Received: 4 February 2023 | Revised: 10 May 2023 | Accepted: 29 June 2023

DOI: 10.1002/pri.2039

REVIEW

WILEY

Devices to measure calf raise test outcomes: A narrative review

Ma. Roxanne Fernandez[1,2] | Kim Hébert-Losier[1]

[1]Division of Health, Engineering, Computing and Science, Te Huataki Waiora School of Health, University of Waikato, Hamilton, New Zealand

[2]Department of Physical Therapy, College of Rehabilitation Sciences, University of Santo Tomas, Manila, Philippines

Correspondence
Ma. Roxanne Fernandez, University of Waikato Adams Centre for High Performance, 52 Miro Street, Mount Maunganui, Tauranga 3116, New Zealand.
Email: mf182@students.waikato.ac.nz

Abstract

Background: The calf raise test (CRT) is commonly administered without a device in clinics to measure triceps surae muscle function. To standardise and objectively quantify outcomes, researchers use research-grade or customised CRT devices. To incorporate evidence-based practice and apply testing devices effectively in clinics, it is essential to understand their design, applicability, psychometric properties, strengths, and limitations. Therefore, this review identifies, summarises, and critically appraises the CRT devices used in science.

Methods: Four electronic databases were searched in April 2022. Studies that used devices to measure unilateral CRT outcomes (i.e., number of repetitions, work, height) were included.

Results: Thirty-five studies met inclusion, from which seven CRT devices were identified. Linear encoder ($n = 18$) was the most commonly used device, followed by laboratory equipment ($n = 6$) (three-dimensional motion capture and force plate). These measured the three CRT outcomes. Other devices used were electrogoniometer, Häggmark and Liedberg light beam device, Ankle Measure for Endurance and Strength (AMES), Haberometer, and custom-made. Devices were mostly used in healthy populations or Achilles tendon pathologies. AMES, Haberometer, and custom-made devices were the most clinician-friendly, but only quantified repetitions were completed. In late 2022, a computer vision mobile application appeared in the literature and offered clinicians a low-cost, research-grade alternative.

Conclusion: This review details seven devices used to measure CRT outcomes. The linear encoder is the most common in research and quantifies all three CRT outcomes. Recent advances in computer-vision provide a low-cost research-grade alternative to clinicians and researchers via a n iOS mobile application.

KEYWORDS
ankle plantar flexors, assessment, endurance, heel-rise test, physiotherapy

1 | INTRODUCTION

The unilateral calf raise test (CRT) is a clinical tool that assesses calf muscle-tendon unit function (Hébert-Losier et al., 2009). The test requires individuals to go up on their toes and back down as many times as possible standing on one leg and therefore requires repetitive concentric-eccentric actions of the ankle plantar flexors until volitional cessation. Clinicians document the total number of repetitions achieved as the primary outcome (Hébert-Losier et al., 2009; Lunsford & Perry, 1995; Ross & Fontenot, 2000). The CRT is therefore considered a clinical-friendly method to assess calf muscle-tendon unit function that requires neither specialised equipment nor much time or space, which is advantageous for field-based and in-clinic testing (André et al., 2016; Haber et al., 2004). The CRT is used across disciplines, such as paediatrics (Maurer et al., 2007), cardiology (Monteiro et al., 2013), neurology (Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998), orthopaedics (J. A. Zellers et al., 2020), and geriatric (André et al., 2018). However, Sman et al. (2014) advanced that using a standardised CRT device and protocol across individuals would be ideal to monitor and replicate outcomes. In science, researchers commonly use devices to assist in standardising the test protocol and quantify additional outcomes other than the number of repetitions completed (Cibulka et al., 2017; Pereira, Schettino, Machado, da Silva, & Neto, 2010; Van Cant et al., 2017).

Indeed, aside from the number of calf raises performed, other CRT measures are considered key outcomes and indicators of function; for example, total (concentric) work and maximum calf raise height during the CRT are markers of functional recovery post Achilles tendon rupture (ATR) (Byrne et al., 2017; Silbernagel et al., 2010; J. A. Zellers et al., 2020). More specifically, the amount of work completed during the CRT has been shown to be a more sensitive metric in the presence of ATR than the number of repetitions (Silbernagel et al., 2010), where work is computed considering calf raise height and body mass displaced during repetitions. Given that work considers the positive displacement of each repetition during the CRT, this measure is deemed scientifically more rigorous and accurate than the number of repetitions to quantify calf muscle-tendon unit endurance (Byrne et al., 2017). In terms of peak height during the CRT, this measure expressed as a ratio of the involved to uninvolved limb (i.e., limb symmetry index) at 6-months post ATR predicted patient-reported symptoms and physical activity levels at 12-months as quantified using the Achilles tendon Total Rupture Score (Olsson et al., 2014). Furthermore, both total work and peak height during the CRT have been identified as more sensitive metrics of residual impairments than the total number of repetitions at 6- and 12-months post ATR. Specifically, repetitions identified the percentage of patients with normal function (defined as the limb symmetry index reaching 90% or higher) as 38% and 63% at 6- and 12-months, respectively. Comparatively, these figures based on the limb symmetry index were 9% and 23% when considering total work, and 6% and 22% when considering peak height (Silbernagel et al., 2010).

Hence, total work and peak height during the CRT are deemed important measures of calf muscle-tendon unit function (Nordenholm et al., 2022; J. A. Zellers et al., 2020). These CRT metrics cannot be quantified clinically without a device. However, clinicians typically only count the number of repetitions as the primary outcome. This narrative review aimed to provide an overview of the CRT devices used in the scientific literature to measure CRT outcomes, namely the number of repetitions, peak height, and total work performed. This review focused on the design, reliability, concurrent validity, and perceived strengths and limitations of these devices for clinical use. It is anticipated that this narrative review will provide practitioners with a clear understanding of CRT devices potentially available to them for quantifying outcomes.

2 | METHODS

2.1 | Search strategy

Even though a narrative review was planned, a systematic process aligning with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Page et al., 2021) was applied. A systematic electronic search was conducted on April 9th, 2022, in the following databases: Cochrane, PubMed®, Scopus, Sports Medicine & Education Index, and SPORTDiscus. The search terms used were "calf raise", "heel raise", "heel rise", "test$", and "eval$". The Boolean operators "OR" and "AND" were used to combine the search terms. When available, search limiters applied included peer-reviewed journal articles, the English language, and humans. The searches implemented for each database are provided in the supplementary materials (Supplemental File S1). References of all studies meeting inclusion criteria were screened to identify additional relevant studies that might have been missed.

The search results were imported into Endnote 20 (Clarivate Analytics, Boston, USA). After removing duplicates, all titles and abstracts were transferred to Rayyan (Qatar Computing Research Institute, Qatar), a free web application for systematic reviews (Ouzzani et al., 2016). In Rayyan, titles and abstracts were screened against the inclusion criteria. The inclusion criteria were scientific peer-reviewed original research of any analytical study design (i.e., observational or experimental) as defined by the Oxford Centre for Evidence-Based Medicine (i) published in English; (ii) that used the unilateral CRT (i.e., repeated concentric-eccentric plantar-flexor actions in unilateral stance to volitional cessation); and (iii) measured CRT outcomes using equipment (e.g., motion capture system, force plate, linear encoder, custom devices, or any other device). Exclusion criteria were (i) editorials, commentaries, discussion papers, conference abstracts, and reviews; (ii) studies that did not describe their methods; and (iii) studies where unilateral CRT outcomes to volitional cessation were not assessed. All studies that met inclusion were retrieved in full text, and their eligibility criteria were assessed. A single reviewer (XXXX) conducted all the screenings, which was verified by a second reviewer (XXXX).

2.2 | Data extraction and synthesis

Relevant information was extracted from each included paper in a custom-made Excel (Microsoft Office, Microsoft, Redmond, WA, USA) data extraction form. The following data were extracted from each study: authors, publication year, study location (based on where data were collected when stated explicitly or institutional ethics approval), study aims, participant characteristics (i.e., healthy or pathologic population, age, gender, body mass, and height), CRT device, and CRT outcomes (i.e., number of repetitions, peak height, and work). In addition, data on the reliability of the identified CRT devices (i.e., test-retest, intra-rater, and inter-rater) and their concurrent validity (i.e., agreement of outcomes between devices) were extracted. Any stated strengths and limitations of devices were also extracted. A single reviewer (RF) extracted all data, and a second reviewer (KHL) verified the completeness of extraction.

