Does stretching increase ankle dorsiflexion range of motion? A systematic review


doi: 10.1136/bjsm.2006.029348

Updated information and services can be found at:
http://bjsm.bmj.com/content/40/10/870.full.html

These include:

References
This article cites 12 articles, 8 of which can be accessed free at:
http://bjsm.bmj.com/content/40/10/870.full.html#ref-list-1

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://journals.bmj.com/cgi/ep
Does stretching increase ankle dorsiflexion range of motion? A systematic review

J A Radford, J Burns, R Buchbinder, K B Landorf, C Cook

Background: Many lower limb disorders are related to calf muscle tightness and reduced dorsiflexion of the ankle. To treat such disorders, stretches of the calf muscles are commonly prescribed to increase available dorsiflexion of the ankle joint.

Hypothesis: To determine the effect of static calf muscle stretching on ankle joint dorsiflexion range of motion.

Study design: A systematic review with meta-analyses.

Methods: A systematic review of randomised trials examining static calf muscle stretches compared with no stretching. Trials were identified by searching Cinahl, Embase, Medline, SportDiscus, and Central and by recursive checking of bibliographies. Data were extracted from trial publications, and meta-analyses performed that calculated a weighted mean difference (WMD) for the continuous outcome of ankle dorsiflexion. Sensitivity analyses excluded poorer quality trials. Statistical heterogeneity was assessed using the quantity $I^2$.

Results: Five trials met inclusion criteria and reported sufficient data on ankle dorsiflexion to be included in the meta-analyses. The meta-analyses showed that calf muscle stretching increases ankle dorsiflexion after stretching for <15 minutes (WMD 0.07˚; 95% confidence interval 0.86 to 3.27), >15–30 minutes (WMD 3.03˚; 95% confidence interval 0.31 to 5.75), and >30 minutes (WMD 2.49˚; 95% confidence interval 0.16 to 4.82). There was a very low to moderate statistical heterogeneity between trials. The meta-analysis results for <15 minutes and >15–30 minutes of stretching were considered robust when compared with sensitivity analyses that excluded lower quality trials.

Conclusions: Calf muscle stretching provides a small and statistically significant increase in ankle dorsiflexion. However, it is unclear whether the change is clinically important.
The search strategy used for all databases was:

1. stretch$\text{.tw}$
2. (ankle or gastrocnemius or soleus or soleal or (triceps and surae) or Achilles or (calf adj muscle$\text{s}$)).tw
3. (motion or range or dorsiflexion or plantarflexion or flexibility or extensibility or stiffness).tw
4. and/1–3
5. limit 4 to human

**Assessment of study quality**

Two reviewers (JR and JB) independently used the PEDro scale to determine the quality of the trials. The PEDro scale is an 11 item scale designed for rating methodological quality of randomised controlled trials. Each satisfied item (except item 1) contributes one point to the total PEDro score (range 0–10 points). The scale items are:

1. Eligibility criteria were specified.
2. Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated in an order in which treatments were received).
3. Allocation was concealed.
4. The groups were similar at baseline regarding the most important prognostic indicators.
5. There was blinding of all assessors who measured at least one key outcome.
6. There was blinding of all therapists who administered the treatment.
7. There was blinding of all assessors who measured at least one key outcome.
8. Measurements of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups.
9. All subjects for whom outcome measurements were available received the treatment or control condition as allocated, or where this was not the case, data for at least one key outcome were analysed by “intention to treat”.
10. The results of between-group statistical comparisons are reported for at least one key outcome.
11. The study provides both point measurements and measurements of variability for at least one key outcome.

The inter-rater reliability of the total PEDro score (obtained by summing “yes” responses to items 2–11) was evaluated using type 2,1 intraclass correlation coefficients (ICCs).

**Data extraction**

Two reviewers (JR and JB) independently extracted data (study population, intervention, outcomes) from the trials using standardised extraction forms. We intended to contact authors for further information if required, but this was not necessary. To assess effectiveness, we extracted raw data for outcomes of interest, means and standard deviations, from published reports.

