Concentric Evertor Strength Differences and Functional Ankle Instability: A Meta-Analysis

Brent L. Arnold, PhD, ATC, FNATA; Shelley W. Linens, PhD, ATC; Sarah J. de la Motte, PhD, ATC; Scott E. Ross, PhD, ATC

Virginia Commonwealth University, Richmond, VA. Dr de la Motte is now at the Uniformed Services University of the Health Sciences, Bethesda, MD.

Objective: To determine whether concentric evertor muscle weakness was associated with functional ankle instability (FAI).

Data Sources: We conducted an electronic search through November 2007, limited to English, and using PubMed, Pre-CINAHL, CINAHL, and SPORTDiscus. A forward search was conducted using the Science Citation Index on studies from the electronic search. Finally, we conducted a hand search of all selected studies and contacted the respective authors to identify additional studies. We included peer-reviewed manuscripts, dissertations, and theses.

Study Selection: We evaluated the titles and abstracts of studies identified by the electronic searches. Studies were selected by consensus and reviewed only if they included participants with FAI or chronic ankle instability and strength outcomes. Studies were included in the analysis if means and SDs (or other relevant statistical information, such as *P* values or *t* values and group n's) were reported for FAI and stable groups (or ankles).

Data Extraction: Data were extracted by the authors independently, cross-checked for accuracy, and limited to outcomes of concentric eversion strength. We rated each study for quality. Outcomes were coded as either fast or slow velocity (ie, equal to or greater than 110°/s or less than 110°/s, respectively).

Data Synthesis: Data included the means, SDs, and group sample sizes (or other appropriate statistical information) for the FAI and uninjured groups (or ankles). The standard difference in the means (SDM) for each outcome was calculated using the pooled SD. We tested individual and overall SDMs using the Z statistic and comparisons between fast and slow velocities using the Q statistic. Our analysis revealed that ankles with FAI were weaker than stable ankles (SDM = 0.224, Z = 4.0, $P < .001$, 95% confidence interval = 0.115, 0.333). We found no difference between the fast- and slow-velocity SDMs (SDM_{Fast} = 0.189, SDM_{Slow} = 0.244, Q = 29.9, *df* = 24, *P* = .187). Because of the small SDM, this method of measuring ankle strength in the clinical setting may need to be reevaluated.

Key Words: ankle sprains, chronic ankle instability, ankle weakness, isokinetic assessment, ankle torque, ankle force

---

Key Points

- Our meta-analysis showed that participants with functional ankle instability had weaker ankles than participants with stable ankles.
- No difference between fast-velocity and slow-velocity strength testing was noted.

Functional ankle instability (FAI) remains a significant problem for both the general and athletic populations. Ankle injuries account for approximately 14% of all injuries in collegiate sports,[1] and 20% to 40% of all ankle sprains result in FAI.[2–4] Functional ankle instability is clinically important because it prevents approximately 6% of patients from returning to their occupations,[5] and 13% to 15% of patients remain occupationally handicapped for at least 9 months and up to 6.5 years after injury.[5,6] Originally defined by Freeman et al,[7] FAI is a condition that results in the sense of giving way after an ankle sprain. A variety of mechanisms have been proposed as the cause of FAI, including delayed neuromuscular response,[8] proprioceptive deficits,[9,10] impaired balance,[11] and ankle evertor weakness.[12]

Examination of the FAI literature indicates that the relationship between ankle muscle strength (especially evertor strength) and FAI is equivocal. Several groups[13–19] have failed to find ankle strength differences in participants with FAI. However, others[20–25] have reported decrements in FAI ankle concentric evertor isokinetic strength. The reason for these discrepancies is unclear. However, we suspect that one potential reason is the variation in velocities used to test ankle strength. The slowest testing reported was an isometric contraction: 0°/s.[22] Alternatively, isokinetic testing has been used and can theoretically range from just above isometric to concentric speeds approaching 300°/s: values ranged from 30°/s[14–17,25–29] to 240°/s.[26] This wide range of velocities inherently increases the variability of results and may explain the literature's lack of clarity.

Whether evertor weakness is associated with FAI is unclear based on the current literature. Thus, our primary objective was to combine study outcomes to determine the association between FAI and evertor weakness. Because of the variety of velocities used to test evertor strength, velocity was a possible confounding factor. Therefore, we also sought to determine whether strength differences were affected by testing speed by categorizing the outcomes in each study into slow and fast velocities to create a velocity moderator variable.

We hypothesized that when studies were combined, ankles with FAI would be weaker than uninjured ankles. We also hypothesized that slow and fast testing velocities would produce different effect sizes but that both the fast and slow effect sizes would indicate that FAI ankles were weaker than uninjured ankles.

METHODS

Published Study Selection

Our search strategy was conducted in 4 stages (Figure 1). For stage 1, we performed an initial electronic search and evaluated studies for inclusion. In stage 2, we used the previously selected studies and performed a forward search. Stage 3 consisted of a hand search of the reference lists of articles selected in stages 1 and 2. The final stage consisted of our contacting the corresponding authors of the previously selected studies to solicit any additional references. Stages 2 and 3 were iterative until no further articles were identified.

Literature Search. Our electronic literature search was restricted to English-language publications found in the following databases through November 2007: PubMed (National Library of Medicine, Bethesda, MD), Pre-CINAHL, CINAHL, and SPORTDiscus. We searched the latter 3 databases simultaneously using EBSCOhost (EBSCO Industries, Inc, Birmingham, AL). The strategy and search results are presented in Table 1. Our senior research team member (B.L.A.) directed the search. He has 14 years of research experience and expertise in the area of FAI and was assisted by 2 doctoral students specializing in FAI research.

Article Inclusion and Exclusion Criteria. Once we completed the initial search, we reviewed the titles and abstracts of each article retrieved by the search. After reading the titles and abstracts, we included articles based on a group consensus. For us to include an article, it had to either have reported means and SDs for an injured group (or ankles) and an uninjured comparison group (or ankles) or have statistics reported in enough detail that we could calculate effect sizes. We also required each study to use an inclusion criterion (or FAI definition) of giving way or frequent sprains or to have described the target condition as FAI.

In addition to journal articles, we also reviewed any theses and dissertations retrieved as part of the initial search. Both types of manuscripts were included in the analysis if they met our inclusion criteria. Abstracts from conferences were not reviewed for inclusion because of their limited availability in the electronic databases.