Data are summarised using tables for the characteristics of the reviewed studies, reliability and concurrent validity properties, and perceived strengths and limitations of the CRT devices for clinical use. The reliability and concurrent validity of devices were deemed excellent, good, moderate, and poor when corresponding intraclass correlation (ICC) values were >0.90, >0.75 to 0.90, between 0.50 and 0.75, and <0.50 (Portney, 2020). Additionally, it was possible to group CRT devices under thematic headings; therefore, devices are presented thematically using a narrative synthesis format. No risk of bias assessment was undertaken as it was not relevant to the aims of this review.

3 | RESULTS

3.1 | Selection of studies and study characteristics

Figure 1 presents a flowchart of the screening and selection processes. Thirty-five articles met inclusion criteria and were included in the narrative synthesis. The individual study characteristics are presented in Table 1.

3.2 | Calf raise test measuring devices

It was possible to thematically group CRT devices into seven categories: Häggmark and Liedberg's light beam electronic device; electrogoniometer; laboratory-based devices—three-dimensional (3D) motion capture and force plate (used separately or together); Haberometer, linear encoder; Ankle Measure for Endurance and Strength (AMES); and a custom-made device. The timeline of the first use of these devices in the scientific literature is presented in Figure 2. Furthermore, the reliability and concurrent validity of the CRT devices are summarised in Table 2, and their perceived strengths and limitations from a clinical perspective are outlined in Table 3.

Of all the measuring devices, the linear encoder was used most frequently in studies (n = 18, 51.4%) (Andreasen et al., 2020; Arch et al., 2018; Annelie Brorsson et al., 2018; A. Brorsson et al., 2021; A. Brorsson et al., 2017; Byrne et al., 2017; Hamrin et al., 2020; Nordenholm et al., 2022; Olsson et al., 2014; Silbernagel et al., 2015; Silbernagel et al., 2006; Silbernagel et al., 2010; Silbernagel et al., 2012; Svensson et al., 2019; Westin et al., 2018; Zellers et al., 2018; J. A. Zellers et al., 2020; J. A. Zellers et al., 2017), followed by a 3D motion capture with (n = 4, 11.4%) (Hébert-Losier et al., 2011; Hébert-Losier & Holmberg, 2013; Nawoczenski et al., 2016; Tengman et al., 2015) or without (n = 2, 5.7%) (Hébert-Losier et al., 2011, 2012) force plate. Four studies (11.4%) used the electrogoniometer alone (Jan et al., 2005; Lunsford & Perry, 1995; Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998; Svantesson, Osterberg, Thomeé, et al., 1998) three studies (8.5%), used the AMES (DeWolf et al., 2018; Sman et al., 2014; Van Cant et al., 2017), and two studies (5.7% each) used the Häggmark and Liedberg light beam electronic device (Häggmark et al., 1986; Möller et al., 2005) and Haberometer (Haber et al., 2004; Pereira et al., 2010). Finally, one study (2.9% each) used the force plate with an electrogoniometer (Österberg et al., 1998), 3D motion capture with a linear encoder (Andreasen et al., 2020), and a custom made CRT device (Sara et al., 2021), as reported in Table 1. These devices were used most often in studies to examine Achilles tendon pathologies (n = 18, 51.4%) (Andreasen et al., 2020; Annelie Brorsson et al., 2018; A. Brorsson et al., 2021; A. Brorsson et al., 2017; Häggmark et al., 1986; Hamrin et al., 2020; Nawoczenski et al., 2016; Nordenholm et al., 2022; Olsson et al., 2014; Silbernagel et al., 2015; Silbernagel et al., 2006; Silbernagel et al., 2010; Silbernagel et al., 2012; Svensson et al., 2019; Tengman et al., 2015; Westin et al., 2018; Zellers et al., 2018; J. A. Zellers et al., 2020) and healthy populations (n = 17, 48.6%) (Arch et al., 2018; Byrne et al., 2017; DeWolf et al., 2018; Haber et al., 2004; Hébert-Losier & Holmberg; Hébert-Losier et al., 2012; Hébert-Losier et al., 2011; Jan et al., 2005; Lunsford & Perry, 1995; Möller et al., 2005; Österberg et al., 1998; Pereira et al., 2010; Sara et al., 2021; Sman et al., 2014; Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998; Svantesson, Osterberg, Thomeé, et al., 1998; J. A. Zellers et al., 2017) but in one study each (2.9%) for specific musculoskeletal [patellofemoral pain (Van Cant et al., 2017)] or medical [stroke (Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998) and deep vein thrombosis (Haber et al., 2004)] conditions.

3.2.1 | Häggmark and Liedberg's light beam electronic device

In 1986, Häggmark and Liedberg reported using a light beam electronic device for measuring fatigue resistance in the calf muscles in ATR individuals (Häggmark et al., 1986). Specifically, they constructed a device with a light beam attached to vertical rods at a fixed height of 5 cm (Figure 2). Participants needed to lift their heels over

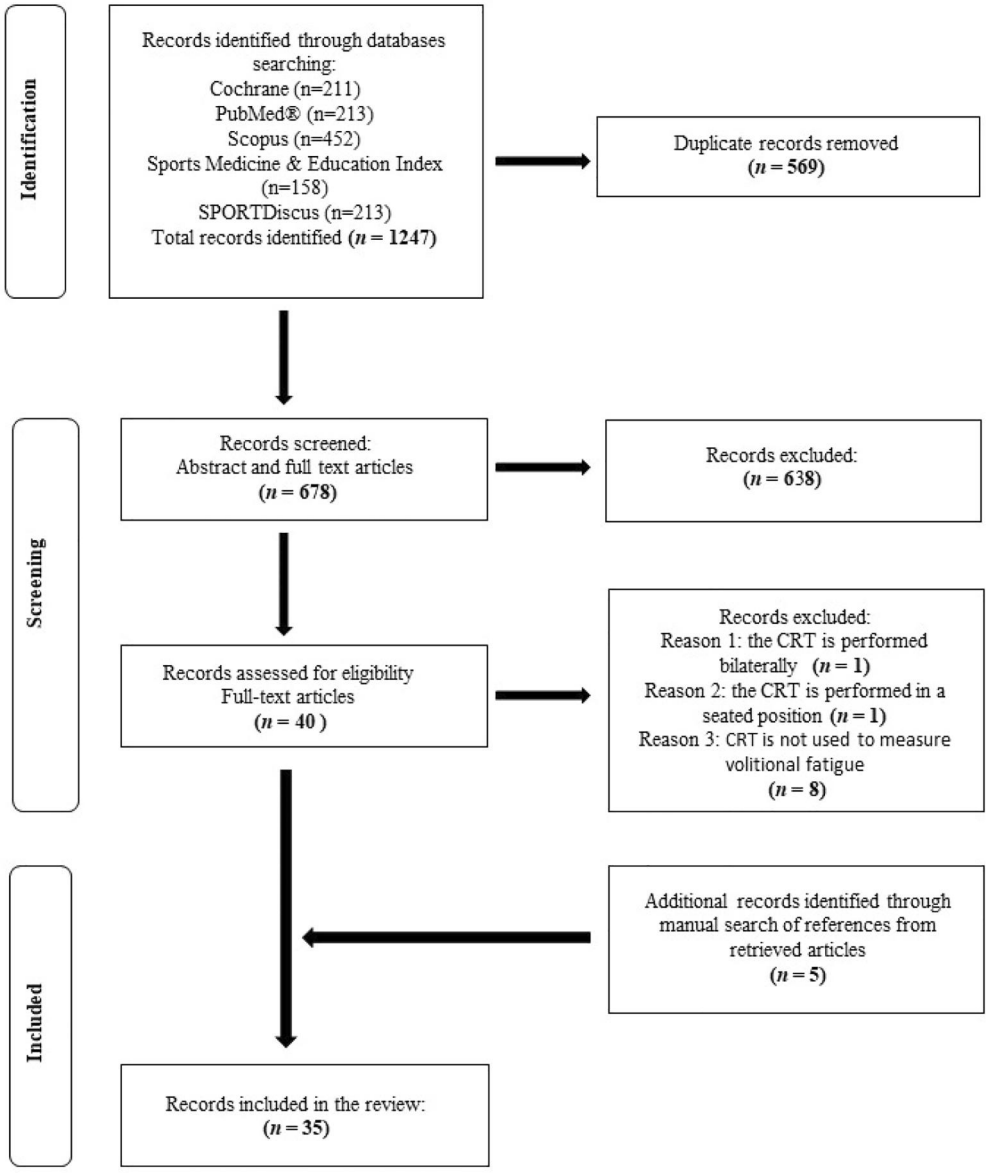

FIGURE 1 Flowchart of the search strategy and the selection process.

the light beam to the sound of a metronome that controlled and monitored their pace. When the 5 cm heel target was reached, the device emitted an audible signal to assist the researchers track the number of repetitions. Möller et al. (2005) developed a modified version of this device and tested it in healthy individuals, adding a foot block to prevent the foot from sliding during testing and to enhance participant safety (Möller et al., 2005). The test-retest reliability of the main outcome (i.e., number of repetitions) was good when performed a week later in healthy individuals (Möller et al., 2005) (Table 2). One limitation of this device is that it requires electricity to function, which makes using the device in remote areas or in field settings challenging (Möller et al., 2005; Sman et al., 2014). Furthermore, this device is not available for purchase, and only monitors the number of repetitions.