**Data analysis**

The result of each randomised trial was plotted as a point—that is, mean and 95% confidence interval. To obtain a pooled estimate of the impact of calf muscle stretching on ankle dorsiflexion, we planned meta-analyses when the data were available. If possible, weighted mean differences (WMDs) were calculated for the continuous outcome of ankle dorsiflexion using Review Manager 4.2.7 (2004). Results were considered significant if $p < 0.05$. Three meta-analyses were planned on the basis of the duration of stretching interventions provided in each trial. For reasons relating to generalisability, we thought it appropriate to conduct three meta-analyses for trials providing data for similar stretching time periods: $\leq 15$ minutes; $>15–30$ minutes; $>30$ minutes. Bohannon et al provided data for five minutes of stretching, Knight et al for 8 and 24 minutes, Peres et al for 10, 20, and 60 minutes, Pratt and Bohannon for nine minutes, and Youdas et al for 15, 30, and 60 minutes of stretching (see fig 2).

Trials were assessed for clinical heterogeneity with respect to their inclusion and exclusion criteria—for example, age, healthy participants, duration of stretching—and meta-analyses performed when they were found to be clinically homogeneous and the data readily available. The statistical heterogeneity (whether there are genuine differences underlying the results of the trials in the review) and homogeneity (whether the variation in findings is compatible with chance alone) of the results of the trials were measured using the quantity $I^2$. The $I^2$ value is calculated as $I^2 = 100\%(Q – df)/Q$, where $Q$ is Cochran’s heterogeneity statistic and $df$ the degrees of freedom (where $n$ is the number of trials and therefore degrees of freedom equals number of studies minus one). The Cochran’s $Q$ is computed by summing the squared deviations of each trial’s estimate from the overall meta-analytical estimate, and a $p$ value obtained by comparing the statistic with a $\chi^2$ distribution with $k–1$ degrees of freedom (where $k$ is the number of trials). The $p$ value was formerly used to determine statistical heterogeneity ($p<0.10$). However, Higgins et al replaced it with the quantity $I^2$ as the $p$ value was known to be poor at detecting true heterogeneity among studies as significant. Trials in the meta-analyses were considered to have low statistical heterogeneity if $I^2<25\%$; in such instances a fixed effects model was used to estimate the pooled effect. A random effects model was used for all trials with $I^2>25\%$. A fixed effect meta-analysis assumes that the true effect of treatment (in both magnitude and direction) is the same value in every trial—that is, fixed across studies. In contrast, a random effects meta-analysis model assumes that the effects being
estimated in the different studies are not identical, but follow a similar distribution.16

There are different approaches to conducting a systematic review, and a sensitivity analysis is required to test how robust the results of the review are relative to key decisions and assumptions that are made in the process of conducting the review.25 A sensitivity analysis changes the method of the review for a secondary analysis to determine whether key decisions or assumptions may conceivably have affected the results for a particular review. If the effect and confidence intervals in the sensitivity analysis lead to the same conclusion as the primary meta-analysis value, the results are deemed robust. We considered it highly important to blind the outcome assessor in the trials included in the review, as some methods of measuring ankle range of motion are open to assessor bias. Therefore sensitivity analyses of meta-analyses were conducted that only included trials that blinded their outcome assessors.

Table 1 Description of studies included in systematic review of calf muscle stretching intervention

