Appendix 1. Search strategy

EMBASE
#1. SPINE/exp
#2. (spine or spinal or vertebra*):ti
#3. Fracture/exp
#4. fractur*:tw.
#5. #1 or #2
#6. #3 or #4
#7. #5 and #6
#8. 'Spine Fracture'/exp
#9. #7 or #8
#10. 'Acupuncture Therapy'/exp OR Acupuncture/exp OR 'Acupuncture Points'/exp OR 'acupuncture needle'/exp
#11. electroacupuncture/exp OR meridians/exp
#12. acupunctur*:ti,ab,kw
#13. needl*:ti,ab,kw
#14. acupoint* or meridian*:ti,ab,kw
#15. electroacupuncture OR electro-acupuncture:ti,ab,kw
#16. pharmaacoacupuncture* OR 'pharmac-o-acupuncture*' OR 'acupoint injection':ti,ab,kw
#17. 'auricular acupuncture' or 'ear acupuncture' or 'auricular needl*' or 'ear needl*':ti,ab,kw
#18. 'fire acupuncture' or 'fire needl*' or 'warm acupuncture' or 'warm needl*:ti,ab,kw
#19. 'scalp acupunctur*:ti,ab,kw
#20. or/#10-#19
#21. #9 and #20
Appendix 2. Summary of findings
Acupuncture plus routine care versus routine care for vertebral compression fracture

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
<th>Relative effect (95% CI)</th>
<th>Anticipated absolute effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain (VAS)</td>
<td>252 (5 studies)</td>
<td>Very lowab,c</td>
<td>The mean VAS was 0</td>
<td>MD 1.05 lower (1.45 lower to 0.65 lower)</td>
</tr>
<tr>
<td>Pain (Number of patients experiencing pain relief)</td>
<td>52 (1 study)</td>
<td>Very lowd</td>
<td>RR 1.32 (1.03 to 1.68)</td>
<td>731 per 1,000</td>
</tr>
<tr>
<td>Function (ODI)</td>
<td>107 (2 studies)</td>
<td>Very lowab,d</td>
<td>Not pooled</td>
<td>234 more per 1,000 (22 more to 497 more)</td>
</tr>
<tr>
<td>Function (Lumbar dysfunction)</td>
<td>56 (1 study)</td>
<td>Very lowe</td>
<td>The mean lumbar dysfunction was 0</td>
<td>MD 1.22 lower (1.42 lower to 1.02 lower)</td>
</tr>
<tr>
<td>Function (Barthel)</td>
<td>45 (1 study)</td>
<td>Very lowf,g</td>
<td>Not pooled (High heterogeneity)</td>
<td></td>
</tr>
<tr>
<td>Quality of life</td>
<td>0 (no study)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient satisfaction</td>
<td>0 (no study)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident vertebral fracture</td>
<td>0 (no study)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not measured</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence.
High certainty : We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty : We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty : Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very low certainty : We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of the effect.

a. Included study had high risk of bias in domains of blinding of participants and personnel and unclear risk of bias in domains of random sequence generation, allocation concealment, blinding of outcome assessment and incomplete outcome data. Therefore, quality of evidence was downgraded by two levels.
b. There is very serious heterogeneity unexplained. Therefore, the quality of evidence was downgrade.
c. Sample size is small and 95% CI passes minimal important difference. So it can be judged that there are uncertainty and imprecision.
d. Sample size is small and 95% CI is so wide. So it can be judged that the precision of the effect size is low, and benefit of effect is uncertain.
e. Included study had high risk of bias in domains of blinding of participants and personnel and unclear risk of bias in domains of allocation concealment, blinding of outcome assessment and incomplete outcome data. Therefore, quality of evidence was downgraded by two levels.
f. Included study had high risk of bias in domains of blinding of participants and personnel and unclear risk of bias in domains of allocation concealment and blinding of outcome assessment. Therefore, quality of evidence was downgraded by two levels.
g. Sample size is small and 95% CI passes zero. So it can be judged that there are uncertainty and imprecision.
Appendix 2. Continued

Acupuncture versus sham acupuncture for vertebral compression fracture

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
<th>Relative effect (95% CI)</th>
<th>Anticipated absolute effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain (VAS)</td>
<td>53 (1 study)</td>
<td>⊕⊕⊕⊕ low&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>-</td>
<td>Risk with sham acupuncture: No exact value was provided (There is no significant difference between verum acupuncture and sham acupuncture.)</td>
</tr>
<tr>
<td>Quality of life (QUALEFFO-41)</td>
<td>53 (1 study)</td>
<td>⊕⊕⊕⊕ low&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>-</td>
<td>Risk with sham acupuncture: No exact value was provided (There is no significant difference between verum acupuncture and sham acupuncture.)</td>
</tr>
<tr>
<td>Function</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Patient satisfaction</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Incident vertebral fracture</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI : Confidence interval, RR : Risk ratio, MD : Mean difference, VAS : Visual analog scale.  
GRADE Working Group grades of evidence.  
High certainty : We are very confident that the true effect lies close to that of the estimate of the effect.  
Moderate certainty : We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.  
Low certainty : Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.  
Very low certainty : We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.  

a. This study has high risk of bias in domains of blinding of participants and personnel and unclear risk of bias in domains of selective reporting and other bias. Therefore, quality of evidence was downgraded.  
b. Sample size is small and 95% CI contain zero. So it can be judged that there are uncertainty and imprecision.
Appendix 2. Continued

Acupuncture versus analgesics for vertebral compression fracture

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
<th>Relative effect (95% CI)</th>
<th>Anticipated absolute effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain (VAS)</td>
<td>135 (1 study)</td>
<td>⊕⊕⊕⊕ very low&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>The mean VAS was 0 MD 0.26 higher (0.44 lower to 0.96 higher)</td>
<td>Risk with analgesics Risk difference with acupuncture</td>
</tr>
<tr>
<td>Quality of life</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Function</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Patient satisfaction</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Incident vertebral fracture</td>
<td>0 (no study)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Cl : Confidence interval, RR : Risk ratio, MD : Mean difference, VAS : Visual analog scale.

GRADE Working Group grades of evidence.

High certainty : We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty : We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty : Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very low certainty : We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

a. Sample size is small and 95% CI contain zero. So it can be judged that there are uncertainty and imprecision.
b. Included study had high risk of bias in domains of blinding of participants and personnel and unclear risk of bias in domains of allocation concealment, blinding of outcome assessment, and selective reporting. Therefore, quality of evidence was downgraded by two levels.