3.2.2 | Electrogoniometer

Researchers and clinics use electrogoniometers to measure joint movement (Bronner et al., 2010; Shamsi et al., 2019). A sensor is positioned over the joint centre of rotation. In comparison to motion analysis, the electrogoniometer has a high level of reliability and validity (Bronner et al., 2010).

The electrogoniometer was first used by Lunsford and Perry (1995) during the CRT in healthy individuals to determine the criterion for normal CRT performance. The recorded plantar-flexion ankle measurements were used to quantify the number of repetitions as well as to determine when the test should end based on the ankle plantar-flexion range of motion decreasing by more than 50% from the initial range (Lunsford & Perry, 1995). Jan et al. (2005) used

TABLE 1 Summary of articles reviewed (n = 35).

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
Andreasen et al. (2020) Denmark	To evaluate the concurrent validity of the heel-rise work test performed with use of the heel as surrogate for centre of body mass.	**Sample**: 45 patients with ATR (36 males, 9 females) **Age**: 41 ± 9 y **Height**: NR **Mass**: NR	Linear encoder Laboratory-based equipment (3D MOCAP)	Work
Arch et al. (2018) USA	To evaluate the relationship between gait and clinical measures of plantar flexor function for individuals with no neuromuscular injuries or diseases.	**Sample**: 24 healthy (15 males, 9 females) **Age**: 43.6 ± 24.5 y **Height**: 1.74 ± 0.08 m **Mass**: 78.7 ± 11.8 kg	Linear encoder	Number of repetitions
A. Brorsson et al. (2021) Sweden	To evaluate the possible differences in foot structure between the injured and the healthy limb and between treatment groups 6 years after an ATR.	**Sample**: 90 patients with ATR Surgery **Sample**: 36 males, 9 females **Age**: 50 ± 9 y **Height**: NR **Mass:** NR Non-surgery **Sample**: 39 males, 6 females **Age**: 48 ± 9 y **Height**: NR **Mass**: NR	Linear encoder	Number of repetitions Peak height Work
Annelie Brorsson et al. (2018) Sweden	To evaluate calf muscle performance and patient reported outcomes at least 5 years after an ATR in patients included in a prospective, randomised controlled trial.	**Sample**: 66 patients with ATR (53 males, 13 females) **Age**: 50 ± 8.5 y **Height**: 178 ± 9.7 cm **Mass**: 85.9 ± 13.5 kg	Linear encoder	Number of repetitions Peak height
A. Brorsson et al. (2017) Sweden	To explore differences in ankle biomechanics, calf muscle recovery, tendon length, and patient-reported outcome measurements at a mean of 6 years after ATR between two groups with less than 15% and greater than 30% differences in heel-rise height at 1-year follow-up, respectively.	**Sample**: 34 patients with ATR ATR <15% difference in heel-rise height **Sample**: 15 males, 2 females **Age**: 40 ± 5 y **Height**: NR **Mass**: NR ATR >30% difference in heel-rise height **Sample**: 16 males, 1 female **Age**: 56 ± 9 y **Height**: NR **Mass**: NR	Linear encoder	Number of repetitions Peak height
Byrne et al. (2017) UK	To measure and compare the intrarater test-retest reliability and measurement agreement of the three heel raise endurance test outcome measures in healthy adult during a standardised and computerised heel raise endurance test employing a linear displacement sensor.	**Sample**: 38 healthy individuals (18 males, 20 females) **Age**: 22.7 ± 3.13 y **Height**: 1.73 ± 0.104 m **Mass**: 74.9 ± 15.1 kg	Linear encoder	Number of repetitions Peak height Work
DeWolf et al. (2018) USA	To objectively compare musculoskeletal attributes of pre pointe and recently end pointe ballet dancers to identify differences between those cohorts and secondarily to investigate ant	**Sample**: 49 healthy females Pre-pointe **Sample**: 28 females **Age**: 10.21 ± 1.17 y **Height**: 124.05 ± 13.45 cm	AMES	Number of repetitions

(Continues)

TABLE 1 (Continued)

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
	relationships between the resulting quantitative measures and a qualitative pointe success appraisal completed by each dancer's experienced ballet teacher.	**Mass**: 39.13 ± 13.18 kg Pointe **Sample**: 21 females **Age**: 11.42 ± 0.81 y **Height**: 136.91 ± 16.04 cm **Mass**: 40.79 ± 8.77 kg		
Haber et al. (2004) Canada	To assess the reliability of a protocol using an apparatus specifically designed to standardised the standing heel raise test for triceps surae muscle fatigability on a healthy group of subjects without a current injury.	Short term test-retest group (30 min) and Intermediate term group (48 h) **Sample**: 40 healthy individuals (19 males, 21 females) **Age**: 24 y (range 17–63) **Height**: NR **Mass**: NR Long term test-retest group (7 days) **Sample**: 38 patients with deep vein thrombosis (21 males, 16 females); unaffected side tested **Age**: 51 y (range 25–76) **Height**: NR **Mass**: NR	Haberometer	Number of repetitions
Häggmark et al. (1986) Sweden	To compare the muscle function in a group of patients with ATR treated with surgery versus a group of patients treated non-operatively with a follow up time of three to 5 years.	Surgical ATR group (10 males, 5 females) **Age**: 35. 5 y (range 23–59) **Height**: NR **Mass**: NR Non- surgical ATR group (6 males, 2 females) **Age**: 34. 9 y (range 25–55) **Height**: NR **Mass**: NR	Häggmark and Liedberg light beam electronic device	Work
Hamrin et al. (2020) Sweden	To determine patient-related and treatment related predictors of superior and inferior function in sport and recreational activities 1 year after an ATR.	**Sample**: 285 patients with ATR (238 males, 47 females) **Age**: 40.0 ± 8.4 y **Height**: 178.4 ± 8.3 cm **Mass**: 83.3 ± 13.1 kg	Linear encoder	Number of repetitions Peak height Work
Hébert-Losier et al. (2012) New Zealand	To investigate with surface EMG the influence of knee flexion angles on the soleus, medial gastrocnemius and lateral gastrocnemius fatigue during the maximal numbers of unilateral heel raises.	**Sample**: 48 healthy individuals Younger males **Sample**: 12 males **Age**: 22.4 ± 1.8 y **Height**: 177.4 ± 5.6 cm **Mass**: 71.7 ± 10.2 kg Younger females **Sample**: 12 females **Age**: 22.7 ± 2.0 y **Height**: 165.1 ± 4.2 cm **Mass**: 61.1 ± 10.7 kg Middle aged males **Sample**: 12 males **Age**: 41.1 ± 3.1 y **Height**: 177.7 ± 5.6 cm **Mass**: 81.7 ± 14.9 kg	Laboratory-based equipment (3D MOCAP)	Number of repetitions Peak height Work

TABLE 1 (Continued)

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
		Middle aged females **Sample**: 12 females **Age**: 41.5 ± 3.4 y **Height**: 166.5 ± 8.1 cm **Mass**: 66.6 ± 10.3 kg		
Hébert-Losier et al. (2011) New Zealand	To provide an estimate of the ability of a healthy population to maintain 0° and a 30° knee flexion angle during knee extension heel raise test and knee flexion heel raise test, by investigating the average knee angle maintained and the absolute angular error in knee flexion position during the two versions.	**Sample**: 17 healthy individuals (9 males, 8 females) **Age**: 25.6 ± 4.6 y **Height**: 172.4 ± 9.3 cm **Mass**: 71.1 ± 10.0 kg	Laboratory-based equipment (3D MOCAP)	Number of repetitions Peak height
Hébert-Losier and Holmberg (2013) Sweden	To characterise and compare the biomechanics and clinical outcomes of the single legged heel raise test performed on an incline with the knee straight and bent while considering age and sex as cofounders.	**Sample**: 48 healthy individuals Males **Sample**: 28 males **Age**: 38 ± 12 y **Height**: 169 ± 7 cm **Mass**: 82 ± 9 kg Females **Sample**: 20 females **Age**: 41 ± 11 y **Height**: 169 ± 8 cm **Mass**: 69 ± 9 kg	Laboratory-based equipment (3D MOCAP and force plate)	Number of repetitions Peak height
Jan et al. (2005) Taiwan	To investigate the number of repetitions of the one-leg heel-rise test required for normal plantar-flexor strength in different groups of subjects categorized by age and sex.	**Sample**: 180 healthy individuals Males 21–40y **Sample**: 30 males **Age**: 29.0 ± 4.8 y **Height**: 169.7 ± 6.1 cm **Mass**: 69.7 ± 8.0 kg Males 41–60y **Sample**: 30 males **Age**: 50.2 ± 4.9 y **Height**: 167.2 ± 5.4 cm **Mass**: 67.0 ± 8.0 kg Males 61-80 **Sample**: 30 males **Age**: 69.0 ± 4.0 y **Height**: 166.3 ± 5.4 cm **Mass**: 66.5 ± 6.5 kg Females 21–40y **Sample**: 30 females **Age**: 30.3 ± 4.9 y **Height**: 160.5 ± 3.9 cm **Mass**: 52.4 ± 5.5 kg Females 41–60y **Sample**: 30 females **Age**: 49.9 ± 1.0 y **Height**: 157.0 ± 6.0 cm **Mass**: 57.9 ± 9.2 kg	Electrogoniometer	Number of repetitions