Forward and Hand Search. We conducted our forward search using the Science Citation Index (The Thomson Corp, New York, NY). This search was conducted on each of the previously included articles. Specifically, we used each article as a search item and searched from its publication date forward to November 2007. The search identified other articles that cited the initial article. Our review and inclusion or exclusion of articles identified in the forward search was identical to that described above. The references from the articles of the initial and forward searches were then hand searched to identify other potential articles.

Contact With Authors. To complete the search, we contacted the corresponding author of each article by letter or e-mail. In the correspondence, we listed which articles of theirs we included and asked that they identify additional articles for us to consider. We then applied our inclusion criteria to determine whether to include the additional studies.

Meta-Analysis

Data Extraction. For each study, 3 investigators independently extracted means and SDs or other appropriate measures, such as participant n and t values, for the FAI and stable ankle groups (or ankles in studies with contralateral comparisons). We resolved any discrepancies by reexamining the data and agreeing by consensus on the final data to be included. For studies using treatments, we selected only the pretreatment data for inclusion. If pretreatment data were not available, the study was excluded.

The outcomes we selected for this study were limited to measures of evertor muscle strength. Within this restriction, any mode of strength assessment (eg, isokinetic, isometric, isotonic) was eligible for inclusion. However, after we applied the inclusion criteria, only isokinetic studies with speeds ranging from 30°/s to 240°/s were included. The outcomes for each study are presented in Table 2.

Our analysis was not restricted to any particular research design or question. To be included, each study had to have at least 1 measure of evertor strength and provide a comparison between injured and uninjured participants or ankles. The typical study selected for review was a case-control study, but randomized controlled trials and ex post facto designs were eligible provided pretreatment data were included for both groups. Study populations were not restricted. Table 3 includes the participant demographics for each study.

Quality Assessment. To assess quality, we developed a 20-item questionnaire[30] to identify possible threats to construct, internal, and external validity as identified by Cook and Campbell.[31] Based on validity threats, we tailored our questions to ankle instability research. Each item was scored as *yes* or *no*, with 3 items potentially scored as *not applicable*. Thus, the overall score was calculated as the percentage of items scored as *yes* after any *not applicable* items were removed. Three reviewers independently scored each study. The mean of the 3 reviewers' scores served as the quality score. Meta-regression was then used to establish the relationship between study quality and the standard difference in the means (SDM).

Statistical Methods. We completed the statistical analysis using Comprehensive Meta-Analysis (version 2.2.034; BioStat International, Inc, Tampa, FL). Depending on the information presented in each study, we entered data either as means and SDs (n = 11) or as paired groups using sample size and the paired t value (n = 1). From these data, we calculated the SDM with the pooled SD for each study and used it for statistical analysis. We used the Z statistic (which follows the normal distribution) to test whether individual and study category SDMs were different from zero.[32] However, before calculating the Z value, we assessed the heterogeneity of effect sizes among the studies
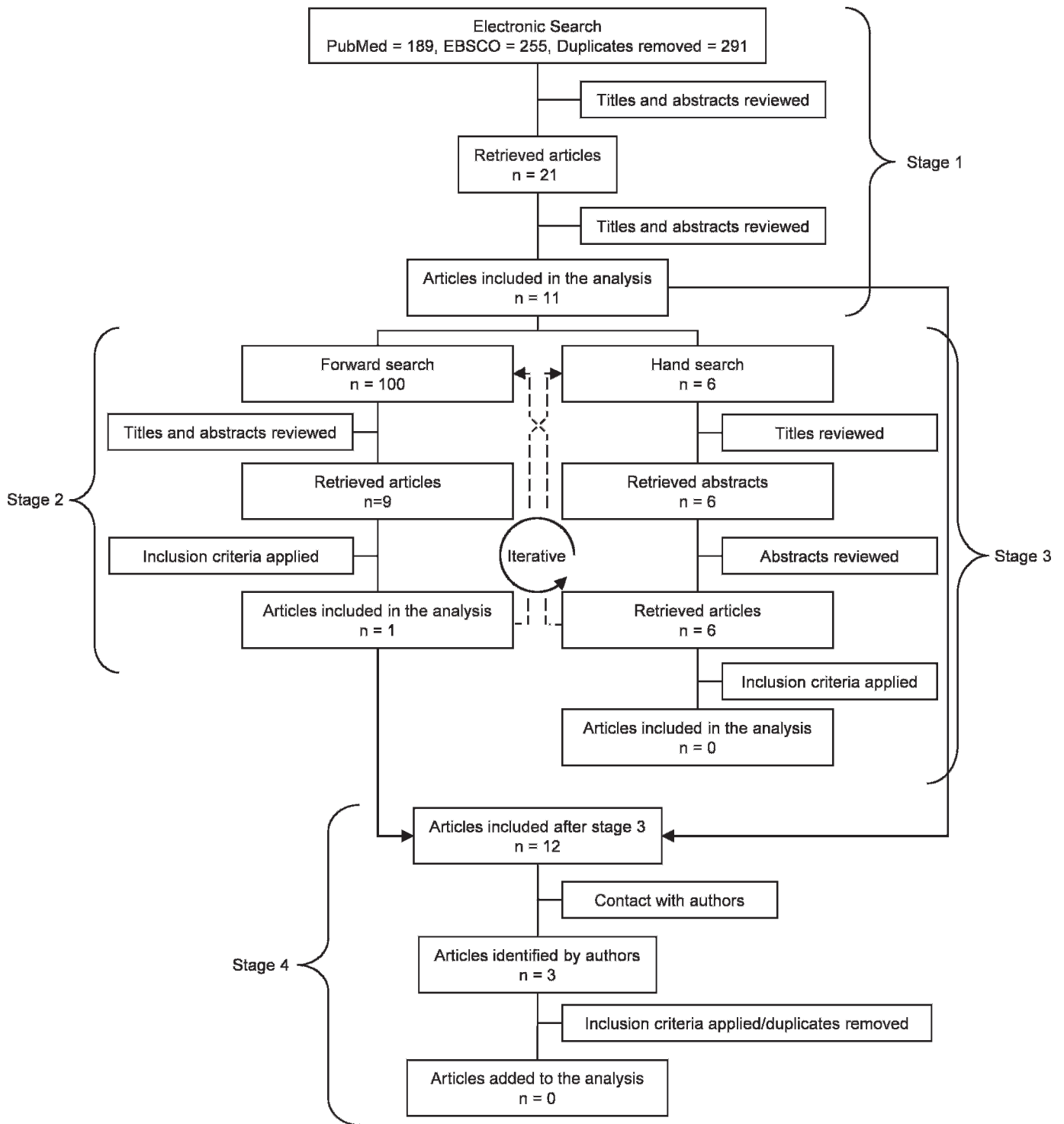
Figure 1. Flow chart for the manuscript review process. Stages 2 and 3 were iterative.

