**Table 1. Methodological quality assessment scheme**

<table>
<thead>
<tr>
<th>Items</th>
<th>Scores</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the assigned treatment adequately concealed prior to allocation?</td>
<td>Y = method did not allow disclosure of assignment.</td>
<td>Cochrane code (see Handbook): Clearly yes = A; Not sure = B; Clearly no = C.</td>
</tr>
<tr>
<td></td>
<td>? = small but possible chance of disclosure of assignment or unclear.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = quasi-randomised, or open list or tables.</td>
<td></td>
</tr>
<tr>
<td>2. Were the outcomes of participants who withdrew described and included in the analysis (intention-to-treat)?</td>
<td>Y = withdrawals well described and accounted for in analysis.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>? = withdrawals described and analysis not possible, or probably no withdrawals.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = no mention, inadequate mention, or obvious differences and no adjustment.</td>
<td></td>
</tr>
<tr>
<td>3. Were the outcome assessors blinded to treatment status?</td>
<td>Y = effective action taken to blind assessors.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>? = small or moderate chance of unblinding of assessors, or some blinding of outcomes attempted.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = not mentioned or not possible.</td>
<td></td>
</tr>
<tr>
<td>4. Were important baseline characteristics reported and comparable?</td>
<td>Y = good comparability of groups.</td>
<td>Although many characteristics including hand dominance are important, the principal confounders are considered to be age, gender, type of fracture.</td>
</tr>
<tr>
<td></td>
<td>? = confounding small, or comparability reported in text without confirmatory data.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = large potential for confounding, or not discussed.</td>
<td></td>
</tr>
<tr>
<td>5. Were the trial participants blind to assignment status after allocation?</td>
<td>Y = effective action taken to blind participants.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>? = small or moderate chance of unblinding of participants.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = not possible, or not mentioned (unless double-blind), or possible but not done.</td>
<td></td>
</tr>
<tr>
<td>6. Were the treatment providers blind to assignment status?</td>
<td>Y = effective action taken to blind treatment providers.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>? = small or moderate chance of unblinding of treatment providers.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = not possible, or not mentioned (unless double-blind), or possible but not done.</td>
<td></td>
</tr>
<tr>
<td>7. Were care programmes, other than the trial options, identical?</td>
<td>Y = care programmes clearly identical.</td>
<td>Examples of clinically important differences in other interventions are: time of intervention, duration of intervention, difference in rehabilitation.</td>
</tr>
<tr>
<td></td>
<td>? = clear but trivial differences, or some evidence of comparability.</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Methodological quality assessment scheme (Continued)

<table>
<thead>
<tr>
<th>Question</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Were the inclusion and exclusion criteria for entry clearly defined?</td>
<td>Y</td>
<td>Y = clearly defined (including type of fracture).</td>
</tr>
<tr>
<td></td>
<td>?</td>
<td>? = inadequately defined.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N = not defined.</td>
</tr>
<tr>
<td>9. Were the outcome measures used clearly defined?</td>
<td>Y</td>
<td>Y = clearly defined.</td>
</tr>
<tr>
<td></td>
<td>?</td>
<td>? = inadequately defined.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N = not defined.</td>
</tr>
<tr>
<td>10. Were the accuracy and precision, with consideration of observer variation, of the outcome measures adequate; and were these clinically useful and did they include active follow up?</td>
<td>Y</td>
<td>Y = optimal.</td>
</tr>
<tr>
<td></td>
<td>?</td>
<td>? = adequate.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N = not defined, not adequate.</td>
</tr>
<tr>
<td>11. Was the timing (e.g. duration of surveillance) clinically appropriate?</td>
<td>Y</td>
<td>Y = optimal. (&gt; 1 year)</td>
</tr>
<tr>
<td></td>
<td>?</td>
<td>? = adequate. (6 months - 1 year)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N = not defined, not adequate. (&gt; 6 months)</td>
</tr>
</tbody>
</table>

**Measures of treatment effect**

Dichotomous outcome data were analysed as risk ratios (RR) with 95% confidence intervals. Continuous outcome data were expressed as mean differences (MD) with 95% confidence intervals. We intended to pool the data as a mean difference (MD); if two or more studies presented data derived from the same validated instrument of evaluation (with the same units of measure); if primary studies measured the same variables through different instruments (and different units of measure) we intended to use the standardised mean difference (SMD).

We planned to assess the heterogeneity of estimate effects between the included studies by visual inspection of the forest plot, and using the chi-squared test and the I² statistic.

