Raising the standards of the calf-raise test: A systematic review

Kim Hébert-Losiera, Richard J. Newsham-Westa, Anthony G. Schneiders

Abstract and keywords

Summary The calf-raise test is used by clinicians and researchers in sports medicine to assess properties of the calf muscle-tendon unit. The test generally involves repetitive concentric-eccentric muscle action of the plantar-flexors in unipedal stance and is quantified by the number of raises performed. Although the calf-raise test appears to have acceptable reliability and face validity, and is commonly used for medical assessment and rehabilitation of injuries, no universally acceptable test parameters have been published to date. A systematic review of the existing literature was conducted to investigate the consistency as well as universal acceptance of the evaluation purposes, test parameters, outcome measurements, and psychometric properties of the calf-raise test. Nine electronic databases were searched during the period May 30th to September 21st 2008. Forty-nine articles met the inclusion criteria and were quality assessed. Information on study characteristics and calf-raise test parameters, as well as quantitative data, were extracted; tabulated; and statistically analysed. The average quality score of the reviewed articles was 70.4%±12.2% (range 44-90%). Articles provided various test parameters; however, a consensus was not ascertained. Key testing parameters varied, were often unstated, and few studies reported reliability or validity values, including sensitivity and specificity. No definitive normative values could be established and the utility of the test in subjects with pathologies remained unclear. Although adapted for use in several disciplines and traditionally recommended for clinical assessment, there is no uniform description of the calf-raise test in the literature. Further investigation is recommended to ensure consistent use and interpretation of the test by researchers and clinicians.

Keywords Review; Lower Extremity; Physical Examination; Sports Medicine; Musculoskeletal System; Calf-Raise Test

Introduction

Clinicians and researchers in sports science and medicine often use the calf-raise test to assess properties of the calf muscle-tendon unit (MTU) [1-6]. The test was originally developed in the 1940s during the poliomyelitis epidemic to grade and detect plantar-flexor muscle weakness [7-9]. Today's calf-raise test generally involves repetitive concentric-eccentric muscle action of the plantar-flexors in unipedal stance and is quantified by the total number of raises performed [7-9]. Several disciplines have adapted the test for use including neurology [8, 10], gerontology [11-13], cardiology [10, 14-16], orthopaedics [17, 18], and sports medicine [19, 20]. The calf-raise test has traditionally been used to assess various calf MTU properties including endurance, strength, fatigue, function, and performance [1-7, 9, 21-36]. The test has also been employed to assist diagnosis, quantify injury, grade impairment, and measure treatment outcomes of the lower extremity [17, 37-39]. A wide range of administrative protocols are currently available and detail multiple parameters, such as starting position, height of raise, pace of execution, balance support, termination criteria, and outcome measurements.

Normative values are often utilised to develop evidence-based clinical references [40]. In studies that employ the calf-raise test, normative values, such as the number of raises, are often reported and used as clinical reference. The research literature commonly recommends 25 raises as norm clinical performance targets for healthy subjects [9, 34], although higher and lower values have also been suggested [1, 2, 5, 14, 35, 36, 41-46]. Conversely, musculoskeletal assessment textbooks generally recommend lower target values ranging from 7 to 15 raises [22, 26, 28].

In sports medicine, it has been suggested that as high as 30-50% of all sporting injuries are related to overuse tendon disorders [23]. Achilles tendonopathies are considered the most common tendon pathology affecting the lower extremity [21]; accounting for 11% of all running injuries [47] and a reported annual incidence of 7-9% in elite runners [48, 49]. Rehabilitation of these disorders often includes eccentric exercises [50, 51]. Eccentric exercises specific to Achilles tendon disorders utilise a modified form of the calf-raise test within exercise prescription protocols. Since both incorporate similar movements, the calf-raise test is frequently used to determine the treatment effects of the eccentric exercise regime [52]. The test is therefore not only used in the initial assessment of Achilles tendon disorders, but also during rehabilitation to quantify treatment outcomes and to monitor evolution of these conditions [4, 6, 17, 19, 30-33, 37, 53-61].

Although the calf-raise test appears to have acceptable reliability and face validity, and is commonly used for medical assessment and rehabilitation of injuries, there are no universally acceptable test parameters to guide clinicians in its administration and interpretation. The aim of this paper is therefore to systematically review the existing published literature relevant to the calf-raise test and to identify the consistency and acceptance of the test's evaluation purposes, test parameters, outcome measurements, available normative values, and reliability and validity values. The paper primarily explores the calf-raise test in an orthopaedic and sports medicine context, with particular clinical consideration given to Achilles tendon pathologies. Investigating this test is vital to promote its uniform description, comprehension, utilisation, interpretation, and standardisation in clinical practice.

Methods

Search strategy

Nine electronic databases were searched on May 30th and monitored until September 21st 2008: Ovid MEDLINE (1950-2008), Scopus (1841-2008), ISI Web of Science (1900-2008), SPORTDiscus (1800-2008), EMBASE (1988-2008), AMED (1985-2008), CINAHL (1981-2008), PEDro, and The Cochrane Library (1991-2008). The calf-raise test has been previously identified by various combinations of the terms *calf* or *heel* or *toe* combined with *raise* or *rise* or *lift*. Therefore, the keywords; calf raise, calf rise, calf lift, heel raise, heel rise, heel lift, toe raise, toe rise, and toe lift were combined by the Boolean OR; as were the truncated keywords evaluation (eval$) and test (test$). These two searches were combined with no limits applied.

Inclusion and exclusion criteria

Articles were included if they addressed evaluation, testing, or assessment of function, endurance, strength, or performance of the calf MTU. Papers were excluded if they had no full-text (complete) versions, such as conference abstracts or dissertations; had no statistical analyses; or referred to gait analysis, orthotic devices, isokinetic measurements, or the single heel-raise test (performance of a single raise) rather than the calf-raise test (performance of repeated raises).