using the Q statistic (which approximates the χ^2 distribution). If the Q value was significant (ie, the between-studies variance was greater than chance expectations), we computed the Z statistic using a random-effects model. Otherwise, the Z statistic was calculated using a fixed-effects model.[33] We also used the standardized residual to identify outcomes that were outliers. Studies with stan-

dardized residuals greater than or equal to 3.0 were deleted from the analysis.[34]

Multiple Isokinetic Velocities Within a Single Study. Several groups included strength assessments at more than 1 velocity (n = 6). Although the literature does not provide a clear distinction between fast and slow test velocities, it is common for authors to test at more than 1 velocity, ranging

Table 1. Stepped PubMed/EBSCO Host Search Strategy With the Number of Studies Returned at Each Step[a]

Step	Strategy[a]	PubMed	EBSCO
#27	Search (#18) AND (#26)	189	255
#26	Search ((((((#19) OR (#20)) OR (#21)) OR (#22)) OR (#23)) OR (#24)) OR (#25)	428 715	108 211
#25	Search power*[TIAB]	158 382	43 558
#24	Search force*[TIAB]	164 599	35 297
#23	Search isomet*[TIAB]	20 238	6091
#22	Search isoton*[TIAB]	10 625	850
#21	Search isokin*[TIAB]	3429	4950
#20	Search torque*[TIAB]	8208	4824
#19	Search strength[TIAB]	99 872	34 248
#18	Search ((#4) AND (#10)) AND (#17)	1133	1108
#17	Search ((((((#11) OR (#12)) OR (#13)) OR (#14)) OR (#15)) OR (#16)	1 770 820	136 088
#16	Search multiple [TIAB]	457 447	39 268
#15	Search repetitive [TIAB]	33 616	3012
#14	Search functional* [TIAB]	545 870	36 451
#13	Search recurren* [TIAB]	239 173	12 473
#12	Search chronic* [TIAB]	566 587	51 564
#11	Search "Recurrence" [MeSH]	117 396	7873
#10	Search (((((#5) OR (#6)) OR (#7)) OR (#8)) OR (#9)	104 954	12 248
#9	Search inversion [TIAB]	17 721	1193
#8	Search instability [TIAB]	44 779	5369
#7	Search sprain* [TIAB]	2516	2708
#6	Search unstable [TIAB]	39 169	2905
#5	Search "Sprains and Strains"[MeSH:NoExp] OR "Joint Instability" [MeSH]	11 935	2395
#4	Search ((#1) OR (#2)) OR (#3)	24 851	11 741
#3	Search ankle* [TIAB]	22 721	11 613
#2	Search "Lateral Ligament, Ankle" [MeSH]	191	21
#1	Search "Ankle Joint" [MeSH]	6854	748

Abbreviations: TIAB, title and abstract; MeSH, medical subject heading.

[a] *Indicates wild card.

from relatively slow to relatively fast velocities.[14,15,25,26,28,35] Depending on the source, the combined inversion-eversion motion ranges from 50°[20] to 55°.[36] By arbitrarily defining *fast* as completing the range of motion in 0.5 seconds or less, we determined that isokinetic velocities at or above 110°/s (55°/0.5 s = 110°/s) would be termed *fast*. Thus, using 110°/s as the cutoff, we created a moderator variable that divided velocities into fast and slow.

Because we expected that our results might differ across the range of velocities used in ankle strength testing, we elected to treat multiple velocities within a single study as independent. However, multiple velocities within a study are not, in fact, independent, so the standard error and confidence interval (CI) for the overall effect are smaller than expected, the statistical test for the overall effect is liberal, and the comparison between fast and slow velocities is conservative. The alternative would have been to average the velocities within each study. Yet doing so would have eliminated the ability to compare fast and slow velocities.

To compare the effects of fast and slow velocities, our first step was to test for heterogeneity among the studies using the Q statistic. If the Q statistic was nonsignificant, no further testing was warranted. With a significant Q statistic, we grouped study outcomes based on test velocity as a moderator variable.

Bias Assessment. Bias was assessed using 3 techniques. First, a funnel plot was created to visually interpret the data (ie, data points should be symmetrical within the funnel). Next, we used the Egger regression intercept method[37] and the Duvall and Tweedie[38] trim-and-fill procedure to confirm the funnel plot. We elected to conduct the bias assessment on the mean effect for each study (ie, the SDM calculated from the average of the study's outcomes). We did this because bias is more appropriately related to studies, not outcomes, and because bias can have multiple causes (eg, study quality[39]) that would be expected to affect all of a study's outcomes.

RESULTS

Identification of Participant Characteristics

We extracted study and participant characteristics from either reported characteristics or reported inclusion and exclusion criteria (Tables 2 and 3). No consistency is evident in the participant characteristics reported. Because of this variability, we did not attempt to compare studies using these characteristics. However, studies that included statistical comparisons between injured and uninjured groups (or ankles) received additional credit as part of our quality assessment.

Quality Assessment

To evaluate our quality assessment, we calculated the intraclass correlation coefficient (2,1) on the independent study ratings given by our 3 evaluators. This produced a coefficient of 0.676 with a 95% CI of 0.360 to 0.882. This result is similar to previous reports on the PEDro quality assessment tool.[40] The average quality score was 26.6% ± 9.4%, with a range of 15% to 42%. No relationship was found between study quality and the SDM (slope = −0.002, P = .78), indicating that study quality did not relate to greater or smaller effect sizes.

