
Design: Meta-analysis of randomized clinical trials

Study question: In the setting of displaced midclavicular fractures, are there differences in outcome between surgical and nonsurgical treatment?

PICOs:

- Patients: adolescents (age >10) and adults diagnosed with middle third clavicle fractures
- Interventions: Surgical intervention including plate-and-screw fixation, Kirschner wires, titanium nail, Knowles pin, and external fixation
- Comparison interventions: Conservative treatments such as slings, strapping, figure-of-eight bandages, or other physical treatments as well as adjunctive therapy such as ultrasound
- Outcomes: Function/disability, pain, therapeutic failure
  o Functional scales such as the Constant score and the DASH, which include patient-reported measures of function, were preferred
  o Pain was preferably measured by validated scales such as the VAS and the Numerical Rating Scale (NRS)
  o Treatment failure was measured by the number of patients who have undergone or are being considered for secondary surgical intervention for nonunion, malunion, and other complications
  o Timing of outcomes was short term (up to six weeks), intermediate term (6 weeks to 6 months), and long term (longer than six months) after the end of treatment
- Study types: Randomized or quasi-randomized clinical trials comparing a surgical versus a conservative intervention for a mid-clavicle fracture

Study selection:

- Databases included MEDLINE, EMBASE, LILACS, and CENTRAL through December 2012
  o Reference lists and proceedings of a number of orthopedic conferences were also searched
  o Registries of ongoing trials (from WHO and clinicaltrials.gov) were also searched for recently completed and ongoing trials
- Two authors independently assessed articles for inclusion and extracted data from the trials, resolving any disagreements through discussion or adjudication from a third author
- Risk of bias was assessed through application of the Cochrane criteria of randomization sequence generation, allocation concealment, blinding of participants, providers, and outcome assessors, incomplete follow-up, selective reporting, inappropriate influence of funding sources, differences in rehabilitation programs, and bias due to inexperience of surgeons and other care providers.

- Heterogeneity in meta-analysis was graded with the $I^2$ statistic: from 0 to 40% might not be important; from 30 to 60% may mean moderate heterogeneity; from 50 to 90% may mean substantial heterogeneity, and from 75 to 100% was considerable heterogeneity.

- The authors planned subgroup analyses for age, type of surgical intervention (intramedullary versus plate fixation), timing of surgery (immediate versus delayed), two fragments versus more than two fragments, and primarily undisplaced versus displaced.

Results:

- The search strategy yielded 42 citations potentially eligible for inclusion; 8 studies with 555 patients were included in the analysis of results.

- All studies included only patients with displaced mid-clavicular fractures.

- All studies were parallel RCTs with two intervention groups; one study had to be translated from German and one from Portuguese by the authors.

- The results were grouped into two basic comparisons: plate fixation versus sling, and intramedullary fixation versus sling or figure-of-eight immobilization.

- All studies had methodological flaws rendering them at a high risk of bias.
  - None of the studies had blinded assessment of outcome.
  - Concealment of allocation was adequately described in four studies.
  - Two studies were considered at high risk of bias because of greater attrition in the conservative group than in the surgical group.

- For function or disability at one year or more, meta-analysis pooled data from seven studies, four involving plate fixation and three involving intramedullary fixation.
  - The outcomes were diverse, but a hierarchy of measures was established by the authors in the following order of preference: Constant score, DASH, UCLA, Single Assessment Numerical Evaluation (SANE), and L’Insalata score (the latter two used in only one study, Judd 2009).
  - The pooling of functional outcome data took the form of a standardized mean difference (SMD), in which group differences are expressed in terms of standard deviations (by convention, less than 0.2 SD is no difference, 0.2 to 0.5 SD is a small difference, 0.5 to 0.8 SD is a moderate difference, and more than 0.8 SD is a large difference).
The pooled SMD from the seven studies was a non-statistically significant 0.46 SD (95% confidence interval from -0.06 to 0.98), and the results were highly heterogeneous ($I^2$ was 85%)