Article selection

Potentially identifiable information, such as authors, affiliations, and source of publication, were removed from all articles to assure blinding of the reviewers and to reduce bias. The article selection process was initiated by excluding duplicates retrieved from the electronic search based on matching titles (see Figure 1). The remaining articles were screened with respect to the inclusion-exclusion criteria by two independent reviewers (KH-L, RN-W), with foreign language articles being translated if not accessible in English. Screening results were compared and if no consensus was reached, a third independent reviewer was consulted (AS). Subsequently, abstracts and full-text articles were sequentially screened using the same screening procedure as for titles. A selection of sports medicine and orthopaedic journals and the reference lists of all retrieved full-text articles were hand-searched (see Figure 1).

Quality assessment

A modified version of the Downs and Black Quality Index [62] was employed to quality assess the articles that met the inclusion criteria. The original index has demonstrated high internal consistency (Kuder-Richardson 20: 0.89), good test-retest (r =0.88) and inter-rater reliability (r= 0.75), and high correlations with other validated quality assessment instruments (r=0.90) used for non-randomised studies [62]. The reliability and validity of the modified quality assessment index used in this review were not assessed since other researchers have successfully applied similar modified versions of the initial quality index [63, 64].

For the purpose of this review, as performed in a prior systematic review [63], eight questions (8, 9, 13, 17, 19, 24, 26, and 27) from the original 27 itemed index were excluded since they were not relevant to non-randomised control trials. The category "not applicable" was added to questions 4, 14, 15, and 23, as the questions were only pertinent to intervention type studies. Age, sex, physical activity level, height, weight, prior lower limb injury, dominance, and health condition were defined as the principal confounders for questions 5 and 25 with the first three considered core confounders. These confounders were selected since they have been documented a priori as confounding factors for the calf-raise test in previous investigated literature [7, 44]. Question 25 was answered *yes* if there was adjustment for any of the confounders in the analysis. Question 5 scored: *yes,* if defining three core confounders and three or more other confounders; *partially,* if defining three core confounders and one or two other confounders; and *no,* if defining three core confounders but no other confounders, or if defining less than two core confounders. The final quality assessment scores were expressed as percentages; higher percentages indicating articles of higher quality. Two reviewers (KH-L, RN-W) independently assessed the methodological quality of the articles (*n*=49). If independent quality scores were within 10% difference of each other, the average score of the two reviewers was used. If scores differed by more than 10% between the reviewers, quality scoring was discussed and, if required, a third independent reviewer was consulted (AS). The 10% value was determined a priori since this value represented 1 quality assessment criterion.

Data extraction

To collect information from the articles included in this review, a data sheet was designed to extract study characteristics and calf-raise test parameters. The latter included (where available): evaluation purposes; outcome measurements; normative values; and reliability and validity values, including sensitivity and specificity. To determine optimal standardisation of the calf-raise test, six key parameters were identified by the researchers prior to the data extraction process. These were selected for frequency of use in clinical practice and in previously investigated literature: start position of the tested ankle and knee; height of raise; pace of execution; balance support; and test termination criteria. One reviewer (KH-L) extracted the data which was then independently verified by a second reviewer (RN-W). Data was tabulated in Microsoft® Office Excel 2007 and descriptive statistics were calculated, including means, weighted means of normative values, standard deviations, and ranges.

Results

Search result and methodological quality

The electronic database search generated 520 articles related to the calf-raise test (see Figure 1). Forty-nine articles met the inclusion criteria and were quality assessed, resulting in an average quality score and standard deviation of 70.4±12.2% (range 44-90%) (see Table 1). Consensus was reached between the independent reviewers for all articles.

Figure 1*: Search strategy and article selection process*

Table 1*: Article characteristics and calf-raise test properties*

Population and sample size

A summary of the articles' (*n*=49) populations and sample sizes is presented in Table1. Twenty-five articles comprised a total of 1422 normal subjects. Seven of these articles described subjects as being athletic, sedentary, or military. The average sample size of these articles (*n*=25) was 55±62 (range 8-203).

Thirty-four articles contained a total of 1524 subjects affected by either orthopaedic or medical pathologies. Their mean sample size was 45±36 (range 8-148). More than half the articles (*n*=18) included Achilles tendon disorders, ranging from acute tendinopathies to complete tendon ruptures. Other orthopaedic conditions included ankle injuries, patellar tendinopathies, chronic plantar heel pain, medial tibial stress syndrome, anterior cruciate ligament reconstructions, and posterior cruciate ligament tibial avulsion fractures. The medical conditions investigated comprised chronic heart failure, chronic venous insufficiency, stroke, and kidney failure patients.

Calf-raise test

Purposes and outcome measurements

Reasons for utilising the calf-raise test varied substantially between articles (see Table 1) and included assessing endurance, strength, fatigue, function, performance, or multiple calf MTU properties. Some studies reported up to three purposes for employing the test (I, II, III in Table 1). The multitude of assessment purposes were mirrored by the multiple outcome measurements reported (see Table 2), with some studies documenting up to four outcome measurements (I to IV in Table 1). These included the number of raises; work performed; reason for stopping; pain; time to complete; plantar-flexion range of motion (ROM); heel-raise or lower limb symmetry indexes; functional scores; and torque, force or impulse. Assessing endurance and reporting the number of raises were respectively the most frequently cited purpose and outcome measurement for the test.

Key testing parameters

The six identified testing parameters for the calf-raise test were inconsistently reported and defined in numerous ways (see Table 2) despite being used to describe and standardise administration protocols. Although frequently employed in clinical practice, only 14 articles reported all six key testing parameters and five failed to identify any (see Table 1). Half the articles reviewed did not describe the starting position of the knee; and a third did not document the termination criteria for the calf-raise test (see not specified in Table 2).

***Table 2**: Key testing parameters and cited frequency in articles*

Reliability and validity values

Reliability values for the calf-raise test were reported as part of the methodology in 13 articles; were cited from other articles in 10; and, not provided in 26 (see reliability in Table 1). Only nine addressed validity, which included sensitivity and specificity (see validity in Table 1). The reported test-retest reliability values ranged from moderate to excellent, and were, where discernable, either intra-rater (same evaluator) or inter-rater (two separate evaluators) measures of test-retest reliability (see Table 1). Intraclass correlation coefficient values ranged from 0.57 to 0.99; and Pearson's correlation coefficient (r), from 0.56 to 0.98.