Meta-Analysis

Our initial overall analysis revealed 1 outlier (ie, a standardized residual equal to or greater than 3.0)[33] in the data set (ie, Willems et al).[28] Thus, we deleted this study (and 2 corresponding outcomes) from the overall analysis. For the overall analysis investigating whether FAI was associated with ankle weakness, we found no heterogeneity across outcomes ($Q = 29.9$, $df = 24$, $P < .187$). Thus, the use of the fixed-effects model was warranted, revealing an SDM of 0.224 ($Z = 4.0$, $P < .001$, 95% CI = 0.115, 0.333). In other words, ankle weakness was associated with functionally unstable ankles (Figure 2).

Differences Between Velocities. Based on the nonsignificant Q value reported above, we determined that a separate analysis for velocity was unnecessary because the lack of heterogeneity meant that velocity was not affecting the data in a meaningful way. The SDMs for fast and slow velocities are presented in Figure 2.

Bias Assessment. The Egger intercept was 0.974 ($P = .28$, 2 tailed) and the Duvall and Tweedie trim and fill identified no studies to be filled in the bias funnel plot (Figure 3), These results indicate that our findings were not biased by unpublished or inaccessible (ie, "fugitive")[41] studies.

DISCUSSION

We conducted this meta-analysis because of conflicts in the literature. Several groups[13–17,26,27,35] have failed to detect concentric evertor weakness in participants with FAI. In contrast, others[22–25,28] have reported ankle weakness with FAI. Thus, our primary intent was to clarify this discrepancy. Our findings indicate that concentric evertor weakness is present in participants with FAI. Specifically, we noted the SDM between stable and unstable ankles to be significant but small (0.224). In other words, on average, the means of the FAI groups and stable groups were separated by 0.224 SDs (Figure 2). Interestingly, of the 27 outcomes included in this analysis, only 4 outcomes (limited to 1 study) were significant (as indicated by the P value associated with the Z value). This finding clearly illustrates the usefulness of meta-analysis: when the studies were combined, a significant result that most individual studies failed to demonstrate was revealed. The most likely reason for the lack of differences in the individual studies was inadequate statistical power. Using the Cohen method for determining sample size,[42] the overall effect size of 0.224 from this analysis, an alpha level of .05, and a target power of 80%, we determined that 313 participants would be needed per group in a 2-group analysis to detect this effect. Of the studies we included, Lentell et al[15] had the largest group (N = 42). Not surprisingly, none of the outcomes of Lentell et al displayed a significant difference.

Based on these findings, we question whether ankle weakness can be detected clinically. One method of determining clinical detectability is by calculating the minimal detectable change (MDC). The MDC is calculated as $MDC = SD\sqrt{1-R} \times \sqrt{2}$ for differences between 2 measures. Using the average intraclass correlation coefficient (0.89) from previous studies[43,44] and the pooled SD of the injured ankles from our analysis (5.62 Nm), the calculated MDC was 2.64 Nm. Similarly, using our overall

SDM (Figure 2) and multiplying it by the pooled SD, the difference between stable and unstable ankles was 1.26 Nm, below the MDC.

The clinical importance of this difference is unknown. The 1.26-Nm difference between stable and unstable ankles seems too small to be clinically important. However, it appears likely to us that the Nm unit may not be the best for ankle strength measurements. The moment arm of the peroneal muscles is only 21 to 22 mm.[45] Thus, 1.26 Nm converts to a 57.3- to 60.0-N (12.9- to 13.5-lb) weakness of the peroneal muscles. We believe this may represent a clinically meaningful weakness of the ankle evertors.

Based on the relatively small overall SDM, we question previous suggestions that ankle strength training may be beneficial to only a minority of patients[16] and that return-to-play criteria not be based on strength.[26] Given the result of our analysis, we suspect that strength is an important factor. Strength might be better assessed using a measure of muscle force rather than joint torque. Finally, we believe that evidence of ankle evertor weakness in FAI is sufficient to warrant randomized control trials focused on strength training at the ankle.

Differences Between Velocities

Six of the 12 groups included in this analysis used more than 1 velocity. Although we are unaware of any previous research suggesting this is necessary, we believe it to be a common practice based on theoretical grounds. Specifically, faster velocities are perceived as being more representative of joint power than strength. We also suspect that clinicians and researchers believe that capturing strength across the spectrum of testing speeds offers a more complete assessment of ankle strength. However, our results suggest that this is not the case. Our analysis (Figure 2) did not reveal a difference between fast and slow velocities. Based on this finding, we would recommend performing clinical testing at only 1 velocity and using any velocity above 30°/s and below 240°/s. However, with the small differences described above and the force-velocity relationship of muscle, testing at slower velocities to maximize the difference would seem wiser to us.

Assessment of Bias

We did not detect any publication bias in our analysis. Publication bias may actually be the result of multiple factors, including true publication bias as well as bias in the retrieval method.[37] The primary concern is that studies may be missing from the data set and that these missing studies may actually cause a false-positive result. One of the easiest ways to detect bias is via the funnel plot (Figure 3). When no bias exists, studies are expected to distribute equally to the left and right of the SDM (ie, the open diamond). When bias exists, the typical pattern is an absence of data points (ie, studies) on the lower, left-hand side of the funnel. This is because the left hand side of the plot represents studies having or approaching either no effect or a negative effect. For example, in our analysis, studies between the open diamond (SDM = 0.19) and zero had no strength differences or small differences, with unstable ankles being weaker. Conversely, studies to the left of zero indicated that stable ankles were weaker than unstable ankles (ie, a negative effect). Studies tend to be

Table 2.   Inclusion and Exclusion Criteria and Outcome Variables for the Included Studies