**Data synthesis**

When considered appropriate, we planned to pool results of comparable groups of trials using the fixed-effect model and 95% confidence intervals. However, the results using the random-effects model will also be inspected where there is diversity in clinical or methodological characteristics.

**Dealing with missing data**

With the purpose of including all patients randomised to any intervention, we performed an intention-to-treat analysis. When there was insufficient information relative to estimate effects, such as number of participants, means, means of uncertainty (standard deviation or error), or number of events and participants, we contacted the main authors of the included trials. When it was impossible to acquire adequate data for the forest plot (e.g., means and standard deviations), we present data in the text and/or appendices.

**Assessment of heterogeneity**

Subgroup analysis and investigation of heterogeneity

If data had been available, we had aimed to perform subgroup analyses by age (adolescent; adult) and type of fracture (primarily: undisplaced, displaced, and foreshortening), using the test for interaction outlined in Altman 2003.

**Sensitivity analysis**

Where data were available, we planned to develop sensitivity analyses examining various aspects of trial and review methodology, including the effects of missing data and study quality (specifically allocation concealment and outcome assessor blinding).
RESULTS

Description of studies
See: Characteristics of included studies; Characteristics of excluded studies; Characteristics of ongoing studies.

Results of the search
The search of five main electronic databases for completed research yielded 159 references (Figure 1). JB and ML screened the titles and abstracts for references and finally excluded 152 citations. Among these, 115 were duplicated citations or not relevant, 31 were excluded because they did not meet the inclusion criteria for participants or interventions, and six were not randomised or quasi-randomised trials. The remaining seven potentially relevant studies were evaluated from full trial reports. Three trials were included (Andersen 1987; Hoofwijk 1988; Lubbert 2008). A translation of one report (Jensen 1985) showed it that was another report of Andersen 1987. One full report and two abstracts, one of which reported interim results (Hoofwijk 1986), were available for Hoofwijk 1988. Finally, one study (Thompson 2005) was excluded. Two trials were translated to English (see Acknowledgements). Of the two further trials identified from Trial Registers, one was excluded (Talbot 2008) and one is ongoing (Roberti 2008).
Figure 1. Search strategy

Algorithm of search strategy results

Total number of references = 159
- MEDLINE (PubMed) = 57
- CENTRAL (Wiley) = 24
- EMBASE (OVID) = 59
- LILACS (Bireme) = 4
- Specialised register (BUMT) = 15

References excluded = 152
Reasons:
- Duplicate references: 49
- Other issues or not adult or adolescent participants: 66
- Studies of lateral or medial clavicle fractures (not middle third clavicle fractures): 4
- Studies of surgical interventions or surgical versus conservative interventions: 27
- Not randomised or quasi-randomised controlled trials: 6

Potentially relevant references retrieved for detailed evaluation = 7

Main reports of randomised clinical trials meeting review inclusion criteria = 3
- Andersen 1987
- Hoofwijk 1988
- Lubbert 2006

References excluded = 1
Reason:
Not randomised or quasi-randomised study

Three references were additional reports of two trials
- Jensen 1985 is another report of Andersen 1987
- There were two conference abstracts for Hoofwijk 1988: one of which (Hoofwijk 1996) presented interim results.
**Included studies**

Three randomised controlled trials are included in this review: Andersen 1987 (reported in English and Danish); Hoofwijk 1988 (reported in German); and Lubbert 2008 (reported in English). All trials were located in MEDLINE (PubMed). We also located two reports of both Andersen 1987 and Hoofwijk 1988 in The Cochrane Library (Wiley), and the two reports of Andersen 1987 in EMBASE (OVID).

**Design of the studies**

Andersen 1987 (Figure 2) and Hoofwijk 1988 (Figure 3) were single-centre randomised controlled trials, conducted respectively in hospitals in Denmark and in the Netherlands. Both trials used a two group design comparing the same interventions (bandage and sling).
Figure 2. Participant flow (Andersen 1987)

Andersen 1987

Randomised
79 participants (fractures)

Allocated Figure of eight bandage
n = 45 participants

Follow-up (10 to 16 months)
Median = 12 months
11 lost to follow-up:
1 refused bandage
1 bandage removal due to pain*
1 had a DVT*
2 had fracture displacement*
8 defaulted follow-up examination

*Had uneventful recovery after change of the treatment to a simple sling.