(Continues)

TABLE 1 (Continued)

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
		Females 61-80 **Sample**: 30 females **Age**: 69.1 ± 4.1 y **Height**: 154.9 ± 5.2 cm **Mass**: 58.8 ± 5.5 kg		
Lunsford and Perry (1995) USA	To further refine the standing heel-rise test by assessing the number of heel-rises that can be accomplished by both male and female.	**Sample**: 203 healthy individuals (122 males, 81 females) Males **Sample**: 122 males **Age**: 34.7 ± 8.5 y **Height**: 178.9 ± 7.9 cm **Mass**: 79.7 ± 11.5 kg Females **Sample**: 81 females **Age**: 29.3 ± 5 y **Height**: 164.8 ± 6 cm **Mass**: 60 ± 8.6 kg	Electrogoniometer	Number of repetitions
Möller et al. (2005) Canada	To evaluate the test-retest intra tester reliability of isokinetic measurements in three different positions for ankle plantar flexion and dorsi flexion torque production and to evaluate calf muscle endurance with a standardised heel raise test.	**Sample**: 10 healthy males **Age**: 37 y, range: 31–43 **Height**: 184 cm, range: 172–195 **Mass**: 88 kg, range: 75–98	Häggmark and Liedberg light beam electronic device (modified)	Number of repetitions
Nawoczenski et al. (2016) USA	To investigate muscle performance (ankle plantarflexion power and endurance) during functional tasks and patient-reported outcomes following an isolated gastrocnemius recession for individuals with recalcitrant achilles tendinopathy and an isolated gastrocnemius contracture.	**Sample**: 24 participants Gastrocnemius recession group 8 males, 6 females **Age**: 52.8 ± 7.9 y **Height**: 1.7 ± 0.7 m **Mass**: 92.3 ± 15.5 kg Control group **Sample**: 5 males, 5 females **Age**: 53.3 ± 3.3 y **Height**: 1.7 ± 1.0 m **Mass**: 84.0 ± 16.1 kg	Laboratory-based equipment (3D MOCAP and force plate)	Work
Nordenholm et al. (2022) Sweden	To evaluate the one-year postoperative outcomes in patients with chronic ATR using a comprehensive battery including several validated tests.	**Sample**: 22 patients with ATR (14 males, 8 females) **Age**: 61 ± 15 y **Height**: 173 ± 9 cm **Mass**: 85 ± 15 kg	Linear encoder	Number of repetitions Peak height Work
Olsson et al. (2014) Sweden	To investigate predictors of both symptomatic and functional outcomes for both symptoms and function after ATR.	**Sample**: 93 patients with ATR (79 males, 14 females) **Age**: 39.7 ± 9.3 y **Height**: 179 ± 8 cm **Mass**: 84 ± 12 kg	Linear encoder	Peak height
Österberg et al. (1998) Sweden	To measure the torque influencing the ankle joint during a standing heel raise test from force plate to calculate work during the test.	**Sample**: 10 healthy males **Age**: 25 ± 3 y **Height**: 179 ± 3 cm **Mass**: 76 ± 7 kg	Laboratory-based equipment (force plate) Electrogoniometer	Number of repetitions Work
Pereira et al. (2010) Brazil	To investigate the amplitude and sub-100 Hz frequency content of	**Sample**: 22 healthy individuals (14 males, 8 females)	Haberometer	Number of repetitions

TABLE 1 (Continued)

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
	surface EMG signals obtained from several muscles during the lowering and raising phases of a heel-raise task performed until failure.	**Age:** 21 ± 1 y **Height**: 171 ± 2 cm **Mass**: 65 ± 2 kg		
Sara et al. (2021) USA	To determine (1) associations between standing heel raise test repetitions and measures of maximal plantar flexion strength, assessed as baseline maximal voluntary isometric contraction, (2) associations between standing heel raise test repetitions and the reduction in maximum voluntary isometric contraction following the standing heel raise test, and (3) whether sex differences exist in performance of the standing heel raise test.	**Sample**: 28 healthy individuals 14 males **Age**: 21.5 ± 8 y **Height**: 1.81 ± 0.08 m **Mass**: 79.4 ± 10.3 kg 14 females **Age**: 21.1 ± 9 y **Height**: 1.66 ± 0.07 cm **Mass:** 64.0 ± 10.8 kg	Custom-made device	Number of repetitions
Silbernagel et al. (2006) USA	To evaluate if achilles tendinopathy caused functional deficits on the injured side compared with the non-injured side in patients.	**Sample**: 42 patients with achilles tendinopathy (23 males, 19 females) **Age**: 26 ± 8 y **Height**: 178 ± 8 cm **Mass**: 74.9 ± 15.1 kg	Linear encoder	Number of repetitions Work
Silbernagel et al. (2010) USA	To examine this heel-rise test (that evaluates the height of each heel-rise along the number of repetitions) to evaluate its validity and ability to detect differences in outcome and to compare this test to the test that will be only measures of ankle range of motion and patient-reported outcome.	**Sample**: 78 patients with ATR (65 males, 13 females) **Age**: 42 ± 9 y **Height**: 178 ± 9 cm **Mass**: 85 ± 13 kg	Linear encoder	Number of repetitions Peak height Work
Silbernagel et al. (2012) USA	To evaluate if differences in heel raise height are associated with differences in achilles tendon length after an ATR.	**Sample**: 18 participants Controls 7 males, 3 females **Age**: 28 ± 8 y **Height**: 177 ± 13 cm **Mass**: 73 ± 16 kg Acute complete ATR 5 males3 females **Age**: 46 ± 13 y **Height**: 176 ± 7.7 cm **Mass**: 83 ± 13 kg	Linear encoder	Number of repetitions Peak height
Silbernagel et al. (2015) USA	To evaluate whether there are any differences in outcome between men and women after an acute ATR.	**Sample**: 182 patients with ATR Surgical **Sample**: 76 males, 18 females **Age**: 40 ± 10 y **Height**: NR **Mass**: NR Nonsurgical **Sample**: 76 males, 12 females **Age**: 39 ± 14 y **Height**: NR **Mass**: NR	Linear encoder	Peak height Work

(Continues)

TABLE 1 (Continued)