Author(s)	Quality Score	Stated Inclusion and Exclusion Criteria	Involved Participants or Ankles	Uninvolved Participants or Ankles	Strength Instrument	Outcome Variables (Velocity of torque production, °/s)
Bernier et al[13]	22	Not specified	9	9	Kin Com II (Chattanooga Group Inc, Hixson, TN)	90
Hubbard et al[29]	44	Previous history of unilateral ankle sprain, frequent giving way of the ankle, pain, feelings of instability, and decreased function. No acute ankle sprain within 6 weeks of participation in study, history of surgery or fracture to either lower extremity, or any previous sprain on limb contralateral to the chronically unstable ankle. Control group: free of previous injury or surgery to either lower extremity.	30	30	Biodex 2 (Biodex, Inc, Shirley, NY)	30
Kaminski et al[14]	37	Participants (1) experienced at least 1 significant lateral ankle sprain of either right or left ankle, but not both, in which the participant was unable to bear weight or was placed on crutches within last year; (2) had no reported history of fracture to either ankle; (3) sustained at least 1 repeated injury or the experience of feelings of ankle instability or "giving way" in either the right or left ankle but not both; (4) were not undergoing any formal or informal rehabilitation of the unstable ankle; and (5) had no evidence of mechanical instability as assessed by a physician using an anterior drawer test.	21	21	Kin Com 125 AP (Chattanooga Group Inc)	30 60 90 120 150 180
Lentell et al[15]	38	Each participant reported a past history of inversion injury to only 1 ankle, which required protected weight bearing and/or immobilization. Participant reported no history of fracture to either lower extremity. The involved ankle was reported to be chronically weaker, more painful, and/or less functional than the other at the time of testing. Current subjective complaints were reported to be secondary to the past history of traumatic inversion sprain. No significant trauma reported 3 months before testing. Participant had been full weight bearing, without a limp, for at least 3 weeks before testing. Participant was currently not undergoing any formal or informal rehabilitation program for the ankle. Participant reported that the functional use of the ankle had plateaued or peaked since the original insult.	42	a	Cybex II+ (Cybex Division of Lumex Inc, Ronkonkoma, NY)	30 90 150 210
Lentell et al[16]	15	Participant reported a past history of inversion injury to only 1 ankle, which required protected weight bearing and/or immobilization. Participant reported no history of fracture to either lower extremity. The involved ankle was reported to be chronically weaker, more painful, and/or less functional than the other at the time of testing. Current subjective complaints were reported to be secondary to the past history of traumatic inversion sprain. No significant trauma reported 3 months before testing. Participant had been full weight bearing, without a limp, for at least 3 weeks before testing. Participant was currently not undergoing any formal or informal rehabilitation program for the ankle. Participant reported that the functional use of the ankle had plateaued or peaked since the original insult.	33	a	Cybex II+	30

Table 2. Continued

Author(s)	Quality Score	Stated Inclusion and Exclusion Criteria	Involved Participants or Ankles	Uninvolved Participants or Ankles	Strength Instrument	Outcome Variables (Velocity of torque production, °/s)
McKnight and Armstrong[27]	15	Not specified	15[b] 14[c] 29 total	14	Biodex (Biodex Medical Systems, Shirley, NY)	30 (reported as % body weight)
Pontaga[25]	28	Athletes who had had only 1 ankle injury were not included. Only handball players with grade II sprains with a partial tear of the lateral ligaments and moderate laxity of the ankle joint were selected for the investigation by a physiotherapist.	23 unstable 33 stable	[a]	REV-9000 (Technogym, Gambettola, Italy)	30 60 90 120
Porter et al[26]	26	Cited Hubbard TJ, Kaminski TW. Kinesthesia is not affected by functional ankle instability status. *J Athl Train.* 2002;37(4):481–486.	15	15	Kin Com 125 AP	120 (reported as % body weight) 240 (reported as % body weight)
Ryan[17]	35	Participants had history of unilateral functional instability of the ankle. At least 6 episodes of giving way of the ankle into inversion, with or without pain, within past 12 months, or a total of at least 3 inversion sprains, including 2 or more within past 18 months, with at least 1 in past 6 months. Sprain defined as an episode of giving way into inversion, followed by more than 2 days of pain and restriction of function. Affected ankle must have recovered to its usual state since last giving way and been fully engaged in sporting activity for at least 2 weeks before testing. Excluded if participant had sustained any other significant injury or suffered from any condition likely to cause a strength or mobility deficit in either lower limb or to cause disturbance of balance or reported pain during upper body ergometer or Cybex II testing.	45	[a]	Cybex II	30
Schrader[35]	15	More than 2 sprains. Subjective functional complaints.	20	20	Kin Com	120 180 60
Sekir et al[19]	29	Participants sustained at least 2 moderate sprains to same ankle that required medical intervention and complained of repeated episodes of "giving way," both within last 6 months. Moderate sprain defined as having moderate pain, swelling, and tenderness over the involved ligaments without significant instability.	24	[a]	Cybex	120

[a] Between-ankles comparison. Unless otherwise indicated, the number of stable ankles equaled the number of unstable ankles.
[b] Functional ankle instability group.
[c] Rehabilitation group (prerehabilitation date used for analysis).

Table 3. Characteristics of Participants in the Included Studies

Author(s)	Injured Group			Uninjured Group		
	Age, y	Height, cm	Mass, kg	Age, y	Height, cm	Mass, kg
Bernier et al[13]	22.89	181	80.25	26.22	170	65.08
Hubbard et al[29]	20.3	172.5	72.9	23.3	172.6	71.9
Kaminski et al[14]	19.3	181.5	84	19.5	179.5	82.5
Lentell et al[15]	22 (overall)	Not specified	Not specified	22 (overall)	Not specified	Not specified
Lentell et al[16]	Males: 26.9	Not specified	Not specified	Males: 26.9	Not specified	Not specified
	Females: 25.3			Females: 25.3		
McKnight and	FAI: 19.6	FAI: 171.7	FAI: 68.55	21.13	170.84	67.68
Armstrong[27]	R: 18.47	R: 172.9	R: 71.16			
Pontaga[25]	21 (overall)	186 (overall)	84 (overall)	21 (overall)	186 (overall)	84 (overall)
Porter et al[26]	22.1	170.3	73.6	21.7	169.5	72.4
Ryan[17]	23 (overall)	Not specified	Not specified	23 (overall)	Not specified	Not specified
Schrader[35]	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified
Sekir et al[19]	21	Not specified	Not specified	21	Not specified	Not specified

Abbreviations: FAI, functional ankle instability group, R, rehabilitation group (prerehabilitation date used for analysis).

missing from the lower portion of the plot because they have smaller sample sizes and, thus, larger standard errors. (It should be noted that, by convention, the funnel plot has a reversed y-axis, with higher values at the origin.) In other words, it is typical for small studies to be published only when they have large (ie, significant) effect sizes.

As can be seen from our funnel plot, studies are equally distributed to the left and right of the overall SDM, reflecting no bias. Although the funnel plot is a useful tool in visually identifying bias, as with regression plots, the funnel plot can be deceptive. Thus, we conducted confirmatory statistical analyses (ie, the Egger intercept and the Duvall and Tweedie trim-and-fill tests). Both of these tests also failed to detect bias.

Assessment of Quality of Included Studies

One concern with any meta-analysis is the effect of study quality on the analysis; that is, does individual study quality influence its effect size (SDM), and does that influence the overall effect? We are not aware of any quality assessment tool that has been specifically designed

Velocity	Study	Outcome: Evertor Strength at Indicated Speed, °/s	Standard Difference in Means	Lower Limit	Upper Limit	Z Value	P Value
Fast	Kaminski et al (1999)[14]	120	0.084	-0.521	0.689	0.272	.785
Fast	Kaminski et al (1999)[14]	150	0.363	-0.246	0.973	1.168	.243
Fast	Kaminski et al (1999)[14]	180	0.464	-0.149	1.077	1.483	.138
Fast	Lentell et al (1995)[15]	150	-0.135	-0.563	0.293	-0.617	.537
Fast	Lentell et al (1995)[15]	210	-0.115	-1.141	0.911	-0.219	.827
Fast	Pontaga (2004)[25]	120	0.667	0.120	1.213	2.391	.017
Fast	Porter et al (2002)[25]	120	0.005	-0.711	0.721	0.014	.989
Fast	Porter et al (2002)[25]	240	0.117	-0.599	0.834	0.321	.748
Fast	Schrader (1992)[35]	120	0.059	-0.561	0.679	0.187	.852
Fast	Schrader (1992)[35]	180	-0.019	-0.638	0.601	-0.059	.953
Fast	Sekir et al (2007)[19]	120	0.487	-0.087	1.061	1.664	.096
Fast	Group summary		0.189	0.006	0.371	2.023	.043
Slow	Bernier et al (1997)[13]	90	0.097	-0.827	1.022	0.206	.836
Slow	Hubbard et al (2007)[29]	30	0.323	-0.187	0.832	1.241	.215
Slow	Kaminski et al (1999)[14]	30	0.300	-0.308	0.908	0.966	.334
Slow	Kaminski et al (1999)[14]	60	0.401	-0.210	1.012	1.285	.199
Slow	Kaminski et al (1999)[14]	90	0.512	-0.103	1.126	1.631	.103
Slow	Lentell et al (1995)[15]	30	-0.073	-0.556	0.409	-0.298	.766
Slow	Lentell et al (1995)[15]	30	-0.091	-0.519	0.337	-0.416	.677
Slow	Lentell et al (1995)[15]	90	-0.038	-0.465	0.390	-0.173	.863
Slow	McKnight, Armstrong (1997)[27]	30	0.118	-0.520	0.757	0.363	.716
Slow	Pontaga (2004)[25]	30	0.958	0.397	1.519	3.346	.001
Slow	Pontaga (2004)[25]	60	0.833	0.279	1.388	2.947	.003
Slow	Pontaga (2004)[25]	90	0.678	0.131	1.225	2.430	.015
Slow	Ryan (1994)[17]	30	0.104	-0.189	0.397	0.698	.485
Slow	Schrader (1992)[35]	60	0.090	-0.530	0.710	0.284	.776
Slow	Group summary		0.244	0.108	0.379	3.533	.000
Overall			0.244	0.115	0.333	4.044	.000

Standard Difference in Means and 95% Confidence Interval
-2.00 -1.00 0.00 1.00 2.00
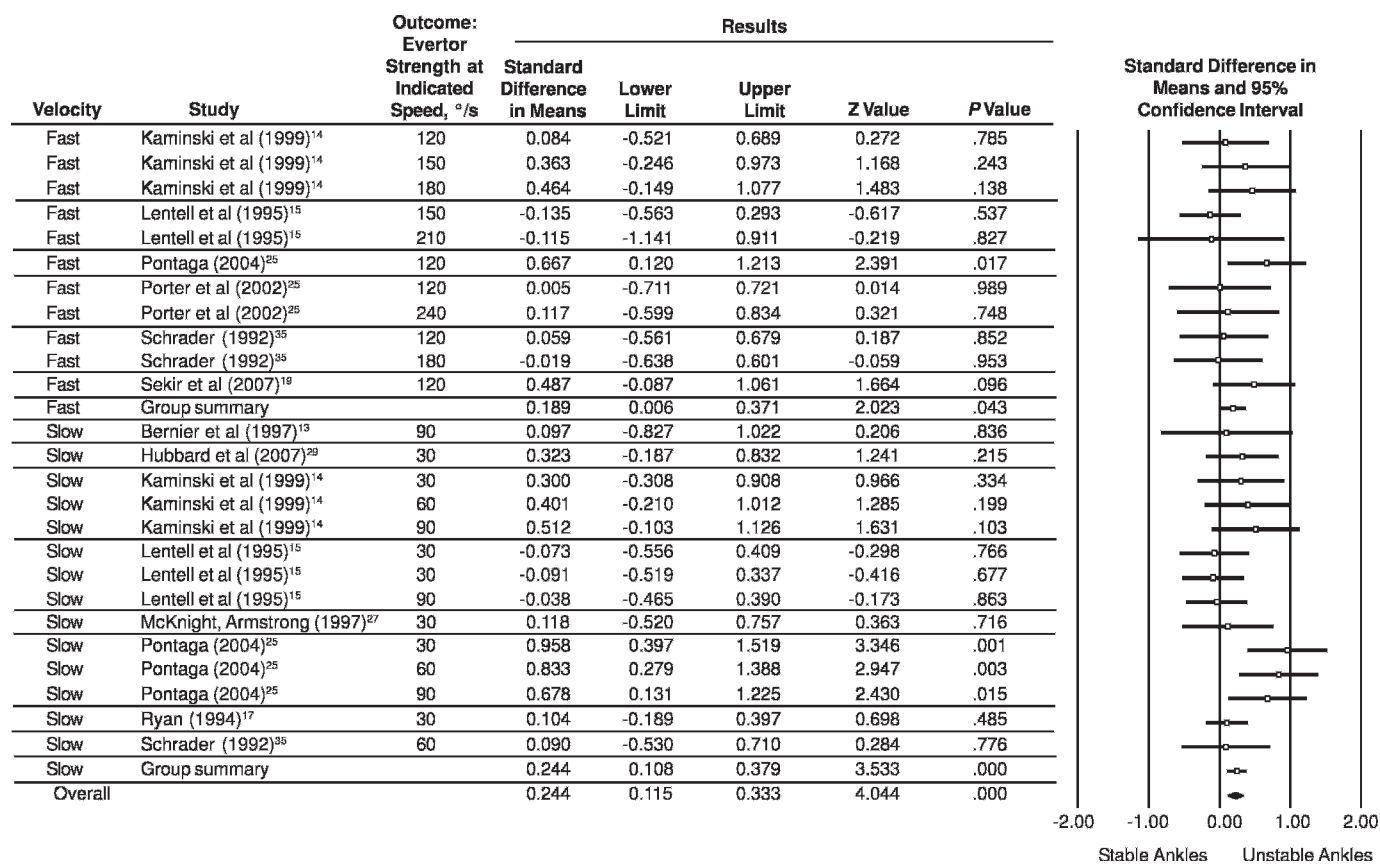Stable Ankles   Unstable Ankles

Figure 2. Forrest plot for strength outcomes. Studies are grouped (column 1) based on velocity of testing. Error bars equal the 95% confidence interval. ◆ Indicates standard difference in the means for the overall effect, with the diamond width representing the standard error.
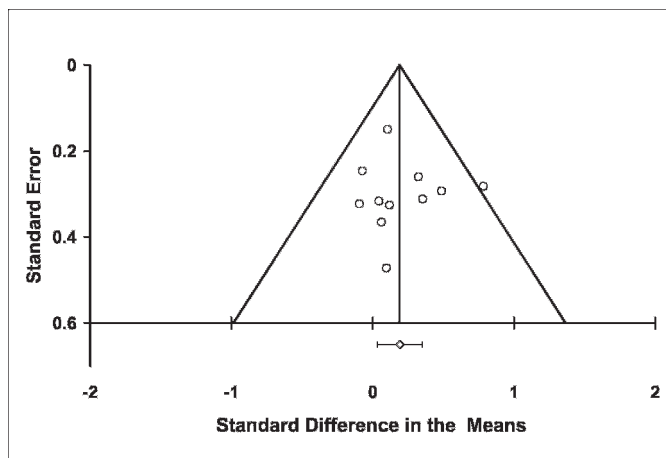
Figure 3. Funnel plot with the standard difference in the means plotted against the standard error. ○ Indicates observed studies; ◇, standard difference in the means for the observed studies (±SE).

for nonexperimental or quasi-experimental designs. Rather, the existing tools are designed for randomized controlled trials. The consequence is that studies that are not randomized or do not have appropriate control groups are necessarily scored lower. Although we agree that this is appropriate for studies comparing treatments, it seems inappropriate for studies that are not specifically designed to compare treatments. Therefore, we created a new quality scoring tool based on the threats of internal, construct, and external validity identified by Cook and Campbell.[31] Our intertester reliability (0.676) was comparable with that of the PEDro.[40] We would consider this intraclass correlation coefficient fair to good. Unlike laboratory measures that are quite stable, parts of this assessment were more subjective and potentially more error prone. The range of the CIs (0.360 to 0.882) suggests that the reliability may be quite high but could also be rather low. This finding suggests to us that users should practice with the tool before engaging in a serious quality assessment of studies.

There was no evidence that study quality affected the results, but the overall quality of the included studies was not high. It is important to note that quality is evaluated based on what the authors reported, not what they did. In our view, the quality score reflects at least 2 facets of the research process: how the research was conducted and how it was reported (which also includes the editorial process). Certain studies may have received lower-than-deserved scores because of how they were reported. For example, if multiple levels of outcome measures (eg, isokinetic velocities) were not reported as counterbalanced or randomly assigned (ie, monomethod bias[31]), studies lost points. This does not mean that levels of an outcome were not counterbalanced, only that such assignments were not reported. Thus, the low quality ratings of the included studies should be viewed cautiously, because they do not solely represent the quality of the research.

CONCLUSIONS

Despite previous authors' failure to demonstrate strength differences between participants with FAI and those with stable ankles, our meta-analysis clearly shows that weakness is associated with FAI. Based on this finding, we disagree with suggestions that strength assessment[27] and strength training[15] are not important parts of return-to-play criteria and rehabilitation, respectively. However, because of the small SDM and MDC calculated for this investigation, it may be necessary to rethink how ankle strength is measured, particularly whether Newtons (force) or Newton-meters (torque) are the preferred units. Finally, our results suggest that the velocity of strength testing is not a relevant factor in ankle strength testing.