Normative values

No definite normative outcome values could be determined. Pooling norms was difficult due to the heterogeneity of the articles' underlying constructs and the variety of structured calf-raise tests used. Therefore, descriptive statistics using weighted means were used to report the results from the available normative values. When the total number of raises performed was compared, regardless of specific testing parameters, the highest number achieved by a healthy subject was 120, and 54 by a subject with pathology. The weighted means and standard deviations of the total number of raises achieved by healthy subjects and by subjects with pathologies were 27.9±11.1 (range 2.7-68.0) and 19.1±11.4 (range 6.4-53.0), respectively.

To compare performance of similar calf-raise tests, articles using the most frequently reported test parameters for starting position of the knee (extended), pace of execution (60 raises/minutes), and outcome measurement (number of raises) were extracted (see Table 2). Eleven articles [2, 5, 9, 11, 13, 16, 41, 46, 53, 65, 66] were subsequently compared. Seven of these included healthy subjects (*n*=572), while six included subjects with pathologies (*n*=176). The weighted mean and standard deviation of the number of raises performed was 25.3±13.9 (range 2.7-68.0) in healthy subjects; and less in subjects with pathologies, with a weighted mean and standard deviation of 17.6±7.4 (range 6.4-26.0). In the latter, subjects with Achilles tendon disorders were not represented.

Discussion

The purpose of this review was to identify the consistency and acceptance of the calf-raise test properties. Despite the apparent face validity and acceptable reliability of the calf-raise test, no consensus was found in relation to the purpose, the optimal test parameters, the adequate outcome measurements, or the appropriate normative values associated to this test (see Table 2).

The calf-raise test is reported to assess calf MTU properties, and its clinical use has been consistently recommended [9, 22, 26, 28, 67]. Some authors have advocated the calf-raise test as the primary method of assessment for plantar-flexion function [9, 67]. However, despite the frequent use of the test by clinicians and researchers, no consensus was found in relation to the underlying properties of the calf-raise test (see Table 2). The test is most frequently used to evaluate endurance of the calf MTU; however, supporting anatomical or physiological evidence is limited. Analogously, the most commonly reported outcome measurement is the number of raises performed, although this is not supported by any scientific rationale. The preference for using the number of raises as prime outcome measurement may be attributed to its practical and “user-friendly” clinical application. Although this may encourage a common language amongst researchers and clinicians, valid justification for using the number of raises as outcome measurement is lacking.

Reliability of a test is crucial to ensure replication and comparison of scientific and clinical results [68]. Over half of the reviewed literature failed to report the reliability of the calf-raise test (see Table 1). Reliability can be increased by standardising administrative protocols, which reduces the variation of test parameters [69]. The importance of standardising parameters of the calf-raise test such as body position, height of raise, pace of execution, and termination criteria, has been strongly advocated [1, 4, 6, 10, 13-15, 34, 37]; however, these same parameters were often not monitored or described (see Table 2). In articles that did monitor calf-raise test parameters, complex devices were typically employed, which included metronomes, light beam sensors, linear encoders, heel counters, and perpendicularly connected parallel uprights. Utilisation of a standardised apparatus assists in monitoring, quantifying, and replicating a test; however, the apparatus itself needs to be simple and reliable, particularly if to be used clinically [69]. For the calf-raise test, two devices [5, 70] were modified and implemented in ten articles [2, 17, 30, 35, 42, 56, 57, 66, 71, 72], but only one study [2] reported reliability values. Consequently, caution should be exercised when interpreting the results of some studies and consideration should be given to the fact that calf-raise test monitoring devices used in research may not be clinically accessible due to associated costs and/or availability.

There is some evidence that demographic, anthropometric, and ethnic factors influence performance of the calf-raise test [15, 41-44, 46]. However, the reviewed literature provided no consensus on the precise significance, contribution, or effect of these factors on calf-raise test performance. For example, when considering the influence of gender (demographic factor) on the number of raises (performance factor), Maurer et al. [43] found no significant difference between the number of raises performed by females compared to males in children (7-9 yrs). In contrast, Jan et al. [41] observed a significantly lower number of raises performed by females versus males in adults (21-80 yrs of age). Although the respective populations' ages may explain the differing influence of gender on the number of raises achieved, it remains unclear which specific factors and variables need to be considered when documenting performance of the calf-raise test, and to which extent they affect test performance.

Normative outcome values may be implemented as evidence-based references by clinicians [40]. However, the reviewed literature provided no definitive normative values for the calf-raise test. Comparing tests and pooling values between articles proved difficult due to the heterogeneity and inconsistency of the article constructs (aims and population), as well as the type of calf-raise tests used (administrative protocols, test parameters, and outcome measurements). Descriptive statistical analysis of the available outcome measurements documented by 25 articles involving healthy subjects and 34 articles including subjects with pathology resulted in weighted means and standard deviations of 27.9±11.1 and 19.1±11.4 raises respectively. Although these results derive from a variety of different calf-raise test variations, they are consistent with the majority of previously suggested clinical norms [9, 42, 44, 46]. These have been reported as: 25 raises for the general population [9]; 22 raises for males and 17 for females [44]; more raises for an active versus sedentary population [42, 44, 46]; and less raises with increasing age [15, 41, 43]. The standard reported by Lunsford and Perry [9] of 25 raises for a general healthy population was cited by many of the reviewed articles (*n*=15) and was often used as comparison criterion. Lunsford and Perry investigated 203 healthy subjects with the aims of refining the calf-raise test, documenting the number of raises performed, and providing standardised clinical references. However, when reviewed, the study scored poorly on the quality assessment criteria and did not report reliability of the employed calf-raise test. These authors investigated normal subjects, and only provided references for healthy populations. Although this study is considered as seminal in the investigation of a clinically applicable calf-raise test, updating norms for both subjects with and without pathology using a standard, revised, and reliable clinical calf-raise test is required. Until universally accepted standardisation of the calf-raise test is achieved and used to determine normative values in healthy individuals and in individuals with pathology, clinicians can only rely on level 5 evidence, as described by the Centre for Evidence Based Medicine [73].