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
Sman et al. (2014) Australia	To document the construction and reliability of the AMES device.	**Sample**: 40 healthy individuals (21 males, 19 females) **Age**: 24 ± 6.2 y **Height**: 174 ± 12.3 cm **Mass**: 68 ± 9.3 kg	AMES	Number of repetitions
Svantesson, Osterberg, Thomeé, and Grimby (1998) Sweden	To investigate the fatigue process of the gastrocnemius and soleus muscles separately in a standard heel raise test.	**Sample**: 10 healthy women. **Age**: 24 ± 3 y **Height**: 167 ± 4 cm **Mass**: 67 ± 8 kg	Electrogoniometer	Number of repetitions Work
Svantesson, Osterberg, Grimby, and Sunnerhagen (1998) Sweden	To investigate the fatigue process in the triceps surae during the heel-rise test (eccentric and concentric phases) in comparison with a walking test and muscle strength.	**Sample**: 16 males Hemiparesis **Sample**: 8 males **Age**: 57 ± 4 y **Height**: NR **Mass**: 82 ± 10 kg Reference (Healthy) **Sample**: 8 males **Age**: 59 ± 3 y **Height**: NR **Mass**: 82 ± 14 kg	Electrogoniometer	Number of repetitions Work
Svensson et al. (2019) Denmark	To examine muscle function, muscle architecture, and tendon length in persons who reported that they experience a functional deficit more than 2 years after an ATR.	**Sample**: 12 patients with ATR (8 males, 3 females) **Age**: 51 ± 12 y **Height**: 178 ± 10 cm **Mass**: 90 ± 19 kg	Linear encoder	Number of repetitions Peak height Work
Tengman et al. (2015) Sweden	To evaluate muscle fatigue and determine whether fatigue could be detected with a limited number of heel raises after total ATR.	**Sample**: 52 patients with ATR (46 males,6 females) **Age**: 47.8 ± 10.2 y **Height**: NR **Mass**: NR	Laboratory-based equipment (3D MOCAP and force plate)	Number of repetitions Peak height Work
Van Cant et al. (2017) Belgium	To evaluate hip abductor, trunk extensor, and ankle plantar flexor endurance in females and without patellofemoral pain, using clinical tests.	**Sample**: 96 females Patellofemoral pain 20 females **Age**: 21.1 ± 2.6 y **Height**: 162.1 ± 5.8 cm **Mass**: 55.9 ± 7.4 kg Controls 76 females **Age**: 20.5 ± 2.8 y **Height**: 165.5 ± 5.8 cm **Mass**: 58.3 ± 7.4 kg	AMES	Number of repetitions
Westin et al. (2018) Sweden	To perform a long-term follow-up of patients with an achilles tendon re-rupture using established outcome measurements for tendon structure, lower extremity function and symptoms, and to compare the results with those for the uninjured side.	**Sample**: 391 patients with ATR (326 males, 65 females) **Age**: 40.4 ± 8.7 y **Height**: 178.5 ± 8.6 cm **Mass**: 83.7 ± 13.1 kg	Linear encoder	Number of repetitions Peak height Work
J. A. Zellers, van Ostrand, and Silbernagel (2017) USA	To describe the achilles tendon structure and plantar flexor function of classical ballet dancers compared to non-dancers using	**Sample**: 20 healthy individuals Non-dancers 2 males, 8 females **Age**: Range: 16 to 35 y	Linear encoder	Number of repetitions Peak height Work

TABLE 1 (Continued)

Authors (year of publication) Country	Purpose	Participants	Device	CRT outcomes
	established, clinical achilles tendon examination methods.	**Height**: NR **Mass**: NR Ballet dancers 2 males, 8 females **Age**: Range 16 to 35 y **Height**: NR **Mass**: NR		
Jennifer A. Zellers et al. (2018) USA	To determine the strength of the relationship of the achilles tendon resting angle in both knee extended and knee flexed positions with tendon length measured using ultrasound as a validation study; and to identify the relationship between the achilles tendon resting angle with tendon material properties and patient functional performance to better understand its clinical utility.	**Sample**: 42 patients with ATR (34 males, 8 females) **Age**: 45.9 ± 16.2 y **Height**: NR **Mass**: NR	Linear encoder	Work
J. A. Zellers et al. (2020) USA	To investigate the relationship between early tendon morphology and mechanical properties to long-term function on heel-rise and jumping tests in individuals after ATR.	**Sample**: 22 patients with ATR (17 males, 5 females) **Age**: 40 ± 11 y **Height**: NR **Mass**: NR	Linear encoder	Peak height Work

Abbreviations: 3D MOCAP, three-dimensional motion capture; AMES, Ankle Measure for Endurance and Strength; ATR, Achilles tendon rupture; EMG, electromyography; NR, not reported.

the device in a similar fashion to establish normative values for different ages and genders.

Studies in 1998 used electrogoniometers during the CRT in healthy individuals (Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998; Svantesson, Osterberg, Thomeé, et al., 1998; Österberg et al., 1998) and stroke patients (Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998). These studies used an electrogoniometer to determine the concentric and eccentric parts of the calf raise motion and inform electromyographic analysis. Furthermore, total (concentric) work was calculated using the mass of individuals, gravitational acceleration constant, length of the foot between the axis of rotation of the ankle and metatarsophalangeal joints, and angular velocity (Svantesson, Osterberg, Grimby, & Sunnerhagen, 1998; Svantesson, Osterberg, Thomeé, et al., 1998).

The main advantage of electrogoniometers over standard goniometers is their increased precision of joint angle measurements (Bronner et al., 2010). Furthermore, the voltage signals recorded during dynamic motion can be immediately transferred to a computer (Österberg et al., 1998) or data logger (Bronner et al., 2010) and provide joint displacement data in real-time to inform CRT termination (e.g., 50% of ankle plantar-flexion range of motion). Although it is more expensive than a standard goniometer, electrogonimoters are still a low-cost alternative to 3D motion capture systems. To apply this device to the CRT, however, requires a certain amount of programming to compute work, as well as the recording of foot length for work computation.

3.2.3 | Laboratory-based devices: Three-dimensional motion capture and force plate

Three-dimensional (3D) motion capture systems (Jakob et al., 2021) and force plates (Peterson Silveira et al., 2017) are considered the gold standards for collecting biomechanical data in laboratory settings (Figure 2). Österberg et al. (1998) were the first to use a force plate during the CRT alongside an electrogoniometer to quantify torque and work during the test in healthy individuals, while Hébert-Losier et al. were the first to use 3D motion capture in isolation (Hébert-Losier et al., 2011) and together with a force plate (Hébert-Losier & Holmberg, 2013) to quantify CRT outcomes in healthy individuals. The main advantages of these devices are their high accuracy in quantifying biomechanical measures and ability to calculate metrics other than those traditionally reported for the CRT, such as joint angles and torques. Force plates also provide an actual force measure, which can

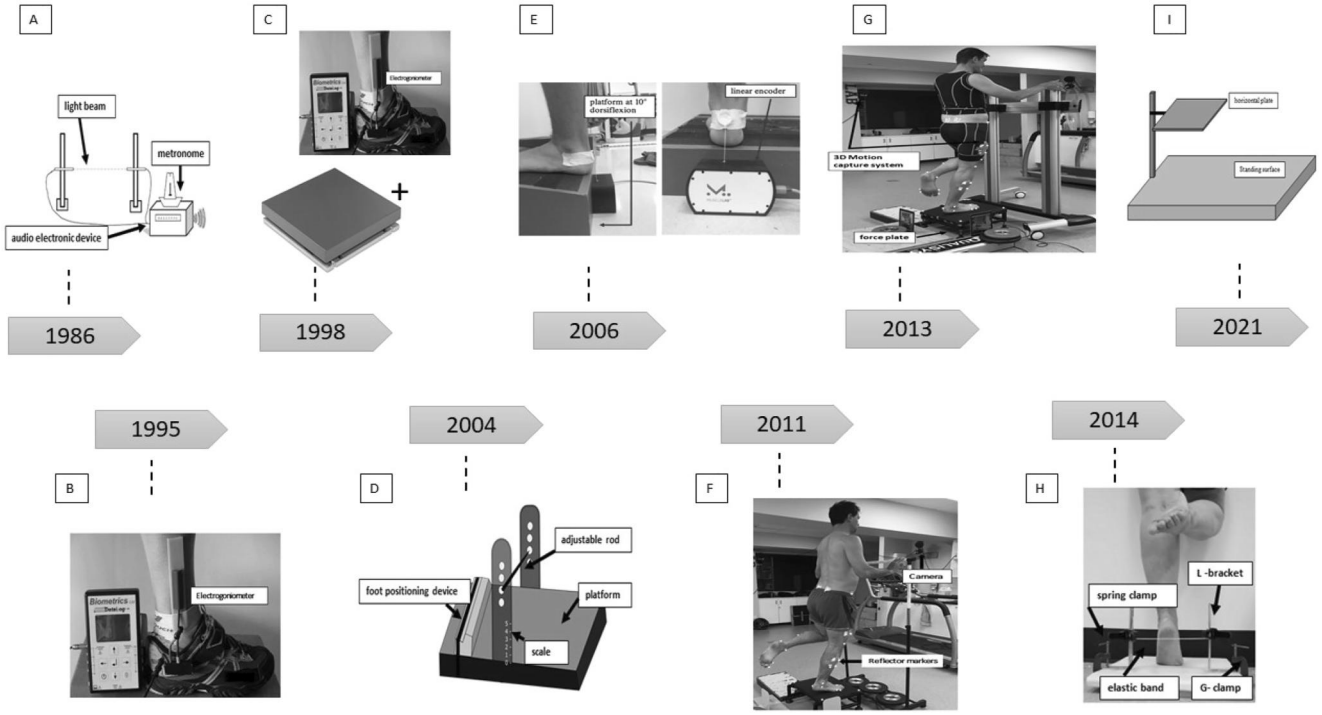

FIGURE 2 Timeline of when various calf raise test (CRT) devices were first used in the scientific literature. (a) Häggmark and Liedberg light beam electronic device (Häggmark et al., 1986), (b) Electrogoniometer (van der Linden, Andreopoulou, Scopes, Hooper, & Mercer, 2018), (c) Electrogoniometer and force plate (Österberg et al., 1998), (d) Haberometer (Haber et al., 2004), (e) Linear encoder (Arch et al., 2018), (f) 3D motion capture system (Hébert-Losier et al., 2011), (g) 3D motion capture and force plate (Hébert-Losier & Holmberg, 2013), (h) Ankle Measure for Endurance and Strength (Sman et al., 2014), (i) Custom-made CRT device (Sara et al., 2021).

be used to calculate work as a product of (actual) force and displacement rather than a (fixed) force based on the mass of individuals and gravitational acceleration constant. Although motion capture systems and force plates are common in research and have been used to assess CRT outcomes in healthy individuals (Hébert-Losier et al., 2011; Hébert-Losier & Holmberg, 2013) as well in patients with ATR (Andreasen et al., 2020; Nawoczenski et al., 2016; Tengman et al., 2015), these devices have limited application in day-to-day clinical practice because of their high costs, limited availability, and time-consuming setup requirements (Schurr et al., 2017).