REFERENCES

1. Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives. *J Athl Train*. 2007;42(2):311–319.
2. Braun BL. Effects of ankle sprain in a general clinic population 6 to 18 months after medical evaluation. *Arch Fam Med*. 1999;8(2): 143–148.
3. Yeung MS, Chan KM, So CH, Yuan WY. An epidemiological survey on ankle sprain. *Br J Sports Med*. 1994;28(2):112–116.
4. Anandacoomarasamy A, Barnsley L. Long term outcomes of inversion ankle injuries. *Br J Sports Med*. 2005;39(3):e14. doi: 10.1136/bjsm.2004.011767.
5. Verhagen RAW, de Keizer G, van Dijk CN. Long-term follow-up of inversion trauma of the ankle. *Arch Orthop Trauma Surg*. 1995;114(2):92–96.
6. Schaap GR, de Keizer G, Marti K. Inversion trauma of the ankle. *Arch Orthop Trauma Surg*. 1989;108(5):273–275.
7. Freeman MA, Dean MR, Hanham IW. The etiology and prevention of functional instability of the foot. *J Bone Joint Surg Br*. 1965;47(4):678–685.
8. Konradsen L, Ravn JB. Ankle instability caused by prolonged peroneal reaction time. *Acta Orthop Scand*. 1990;61(5):388–390.
9. Arnold BL, Docherty CL. Low-load eversion force sense, self-reported ankle instability, and frequency of giving way. *J Athl Train*. 2006;41(3):233–238.
10. Garn SN, Newton RA. Kinesthetic awareness in subjects with multiple ankle sprains. *Phys Ther*. 1988;68(11):1667–1671.
11. Tropp H, Ekstrand J, Gillquist J. Stabilometry in functional instability of the ankle and its value in predicting injury. *Med Sci Sports Exerc*. 1984;16(1):64–66.
12. Kaminski TW, Hartsell HD. Factors contributing to chronic ankle instability: a strength perspective. *J Athl Train*. 2002;37(4):394–405.
13. Bernier JN, Perrin DH, Rijke A. Effect of unilateral functional instability of the ankle on postural sway and inversion and eversion strength. *J Athl Train*. 1997;32(3):226–232.
14. Kaminski TW, Perrin DH, Gansneder BM. Eversion strength analysis of uninjured and functionally unstable ankles. *J Athl Train*. 1999;34(3):239–245.
15. Lentell G, Baas B, Lopez D, McGuire L, Sarrels M, Snyder P. The contributions of proprioceptive deficits, muscle function, and anatomic laxity to functional instability of the ankle. *J Orthop Sports Phys Ther*. 1995;21(4):206–215.
16. Lentell GL, Katzman LL, Walters MR. The relationship between muscle function and ankle stability. *J Orthop Sports Phys Ther*. 1990;11(12):605–611.
17. Ryan L. Mechanical stability, muscle strength, and proprioception in the functionally unstable ankle. *Aust Physiother*. 1994;40(1):41–47.
18. Wilkerson GB, Pinerola JJ, Caturano RW. Invertor vs evertor peak torque and power deficiencies associated with lateral ankle ligament injury. *J Orthop Sports Phys Ther*. 1997;26(2):78–86.
19. Sekir U, Yildiz Y, Hazneci B, Ors F, Aydin T. Effect of isokinetic training on strength, functionality and proprioception in athletes with functional ankle instability. *Knee Surg Sports Traumatol Arthrosc*. 2007;15(5):654–664.

20. Tropp H. Pronator muscle weakness in functional instability of the ankle joint. *Int J Sports Med*. 1986;7(5):291–294.

21. American Academy of Orthopaedic Surgeons. *Joint Motion: Method of Measuring and Recording*. Chicago, IL: American Academy of Orthopaedic Surgeons; 1965.

22. Bosien WR, Staples OS, Russell SW. Residual disability following acute ankle sprains. *J Bone Joint Surg Am*. 1955;37(6):1237–1243.

23. Staples OS. Result study of ruptures of lateral ligaments of the ankle. *Clin Orthop Rel Res*. 1972;85:50–58.

24. Staples OS. Ruptures of the fibular collateral ligaments of the ankle: result study of immediate surgical treatment. *J Bone Joint Surg Am*. 1975;57(1):101–107.

25. Pontaga I. Ankle joint evertor–invertor muscle torque ratio decrease due to recurrent lateral ligament sprains. *Clin Biomech (Bristol, Avon)*. 2004;19(7):760–762.

26. Porter GK Jr, Kaminski TW, Hatzel B, Powers ME, Horodyski MB. An examination of the stretch-shortening cycle of the dorsiflexors and evertors in uninjured and functionally unstable ankles. *J Athl Train*. 2002;37(4):494–500.

27. McKnight CM, Armstrong CW. The role of ankle strength in functional ankle instability. *J Sport Rehabil*. 1997;6(1):21–29.

28. Willems T, Witvrouw E, Verstuyft J, Vaes P, De Clercq D. Proprioception and muscle strength in subjects with a history of ankle sprains and chronic instability. *J Athl Train*. 2002; 37(4):487–493.

29. Hubbard TJ, Kramer LC, Denegar CR, Hertel J. Contributing factors to chronic ankle instability. *Foot Ankle Int*. 2007;28(3): 343–354.

30. Arnold BL, De La Motte SJ, Linens SW, Ross SE. Functional ankle instability is associated with balance impairments: a meta-analysis. *Med Sci Sports Exerc*. 2009;41(5):1048–1062.

31. Cook TD, Campbell DT. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. 1st ed. Boston, MA: Houghton Mifflin; 1979.

32. Hedges LV. Fixed effects models. In: Cooper H, Hedges LV, eds. *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994:285–299.

33. Shadish WR, Haddock CK. Combining estimates of effect size. In: Cooper H, Hedges LV, eds. *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994:261–281.

34. Hedges LV, Olkin I. Diagnostic procedures for research synthesis models. *Statistical Methods for Meta-Analysis*. San Diego, CA: Academic Press; 1985:247–263.

35. Schrader JW. *Concentric and Eccentric Muscle Function in Normal and Chronically Sprained Ankles: Prevention Implications* [dissertation]. Bloomington: Indiana University; 1994.

36. Kendall HO, McCreary EK. *Muscles: Testing and Function*. Baltimore, MD: Williams & Wilkins; 1983.

37. Sterne JAC, Egger M. Regression methods to detect publication bias and other bias in meta-analysis. In: Rothstein HR, Sutton AJ, Borenstein M, eds. *Publication Bias in Meta-analysis: Prevention, Assessment, and Adjustments*. Chichester, United Kingdom: John Wiley & Sons; 2005:99–110.

38. Duval S. The trim and fill method. In: Rothstein HR, Sutton AJ, Borenstein M, eds. *Publication Bias in Meta-Analysis: Prevention, Assessment, and Adjustments*. Chichester, United Kingdom: John Wiley & Sons; 2005:127–144.

39. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109): 629–634.

40. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther*. 2003;83(8):713–721.

41. Rosenthal MC. The fugitive literature. In: Cooper H, Hedges LV, eds. *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994:85–94.

42. Cohen J. The *t* test for means. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988:567.

43. Amaral De Noronha M, Borges NG Jr. Lateral ankle sprain: isokinetic test reliability and comparison between invertors and evertors. *Clin Biomech (Bristol, Avon)*. 2004;19(8):868–871.

44. Aydog E, Aydog ST, Cakci A, Doral MN. Reliability of isokinetic ankle inversion- and eversion-strength measurement in neutral foot position, using the Biodex dynamometer. *Knee Surg Sports Traumatol Arthrosc*. 2004;12(5):478–481.

45. Klein P, Mattys S, Rooze M. Moment arm length variations of selected muscles acting on talocrural and subtalar joints during movement: an in vitro study. *J Biomech*. 1996;29(1):21–30.