Differential diagnosis can be defined as distinguishing a condition from others presenting similar characteristics [74]. Recognising a test's sensitivity and specificity values may help to establish an appropriate diagnosis in the case of underlying pathology. The documentation of validity, including sensitivity and specificity, values for the calf-raise test was limited (see Table 1). From the pooling of the available normative values, the calf-raise test appears to have discriminative validity when comparing healthy subjects to subjects with pathology, but this necessitates further investigation. Additionally, sensitivity and specificity values are only relevant to populations in which these values have been specifically investigated and documented. For instance, the test has demonstrated the ability to detect endurance and functional deficits in subjects with medial tibial stress syndrome [2], ankle injuries [39], and chronic heart failure [14]. However, these abilities remain questionable in subjects with Achilles[4] or patellar tendinopathies [65], with other conditions yet to be explored.

Achilles tendon pathologies affect the integrity of the calf MTU and the entire function of the lower extremity [32, 52]. The calf-raise test is considered to be an important tool in assessing Achilles tendon disorders [52], which is reflected by the frequent inclusion of subjects with these pathologies in the literature pertaining to the test. However, the relevance of the test in evaluating Achilles tendinopathies is still debated, with some authors questioning whether the test should be used at all [4]. Validating the use of the calf-raise test for healthy subjects and subsequently for Achilles tendon disorders are important steps in resolving the existing debate. This may result in the test being justifiably employed to assist diagnosis and to monitor the natural history and treatment efficacy of Achilles tendon pathologies.

The calf-raise test is not a MeSH term and therefore, our systematic review required the use of a less efficient keyword search strategy. To counter this limitation, our review included an expansive selection of databases which was supplemented by thorough hand-searching. Furthermore, while an all-inclusive approach was implemented to ensure the review was representative of all published literature relating to the calf-raise test, it is acknowledged that the exclusion of non-published data may be considered a limitation of the study. A further possible limitation of the all-inclusive approach is that all articles meeting the inclusion criteria were analysed, despite many scoring low in the quality assessment. Lower quality scores were commonly associated with: inadequate description of the main outcome measurements; omission of factors influencing outcome measurements in the analysis; and the failure to adequately address external validity, such as population source, representativeness, or recruitment.

Conclusion

A systematic review of the literature was conducted to investigate calf-raise test properties. Key testing parameters varied between articles and were often unstated, with few studies reporting reliability or validity values, including sensitivity and specificity. No definitive normative values could be determined and the utility of the test in subjects with pathology remained unclear. Although adapted for use in several disciplines and traditionally recommended as clinical assessment and rehabilitation tool, there is no uniform description of the calf-raise test. The precise evaluative purpose of the calf-raise test (endurance, strength, fatigue, function, or performance) should be identified and validated by anato-physiological evidence in future studies to ensure a consistent definition, comprehension, utilisation, standardisation, and interpretation of the calf-raise test and derived data. Development of a standardised calf-raise test is vital for researchers and clinicians in sports medicine.

Practical implications

- The calf-raise test assesses the properties of the calf-muscle tendon unit through repetitive concentric-eccentric muscle action of the plantar-flexors in unipedal stance.
- No consistent evaluation purpose; test parameters; outcome measurements; normative values; or reliability and validity values are currently documented for the calf-raise test.
- Clinicians can only rely on level 5 evidence until universally accepted standardisation of the calf-raise test is achieved and used to determine normative values in healthy individuals and in individuals with pathology.

Acknowledgments

The authors acknowledge the support of the Centre for Physiotherapy Research in this review. No financial assistance was provided for this project.

References

1. Haber M, Golan E, Azoulay L, et al. Reliability of a device measuring triceps surae muscle fatigability. *Br J Sports Med* 2004; 38(2):163-7.

2. Madeley LT, Munteanu SE, Bonanno DR. Endurance of the ankle joint plantar flexor muscles in athletes with medial tibial stress syndrome: a case-control study. *J Sci Med Sport* 2007; 10(6):356-62.

3. Möller M, Lind K, Styf J, et al. The reliability of isokinetic testing of the ankle joint and a heel-raise test for endurance. *Knee Surg Sports Traumatol Arthrosc* 2005; 13(1):60-71.

4. Neeter C, Thomee R, Silbernagel KG, et al. Iontophoresis with or without dexamethazone in the treatment of acute Achilles tendon pain. *Scand J Med Sci Sports* 2003; 13(6):376-82.

5. Ross MD, Fontenot EG. Test-retest reliability of the standing heel-rise test. *J Sport Rehabil* 2000; 9(2):117-23.

6. Silbernagel KG, Thomee R, Thomee P, et al. Eccentric overload training for patients with chronic Achilles tendon pain - a randomised controlled study with reliability testing of the evaluation methods. *Scand J Med Sci Sports* 2001; 11(4):197-206.

7. Beasley WC. Quantitative muscle testing: principles and applications to research and clinical services. *Arch Phys Med Rehabil* 1961; 42:398-425.

8. Florence JM, Pandya S, King WM, et al. Intrarater reliability of manual muscle test (Medical Research Council scale) grades in Duchenne's muscular dystrophy. 1992; 72(2):115-22.

9. Lunsford BR, Perry J. The standing heel-rise test for ankle plantar flexion: criterion for normal. *Phys Ther* 1995; 75(8):694-8.

10. Svantesson U, Österberg U, Grimby G, et al. The standing heel-rise test in patients with upper motor neuron lesion due to stroke. *Scand J Rehabil Med* 1998; 30(2):73-80.

11. Cider A, Schaufelberger M, Sunnerhagen KS, et al. Hydrotherapy - a new approach to improve function in the older patient with chronic heart failure. *Eur J Heart Fail* 2003; 5(4):527-35.

12. Ferri A, Narici M, Grassi B, et al. Neuromuscular recovery after a strength training session in elderly people. *Eur J Appl Physiol* 2006; 97(3):272-9.

13. Li Y, Devault CN, Van Oteghen S. Effects of extended Tai Chi intervention on balance and selected motor functions of the elderly. *Am J Chin Med* 2007; 35(3):383-91.

14. Cider A, Carlsson S, Arvidsson C, et al. Reliability of clinical muscular endurance tests in patients with chronic heart failure. *Eur J Cardiovasc Nurs* 2006; 5(2):122-6.

15. van Uden CJT, van der Vleuten CJM, Kooloos JGM, et al. Gait and calf muscle endurance in patients with chronic venous insufficiency. *Clin Rehabil* 2005; 19(3):339-44.