Allocated Arm sling
n = 34 participants

Follow-up (10 to 17 months)
Median = 13 months
7 lost to follow-up:
1 bed rested (5 weeks)
1 Velpeau bandage used because of pain (1/wk)
1 hemiplegic attack
4 defaulted follow-up examination

Analysed
n = 34
age 14-81 years (median 19)
Classification
26 had two fragments; 7 had one intermediary fragment; 2 had two or more intermediary fragments;
3 were undisplaced; 15 had minor displacement; 16 had major displacement.

Analysed
n = 27
age 14-66 years (median 19)
Classification
20 had two fragments; 6 had one intermediary fragment; 2 had two or more intermediary fragments;
3 were undisplaced; 12 had minor displacement; 12 had major displacement.
Figure 3. Participant flow (Hoofwijk 1988)

Hoofwijk 1988

Assessed for eligibility 488 clavicle fractures
Dec 1983 – May 1987

Excluded (n = 331)
159 people were too young
77 were taken as inpatients
52 with metaphyseal fracture
4 people refused to participate
40 for other reasons (concomitant injuries of the same extremity,
pathological fractures, subsequent treatment took place elsewhere).

Randomised
155 participants
157 fractures

Allocated
Figure of eight bandage
n = 76 (fractures)
age (mean (SD)): 24.4 (12.5) years
gender (female/male): 22 / 56

Follow-up (6 to 36 months)
4 lost to follow-up

Analysed
n = 74 (fractures)

Allocated
Arm sling
n = 79 (fractures)
age (mean (SD)): 25.4 (14.5) years
gender (female/male): 22 / 57

Follow-up (6 to 36 months)
1 lost to follow-up

Analysed
n = 78 (fractures)

Lubbert 2008 (Figure 4) was a multi-centre double blind
randomised controlled trial, conducted in six hospitals in The Nether-
lands. This trial used a two group design comparing low intensity
pulsed ultrasound (LIPUS) and placebo.
Figure 4. Participant flow (Lubbert 2006)

Assessed for eligibility 1050 clavicle fractures
Mar 2001 – Dec 2003

Excluded (n = 930)
910 didn’t meet inclusion criteria / refused to participate / or no time to arrange study inclusion
20 no transducer available

Randomised
120 participants (fractures)

Allocated
LIFUS
n = 61 participants

Allocation

Follow-up (12 to 43 months)
Mean = 29.6 months
9 lost to follow-up:
All because diary not fully filled in

Allocated
Placebo
n = 59 participants

Follow-up (12 to 43 months)
Mean = 30.1 months
10 lost to follow-up:
7 because diary not fully filled in
3 transducer failure (too much pain)

Analysed
n = 52
age (mean (SD)): 37.7 (12.9) years
gender (female/male): 16 / 46
side (left / right): 32 / 20
AO classification
A1: 6 / A2: 10 / A3: 5
B1: 4 / B2: 2 / B3: 10
C1: 1 / C2: 3 / C3: 0

Analysed
n = 49
age (mean (SD)): 36.9 (12.3) years
gender (female/male): 10 / 39
side (left / right): 22 / 27
AO classification
A1: 4 / A2: 16 / A3: 3
C1: 0 / C2: 1 / C3: 1
Participants
The three included trials performed a total of 354 participants.

Age and gender
 Andersen 1987 did not report the proportion of males and females. Hoofwijk 1988 reported that 72% of trial participants were male; and in Lubbert 2008, 84% were male.
 Participants of Andersen 1987 were aged between 14 and 81 years, the median age of both groups was 19 years. All participants of Hoofwijk 1988 were older than 14 years, with a mean age for the trial population of 24.9 years. Participants of Lubbert 2008 were aged between 19 and 74 years; the mean age for the trial population was 37.3 years.

Types of fractures
All participants of all trials had acute middle third clavicle fractures and were treated just after their diagnosis. Two trials (Andersen 1987; Hoofwijk 1988) did not use a specific classification for fractures. Andersen 1987 divided the fractures in types (two-fragments, one intermediacy fragment and two or more intermediary fragments) and dislocations (undisplaced, minor displacement, major displacement); Hoofwijk 1988 divided the fractures according to displacement and shortening. Lubbert 2008 classified fractures using the AO system (Muller 1991).

Interventions
The included studies were grouped according to the interventions studied.

Comparison 1: Immobilisation bandage (figure-of-eight and backpack bandage) versus sling
Two trials (Andersen 1987; Hoofwijk 1988) compared the use of figure-of-eight bandage versus sling immobilisation in 234 participants.