- A test for subgroup differences between intramedullary and plate fixation does not support a difference for the two surgical interventions ($I^2$ was 0%)

- Only two studies reported pain scores, and data could only be extracted from one study, which reported no significant pain differences between conservative and surgical treatment at 3 weeks, 6 weeks, 3 months, and one year

- Treatment failure defined as symptomatic nonunion, symptomatic malunion, early mechanical failure, or brachial plexus irritation (reported in Analysis 1.8), was marginally in favor of surgical treatment (9/232 versus 24/205, relative risk of 0.38 with 95% confidence interval from 0.15 to 0.99); 9 of the 11 symptomatic malunion failures were reported in only one study

- There were no differences in treatment failures between plate fixation and intramedullary fixation

- Some adverse events, such as dehiscence, hardware irritation, and infection, were specific to the surgical groups

- Adverse outcomes which could occur in both groups, such as stiffness and restriction of shoulder movement, occurred more frequently in the conservative (24/73) than in the surgical groups (5/85)

Authors’ conclusions:

- Low-quality evidence from the included trials indicate that surgery may not result in a significant functional improvement or a significant pain reduction in comparison to conservative treatment

- All trials failed to report at least some important outcomes; only one study presented data on long term pain

- The data on malunions indicated that they were more common in conservatively treated patients, but most of the malunions were reported in only one study; because malunions are generally not symptomatic, these adverse event differences should be viewed with caution

- The overall quality of the evidence was low because of methodological flaws such as failure of blinding of outcome assessment; it was also hampered by inconsistency across trials and imprecision of adverse events measured across trials

- McKee 2012 combined six RCTs showing that surgical intervention provided lower rates of nonunion and symptomatic malunion and an earlier functional return compared to conservative treatments

- McKee did not include three RCTs which were included in the present meta-analysis
- McKee also used a numerical scale to assess trial quality; these scales are not recommended (Juni 1999) because they provide unreliable assessments of validity.

- Until conclusive evidence becomes available, treatment options must be made on the basis of individual patient characteristics with careful consideration of the harms and benefits and patient preferences.

Comments:

- The major contrast with McKee 2012 is not so much on the assessment of study quality as on the outcomes which are being pooled.

- McKee focused on the complications of symptomatic nonunion and malunion, and not on pooling of scores from functional scales.
  - The three studies in McKee which reported Constant scores showed a very small difference (only 4 points on a 100 point scale) in favor of surgery.
  - The current meta-analysis does not directly conflict with McKee, but the emphasis is on a different set of outcome measures.

- There was heterogeneity of outcomes for the functional scores at one year, which the authors intended to explore by comparing the characteristics of individual studies.
  - In Analysis 1.1 on page 57, the pooled STD is 0.46 SD with 95% confidence interval from -0.06 to 0.98 with I² of 85%.
  - A more recent study of plate fixation, Robinson 2013, would have met the inclusion criteria for this meta-analysis, but changes the pooled results only very slightly from a pooled STD of 0.46 to 0.43 (signs of the treatment effects are reversed from Analysis 1.1 but the comparisons are the same):