16. Cider A, Tygesson H, Hedberg M, et al. Peripheral muscle training in patients with clinical signs of heart failure. *Scand J Rehabil Med* 1997; 29(2):121-7.

17. Möller M, Movin T, Granhed H, et al. Acute rupture of tendon Achillis: a prospective randomised study of comparison between surgical and non-surgical treatment. *J Bone Joint Surg Br* 2001; 83(6):843-8.

18. Paoloni JA, Appleyard RC, Murrell GA. The Orthopaedic Research Institute-Ankle Strength Testing System: inter-rater and intra-rater reliability testing. *Foot Ankle Int* 2002; 23(2):112-7.

19. Robinson JM, Cook JL, Purdam C, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. *Br J Sports Med* 2001; 35(5):335-41.

20. Dennis RJ, Finch CF, Elliott BC, et al. The reliability of musculoskeletal screening tests used in cricket. *Phys Ther* 2008; 9(1):25-33.

21. Braddom RL, Peterson AT. *Handbook of Physical Medicine and Rehabilitation*, 2nd ed. Philadelphia. Saunders. 2004.

22. Clarkson HM. *Musculoskeletal Assessment : Joint Range of Motion and Manual Muscle Strength*, 2nd ed. Philadelphia. Lippincott Williams & Wilkins. 2000.

23. DiGiovanni CW, Greisberg J. *Overuse Achilles Tendon Disorders*, Chap 16 in *Foot and Ankle: Core Knowledge in Orthopaedics,* 1st ed. Philadelphia. Elsevier Mosby. 2007. p.200-208.

24. Gross J, Fetto J, Rosen E. *Musculoskeletal Examination*. Cambridge Mass. Blackwell Science. 1996.

25. Handcock PJ, Knight B. *Field Testing Manual*. Wellington. New Zealand Sport Science and Technology Board in conjunction with Coaching New Zealand. 1994.

26. Magee DJ. *Orthopedic Physical Assessment*, 5th ed. Saint Louis. Saunders Elsevier. 2008.

27. Palmer ML, Epler ME. *Clinical Assessment Procedures in Physical Therapy*. Philadelphia. Lippincott. 1990.

28. Palmer ML, Epler MF. *Fundamentals of Musculoskeletal Assessment Techniques*, 2nd ed. Philadelphia. Lippincott. 1998.

29. Petty NJ, Moore AP. *Neuromusculoskeletal Examination and Assessment : A Handbook for Therapists*, 2nd ed. Edinburgh. Churchill Livingstone. 2001.

30. Möller M, Lind K, Movin T, et al. Calf muscle function after Achilles tendon rupture: a prospective, randomised study comparing surgical and non-surgical treatment. *Scand J Med Sci Sports* 2002; 12(1):9-16.

31. Silbernagel KG, Thomee R, Eriksson BI, et al. Full symptomatic recovery does not ensure full recovery of muscle-tendon function in patients with Achilles tendinopathy. *Br J Sports Med* 2007; 41(4):276-80.

32. Silbernagel KG, Gustavsson A, Thomee R, et al. Evaluation of lower leg function in patients with Achilles tendinopathy. *Knee Surg Sports Traumatol Arthrosc* 2006; 14(11):1207-17.

33. Silbernagel KG, Thomee R, Eriksson BI, et al. Continued sports activity, using a pain-monitoring model, during rehabilitation in patients with Achilles tendinopathy: a randomized controlled study. *Am J Sports Med* 2007; 35(6):897-906.

34. Svantesson U, Österberg U, Thomee R, et al. Muscle fatigue in a standing heel-rise test. *Scand J Rehabil Med* 1998; 30(2):67-72.

35. Buchgraber A, Passler HH. Percutaneous repair of Achilles tendon rupture. Immobilization versus functional postoperative treatment. *Clin Orthop Relat Res* 1997(341):113-22.

36. Österberg U, Svantesson U, Takahashi H, et al. Torque, work and EMG development in a heel-rise test. *Clin Biomech* 1998; 13(4-5):344-50.

37. Schepull T, Kvist J, Andersson C, et al. Mechanical properties during healing of Achilles tendon ruptures to predict final outcome: a pilot Roentgen stereophotogrammetric analysis in 10 patients. *BMC Musculoskelet Disord* 2007; 8(1):116-26.

38. Segal RL, Song AW. Nonuniform activity of human calf muscles during an exercise task. *Arch Phys Med Rehabil* 2005; 86(10):2013-7.

39. Kaikkonen A, Kannus P, Jarvinen M. A performance test protocol and scoring scale for the evaluation of ankle injuries. *Am J Sports Med* 1994; 22(4):462-9.

40. Herbert R. *Practical Evidence-Based Physiotherapy*. Edinburgh. Elsevier Butterworth-Heinemann. 2005.

41. Jan M-H, Chai H-M, Lin Y-F, et al. Effects of age and sex on the results of an ankle plantar-flexor manual muscle test. *Phys Ther* 2005; 85(10):1078-84.

42. Irving DB, Cook JL, Young MA, et al. Obesity and pronated foot type may increase the risk of chronic plantar heel pain: a matched case-control study. *BMC Musculoskelet Disord* 2007; 8(1):41-8.

43. Maurer C, Finley A, Martel J, et al. Ankle plantarflexor strength and endurance in 7-9 year old children as measured by the standing single leg heel-rise test. *Phys Occup Ther Pediatr* 2007; 27(3):37-54.

44. Sunnerhagen KS, Hedberg M, Henning G-B, et al. Muscle performance in an urban population sample of 40- to 79-year-old men and women. *Scand J Rehabil Med* 2000; 32(4):159-67.

45. Nakao M, Inoue Y, Matsuhita K, et al. Simple methods for measuring leg muscular endurance. *J Sports Med Phys Fitness* 1986; 26(3):285-91.

46. Thomas KS. Functional élevè performance as it applied to heel-rises in performance-level collegiate dancers. *J Dance Med Sci* 2003; 7(4):115-20.