<table>
<thead>
<tr>
<th>Trial</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bohannon et al 199410</td>
<td>36 women volunteers with mean (SD) age 22.6 (4.2) years. Inclusion: no history of orthopaedic or neurological problems affecting lower limb.</td>
<td>Weight-bearing static stretch for 5 min</td>
<td>Non-weight-bearing active and passive ankle dorsiflexion with knee extended. Measured by taking digital photographs, marking lines on the photographs, and then using a protractor for calculation of angles. Active ankle dorsiflexion ICC = 0.93. Duration between stretching and measurement was 1 min.</td>
</tr>
<tr>
<td>Knight et al 200111</td>
<td>97 volunteers (59 women and 38 men) with mean (SD) age 27.6 (7.7) for women and 26.8 (6.9) for men. Inclusion: active ankle dorsiflexion less than 20°. Exclusion: pregnancy; impaired sensation; bleeding disorders; previous neuromuscular disorders; hip, knee or ankle pathologies in past 2 years, lower extremity malignancies.</td>
<td>Weight-bearing static stretch for 20 s with 10 s rest between stretches, repeated 4 times. Performed 3 times a week for 6 weeks.</td>
<td>Non-weight-bearing active and passive ankle dorsiflexion with knee flexed. Measured using a goniometer. Active ankle dorsiflexion ICC = 0.91. Duration between stretching and measurement was minimal as stretching supervised by researchers and measurements taken after stretching.</td>
</tr>
<tr>
<td>Peres et al 200212</td>
<td>60 volunteers (23 women and 21 men) with mean (SD) age 22.5 (2.0) years. Exclusion: involved in any flexibility or strength training for the calf; recent ankle injury or history of ankle injury; metal plates or screws in right leg; pregnancy; any allergies to cold.</td>
<td>Non-weight-bearing static stretch assisted by weight (one third of participant’s body weight) via a pulley for 10 min daily for 14 days over 3 weeks.</td>
<td>Non-weight-bearing passive ankle dorsiflexion using weight and pulley to apply force with knee extended. Measured using a digital inclinometer. Ankle dorsiflexion ICC = 0.99. Duration between stretching and measurement was &lt;30s.</td>
</tr>
<tr>
<td>Pratt &amp; Bohannon 200313</td>
<td>24 volunteers (12 women and 12 men) with mean (SD) age 24.7 (4.5) years. Inclusion: free of injury; not currently stretching.</td>
<td>Weight-bearing stretch by lowering heels from a platform for 3 min for 3 days.</td>
<td>Weight-bearing active ankle dorsiflexion with force from participant lowering their heels down from a platform while keeping metatarsals on platform with knees extended. Measured by taking digital photographs of lines marked on the foot and using a protractor for calculation of angles. We referenced another trial for reliability, with ICC = 0.92 reported. Duration between stretching and measurement was immediate as photographs taken during stretching.</td>
</tr>
<tr>
<td>Youdas et al 200314</td>
<td>101 volunteers (63 women and 38 men) with mean (SD) age 40.0 (10.9) years. Exclusion: block at the talocrural joint that would limit ankle motion; limitation of subtalar joint mobility; previous history of trauma to the calf that required surgery; evidence of lower extremity dysfunction assessed by visual observation of gait.</td>
<td>(1) Weight-bearing static stretch for 30 s daily, 5 days a week for 6 weeks. (2) Weight-bearing static stretch for 1 min daily, 5 days a week for 6 weeks. (3) Weight-bearing static stretch for 2 min daily, 5 days a week for 6 weeks.</td>
<td>Non-weight-bearing active ankle dorsiflexion with knee extended. Measured using a goniometer. Ankle dorsiflexion ICC = 0.95. Duration between stretching and measurement was 60–72 hours.</td>
</tr>
</tbody>
</table>

 ICC, Intraclass correlation coefficient.

RESULTS

Search results

Eighteen papers were identified through electronic searching (fig 1). Five trials met the inclusion criteria for the review with 161 participants (table 1)10–14 Three trials were identified for inclusion in the sensitivity analyses with 118 participants.10 11 14 No trials included participants with lower limb injuries. Ankle dorsiflexion was measured both actively10 11 13 14 and passively,10–12 weight-bearing13 and non-weight-bearing,10–12 14 and with the knee extended10 12 14 and flexed.11 We found excellent reliability for measurement in four trials,10–12 14 and one trial15 referenced another study for excellent reliability of their measurement technique. Duration between stretching and measurement was immediate in all trials, except for that of Youdas et al14 where participants were measured 60–72 hours after stretching. Four of the trials ensured 100% compliance by supervising the stretches.10–13 Youdas et al14 used log books instead of

Table 2 Trial quality assessed by the PEDro scale

<table>
<thead>
<tr>
<th>Trial</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bohannon et al 199410</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>6/10</td>
</tr>
<tr>
<td>Knight et al 200111</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>6/10</td>
</tr>
<tr>
<td>Peres et al 200212</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>4/10</td>
</tr>
<tr>
<td>Pratt &amp; Bohannon 200313</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>6/10</td>
</tr>
</tbody>
</table>

Note: Column numbers correspond to the PEDro scale criteria.
assess compliance, and all participants showed 95% or greater compliance to the stretching programme. Table 1 presents the stretching techniques used by included trials. Only one trial included multiple groups with varying intensity of the stretches. Two of the five trials also compared the simple static stretching with other stretching programmes: stretching with active heel raises; stretching with superficial moist heat to the calf muscle; stretching with continuous ultrasound; stretching with pulsed diathermy; stretching with pulsed diathermy and ice. However, in accordance with our pre-specified aims, these data were not included in the review. All results presented in this review are the effect of stretching alone compared with no stretching.

The range of PEDro quality scores assessing methodological quality of the included trials was 3–6 (median 6) out of 10 (table 2). The ICC 2,1 for the reviewer's reliability was 0.84 (95% confidence interval −0.01 to 0.98).

More trials provided active measurements of ankle dorsiflexion than passive measurements, so the active measurements from trials that provided both passive and active data were pooled. Only one passive ankle dorsiflexion measurement was pooled.