Comparison 2: Therapeutic ultrasound versus placebo
Lubbert 2008 compared low-intensity pulsed ultrasound (LIPUS) versus placebo in 120 participants treated conservatively using a collar and cuff for passive support.

Outcome measures

Primary outcomes
- Pain was evaluated in all studies. Andersen 1987 used an unvalidated score; Hoofwijk 1988 and Lubbert 2008 measured pain by applying a visual analogue scale and recording analgesics consumption.
- Shoulder function was evaluated in all studies. Andersen 1987 and Hoofwijk 1988 used unvalidated scores. Hoofwijk 1988 and Lubbert 2008 measured clinical fracture consolidation; this is treated as a proxy for recovery of function in this review.
- Health-related quality of life was evaluated indirectly without a specific validated questionnaire. Andersen 1987 used an unvalidated score system and measured discomfort from treatment, severity of discomfort, duration of discomfort.
- Time to return to previous activities was evaluated by Hoofwijk 1988 and Lubbert 2008; both reported continuous data on resumption of household/ work/ school and sports.

Secondary outcomes
- Functional impairment and clinical outcomes were appraised by Andersen 1987 with an unvalidated score system. Hoofwijk 1988 evaluated cosmetic appearance using subjective data.
- Radiographic outcomes were reported by Andersen 1987 and Hoofwijk 1988.
- Lubbert 2008 evaluated adverse effects of interventions and subsequent fracture-related surgery.

Excluded studies
Two studies (Talbot 2008; Thompson 2005) were excluded for reasons stated in the Characteristics of excluded studies.

Ongoing studies
One ongoing study was identified (Robert 2008). Further details of this are given in the Characteristics of ongoing studies.

Risk of bias in included studies

Table 2 shows the results for each of the three trials for the 11 items of the Bone, Joint and Muscle Trauma Group’s former quality assessment tool (Table 1). The first seven items of this tool relate to bias (internal validity), and the remaining four items relate to external validity and outcome measurement. Details of the method of randomisation, assessor blinding, intention-to-treat analysis / loss to follow-up, and length of follow-up are presented in the Characteristics of included studies. A summary of the results and impressions of the likelihood of bias are presented below.
Table 2. Methodological quality assessment results for individual trials

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Allocation concealment</td>
<td>?</td>
<td>?</td>
<td>Y</td>
</tr>
<tr>
<td>2. Intention-to-treat analysis</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3. Assessor blinding</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>4. Baseline characteristics comparability</td>
<td>?</td>
<td>?</td>
<td>Y</td>
</tr>
<tr>
<td>5. Participant blinding</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>6. Treatment provider blinding</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>7. Care programme comparability</td>
<td>?</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>8. Inclusion and exclusion criteria</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>9. Well defined outcome measures</td>
<td>?</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>10. Clinically useful diagnostic texts</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>11. Adequate duration of follow-up</td>
<td>N</td>
<td>?</td>
<td>Y</td>
</tr>
</tbody>
</table>

Y = yes; ? = partial/unknown/unclear; N = no

Allocation (item 1)

All three studies were randomised. Neither Andersen 1987, which used a random numbers table, nor Hoofwijk 1988, which used pre-numbered envelopes, gave sufficient details to ascertain that allocation was concealed. Allocation was concealed in Lubbert 2008, which used a double-blind randomised method involving central randomisation by a third party (the manufacturer) and supply of identical packs containing either an active or placebo transducer.

Intention-to-treat analysis and handling of withdrawals/losses to follow-up (item 2)

All three trials described their withdrawals. Participant flow diagrams of the data available for the three trials are presented in Figure 2, Figure 3 and Figure 4 respectively. None of the trials, however, presented outcome data for participants who had withdrawn from the trial or were lost to follow-up. Lubbert 2008 confirmed that data from participants who were lost to follow-up were not available. Not rated in the quality assessment tool was loss to follow-up. This was 23% in Andersen 1987; 3% in Hoofwijk 1988; and 16% in Lubbert 2008. There were comparable losses in the treatment groups of each of the three trials.

Blinding (items 3, 5 and 6)

Only Lubbert 2008 blindered the outcome assessors (item 3), as well as participants (item 5) and care providers (item 6). Blinding of the assessment of most outcomes is impractical for the other two trials (Andersen 1987; Hoofwijk 1988) due to the type of intervention. Similarly, the type of intervention (bandage and sling) precluded...