47. Brotzman SB, Wilk KE. *Handbook of Orthopaedic Rehabilitation*, 2nd ed. Philadelphia PA. Mosby. 2007.

48. Johansson C. Injuries in elite orienteers. *Am J Sports Med* 1986; 14(5):410-415.

49. Lysholm J, Wiklander J. Injuries in runners. *Am J Sports Med* 1987; 15(2):168-171.

50. Alfredson H, Pietila T, Jonsson P, et al. Heavy-load eccentric calf muscle training for the treatment of chronic achilles tendinosis. *Am J Sports Med* 1998; 26(3):360-366.

51. Mafi N, Lorentzon R, Alfredson H. Superior short-term results with eccentric calf muscle training compared to concentric training in a randomized prospective multicenter study on patients with chronic Achilles tendinosis. *Knee Surg Sports Traumatol Arthrosc* 2001; 9(1):42-47.

52. Kountouris A, Cook J. Rehabilitation of Achilles and patellar tendinopathies. *Best Pract Res Clin Rheumatol* 2007; 21(2):295-316.

53. Brodin E, Ljungman S, Sunnerhagen KS. Rising from a chair: a simple screening test for physical function in predialysis patients. *Scand J Urol Nephrol* 2008; 42(3):293-300.

54. Fortis AP, Dimas A, Lamprakis AA. Repair of Achilles tendon rupture under endoscopic control. *Arthroscopy* 2008; 24(6):683-8.

55. McComis GP, Nawoczenski DA, DeHaven KE. Functional bracing for rupture of the Achilles tendon. Clinical results and analysis of ground-reaction forces and temporal data. *J Bone Joint Surg Am* 1997; 79(12):1799-808.

56. Moberg A, Nordgren B, Solveborn SA. Surgically repaired Achilles tendon ruptures with postoperative mobile ankle cast: a 12-month follow-up study with an isokinetic and a dynamic muscle function test. *Scand J Med Sci Sports* 1992; 2(4):231-3.

57. Mortensen NHM, Skov O, Jensen PE. Early motion of the ankle after operative treatment of a rupture of the Achilles tendon: a prospective, randomized clinical and radiographic study. *J Bone Joint Surg Am* 1999; 81A(7):983-90.

58. Nilsson-Helander K, Sward L, Silbernagel KG, et al. A new surgical method to treat chronic ruptures and reruptures of the Achilles tendon. *Knee Surg Sports Traumatol Arthrosc* 2008; 16(6):614-20.

59. Paavola M, Kannus P, Orava S, et al. Surgical treatment for chronic Achilles tendinopathy: a prospective seven month follow up study. *Br J Sports Med* 2002; 36(3):178-182.

60. Paavola M, Kannus P, Paakkala T, et al. Long-term prognosis of patients with achilles tendinopathy: An observational 8-year follow-up study. *Am J Sports Med* 2000; 28(5):634-42.

61. Weber M, Niemann M, Lanz R, et al. Nonoperative treatment of acute rupture of the Achilles tendon: results of a new protocol and comparison with operative treatment. *Am J Sports Med* 2003; 31(5):685-91.

62. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998; 52(6):377-384.

63. Irving DB, Cook JL, Menz HB. Factors associated with chronic plantar heel pain: a systematic review. *J Sci Med Sport* 2006; 9(1-2):11-22.

64. Lynch ME, Campbell F, Clark AJ, et al. A systematic review of the effect of waiting for treatment for chronic pain. *Pain* 2008; 136(1-2):97-116.

65. Crossley KM, Thancanamootoo K, Metcalf BR, et al. Clinical features of patellar tendinopathy and their implications for rehabilitation. *J Orthop Res* 2007; 25(9):1164-75.

66. Ross MD, Hooten S, Moore D. Lower leg girth and ankle plantar-flexor endurance after anterior cruciate ligament reconstruction. *J Sport Rehabil* 2002; 11(2):128-38.

67. Nilsson G, Nyberg P, Ekdahl C, et al. Performance after surgical treatment of patients with ankle fractures - 14-month follow-up. *Physiother Res Int* 2003; 8(2):69-82.

68. Hopkins WG. Measures of Reliability in Sports Medicine and Science. *Sports Med* 2000; 30(1):1-15.

69. Roebroeck ME, Harlaar J, Lankhorst GJ. The application of generalizability theory to reliability assessment: an illustration using isometric force measurements. 1993; 73(6):386-395.

70. Häggmark T, Liedberg H, Eriksson E, et al. Calf muscle atrophy and muscle function after non-operative vs operative treatment of achilles tendon ruptures. *Orthopedics* 1986; 9(2):160-4.

71. Malliaras P, Cook JL, Kent P. Reduced ankle dorsiflexion range may increase the risk of patellar tendon injury among volleyball players. *J Sci Med Sport* 2006; 9(4):304-9.

72. Young MA, Cook JL, Webster KE. The effect of topical wheatgrass cream on chronic plantar fasciitis: a randomized, double-blind, placebo-controlled trial. *Complement Ther Med* 2006; 14(1):3-9.

73. Straus S, Richardson W, Glasziou P, et al. *Evidence-based Medicine: How to Practice and Teach EBM*, 3rd ed. Edinburgh. Churchill Livingstone. 2005.

74. Webster M. *Merriam Webster's Medical Desk Dictionary*, 2nd ed. Clifton Park. Delmar Cengage Learning. 2002.

75. Gu MO, Jeon MY, Eun Y. [The development & effect of an tailored falls prevention exercise for older adults]. 2006; 36(2):341-52.

76. Nicandri GT, Klineberg EO, Wahl CJ, et al. Treatment of posterior cruciate ligament tibial avulsion fractures through a modified open posterior approach: operative technique and 12- to 48-month outcomes. *J Orthop Trauma* 2008; 22(5):317-24.

77. Shaffer MA, Okerehe E, Esterhai Jr. JL, et al. Effects of immobilization on plantar-flexion torque, fatigue resistance, and functional ability following an ankle fracture. *Phys Ther* 2000; 80(8):769-80.

78. Svantesson U, Carlsson U, Takahashi H, et al. Comparison of muscle and tendon stiffness, jumping ability, muscle strength and fatigue in the plantar flexors. *Scand J Med Sci Sports* 1998; 8(5 Pt 1):252-6.

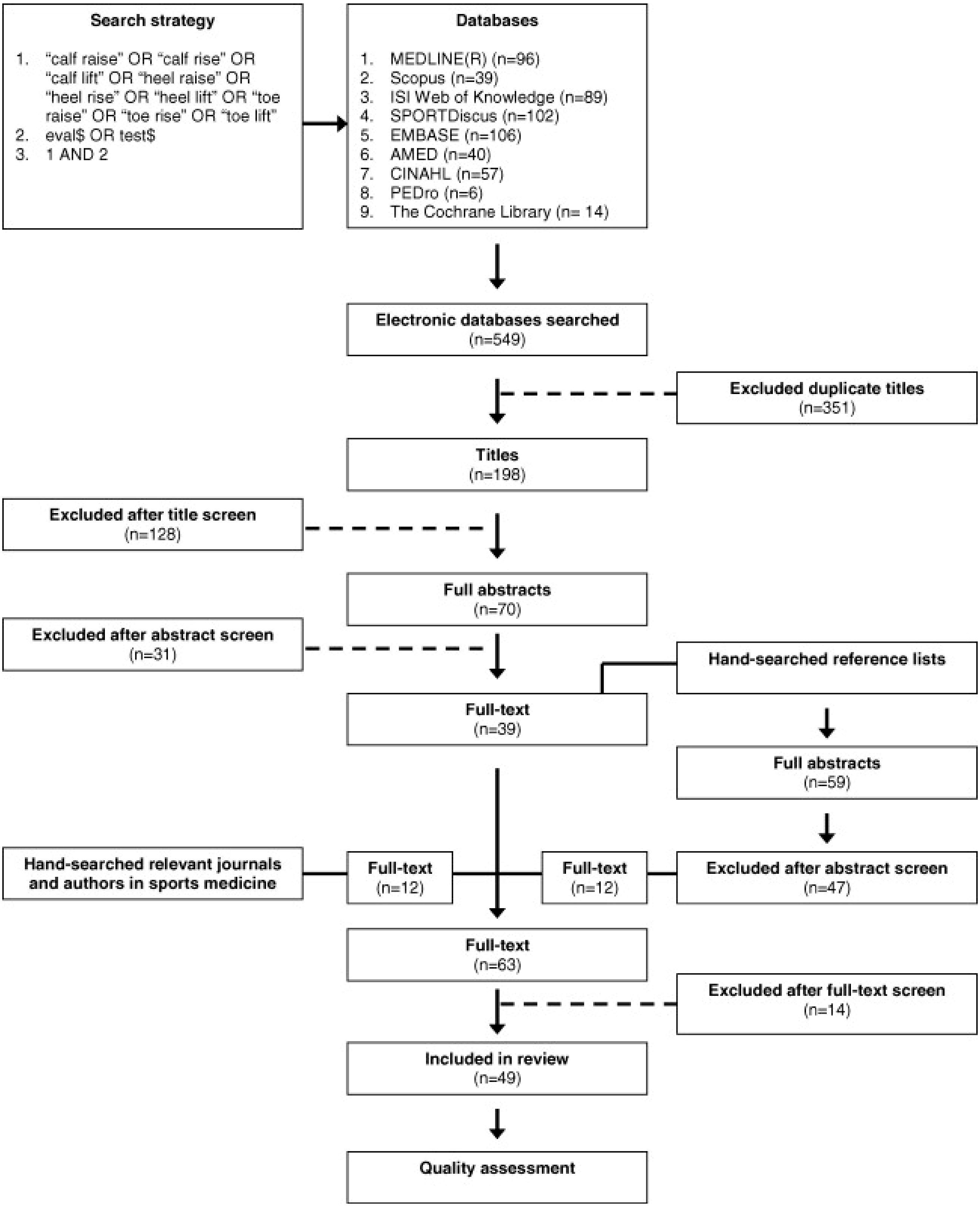

Figure 1: Flow diagram of the search strategy and article selection process

Table 1: Article characteristics and calf-raise test properties

Authors	Quality score	Population[a] and sample size (*n*=)	Purposes[b]			Outcome measurements[c]				Key testing parameters (*n*=6)[d]						Reliability and validity[d]			
																Reliability[e]			Validity[f]
			I	II	III	I	II	III	IV	Ankle	Knee	Height	Pace	Balance	Termination	Intra	Inter	N/S	
Young[72]	90	OC(80)	2	0	0	1	0	0	0	+	-	+	-	-	+	-	-	-	-
Paavola[59]	88	AT(42)	3	0	0	1	0	0	0	+	-	-	+	-	+	-	-	-	-
Maurer[43]	88	N(95)	1	2	0	1	0	0	0	+	+	+	-	+	+	+	+	-	-
Brodin[53]	88	MC(55)	2	0	0	1	0	0	0	+	+	+	+	+	+	-	-	-	-
Nilsson[67]	88	OC(54)	2	0	0	1	0	0	0	+	+	+	-	+	+	-	-	-	-
Fortis[54]	88	AT(20)	1	0	0	1	0	0	0	+	-	+	-	-	+	-	-	-	-
van Uden[15]	84	N(19) MC(15)	1	2	0	1	0	0	0	+	+	+	+	+	-	-	-	-	-
Malliaras[71]	82	OC(113)	2	0	0	1	0	0	0	+	-	-	+	-	-	-	-	-	+
Cider[16]	82	MC(24)	4	0	0	1	0	0	0	+	+	+	+	+	-	-	-	-	-
Paavola[59]	82	AT(83)	5	0	0	1	8	0	0	-	-	-	-	-	-	-	-	-	-
Neeter[4]	82	AT(25)	1	0	0	1	3	4	0	+	-	+	+	+	+	+	-	-	+
Madeley[2]	81	N(30) OC(30)	1	0	0	1	0	0	0	+	+	+	+	+	+	-	-	+	+
Ross[5]	81	N(17)	5	0	0	1	6	0	0	+	+	+	+	+	+	+	-	-	-
Mortensen[57]	80	AT(71)	4	0	0	7	0	0	0	+	-	+	+	-	+	-	-	-	-
Silbernagel[33]	80	AT(38)	1	0	0	2	0	0	0	-	-	-	-	-	-	-	-	-	-
Jan[41]	79	N(180)	2	0	0	1	0	0	0	+	+	+	+	+	-	+	-	-	-
Sunnerhagen[44]	79	N(144)	1	2	5	1	0	0	0	+	-	-	+	+	+	-	-	-	-
Schepull[37]	78	AT(10)	4	5	0	1	7	0	0	+	-	+	+	-	-	-	-	-	-
Silbernagel[6]	75	AT(49)	1	0	0	1	3	4	0	+	+	+	+	+	+	+	-	-	+
Möller[17]	75	AT(112)	1	0	0	1	7	0	0	+	-	+	+	-	+	-	-	-	-
Thomas[46]	75	N(97)	1	0	0	1	0	0	0	+	+	+	+	+	+	-	-	-	-
Nilsson-Helander[58]	75	AT(28)	1	0	0	2	6	0	0	-	-	-	-	-	-	-	-	-	-
Möller[3]	72	N(10)	1	0	0	1	0	0	0	+	+	+	+	+	+	+	-	-	-
Möller[30]	71	AT(112)	1	0	0	1	5	0	0	+	+	+	+	+	+	-	-	-	-
Silbernagel[32]	71	AT(42)	1	0	0	1	2	4	7	+	-	+	+	+	+	-	-	-	+

Irving[42]	71	N(80) OC(80)	1	5	0	1	0	0	0	+	-	+	-	-	-	-	-	-	-
Ross[66]	71	OC(15)	1	0	0	1	0	0	0	+	+	+	+	+	+	-	-	-	-
McComis[55]	71	N(15) AT(15)	1	4	0	1	8	0	9	+	-	-	+	+	+	-	-	-	-
Crossley[65]	69	N(31) OC(27)	1	5	0	1	0	0	0	+	+	+	+	+	+	+	-	-	+
Weber[61]	66	AT(38)	1	0	0	7	0	0	0	+	+	+	+	-	-	-	-	-	-
Cider[14]	65	N(20) MC(20)	1	0	0	1	0	0	0	+	-	+	+	+	-	+	-	-	-
Gu[75]	65	N(59)	2	0	0	1	0	0	0	-	+	+	-	+	+	-	-	-	-
Silbernagel[31]	64	AT(37)	1	0	0	2	0	0	0	-	-	-	-	-	-	-	-	-	-
Nicandri[76]	63	OC(10)	4	0	0	1	0	0	0	-	-	-	-	-	-	-	-	-	-
Svantesson[34]	61	N(10)	3	4	0	1	2	0	0	+	-	+	+	+	+	-	-	-	-
Moberg[56]	61	AT(17)	3	0	0	1	0	0	0	+	-	+	+	-	-	-	-	-	-
Shaffer[77]	61	OC(10)	4	0	0	1	8	0	0	+	+	+	-	+	+	-	-	+	-
Buchgraber[35]	60	N(25) AT(48)	3	0	0	1	2	0	0	+	+	+	+	-	+	-	-	-	-
Cider[11]	60	MC(25)	1	0	0	1	0	0	0	+	+	+	+	+	-	-	-	-	-
Kaikkonen[39]	59	N(100) OC(148)	1	2	3	1	0	0	0	+	-	+	+	+	-	-	-	+	+
Svantesson[10]	59	N(8) MC(8)	1	0	0	1	2	0	0	+	-	+	+	+	-	-	-	-	+
Lunsford[9]	56	N(203)	1	2	4	1	6	0	0	+	+	+	+	+	+	-	-	-	-
Svantesson[78]	56	N(10)	1	0	0	1	0	0	0	+	-	+	+	+	+	-	-	-	-
Häggmark[70]	56	N(10) AT(23)	3	4	0	1	2	0	0	+	-	+	+	-	+	-	-	-	-
Haber[1]	53	N(40)	3	0	0	1	5	0	0	+	+	+	+	+	+	+	-	-	+
Dennis[20]	53	N(10)	1	0	0	1	0	0	0	+	+	+	+	-	+	+	+	-	-
Li[13]	50	N(20)	1	2	0	1	0	0	0	+	+	+	+	-	+	-	-	-	-
Österberg[36]	47	N(10)	4	0	0	1	2	6	9	+	+	+	+	+	+	-	-	-	-
Nakao[45]	44	N(185)	1	0	0	5	0	0	0	+	+	+	+	+	+	-	-	-	-

a N=normal, AT=Achilles tendon pathologies, OC=orthopaedic condition, MC=medical condition

b 0=none stated, 1=endurance; 2=strength, 3=fatigue, 4=function, 5=performance

c 0=none stated, 1=number of raises, 2=work (J), 3=reason for stopping, 4=pain, 5=time to complete, 6=plantar-flexion ROM (º), 7=heel raise or lower limb symmetry indexes, 8=functional scores, 9=torque, force or impulse

d +=reported, -=not reported

e Intra=intra-rater (same evaluator); Inter=inter-rater (two independent evaluators); N/S=not specified (evaluators not specified)

f Validity, including sensitivity and specificity

Table 2: Key testing parameters and citation frequency in articles*		
Parameters	Description	*n*
Ankle starting position	Not specified	6
	0º DF or neutral	31
	10º DF or 10º tilted wedge	11
	Other	1
Knee starting position	Not specified	24
	Extension	25
Height	Not specified	9
	As high as possible	21
	5cm	11
	50% of maximal height	4
	90% of maximal height	2
	Other	2
Pace (raises/min)	Not specified	12
	40	6
	46	2
	50	2
	60	18
	92	1
	120	4
	Other	4
Balance support	Not specified	19
	Fingers	20
	One finger	1
	One hand	6
	Use of wall to maintain or regain balance	3
Outcome measurements*	Number of raises	43
	Work (J)	9
	Reason for stopping	2
	Pain (VAS)	3
	Time to complete	3
	Plantar-flexion ROM (º)	4
	Heel-Raise or Lower Limb Symmetry indexes	5
	Functional scores	3
	Torque, force or impulse	2
Termination criteria*	Not specified	16
	Fatigue or exhaustion	14
	Not proper calf-raise	7
	Participant could not continue	11
	Height not reached	15
	Pace not maintained	6
	Supporting knee flexed	12
	Forward lean or use of upper limbs	8
	Loss of balance	3
	Pain	3
	Time	1