The efficacy of stretching for prevention of exercise-related injury: a systematic review of the literature

S.M. Weldon, R.H. Hill
Osteopaths, Private Practice, London, UK

SUMMARY. The objective of this study was to conduct a systematic analysis of the literature to assess the efficacy of stretching for prevention of exercise-related injury. Randomized clinical trials (RTCs) and controlled clinical trials (CCTs) investigating stretching as an injury prevention measure were selected. A computer-aided search of the literature was conducted for relevant articles, followed by assessment of the methods of the studies. The main outcome measures were scores for methodological quality based on four main categories (study population, interventions, measurement of effect, and data presentation and analysis) and main conclusions of authors with regard to stretching. One RCT (25%) and three CCTs (100%) concluded that stretching reduced the incidence of exercise-related injury. Three RCTs (75%) concluded that stretching did not reduce the incidence of exercise-related injury. Only two studies scored more than 50 points (maximum score = 100 points) indicating that most of the studies selected were of poor quality. Neither of the two highest scoring RCTs showed positive effects for stretching. Due to the paucity, heterogeneity and poor quality of the available studies no definitive conclusions can be drawn as to the value of stretching for reducing the risk of exercise-related injury.
increase injury risk, citing an in vitro study by Noonan et al. (1993) in which the increase in compliance seen when muscles were warmed to 40°C was associated with a reduction in their energy-absorbing capabilities, although the authors interpreted this as a protective effect.

The majority of muscle strain injuries occur when muscles are active and functioning in an eccentric manner (Ciullo & Zarins 1983; Noonan & Garrett 1992). The ability of a muscle to absorb energy is dependent on both the active (contractile) components and its passive (connective tissue) components (Safran et al. 1989) and is significantly greater in active as opposed to passive muscle activity (Garrett 1987). A critical point is reached when the muscle is unable to prevent excessive sarcomere lengthening and the actin–myosin filaments are stretched beyond overlap (Morgan 1990). The ability of an active muscle to resist lengthening and hence injury, is therefore largely dependent on contractile strength and is substantially reduced when a muscle is fatigued (Mair et al. 1996; Safran et al. 1989). It is therefore the compliance of active muscle that is most relevant when looking at injury risk, which bears little relation to passive compliance, except at the extremes of stretch (Hawkins & Bey 1997). Whether passive stretching can influence the compliance of active muscle has been questioned (Shrier 1999).

Various authors have investigated stretching as an injury prevention measure (Kerner & D’Amico 1983; Howell 1984; Jacobs & Berson 1986; Blair et al. 1987; Macera et al. 1989; Brunet et al. 1990). A number of reviews of the stretching literature exist (Shellock & Prentice 1985; Safran et al. 1989; Wilkinson 1992; Smith 1994), in which authors advocate stretching as an important part of an injury prevention programme, although these conclusions are not based on any clinical evidence. The poor scientific quality of such ‘narrative’ or ‘unsystematic’ literature reviews was highlighted by Mulrow (1987). In as much as primary research takes steps to avoid bias and random error, so too should the review article. In other words, a review should be subject to the same standards of scientific rigour as primary research.

A recent systematic review by Shrier (1999) concluded that pre-exercise stretching did not reduce the incidence of local muscle injury. However, the cross-sectional design of five of the articles that concluded that stretching prior to exercise did not reduce injury meant it was impossible to determine whether subjects stretched before injury or because of injury (Kerner & D’Amico 1983; Howell 1984; Jacobs & Berson 1986; Blair et al. 1987; Brunet et al. 1990). In the cohort study that failed to show any evidence of effect (Macera et al. 1989) the authors failed to control for previous injury and weekly running distance in their analysis, both variables having been shown to be significant predictors of injury risk (Ekstrand & Gillquist 1983; Blair et al. 1987; Brunet 1990).

Although the randomized clinical trial (RCT) is considered to be the gold standard by which we judge the benefits of therapy (Riegelman & Hirsch 1996; Greenhalgh 1997), flaws in their design and conduct can result in overestimation or underestimation of effect, leading to false-positive or false-negative conclusions (van der Heijden et al. 1995). Two RCTs (Ekstrand et al. 1983; van Mechelen et al. 1993) and one controlled clinical trial (CCT) (Bixler & Jones 1992) were included in the review by Shrier (1999), but there was no qualitative analysis of study quality. Therefore a systematic review of published RCTs and CCTs on stretching for injury prevention is presented, with all trials scored for methodological quality and their results interpreted in light of the quality scores thereof.

METHOD

Relevant studies were retrieved by means of a computer-aided literature search using MEDLINE, EMBASE, AMED, SPORT Discus, CINAHL and SIGLE databases, using the following Medical Subject Headings terms or text words: stretch, injury, clinical trial, controlled trial, muscles, sport, exercise. In addition, references given in the studies retrieved were further examined and key journals were handsearched for any relevant studies not recovered by other methods. To be included in the review, studies had to meet the following criteria: (1) randomized clinical trials (RCTs) or controlled clinical trials (CCTs) investigating stretching as an injury prevention measure (additional interventions were allowed); (2) study published from 1970 onwards; (3) abstracts and unpublished studies were excluded.

Randomized clinical trials are considered the gold standard by which the benefits of therapy are judged (Riegelman & Hirsch 1996; Greenhalgh 1997); potential selection bias and confounding making it impossible to draw conclusive determinations of efficacy using non-randomized controlled clinical trials. However, where trials of this quality are lacking, it has been suggested that it would be foolish to ignore the potential for gaining information from other sources (de Bie 1996). Therefore, due to the paucity of relevant RCTs, relevant CCTs were included in this review in agreement with the method guidelines for systematic reviews (van Tulder et al. 1997). As such, all RCTs and CCTs meeting the criteria were retrieved.

The quality of design and conduct of the selected studies were assessed using a modification of the method guidelines for systematic reviews by van Tulder et al. (1997) and the criteria used by Koes in
his study on the efficacy of spinal manipulation and mobilization (Koes et al. 1991). These criteria are based on generally accepted principles of intervention research as used by Ter Riet et al. (1990). These methodological principles are grouped into four categories; study population; interventions; effect; and data analysis. These four categories comprise 14 criteria (Table 1, A–N), which have been further divided to create a 41 item checklist (Appendix A).

Each checklist item is weighted numerically, with a maximum score of 100 possible. The weights given to the criteria were arbitrarily defined but believed to reflect their relative importance for validity and precision. Higher weighting was given to those checklist items (C, I, K, L, M) considered as minimum criteria by the Editorial Board of the Cochrane Back Review Group (van Tulder et al. 1997), although empirical evidence of an association with bias (Schulz et al. 1995; Altman 1999) only exists for two of these criteria (C, I). All studies were subjected to the same methodological assessment, but the scores and results of the controlled clinical trials were tabulated separately in view of the caution with which their results should be interpreted.

All trials were then scored by the authors independently, resulting in a hierarchical list according to methodological quality. Where disagreements in scoring occurred, these were solved by consensus. Trials were deemed positive if the authors concluded that stretching resulted in a reduction in injury risk. A study was deemed to be negative if the authors concluded that stretching failed to reduce injury risk, or increased injury risk.

### Table 1. Criteria list for a methodological assessment of clinical trials of stretching for preventing injury

<table>
<thead>
<tr>
<th>Criterion*</th>
<th>Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study population</strong></td>
<td></td>
</tr>
<tr>
<td>A. Homogeneity</td>
<td>30</td>
</tr>
<tr>
<td>B. Comparability of relevant baseline characteristics</td>
<td>12</td>
</tr>
<tr>
<td>C. Randomization procedure</td>
<td>6</td>
</tr>
<tr>
<td>D. &gt;100 subjects in smallest group; &gt;200 subjects in smallest group</td>
<td>10</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td></td>
</tr>
<tr>
<td>E. Stretching procedure explicitly described</td>
<td>20</td>
</tr>
<tr>
<td>F. Reference procedure explicitly described</td>
<td>5</td>
</tr>
<tr>
<td>G. Co-interventions avoided or comparable</td>
<td>5</td>
</tr>
<tr>
<td>H. Compliance reported and acceptable</td>
<td>5</td>
</tr>
<tr>
<td><strong>Effect</strong></td>
<td></td>
</tr>
<tr>
<td>I. Assessor blinded</td>
<td>35</td>
</tr>
<tr>
<td>J. Relevant outcome measures</td>
<td>10</td>
</tr>
<tr>
<td>K. Drop outs described for each study group separately</td>
<td>12</td>
</tr>
<tr>
<td>L. Attrition rate acceptable</td>
<td>5</td>
</tr>
<tr>
<td><strong>Data Analysis</strong></td>
<td></td>
</tr>
<tr>
<td>M. Intention-to-treat analysis</td>
<td>15</td>
</tr>
<tr>
<td>N. Frequencies presented for each group</td>
<td>10</td>
</tr>
</tbody>
</table>

### RESULTS

Seven studies met the inclusion criteria, four randomized clinical trials and three controlled clinical trials. The RCTs ranged in quality from 12 to 68, of a possible 100 points. The CCTs achieved quality scores ranging from 16 to 30, of a possible 100 points. There were three negative RCTs, and only one positive RCT. All three CCTs were positive. Table 2 presents these studies arranged in hierarchical order based on their methodological scores. Only two studies (Pope et al. 1998, 1999) scored more than 50 points, indicating the poor overall methodological quality of most of the studies. To allow comparison between methodological quality of RCTs and CCTs, the same scoring system was used for both trial designs. However, it should be noted that by definition a CCT is unable to fulfil the criterion relating to random treatment allocation (criterion C).

Common methodological flaws amongst RCTs concerned incomparability of subjects at baseline (criterion B), inadequate treatment allocation (criterion C), inadequate description of reference procedure (criterion F), failure to avoid co-interventions (criterion G) and failure to blind the assessor (criterion I).

Common methodological flaws amongst CCTs concerned incomparability of subjects at baseline (criterion B), inadequate treatment allocation (criterion C), failure to describe dropouts (criterion K), high attrition rate (criterion L) and lack of intention-to-treat analysis (criterion M). Despite an attempt at randomization by two authors (Bixler & Jones 1992, Hartig & Henderson 1999), the pseudo-random procedure used for treatment allocation is not a reliable method of eliminating selection bias and therefore these studies are classified as CCTs.

Despite some incomplete information, the studies were generally methodologically sound in the areas of homogeneity (criterion A), description of stretching procedure (criterion E), avoidance of co-interventions (criterion G) and adequate data presentation (criterion N).

A sensitivity analysis of the checklist and the distribution of weights was undertaken, as utilized by van der Heijden et al. (1995), by recalculation of the weighted and unweighted method scores for the 14 criteria (A–N) and the 41 checklist items. The results of these recalculations, presented in Table 3, revealed no change in the hierarchical order of the studies. These recalculations show the robustness of the scoring system.

Table 4 presents a summary of the four RCTs in hierarchical order. Two studies received method scores that exceeded 50 points (Pope et al. 1998, 1999). The three highest scoring RCTs showed no significant reduction in injury risk in subjects following a stretching programme. Protocols varied widely.
across studies, both in duration of stretch and number of sessions.

Pope et al. (1999) showed no significant effect of pre-exercise stretching on all-injuries risk (Hazard ratio = 0.95, 95% CI 0.77–1.18) or soft tissue injuries (HR = 0.83, 95% CI 0.63–1.09). Multivariate analysis showed fitness as a significant predictive indicator for injury (P < 0.001). In the 1998 study by Pope et al., flexibility was shown to be a significant predictor of injury risk (LR = 4.97; df = 1; P = 0.03) in agreement with other authors (Seto cited in Hartig & Henderson 1999). No significant difference in incidence of injury was found between groups, although a small but clinically significant reduction in injury risk could not be ruled out due to low statistical power. In the home-based stretching programme evaluated by van Mechelen et al. (1993), attrition rate was high (22.3%) and compliance with the prescribed programme was low (46.6%). Results were further complicated by differences in data collection methods between control and treatment group. In the only positive RCT (Ekstrand et al. 1983), the multi-faceted nature of the prophylactic programme and lack of similarity between control and intervention procedures made separate analysis of stretching effect impossible.

Table 5 presents a summary of the three CCTs in hierarchical order. All CCTs showed a significant reduction in rates of injury in the intervention group. Hartig and Henderson (1999), who utilized the greatest total stretching stimulus of all the studies, showed a significant (P = 0.02) reduction in lower extremity injuries in the intervention group (RR = 0.63, 95% CI 0.41–0.99). Relative risk could not be calculated for either of the two lowest scoring CCTs (Bixler & Jones 1992; Cross & Worrell 1999) due to lack of information regarding changes in exposure.

To assess whether the quality of published trials has increased over the past decades, a graph of methodological quality scores (after Koes et al. 1995) against the year of publication was plotted (Fig. 1). A linear trend line shows that there has been a gradual increase in the quality of RCTs and CCTs over the past two decades. However, overall the quality of the studies is still poor, there being only two studies in the past decade that attained quality scores of more than 50 points.

**DISCUSSION**

Due to the heterogeneity of the studies reviewed no meta-analysis was undertaken, however a vote count was performed in accordance with the recommendations in the method guidelines for systematic reviews (van Tulder et al. 1997). A vote count of positive and negative RCTs would suggest that stretching does not reduce the incidence of injury, there being three
negative RCTs (75%) and only one positive RCT (25%). All three CCTs (100%) concluded that stretching did reduce injury risk, but due to the weaker trial design, less emphasis is placed on these results. It is of note that the four positive trials also received the lowest methodological quality scores. However, de Bie (1996) emphasises that vote counting may result in small but clinically important effects being overlooked, particularly if studies with statistically non-significant results are counted as negative.

The two highest scoring RCTs (Pope et al. 1998, 1999) both showed a small but non-significant

Table 4. Randomised clinical trials of the efficacy of stretching for preventing injury

<table>
<thead>
<tr>
<th>Author</th>
<th>Intervention</th>
<th>Control</th>
<th>Results</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pope et al. (1999)</td>
<td>One 20s static stretch for each of 6 lower extremity muscle groups during warm-up. 40 sessions over 12 weeks.</td>
<td>Warm-up only.</td>
<td>No significant effect of pre-exercise stretching on all-injuries risk (HR = 0.95, 95% CI 0.77–1.18), soft-tissue injury risk (HR = 0.83, 95% CI 0.63–1.09) or bone injury risk (HR = 1.22, 95% CI 0.86–1.76).</td>
<td>68</td>
</tr>
<tr>
<td>Pope et al. (1998)</td>
<td>Two 20s static stretches for each of their soleus and gastrocnemius muscles during warm-up. 11 week programme of stretching prior to intense physical activity</td>
<td>Two 20s static stretches for wrist flexors and triceps muscles during warm-up.</td>
<td>No significant effect of pre-exercise stretching on all-injuries risk (HR = 0.92, 95% CI 0.52–0.61).</td>
<td>55</td>
</tr>
<tr>
<td>Van Mechelen et al. (1993)</td>
<td>Three 10s static stretches for each of five lower extremity muscle groups following warm-up. Cool-down also performed. Twice daily for 16 weeks. 10min of contract-relax stretching for five lower extremity muscle groups performed as part of warm-up. Stretching performed as part of a seven-part prophylactic programme. Stretches performed prior to every practice session or game over 6 months.</td>
<td>Subjects continued with normal programme of stretching, warm-up and cool-down.</td>
<td>No significant difference (P &gt; 0.05) in injury incidence between groups. Relative risk for injury was 1.12 (95% CI 0.56–2.72).</td>
<td>42</td>
</tr>
<tr>
<td>Ekstrand et al. (1983)</td>
<td>Five 30s static stretches for the hamstrings performed three times daily in addition to normal pre-exercise stretching. 13 week programme.</td>
<td>Subjects continued with normal pre-exercise stretching.</td>
<td>No significant difference (P &gt; 0.05) in injury incidence between groups. Relative risk for injury was 1.12 (95% CI 0.56–2.72).</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 5. Controlled clinical trials of the efficacy of stretching for preventing injury

<table>
<thead>
<tr>
<th>Author</th>
<th>Intervention</th>
<th>Control</th>
<th>Results</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hartig and Henderson (1999)</td>
<td>Five 30s static stretches for the hamstrings performed three times daily in addition to normal pre-exercise stretching. 13 week programme.</td>
<td>Normal pre-exercise stretching.</td>
<td>Subjects in the intervention group had significantly fewer lower extremity overuse injuries (P = 0.02).</td>
<td>30</td>
</tr>
<tr>
<td>Cross and Worrell (1999)</td>
<td>Three 15s stretches for each of four lower extremity muscle groups as part of pre-practice schedule. Stretches performed as part of practice schedule over course of a year.</td>
<td>General pre-practice stretching for upper and lower extremities.</td>
<td>A significantly lower incidence of injury was found in the intervention group (P &lt; 0.05).</td>
<td>26</td>
</tr>
<tr>
<td>Bixler and Jones (1992)</td>
<td>90s static stretching routine performed following 90s warm-up at the end of half-time. Programme lasted one football season.</td>
<td>Normal half-time activities.</td>
<td>Subjects in the intervention group had significantly lower third-quarter sprains and strains per game (P &lt; 0.05).</td>
<td>16</td>
</tr>
</tbody>
</table>
reduction in all-injuries risk. In the 1999 study by Pope et al., statistical power was sufficient to rule out a clinically useful reduction in injury risk, although the authors reduced the stretching duration to one 20 s stretch per muscle group, despite the previous study having shown no significant reduction in injury risk with greater total stretch stimulus. The best estimate of the effect of pre-exercise stretching was that it reduces all-injuries risk by 5%, with a 23% or greater reduction in all-injuries risk being ruled out with 95% confidence (HR = 0.95, 95% CI 0.77–1.18). When this is expressed in absolute terms, the authors calculated that on the basis of the protocol used subjects would need to stretch for an average of 260 h to prevent one injury. It should be noted however, that in both the 1998 study by Pope et al. (RR = 0.76, 95% CI 0.37–1.54) and the 1999 study (HR = 0.83, 95% CI 0.63–1.09) there was a greater, but still statistically insignificant, reduction in risk of soft-tissue injuries. This translates to a best estimate of a 17% reduction in risk of soft-tissue injuries. When expressed in absolute terms, based on the protocol utilized (Pope et al. 1999), subjects would have to stretch for an average of 212 h to prevent one soft-tissue injury. On this basis it would seem that the practice of pre-exercise stretching for reducing injury is of limited clinical significance.

Although the protocol used was based on established methods of stretching (St George 1989) that result in significant increases in flexibility, it has been shown (Halbertsma et al. 1996; Magnusson et al. 1996) that typically advocated protocols only produce a transient change in viscoelastic properties, maintained increases in flexibility being attributed to an increased stretch tolerance. Nevertheless, even transient increases in biomechanical properties of muscle should theoretically result in a reduction in injury for the duration of that effect (Glick 1980; Ciullo & Zarins 1983; Noonan & Garrett 1992; Gleim & McHugh 1997). As ethical reasons prevent the testing of failure properties of human skeletal muscle, research findings are drawn from the animal model in which controlled strain injury has been extensively studied (Gleim & McHugh 1997). However, although some research on the effect of stretching on passive failure properties exists (Best et al. 1989), as yet the effect of passive stretching on active failure properties remains undocumented. As the majority of injuries occur when muscle is active and eccentrically contracting (Noonan & Garrett 1992), it is in this area that further research is required.

It is however feasible that passive stretching of sufficient duration could reduce injury risk. In vitro research (Goldspink et al. 1995; Yang et al. 1997) has demonstrated hypertrophy of muscle fibres and an increase in the expression of Insulin-like Growth Factor (IGF-1), following prolonged immobilization in a lengthened position. Although no evidence exists as to whether similar effects can be achieved in vivo, if similar hypertrophy can be achieved following passive stretching, it is reasonable to assume that this would result in an increase in energy absorption and reduction in injury risk. Indeed, the CCT by Hartig and Henderson (1999) in which subjects stretched four times daily, amounting to a total stretching duration of greater than 40,950 s per muscle group (11.375 h), showed a significant reduction (P = 0.02) in incidence of lower extremity overuse injuries. Although the protocol utilized by van Mechelen et al. (1993) was also of significant duration (6720 s per muscle group, 1.867 h), compliance with the prescribed procedure was extremely low (47%), making evaluation of efficacy difficult.

In one of the negative RCTs (van Mechelen et al. 1993), subjects who stretched were actually found to be at higher risk of injury, in agreement with previous authors (Kerner & D’Amico 1983; Howell 1984; Jacobs & Berson 1986). Basic science evidence would
indicate that injudicious stretching might indeed increase risk of injury. Noonan et al. (1994) showed a significant decrease in maximal contractile force in muscles passively stretched to 30% of force to failure, accompanied by histologic evidence of focal areas of muscle fibre rupture and haemorrhage near the distal musculo-tendinous junction. Hasselman et al. (1995) also found a significant reduction in maximal contractile force in muscles actively stretched to 70% of passive force to failure. Similar reductions in maximal contractile force, seen following repeated muscular contractions, have been attributed to fatigue (Taylor et al. 1993). Fatigue has been shown to significantly reduce the ability of muscle to absorb energy (Mair et al. 1996) and is widely believed to be a predisposing factor in muscle injury (van Mechelen et al. 1992).

In two of the positive trials (Ekstrand et al. 1983; Bixler & Jones 1992) the stretching procedure was performed as part of a routine which included warmup and therefore efficacy of the stretching procedure itself could not be evaluated. In vitro studies by Strickler et al. (1990) demonstrated significant increases in force and length to failure following passive warming of muscle, as well as a non-significant increase in energy absorbed prior to failure. Safran et al. (1988) showed significant ($P<0.01$) increases in length and force to failure following submaximal contraction, implying that a protective effect may be gained from the warmup procedure. However, it is important to ensure warmup intensity is not too vigorous as a 37% decrease in passive stiffness, as well as an increase in energy absorbing capabilities, would therefore seem necessary to reduce injury risk.

Although Magnusson et al. (1997) found that endurance athletes with reduced flexibility had significantly ($P<0.01$) stiffer hamstring muscles, viscoelastic stress relaxation was similar in both tight and normal subjects. This is consistent with the findings by Pope et al. (1998) that there was no difference in the flexibility gains between tight and normal subjects. It would therefore appear that passive muscle stiffness is a function of the number of cross-links or collagen content, which are unlikely to be affected by non-disruptive stretching. It is proposed that in order to effect a change in normal, but short, healthy tissue it is therefore necessary to initiate the plastic change only achieved by creating minor muscle damage (Lederman 1997). From the basic science literature reviewed it is obvious that such stretching performed prior to exercise has the potential to increase the risk of injury.

Evidence would therefore suggest that if stretching is performed as part of a warmup procedure it is necessary to avoid the potential damage or fatigue that can be caused by overstretching. The problem exists as to how to quantify the point at which a muscle enters the plastic region. Typically advocated stretching protocols define the end point as a point creating a sensation of stretch, not pain (Hartig & Henderson 1999, Pope et al. 1999). However, stretching has been shown to increase the pain threshold (Halbertsma et al. 1996; Magnusson et al. 1996) and may therefore adversely affect the ability to perceive the point at which damage occurs. PNF stretching, although potentially offering greater injury preventing effects, should also be used with caution due to the increased possibility for overstretching when enlisting the help of a third party. Determination of a reliable and pragmatic method of achieving the desired end point requires further research.

There is some clinical evidence to suggest that prolonged stretching performed outside of the pre-exercise period can reduce incidence of injury (Hartig & Henderson 1999). Prolonged low load stretching in animals has been shown to increase muscle length and hypertrophy (Goldspink et al. 1995; Lederman 1997; Yang et al. 1997) as well as permanently lengthening connective tissue (Saepa et al. 1981). Similar effects have been achieved in human subjects.
with osteoarthritic hips (Leivseth et al. 1989) and joint contractures (Wessling et al. 1987). Evidence of similar effects in healthy subjects with foreshortened muscles has yet to be documented.

It is apparent that a need exists for carefully designed and conducted RCTs in this field if we are to make informed and unbiased decisions as to whether stretching can reduce injury risk and further our understanding of the mechanisms involved. The development by the Cochrane Collaboration of a database of RCTs in the area of physical therapy (Newham 1995) will hopefully encourage researchers in the field of complementary medicine to recognize the need for quality trials if our sphere of medicine is to continue gaining credibility amongst orthodox practitioners. With further evidence from high-quality RCTs it will be possible to implement an effective injury prevention programme founded on valid and reliable evidence-based research.

CONCLUSION

No definitive conclusions could be drawn as to whether stretching reduces the incidence of exercise-related injury due to the heterogeneity and poor quality of the selected studies. A need exists for carefully controlled clinical trials of sufficient power to identify a clinically significant effect, and with much more attention paid to the proper design and conduct of such studies.

Available evidence would suggest that pre-exercise stretching may increase the risk of injury. However, basic science and preliminary clinical evidence would indicate that prolonged stretching in the post-exercise period may increase the energy absorbing capabilities of muscle thereby reducing the risk of injury. Further research is required to clarify these findings, although a rethinking of current practices are indicated.

References

Altman DG 1999 What randomised trials and systematic reviews can offer decision makers. Hormone Research 51 (Suppl 1): 36–43
Greenhalgh T 1997 How to read a paper: The Basics of Evidence Based Medicine, 2nd edn. BMJ Publishing Group, London
Jacobs SJ, Berson BL 1986 Injuries to runners: a study of entrants to a 10 000 meter race. American Journal of Sports Medicine 14: 151–155

© 2003 Elsevier Science Ltd. All rights reserved.
Efficacy of stretching

McHugh MP, Magnusson SP, Gleim GW, Nicholas JA 1992
Salter RB 1983. Textbook of disorders and injuries of the musculoskeletal system. Williams & Wilkins, Baltimore

Appendix A. Checklist for Assignment of Methodological Scores

Details of criteria listed in Table 1. Each criterion must be applied independently of the other criteria.

A. Description of inclusion and exclusion criteria (1 point). Restriction to a homogenous study population, ie same athletic population, previous injuries (1 point).

B. Comparability for age, sex, previous injuries, hours of training, intensity of training and value of outcome measures (2 points each).

C. Randomization procedure explicitly described (3 points); randomization procedure which excludes bias—for example, sealed envelopes (3 points).
D. Smallest group immediately after randomization.
E. Stretching procedure explicitly described, ie type of stretch used, description of procedure, duration of stretch, number of repetitions, number of times daily/weekly during experimental period (1 point each).
F. Reference procedure explicitly described, ie type of reference activity, description of procedure, duration of procedure, number of repetitions, number of times daily/weekly during experimental period (1 point each).
G. Other mobilizing exercises avoided or comparable between groups.
H. Compliance with allocated procedure reported (2 points); compliance greater than 90% in each group after randomization (3 points)
I. Assessor blind to subjects experimental grouping
J. Relevance of outcome measures—occurrence of injury, type of injury, flexibility, passive muscle stiffness, strength, fitness/ gait economy (2 points each).

K. Number of subjects who withdraw given for each group without reasons for withdrawal (2 point); no dropouts or number of patients for each group with reasons for withdrawal (5 points).
L. All randomized subjects minus the number of subjects at the main point of measurement for the most important outcome measure, as a proportion of all randomized subjects. Less than 20% attrition rate in each group (4 points), less than 10% attrition rate in each group (8 points).
M. When less than 10% attrition rate: analysis on all randomized subjects for main outcome measure and on the most important points of measurement minus missing values, regardless of non-compliance (10 points). When greater than 10% attrition rate: alternative analysis such as baseline comparison for differences between groups at baseline measurement (10 points).
N. Means and standard deviations presented for each group (5 points); means only presented (3 points).