The meta-analyses (fig 2) found that static stretching increases ankle dorsiflexion compared with no stretching after 15 minutes (WMD 2.07˚; 95% confidence interval 0.86 to 3.27; p = 0.0008), 15–30 minutes (WMD 3.03˚; 95% confidence interval 0.31 to 5.75; p = 0.03), and

<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>Stretching mean (SD)</th>
<th>Control mean (SD)</th>
<th>WMD (fixed) 95% CI</th>
<th>WMD (fixed) 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤15 minutes of stretching</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bohannon - 5 minutes</td>
<td>1.40 (3.30)</td>
<td>−0.20 (2.00)</td>
<td>1.60 [-0.18, 3.38]</td>
<td></td>
</tr>
<tr>
<td>Knight - 8 minutes</td>
<td>11.79 (4.21)</td>
<td>9.56 (3.54)</td>
<td>2.23 [-0.27, 4.73]</td>
<td></td>
</tr>
<tr>
<td>Peres - 10 minutes</td>
<td>101.60 (4.00)</td>
<td>96.60 (4.60)</td>
<td>5.00 [-0.15, 10.15]</td>
<td></td>
</tr>
<tr>
<td>Pratt - 9 minutes</td>
<td>24.30 (9.90)</td>
<td>21.40 (8.40)</td>
<td>2.90 [-0.80, 6.60]</td>
<td></td>
</tr>
<tr>
<td>Vousdais - 15 minutes</td>
<td>9.50 (4.60)</td>
<td>7.50 (4.30)</td>
<td>2.00 [-0.58, 4.58]</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>82</td>
<td>80</td>
<td>2.07 [0.86, 3.27]</td>
<td></td>
</tr>
<tr>
<td>Test for heterogeneity: Chi²= 1.59, df = 4 (p = 0.81), P = 0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 3.36 (p = 0.0008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| >15–30 minutes of stretching |                      |                    |                   |                   |
| Knight - 24 minutes        | 13.89 (4.63)         | 10.83 (4.12)       | 3.06 [0.24, 5.88] |
| Peres - 20 minutes         | 103.30 (4.00)        | 96.50 (4.60)       | 6.80 [1.84, 11.76]|
| Vousdais - 30 minutes      | 8.60 (4.70)          | 7.50 (4.30)        | 1.10 [-1.51, 3.71]|
| Subtotal (95% CI)          | 52                   | 50                 | 3.03 [0.31, 5.75]  |
| Test for heterogeneity: Chi²= 4.12, df = 2 (p = 0.13), P = 51.4% |
| Test for overall effect: Z = 2.18 (p = 0.03) |

| >30 minutes of stretching |                      |                    |                   |                   |
| Peres - 60 minutes         | 101.90 (4.30)        | 98.30 (5.50)       | 3.60 [-0.98, 8.18]|
| Vousdais - 60 minutes      | 9.60 (4.90)          | 7.50 (4.30)        | 2.10 [-0.61, 4.81]|
| Subtotal (95% CI)          | 32                   | 32                 | 2.49 [0.16, 4.82]  |
| Test for heterogeneity: Chi²= 0.31, df = 1 (p = 0.58), P = 0% |
| Test for overall effect: Z = 2.09 (p = 0.04) |

Figure 2  Meta-analyses of the effect of stretching on ankle dorsiflexion.

Figure 3  Sensitivity analyses excluding trials that did not blind outcome assessors.
>30 minutes of stretching (WMD 2.49°; 95% confidence interval 0.16 to 4.82; p = 0.04). The sensitivity analyses of trials that blinded the assessor of ankle dorsiflexion to group allocation (fig 3) also found a similar increase in ankle dorsiflexion with stretching compared with no stretching after ≤15 minutes (WMD 1.86°; 95% confidence interval 0.59 to 3.12; p = 0.004) and >15–30 minutes (WMD 2.00°; 95% confidence interval 0.09 to 3.92; p = 0.04). This indicates that the results for these meta-analyses are robust.

The sensitivity analysis of stretching for >30 minutes found no significant difference (WMD 2.10°; 95% confidence interval −0.61 to 4.81; p = 0.13), which does not support the statistically significant result of the meta-analysis. The statistical heterogeneity between trials for the meta-analyses was very low (I² = 0%), except for the meta-analysis of >30 minutes of stretching, which showed moderate statistical heterogeneity (I² values of 25%, 50%, and 75% represent low, moderate, and high statistical heterogeneity respectively).

**DISCUSSION**

This review found five randomised trials (n = 161) that examined whether static stretching increased ankle dorsiflexion flexibility of the calf musculature. Although the search strategy allowed the inclusion of trials examining both normal subjects and people with disorders—for example, plantar fasciitis—the results of the review are only generalisable to normal subjects, as no trials were found that included people with disorders. Pooled data from included trials showed an increase of 2.1–3.0° after 5–60 minutes of stretching when compared with no stretching. The statistical heterogeneity was generally very low, indicating that there are no underlying differences in the trials—that is, all trials examined the same effect. The results of the sensitivity analyses were also generally similar to the meta-analyses, providing confidence that the results of the meta-analyses are robust. The only result that was not supported by the sensitivity analyses was the one for >30 minutes of stretching. As only one trial was included in the sensitivity analysis of this intervention, it may be that the sensitivity analysis sample was underpowered (n = 45) to detect the effect that the primary analysis detected with greater sample power (n = 64).

There was a general trend that the longer the stretch, the greater the increase in ankle dorsiflexion: ≤15 minutes of stretching resulted in a 2.07° increase (95% confidence interval 0.86 to 3.27), whereas >15–30 minutes of stretching resulted in a 3.03° increase (95% confidence interval 0.31 to 5.75). Stretching for >30 minutes resulted in a 2.49° increase (95% confidence interval 0.16 to 4.82), which was still superior to ≤15 minutes of stretching, although provided less of a gain than >15–30 minutes of stretching. Overall it should be noted that the differences between stretching durations are small (clinically non-significant), and the results of this systematic review indicate that stretching for a short duration produces a similar result to that of a longer duration. This may be due to the lack of a dose-response relation or differences in the interventions of the individual trials. Further randomised trials are needed to resolve this.

The PEDro scores of the trials included in the review are relatively high (median 6) considering that blinding of participants and therapists is impossible. The PEDro scale was therefore a useful tool for assessing the quality of the trials except for the two items requiring blinding of participants and therapists, which are inherently difficult for such interventions. Future trial quality could be improved by ensuring that treatment allocation concealment and intention to treat analyses are performed and reported in accordance with the CONSORT statement.

Although our review shows that static calf muscle stretching provides a small and statistically significant increase in ankle dorsiflexion, we do not know whether the result is of clinical importance from the patient’s perspective or whether it may prevent further injury. Symptomatic pain improvement seen in a trial of stretching for plantar fasciitis may be explained by an increase in dorsiflexion²⁰; and in a recently completed placebo controlled trial of adhesive capsulitis, manual techniques and a directed exercise programme after arthrographic glenohumeral joint distension resulted in improved shoulder range of motion accompanied by greater participant perceived success.²¹ These findings suggest that improvement in range of motion may be of clinical importance. Future trials evaluating the effectiveness of static stretching should include patient centred outcome measures such as pain, function, and perceived success so that this can be explored further.

**CONCLUSION**

Calf muscle stretches provide a small but statistically significant increase in ankle dorsiflexion, particularly after 5–30 minutes of stretching. However, it is unclear whether the change that occurs with stretching is clinically important. Therefore calf muscle stretching is recommended where a small increase in ankle range of motion is thought to be beneficial.

**Authors’ affiliations**

J A Radford, C Cook, School of Biomedical and Health Sciences, University of Western Sydney, Campbelltown, Australia

J Burns, Institute for Neuromuscular Research, Discipline of Paediatrics and Child Health Faculty of Medicine, The University of Sydney, Sydney, Australia

R Buchbinder, Monash Department of Clinical Epidemiology at Cabrini Hospital and Monash University, Malvern, Australia

K B Landorf, Department of Podiatry, School of Human Biosciences, La Trobe University, Bundoora, Australia

Competing interests: none declared

**REFERENCES**


This article will be of interest to clinicians, sports physicians, and those interested in rehabilitation and injury prevention. The meta-analysis showed only relatively small increases in ankle joint dorsiflexion during the different time periods. From a clinical perspective, only small significant changes suggest that static stretching for only 5–15 minutes is equivalent to stretching for 60 minutes. However, there is no indication from this review about repetitive or bouts of stretching. As the authors stated, the interpretation of an increase of 2˚ needs further exploration. From a research point of view, a significant increase in ankle joint dorsiflexion from the systematic review does suggest the use of stretching in overuse lower limb conditions.

K Rome
University of Teesside, UK; k.rome@tees.ac.uk

Website of the month

www.physics.usyd.edu.au/~cross

This is an excellent website on the physics of ball sports, and is primarily aimed at secondary school physics students.