3.2.4 | Haberometer

To aid in CRT standardisation, Haber et al. (2004) developed the Haberometer (Figure 2), a simple portable device that measures the number of repetitions. The Haberometer is similar to the Häggmark and Liedberg light beam electronic devices but does not rely on electric components. The Haberometer consists of two vertical rods that set the height of calf raise repetitions to 5 cm and a horizontal block that prevents the foot from sliding forward, which are all attached to a base platform (Haber et al., 2004). The device was used in both healthy individuals (Haber et al., 2004; Pereira et al., 2010) and those with deep vein thrombosis (Pereira et al., 2010).

The Haberometer demonstrated good short-, medium- and long-term test-retest reliability for quantifying the number of repetitions based on ICC measures (Table 2). Haber et al. (2004) recommended the device for clinics and research because of its simplicity and reliable outcomes. However, one of the perceived drawbacks of the device is the rod placement over the foot, which may compromise safety if a loss of balance occurs during testing (Haber et al., 2004). Furthermore, this device is not available for purchase and only monitors the number of repetitions.

3.2.5 | Linear encoder

Silbernagel et al. (2006) were the first researchers to introduce the use of a linear encoder for measuring CRT outcomes, which was in ATR patients. The linear encoder (Figure 2) contains a spring-loaded displacement sensor which is attached to the heel and tracks vertical displacement over time. The linear displacement data can be used to calculate work and velocity. Typically, the linear encoder is used to measure the three main CRT outcomes: number of repetitions, peak height, and total work (Byrne et al., 2017; Silbernagel et al., 2010). These linear encoder-derived outcomes have shown good test-retest reliability (Byrne et al., 2017) (Table 2). Furthermore, outcomes from the linear encoder are the only ones which have been validated against 3D motion capture, with the work from the linear encoder almost perfectly correlated with the work from a 3D marker placed on the heel (Andreasen et al., 2020). Since linear encoders can provide the three main outcomes of repetitions, peak height, and work, these

TABLE 2 Summary of reported reliability or concurrent validity of calf raise test (CRT) devices.

Authors (year of publication) CRT devices	Participants	Reliability or concurrent validity of outcomes	Interpretation[a]
Andreasen et al. (2020) Linear encoder 3D MOCAP	Patients with acute achilles tendon rupture (n = 45, 36 males and 9 females)	**Total work measurement error (%)** *Linear encoder heel* versus *MOCAP heel* Injured side = 1.5% (p = 0.163) Non-injured side = 2.9% (p < 0.0001) *Linear encoder heel* versus *MOCAP pelvis* injured side = 21% (p < 0.0001) Non-injured side = 24.7% (p < 0.0001) **Total work concurrent validity (linear regression slope [95% CI])** *Linear encoder heel* versus *MOCAP heel* Injured side = 0.95 [0.91, 1.00] Non-injured side = 1.00 [0.98, 1.03] *Lienar encoder heel* versus *MOCAP pelvis* Injured side = 0.79 [0.70, 0.87] Non-injured side = 0.92 [0.86, 0.97] **Limb symmetry index (linear regression slope [95% CI])** *Linear encoder heel* versus *MOCAP heel* 0.98 [0.92; 1.02] *Lienar encoder heel* versus *MOCAP pelvis* 1.03 [0.90, 1.09]	The CRT performed using the heel as a surrogate for centre of body mass overestimates the total work by 21.0%–24.7% versus the gold standard (MOCAP pelvis), but can precisely detect the relative difference between limbs. Using the heel is considered valid for assessing relative differences between limbs.
Byrne et al. (2017) Linear encoder	Healthy individuals (n = 38, 18 males, 20 females)	**Test-retest reliability (average 9 days)** *Number of repetitions (n)* ICC = 0.77 SEM = 6.7 CV = 13.9% - *Work (J)* ICC = 0.84 SEM = 419 CV = 13.1% *Peak height (cm)* ICC = 0.85 SEM = 0.8 CV = 6.6%	**Based on the ICC estimates** Linear encoder has "good" test-retest reliability for measuring number of repetitions, work, and peak height when tested in healthy individuals.
Haber et al. (2004) Haberometer	Healthy individuals (n = 40, 19 males, 21 females) Patients with deep vein thrombosis (n = 38, 21 males, 16 females)	**Short term test-retest reliability (30 min)** *Number of repetitions (n)* $ICC_{2,1}$ = 0.85 SEM = 2.3 CV = 9% **Intermediate term test-retest reliability (48 h)** *Number of repetitions (n)* $ICC_{2,1}$ = 0.79 SEM = 3.1 CV = 9%	**Based on the ICC estimates** Haberometer has "good" short, intermediate, and long-term test-retest reliability for measuring the number of repetitions when tested in healthy individuals and uninjured side of patients with deep vein thrombosis.

(Continues)

TABLE 2 (Continued)

Authors (year of publication) CRT devices	Participants	Reliability or concurrent validity of outcomes	Interpretation[a]
		Long term test-retest reliability (7 days) *Number of repetitions (n)* $ICC_{2,1}$ = 0.88 SEM = 3.4 CV = 15%	
Möller et al. (2005) Häggmark and Liedberg light beam electronic device (modified)	Healthy individuals (n = 10 males)	**Test-retest reliability (5–7 days)** *Number of repetitions (n) right side* Difference in mean (right) = 1.2 Limits of agreement = −15.3, 17.7 ICC = 0.84 CV = 19.1% *Number of repetitions (n) left side* Difference in mean (left) = 1.7 Limits of agreement = −8.9, 12.3 ICC = 0.78 CV = 13.5%	**Based on the ICC estimates** Häggmark and Liedberg light beam electronic device (modified) has "good" test-retest reliability for measuring number of repetitions when tested in healthy individuals.
Sman et al. (2014) Ankle measure for endurance and strength	Healthy individuals (n = 40, 21 males, 19 females)	**Inter-rater reliability** *Number of repetitions (n)* $ICC_{2,1}$ = 0.97 SEM = 10.4	**Based on the ICC estimates** AMES has "excellent" inter-rater reliability for measuring number of repetitions when the CRT is assessed simultaneously by two examiners when tested in healthy individuals.

Abbreviations: 3D MOCAP, three-dimensional motion capture; AMES, Ankle Measure for Endurance and Strength; CI, confidence interval; CRT, calf raise test; CV, coefficient of variation; ICC, intraclass correlation; SEM, standard error of the mean.

[a]Reliability and concurrent validity "excellent", "good", "moderate", and "poor" when the corresponding ICC was >0.90, >0.75 to 0.90, between 0.50 and 0.75, and <0.50 (Portney, 2020).

devices have been used the most in research to monitor CRT outcomes in healthy individuals (Arch et al., 2018; Byrne et al., 2017; J. A. Zellers et al., 2017) and patients with Achilles tendon pathologies (Andreasen et al., 2020; Annelie Brorsson et al., 2018; A. Brorsson et al., 2016; A. Brorsson et al., 2021; A. Brorsson et al., 2017; Hamrin et al., 2020; Nordenholm et al., 2022; Olsson et al., 2014; Silbernagel et al., 2015; Silbernagel et al., 2006; Silbernagel et al., 2010; Silbernagel et al., 2012; Svensson et al., 2019; Westin et al., 2018; Zellers et al., 2018; J. A. Zellers et al., 2020).

Although CRT outcomes derived from linear encoders provide meaningful information on ankle plantar-flexion function that can assist in the assessment and management of individuals (Byrne et al., 2017), the associated cost of purchasing linear encoder hardware and software prohibits their clinical use. Nonetheless, linear encoders are more affordable than 3D motion capture or force plate systems and are considered a good option for research-compatible outcomes at a modest cost.

3.2.6 | Ankle Measure for Endurance and Strength

Sman et al. (2014) introduced the AMES to address some of the shortcomings of other CRT devices that were used to date, such as the need for electric current, computers, specialised software, or specialised hardware (e.g., light beams and linear encoders). The AMES (Figure 2) consists of a platform, two blocks, two L-shaped brackets, and an elastic band. The elastic band is attached horizontally to the brackets using two spring clamps on either side. In addition, the elastic band height is fully adjustable. To track the number of repetitions, individuals place their heels on the elastic band between the brackets and raise the heel as high as possible during testing (Sman et al., 2014). Hence, the height of the calf raise can be individually set and is not fixed to a certain threshold, such as 5 cm. The AMES was originally tested in healthy individuals (Sman et al., 2014), and later used in patients with patellofemoral pain (Van Cant et al., 2017) as well as in youth ballet dancers (DeWolf et al., 2018).

The AMES presented excellent inter-rater reliability for the number of repetitions completed (Table 2) when simultaneously assessed by two examiners (Sman et al., 2014). Sman et al. (2014) advanced that AMES is ideal for assessing CRT outcomes in clinical and research settings due to its simplicity. The authors further recommended modifications to the AMES to ensure safety while using the apparatus, such as replacing the L-shaped brackets with curved brackets and adding a foot fixation to minimise foot slippage during testing, which could affect the CRT outcome (Sman et al., 2014).

TABLE 3 Perceived strengths and limitations of the various calf raise test (CRT) devices for use in clinical practice.

Devices	Strengths	Limitations	Outcomes measured	Clinical friendliness[a]
Häggmark and Liedberg light beam electronic device	Simple Portable Good test-retest reliability (healthy) Used in ATR and healthy	Requires electricity Set height of 5 cm Used to record repetitions only Not commercially available	Repetitions	2
Electrogoniometer	Simple Portable Used in stroke and healthy	Requires electricity Requires specific hardware and software	Repetitions Work	2
Haberometer	Simple Portable Good test-retest reliability (healthy) Used in DVT and healthy	Requires electricity Set height of 5 cm Rod over foot may affect balance Used to record repetitions only Not commercially available	Repetitions	1
Linear encoder	Relatively simple Relatively portable Measures three CRT outcomes Good test-retest reliability Most frequently used Valid versus 3D MOCAP Used in ATR, AT, and healthy	Medium cost Requires electricity Requires specific hardware and software Requires programming	Repetitions Peak height Work	2
Laboratory-based equipment (3D MOCAP and force plate)	Measures three CRT outcomes Gold standard for measuring biomechanical variables related to the CRT High accuracy Used in ATR and healthy Can be used to record other biomechanical measure	High-cost Requires electricity Requires specific hardware and software Requires programming Requires user expertise Time consuming to set-up	Repetitions Peak height Work	3
Ankle measure for endurance and strength	Simple Portable Low-cost Adjustable height Excellent inter-rater reliability (healthy) Used in PFP and healthy	Used to record repetitions only Not commercially available	Repetitions	1
Custom-made CRT	Simple Portable Low-cost Adjustable height Used in healthy	Used to record repetitions only Not commercially available No studies on reliability and validity findings	Repetitions	2

Abbreviations: 3D MOCAP, three-dimensional motion capture; ATR, Achilles tendon rupture; CRT, calf raise test; DVT, deep vein thrombosis; PFP, patellofemoral pain.

[a]Rank-ordered from most (1) to least (3) clinical-friendly.

Noteworthy is that this device is not available for purchase and only monitors the number of repetitions.

3.2.7 | Custom-made device

Sara et al. (2021) investigated the correlation between the number of repetitions performed during the CRT and maximal plantar-flexor strength in males and females. A custom-made CRT device was used to aid in standardising the test. The device consists of a horizontal plate affixed above a standing surface by an upright support bar. This horizontal plate acts as a visual and tactile guide that is adjusted (vertical and anteroposterior) to the dorsal ankle crease at the end-range of a maximal single-leg calf raise. Conceptually, this device is similar to the AMES device. Although the device is easy to use, portable, and simple, there are no existing studies to support its validity and reliability, the device only monitors the number of repetitions, and it is not available for purchase.

4 | DISCUSSION

In clinical practice and research, the CRT is used to assess the strength-endurance of the triceps surae muscles (Lunsford & Perry, 1995; Svantesson, Osterberg, Thomeé, et al., 1998). Despite being considered a reliable and valid clinical tool, there are concerns regarding the standardisation of its protocols (Hébert-Losier et al., 2009; Sman et al., 2014) and that key clinical parameters are omitted when only counting the number of repetitions (Byrne et al., 2017). To address these shortcomings, a range of devices have been developed and used to standardise and objectivise CRT performance. This review critically appraised 35 relevant studies that used measuring devices to evaluate CRT outcomes in healthy individuals as well as those with medical conditions to inform evidence-based practice. Among the 35 studies included, the Haberometer, AMES, and custom-made devices were considered the most clinical-friendly, but these only recorded the number of repetitions. Laboratory-based 3D motion capture and force plate systems are considered to provide the greatest precision of measurement and offer the advantage of quantifying the three main CRT outcomes but are the least clinical-friendly and most costly devices. The Häggmark and Liedberg light beam electronic device, electrogoniometer, and linear encoder were all considered moderately clinically friendly from a practical and cost perspective, with the linear encoder being the most often used in the scientific literature and the only device reported to quantify the three main outcomes. As such, the linear encoder method appears to offer the best compromise for clinicians seeking research-grade outcomes for the CRT at a modest cost.

Although the number of calf raises performed is the primary test outcome evaluated in clinics, Svantesson, Osterberg, Thomeé, et al. (1998) suggested assessing calf raise height during the test as shorter height ranges could lead to more repetitions since less work is required per repetition. Furthermore, it is noteworthy from a clinical perspective that the number of repetitions and height are related to different physiological and structural factors (Svensson et al., 2019). The number of repetitions is determined by contractile tissue and muscle endurance metabolism (Holloszy, 1967), while the height of the calf raise is determined by tendon and muscle fiber length (Baxter et al., 2018). These triceps surae muscle properties can all affect the total work performed during the CRT as both the number of repetitions and height are used to compute work (Svantesson, Osterberg, Thomeé, et al., 1998). Furthermore, research supports that peak height and work are more sensitive metrics in the presence of pathology and functional deficits (Baxteret al., 2018; Svensson et al., 2019; J. A. Zellers et al., 2020). For these reasons, several researchers have advocated using peak height and work in addition to the number of repetitions as objective measures of triceps surae muscle-tendon unit function during the CRT (Byrne et al., 2017; Fernandez et al., 2022; Silbernagel et al., 2006). Of the seven thematically grouped devices sourced from the literature (Figure 1), only the linear encoder and laboratory-based 3D motion capture and force plate systems were used to quantify all three CRT outcomes. Of the two methods, the linear encoder is the most affordable option for clinical use and clinically based research. Indeed, the linear encoder was the most frequent device used in the scientific literature, likely due to its ability to provide reliable (Byrne et al., 2017) research-grade outcomes (Andreasen et al., 2020) at a moderate financial cost compared to 3D motion and force plate systems. However, this device still requires specialised software and knowledge to extract data, making the linear encoder less suitable for everyday clinical applications and explaining its lack of general uptake from a clinical standpoint.

The Haberometer, AMES, and custom-made devices were considered as the most clinically friendly CRT devices (Table 3), followed by Häggmark and Liedberg's light beam electronic device because of their simplicity, portability, affordability, and no requirement of specialised hardware or software. These devices can assist in standardising CRT parameters, with the Haberometer and AMES being reliable for measuring the number of repetitions (Haber et al., 2004; Möller et al., 2005; Sman et al., 2014). Though none of the reviewed literature sought to quantify the work performed when using these devices, because of the fixed calf raise height, the work completed can be calculated based on the number of repetitions, known calf raise height, and mass of individuals similar to the work computations used for the linear encoder, electrogoniometer, and motion capture systems. The one drawback, however, is that peak height during each raise is not quantified and the proposed work computations from the set height would therefore underestimate the actual work performed. Furthermore, these devices are not readily available for purchase, again limiting their widespread uptake in clinical practice.