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Conservative Mean</th>
<th>SD</th>
<th>Total</th>
<th>Surgery Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chm 2011</td>
<td>84</td>
<td>5.2</td>
<td>30</td>
<td>42.1</td>
<td>4.2</td>
<td>30</td>
<td>12.2</td>
<td>-0.46 (1.13, -0.00)</td>
<td></td>
</tr>
<tr>
<td>Cole 2007</td>
<td>90.8</td>
<td>14</td>
<td>48</td>
<td>96.3</td>
<td>26</td>
<td>52</td>
<td>13.3</td>
<td>-0.92 (0.84, -0.12)</td>
<td></td>
</tr>
<tr>
<td>Figueiredo 2009</td>
<td>30.8</td>
<td>5.3</td>
<td>16</td>
<td>32.9</td>
<td>2.6</td>
<td>24</td>
<td>11.3</td>
<td>-0.63 (1.17, 0.12)</td>
<td></td>
</tr>
<tr>
<td>Judd 2009</td>
<td>97</td>
<td>3.3</td>
<td>28</td>
<td>93.5</td>
<td>4.2</td>
<td>29</td>
<td>12.2</td>
<td>0.31 (0.36, 1.46)</td>
<td></td>
</tr>
<tr>
<td>Marabin 2011</td>
<td>78.8</td>
<td>6</td>
<td>24</td>
<td>84.8</td>
<td>6</td>
<td>26</td>
<td>11.4</td>
<td>-1.54 (-2.10, -0.98)</td>
<td></td>
</tr>
<tr>
<td>Robinson 2013</td>
<td>97.8</td>
<td>12.23</td>
<td>92</td>
<td>92</td>
<td>18.55</td>
<td>86</td>
<td>14.4</td>
<td>-0.27 (0.56, 0.03)</td>
<td></td>
</tr>
<tr>
<td>Smeik 2009</td>
<td>93.7</td>
<td>6</td>
<td>30</td>
<td>97.9</td>
<td>1.7</td>
<td>30</td>
<td>12.3</td>
<td>-0.94 (1.48, -0.40)</td>
<td></td>
</tr>
<tr>
<td>Virtanen 2013</td>
<td>86.1</td>
<td>8.9</td>
<td>25</td>
<td>85.5</td>
<td>11.5</td>
<td>26</td>
<td>12.2</td>
<td>-0.04 (0.59, 0.95)</td>
<td></td>
</tr>
</tbody>
</table>

| Total (95% CI)    | 294               | 313| 100.0%| -0.43 (-0.84, -0.01) |

- One study appears to be an outlier in showing a significant advantage of nonoperative treatment (Judd 2009):
  - Judd was the only study to use a modified Hagie pin for the surgical group, and Judd’s conclusions were specific to the performance of that pin.
  - McKee mentions the study by Strauss 2007 discussing the problems with the Hagie pin and its need for frequent hardware removal.
When Judd 2007 is removed from the analysis, the heterogeneity is significantly reduced from $I^2$ of 85% to 67%, and the pooled SMD is increased to 0.59 (95% CI from 0.28 to 0.91).

Because Judd’s conclusions were specific to the Hagie pin, it would be appropriate to comment on that fact and to note that Strauss 2007 had discussed the removal rate of that device, whose difficulties may not apply to other types of intramedullary pin.

The exploration of heterogeneity is therefore not of high quality.

- Of the three studies not included by McKee, two were published later and one, published in Portuguese, was not translated by the authors, but was also at a high risk of bias.
- The authors are justified in lamenting the lack of quality of most of the included trials and the failure to report on pain scores.
- Robinson 2013 also failed to report results on pain.
- Mirzatolooei 2011 has a larger effect size than any of the other studies, and was restricted to comminuted fractures rather than displaced fractures in general; this may or may not be a reason for the larger effect size of surgery.
- Nonunion is not necessarily associated with functional deficits; for example, there were 6 nonunions in the conservative treatment group of Virtanen 2012, but because of minor disability associated with nonunion, none of the 6 elected to undergo surgical treatment when offered.
- Robinson 2013 is the only study to have blinded assessment of range of motion and strength (a physiotherapist examined the patient who was wearing a T-shirt), and probably has the best control of bias of any of the studies.
- The evidence overall remains inconclusive with respect to the functional advantages of surgical repair of mid-clavicular fractures, but a large treatment effect cannot be ruled out.
- One appropriate conclusion is that the decision whether to operate should be shared between physician and patient.
Assessment: Adequate for evidence that surgical treatment of displaced midclavicular fractures reduces the rate of nonunion, but the functional benefits of surgery remain uncertain at one year, since nonunions are frequently neither symptomatic nor disabling

References:


