if long term function favoured operatively treated patients what does this modifier really mean in terms of patient outcomes.<br>Page 12<br>I know I am repeating myself but once again you state "there is modest functional improvement at one year in operatively treated patients, however, this finding did not reach clinical significance". I'm not sure the methodology that you are using does in fact support that.<br>Page 14<br>Leroux clearly demonstrated a re-operation rate that was substantially higher than you have found in your clinical trials. This implies to me a bias in the clinical trial population towards lower re-operation rate for surgical patients. You don't take this into account other than to comment on its possibility. This leaves open the question of the value of database mining (either insurance databases or registries) versus randomized trials. As a general rule in my experience registry data invariably shows a higher complication rate than that demonstrated in any type of randomized trial primarily because of the level of expertise of those individuals performing the trial. I think this is relevant and deserves more than a passing comment in your manuscript. |

| | |
|---|---|
| Author response | 1. How did you define "routine hardware removal"? I'm not sure how this would be included in the database since I'm sure not all studies would have defined it as such.<br>RESPONSE:<br>We defined routine hardware removal as the need for implant removal following fracture healing. Only studies comparing pin fixation to non-operative treatment reported routine hardware removal, as pins/nails are routinely removed. This is a procedure that is usually done under local anesthesia, minimal sedation, and a tiny incision over the tip of nail, and not likely to result in any complications. The need for plate removal is typically indicated as a result of discomfort and necessitates new admissions, general anesthesia, and an additional large-sized incision.<br>We took a more conservative approach by excluding the routine pin/nail removals. Had we included these as events in our outcome of 'secondary operations', the relative risk of having a secondary operation would have been much higher in the operative group, as not all pooled events would carry the same health risks.<br>2. Your decision to treat those that were lost to follow-up as not having an event is contrary to the generally accepted practice of assuming they did have an event. I'm not sure why you would assume that people that didn't come back did not have an event – they may well have gone to another doctor because they were unhappy about the event occurring in the hands of their primary orthopaedic surgeon. I will need a little explanation of why you did that.<br>RESPONSE:<br>The nature and criteria of the primary outcomes (secondary operations and all complications) selected for this review were such that patients experiencing an event would require surgical intervention or additional medical management. We felt confident in the statistical approach to handle missing patient data, as it is highly plausible that most patients would return for follow-up if unsatisfied or experienced an adverse event. However, we acknowledge that it is possible that patients could have gone elsewhere to another surgeon to seek treatment.<br>Thus, in addition to this primary analysis, we conducted two sensitivity analyses:<br>We performed a complete case analysis, which excludes missing data from both the numerator and denominator when calculating the relative risk (RR) of a trial. A complete case analysis would effectively increase the RR of an event occurring, as fewer individuals comprise the denominator.<br>To further test the robustness of the assumption made in our primary analysis, we conducted arm-level assumption analyses, where the relative incidence among those with missing data were assigned the same incidence as those followed-up in the same arm (RILTFU/FU = 1).<br>Neither the complete-case nor the RILTFU/FU = 1 produced a statistically different result from the primary analysis performed.<br>3. I'm not sure why you included trials that compared one operative implant to another if the purpose of your study was to determine the value of operative treatment versus non-operative treatment. I don't see that including those trials of one device versus another would add anything to that debate.<br>RESPONSE:<br>Several trials have made the comparison between operative versus nonoperative treatment for midshaft clavicle fractures, as a large debate persists as to which is the most optimal method of |

| | |
|---|---|
| | management. However, in the last 5 years, a number of trials have also investigated various surgical techniques and the use of different implants to treat clavicle fractures. Previous reviews on this clinical topic have only focused on the operative versus nonoperative debate. Our review adds to this body of literature by providing data from the largest and most recent trial, but also generates a summary of the evidence on surgical options for these injuries, as well as nonsurgical techniques. This, in part adds to the uniqueness of our review amongst the other strengths mentioned in the introduction of our manuscript.<br>4. You state that long term function favoured operatively treated patients which seems clear enough (my bias personally is towards non-operative care) but then you state the pooled estimate did not exceed the threshold of + 1.33 SD for the MID. You will have to explain this a little more clearly for the statistically challenged such as myself – if long term function favoured operatively treated patients what does this modifier really mean in terms of patient outcomes.<br>RESPONSE:<br>The minimal important difference (MID) describes the smallest change in the outcome of interest that informed patients perceive as important, either beneficial or harmful, and which would lead the patient or clinician to consider a change in management. Knowledge of the MID facilitates the interpretation of the magnitude of treatment effects, placing a greater emphasis on clinical significance as opposed to statistical significance.<br>To improve interpretability, we converted Standardized Mean Difference (SMD) results to the Disabilities of the Arm, Shoulder, and Hand (DASH) score. The approximation of the SMD as a DASH score is carried out through the following formula: Mean Difference (DASH units) = SMD x (median SD of DASH of included trials).<br>We have revised the way in which we have reported results of long-term functional outcome scores to ensure this concept is clearer:<br>"This is equivalent to an estimated DASH mean difference of 3.5 (95% CI 0.00 to 6.85). This treatment effect failed to exceed the threshold of patient importance based on the MID (10.2 points) (Figure 6)."<br>As mentioned in the methods, the MID for the DASH questionnaire is estimated to be 10.2 points. Since the converted SMD to MD in DASH units reported in the above paragraph (3.5; 95% CI 0.00 to 6.85) is less than 10.2, we can conclude that there was a statistically significant but clinically unimportant difference in function at 1-year.<br>5. I know I am repeating myself but once again you state "there is modest functional improvement at one year in operatively treated patients, however, this finding did not reach clinical significance". I'm not sure the methodology that you are using does in fact support that.<br>RESPONSE:<br>Please see explanation above.<br>6. Leroux clearly demonstrated a re-operation rate that was substantially higher than you have found in your clinical trials. This implies to me a bias in the clinical trial population towards lower re-operation rate for surgical patients. You don't take this into account other than to comment on its possibility. This leaves open the question of the value of database mining (either insurance databases or registries) versus randomized trials. As a general rule in my experience registry data invariably shows a higher complication rate than that demonstrated in any type of randomized trial primarily because of the level of expertise of those |

| | |
|---|---|
| | individuals performing the trial. I think this is relevant and deserves more than a passing comment in your manuscript.<br>RESPONSE:<br>Thank you for providing insight and highlighting the need to discuss this comparison of observational and trial data further. We have revised our discussion around the discrepancy in results:<br>"There are important differences in design characteristics between observational studies and randomized trials that may be responsible for contradictory estimates of treatment effects. Firstly, infrequent events and long-term clinical outcomes are often difficult to study in randomized trials and may be more suitably investigated in large observational studies47. Secondly, it is plausible that surgeons involved in the majority of surgical trials may have substantial generic surgical expertise and expertise in the intervention under evaluation, which may not represent the skill level of the surgical community in which the intervention will be implemented48. Despite the obvious discrepancy between this observational data and the current RCT literature in terms of re-operations, it is incumbent upon us to recognize the complementary roles of both sources of information and understand that the complete body of evidence could have profound clinical implications." |
| Reviewer 2 | |
| Name | Desy, Nicholas |
| Position | Fellow, Orthopaedic Trauma Surgery |
| Institution | McGill University Health Center, Orthopaedic Surgery |
| General comments | Clavicle fractures are common injuries and are seen by many types of clinicians. In recent years, surgical management has been increasingly utilized for midshaft clavicle fractures as a result from the direction of several randomized controlled trials (RCTs) that have shown superior outcomes compared to nonoperative treatment for certain fracture criteria. Devji et al. report a systematic review and meta-analysis of randomized controlled trials on the treatment of clavicle fractures. They found that surgery does not improve the reoperation risk or patient function compared to nonoperative management. This is in contrast to some of the RCTs and meta-analyses that have shown better function with surgical treatment. They conclude that it is still uncertain which fracture characteristics would benefit most from surgery and current evidence remains inconclusive for treatment guidelines.<br>The strengths of this study include: it is well written with strong methodology, the reporting is consistent with the PRISMA guidelines, the use of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) to rate the quality of evidence and a recent RCT was added to the study pool.<br>Recommendations for the paper are:<br>Major comments<br>1. This study concludes that surgery does not improve the reoperation risk for patients with clavicle fractures. However, the most common reason for a secondary procedure amongst the operative group was hardware irritation (54.8%) and amongst the nonoperative group was symptomatic nonunion (57.1%). Hardware removal for hardware irritation is quite different from open reduction and internal fixation with or without bone graft for symptomatic nonunion. It could be useful if the authors discussed this difference, especially that even though the reoperation risk is the same, the types reoperations for the most common reasons in each group are quite different with very different risks and rehabilitation.<br>2. Discussion / Table 4 (GRADE summary): Given that the |

| | |
|---|---|
| | current evidence is of low to very low quality, it would be useful to read some recommendations in the discussion on how to improve the quality of evidence in the literature to fully determine which patients or fracture characteristics would benefit most from surgery.<br>Minor comments<br>1. In Background, page 5, line 48, the authors should reference the prior reviews, including previous meta-analyses and systematic reviews that have been published on clavicle fractures<br>2. Figure 1 (Study flow diagram): A number may be incorrect in the flow diagram. It shows that a full-text screening was done for 36 studies and another 7 articles were also screened by hand searching. Twenty seven articles were then excluded (20 for reasons listed and seven ongoing studies) which would give 16 studies included in qualitative synthesis (not 15).<br>3. Results, Study characteristics, page 10, line 56. The authors should reference the nine studies that compared operative to nonoperative treatment, similar to the references that were included for the studies that compared different implants and the study that managed all fractures nonoperatively.<br>4. Page 11, line 34: the number 8 should be written as eight<br>5. Page 12, line 34: the number 2 should be written as two<br>6. Figures 4, 5 and 6 (Results/Analyses for operative versus nonoperative treatment): How come the paper Koch 2008 was included in Figure 4 (secondary surgery) and Figure 5 (complications) but not in Figure 6 (long-term function)? Similarly, why was Chen 2011 included in Figure 6 but not in Figures 4 and 5?<br>7. Page 13, line 48: a hyphen is missing between 1 and year (to make it consistent) |
| Author response | Clavicle fractures are common injuries and are seen by many types of clinicians. In recent years, surgical management has been increasingly utilized for mid-shaft clavicle fractures as a result from the direction of several randomized controlled trials (RCTs) that have shown superior outcomes compared to non-operative treatment for certain fracture criteria. Devji et al. report a systematic review and meta-analysis of randomized controlled trials on the treatment of clavicle fractures. They found that surgery does not improve the reoperation risk or patient function compared to non-operative management. This is in contrast to some of the RCTs and meta-analyses that have shown better function with surgical treatment. They conclude that it is still uncertain which fracture characteristics would benefit most from surgery and current evidence remains inconclusive for treatment guidelines.<br>The strengths of this study include: it is well written with strong methodology, the reporting is consistent with the PRISMA guidelines, the use of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) to rate the quality of evidence and a recent RCT was added to the study pool.<br>RESPONSE:<br>Thank you for this comment and your thoughtful review of our manuscript.<br>Recommendations for the paper are:<br>Major comments<br>1. This study concludes that surgery does not improve the reoperation risk for patients with clavicle fractures. However, the most common reason for a secondary procedure amongst the operative group was hardware irritation (54.8%) and amongst the nonoperative group was symptomatic nonunion (57.1%). Hardware removal for hardware irritation is quite different from open reduction and internal fixation with or without bone graft for |

| | |
|---|---|
| | symptomatic nonunion. It could be useful if the authors discussed this difference, especially that even though the reoperation risk is the same, the types reoperations for the most common reasons in each group are quite different with very different risks and rehabilitation.<br>RESPONSE:<br>Thank you for this comment and highlighting the need for this distinction.<br>We have added to the discussion as suggested:<br>"Although the risk of a secondary procedure was similar between both treatment groups, the reasons for delayed intervention were quite different. Hardware removal for hardware irritation was the most common indication for a secondary procedure amongst operatively treated patients, whereas nonoperatively treated patients experienced symptomatic nonunion. The latter indication would typically require open reduction and internal fixation with or without bone graft, which may be associated with greater risk for complications and the need for a longer rehabilitation period."<br>2. Discussion / Table 4 (GRADE summary): Given that the current evidence is of low to very low quality, it would be useful to read some recommendations in the discussion on how to improve the quality of evidence in the literature to fully determine which patients or fracture characteristics would benefit most from surgery.<br>RESPONSE:<br>We have revised 'implications for research' in our discussion to incorporate the following sentences:<br>"Recurrent study design limitations, including small sample sizes, lack of blinding, and loss to follow up must be overcome to improve the quality of evidence from future randomized trials. Unified evaluation criteria for outcomes such as nonunion and malunion should be applied to all trials evaluating treatment interventions for these fractures."<br>"Future trials should aim to better identify the subgroup of patients who may benefit from primary surgical intervention and establish optimal surgical indications."<br>Minor comments<br>1. In Background, page 5, line 48, the authors should reference the prior reviews, including previous meta-analyses and systematic reviews that have been published on clavicle fractures<br>RESPONSE:<br>We have revised the following sentence to include the appropriate references:<br>"Our study advances prior reviews1, 9, 11, 12 by including new evidence from randomized controlled trials (RCTs), our focus on major health outcomes such as secondary operations within 1-year, and improved summary of evidence using Grading of Recommendations Assessment, Development, and Evaluation (GRADE) to rate the quality of evidence available for each patient-focused outcome."<br>2. Figure 1 (Study flow diagram): A number may be incorrect in the flow diagram. It shows that a full-text screening was done for 36 studies and another 7 articles were also screened by hand searching. Twenty seven articles were then excluded (20 for reasons listed and seven ongoing studies) which would give 16 studies included in qualitative synthesis (not 15).<br>RESPONSE:<br>We regret this error. One article was mistakenly not accounted for in the study flow diagram. Schemitsch et al., 2011 published 2-year follow-up data from the COTS 2007 trial in a subsequent |

| | |
|---|---|
| | publication.<br>The number of articles excluded in the full text review phase is in fact 21 and not 20. We have updated the study flow diagram to reflect this change.<br>We have also added a sentence in the results section to capture this information: "Two-year follow-up data from one trial10 was reported in a separate publication36."<br>3. Results, Study characteristics, page 10, line 56. The authors should reference the nine studies that compared operative to nonoperative treatment, similar to the references that were included for the studies that compared different implants and the study that managed all fractures nonoperatively.<br>RESPONSE:<br>We regret this oversight and thank you for your attention to this detail. We have referenced the respective operative versus nonoperative trials accordingly.<br>4. Page 11, line 34: the number 8 should be written as eight<br>RESPONSE:<br>Thank you for noting this. We have revised as suggested.<br>5. Page 12, line 34: the number 2 should be written as two<br>RESPONSE:<br>Thank you for noting this. We have revised as suggested.<br>6. Figures 4, 5 and 6 (Results/Analyses for operative versus nonoperative treatment): How come the paper Koch 2008 was included in Figure 4 (secondary surgery) and Figure 5 (complications) but not in Figure 6 (long-term function)? Similarly, why was Chen 2011 included in Figure 6 but not in Figures 4 and 5?<br>RESPONSE:<br>Koch 2008 did not evaluate physical function, and thus, was not included in the pooled analysis for long-term function.<br>Chen 2011 provided efficacy data on physical function at 1-year, as assessed by the DASH and Constant instruments. Chen 2011 incompletely reported data on secondary procedures and complications. Abstraction of this data was difficult and nonsensical; thus, to avoid the inclusion of inaccurate information in our pooled analysis, we erred on the side of caution and excluded this study.<br>We regret that it may be unclear as to which studies evaluated outcomes of interest for this review. We have revised a sentence in the results section under 'functional scores':<br>"All eight studies included in the pooled analysis evaluated function at 1-year with the exception of one trial42, which assessed shoulder function at 2-year follow-up."<br>It is now more apparent that only eight of the nine studies included in our quantitative synthesis were evaluated for long-term function.<br>7. Page 13, line 48: a hyphen is missing between 1 and year (to make it consistent)<br>RESPONSE:<br>We regret this error and thank you for your attention to this detail. We have ensured that 1 and year is written with a hyphen as "1-year" consistently throughout the manuscript. |