Since the systematic search in April 2022, CRT performance has also been quantified using a mobile iOS application (Figure 3) that relies on computer-vision algorithms to track the displacement of a marker placed on the foot (Fernandez et al., 2022; Hébert-Losier & Balsalobre-Fernández, 2020; Hébert-Losier et al., 2022). Specifically, the vertical position of a circular marker is tracked from a video recording via computer vision after calibration to a known distance. The application has been used in athletes (Hébert-Losier et al., 2022) and healthy individuals (Fernandez et al., 2022) and demonstrated good-to-excellent validity of CRT outcomes against 3D motion capture and force plate (ICC $\geq$0.878) and inter-rater, intra-rater, and test-retest reliability. The Calf Raise application (Hébert-Losier & Balsalobre-Fernández, 2020) hence provides a valid and reliable method for assessing the three main CRT outcomes, and an innovative clinical-friendly low-cost option to iOS users. Given that clinicians (Galetsi et al., 2022) and practitioners (Shaw et al., 2021) are increasingly using mobile applications and digital technologies, the computer-vision-based mobile application is an appealing and accessible solution for quantifying CRT outcomes in clinics, although it still requires access to an iOS device.

This literature has a few limitations to acknowledge. First, this review focused on the single-leg CRT performed to fatigue and did not consider other variations of this task, such as when calf raises are done bilaterally or for a set duration (André et al., 2016; Aruje Zahid et al., 2022). Different devices have been used for these task variations, including an overhead bar to set calf raise height (André et al., 2016)

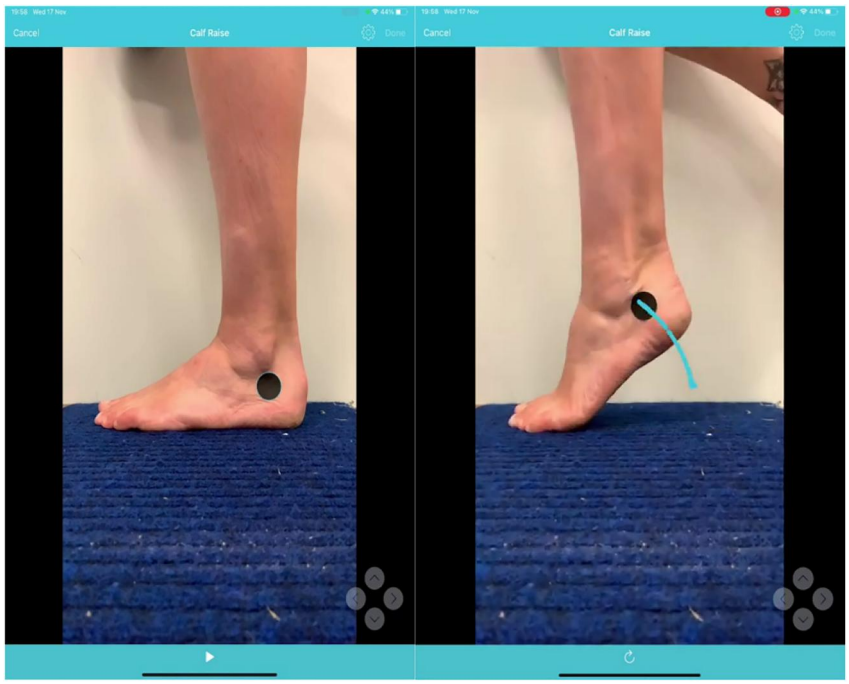

FIGURE 3 A computer-vision algorithm is used to track the vertical displacement of a marker placed on the foot to calculate calf raise test (CRT) outcomes, first introduced in 2022 (Hébert-Losier et al., 2022).

and inertial measurement units (Aruje Zahid et al., 2022), which could be applicable to the single-leg CRT. The former method would have similar strengths and limitations than the Haberometer or AMES, whereas the latter still needs development and validation for the single-leg CRT. This review also focused on the CRT devices and their design, reliability, concurrent validity, and perceived strengths and limitations for clinical use, not on other psychometric properties of the assessment procedures, such as the responsiveness of outcomes, sensitivity, or specificity. Despite our narrative review following a rigorous and systematic process in accordance with the PRISMA guidelines, no critical appraisal of the included studies was undertaken due to the varied methods used in the studies. Furthermore, no risk of bias assessment was completed as it was not relevant to the review aims. Therefore, this review was limited to a narrative synthesis of the findings and conclusions drawn from the studies included. Nonetheless, this approach was deemed suitable for the aims of the review and to provide a comprehensive overview of the current CRT devices used, their strengths, and their limitations.

5 | CONCLUSION

This review provides clinicians and researchers insight into the devices that have been used to assess the CRT, and the strengths and limitations of these devices. The use of devices for the CRT has a dual purpose: to enhance the standardisation of procedures and to further objective CRT outcomes beyond the number of repetitions. The linear encoder and computer-vision mobile application offer the best compromise for clinicians seeking research-grade outcomes for the CRT at a modest cost.

5.1 | Clinical message

- Research-grade or customised devices are used to standardise and quantify CRT outcomes in science. Seven different devices were identified.
- By understanding the design and properties of these CRT devices, practitioners will be able to determine which ones are the most appropriate for their clinical needs and to use these devices to implement evidence-based practice.
- Linear encoders appear to provide the best compromise for clinicians seeking research-grade CRT outcomes at a modest cost.
- Advances in computer vision technology have led to iOS mobile applications that provide low-cost research-grade alternatives for clinicians and researchers to quantify clinical outcomes, including the CRT.

AUTHOR CONTRIBUTIONS

Ma. Roxanne Fernandez and Kim Hébert-Losier: Conceptualization, methodology, formal analysis, investigation, data curation, writing—original draft, writing—review & editing, visualization, project administration. Kim Hébert-Losier: Supervision.

ACKNOWLEDGEMENTS

The authors wish to thank Prof Masayoshi Kubo, Department of Physical Therapy, Niigata University of Health and Welfare, and Dr Josie Athens, Department of Preventive and Social Medicine, University of Otago, for their valuable comments on drafts of the manuscript and exceptional support during the research process. Their valuable contributions to this paper are greatly appreciated.

Open access publishing facilitated by The University of Waikato, as part of the Wiley - The University of Waikato agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Ma. Roxanne Fernandez https://orcid.org/0000-0003-0417-3817
Kim Hébert-Losier https://orcid.org/0000-0003-1087-4986

REFERENCES

André, H.-I., Carnide, F., Borja, E., Ramalho, F., Santos-Rocha, R., & Veloso, A. P. (2016). Calf-raise senior: A new test for assessment of plantar flexor muscle strength in older adults: Protocol, validity, and reliability. *Clinical Interventions in Aging*, *11*, 1661–1674. https://doi.org/10.2147/CIA.S115304

André, H.-I., Carnide, F., Moço, A., Valamatos, M.-J., Ramalho, F., Santos-Rocha, R., & Veloso, A. (2018). Can the calf-raise senior test predict functional fitness in elderly people? A validation study using electromyography, kinematics and strength tests. *Physical Therapy in Sport*, *32*, 252–259. https://doi.org/10.1016/j.ptsp.2018.05.012

Andreasen, K. R., Hansen, M. S., Bencke, J., Hölmich, P., & Barfod, K. W. (2020). The heel-rise work test overestimates the performed work with 21–25% after an Achilles tendon rupture. *Knee Surgery, Sports Traumatology, Arthroscopy*, *29*(5), 1604–1611. https://doi.org/10.1007/s00167-020-06369-8

Arch, E. S., Colon, S., Silbernagel, K. G., & Crenshaw, J. R. (2018). Evaluating the relationship between gait and clinical measures of plantar flexor function. *Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology*, *43*, 41–47. https://doi.org/10.1016/j.jelekin.2018.09.003

Aruje Zahid, S., Celik, Y., Godfrey, A., & Buckley, J. G. (2022). Use of 'wearables' to assess the up-on-the-toes test. *Journal of Biomechanics*, *143*, 111272. https://doi.org/10.1016/j.jbiomech.2022.111272

Baxter, J. R., Hast, D. C., & Hast, M. W. (2018). Plantarflexor fiber length and tendon slack length are the strongest determinates of simulated single-leg heel raise function. *Journal of Biomechanics*, 415679. https://doi.org/10.1101/415679

Bronner, S., Agraharasamakulam, S., & Ojofeitimi, S. (2010). Reliability and validity of electrogoniometry measurement of lower extremity movement. *Journal of Medical Engineering & Technology*, *34*(3), 232–242. https://doi.org/10.3109/03091900903580512

Brorsson, A., Grävare Silbernagel, K., Olsson, N., & Nilsson Helander, K. (2018). Calf muscle performance deficits remain 7 years after an Achilles tendon rupture. *The American Journal of Sports Medicine*, *46*(2), 470–477. https://doi.org/10.1177/0363546517737055

Brorsson, A., Olsson, N., Nilsson-Helander, K., Karlsson, J., Eriksson, B. I., & Silbernagel, K. G. (2016). Recovery of calf muscle endurance 3 months after an Achilles tendon rupture. *Scandinavian Journal of Medicine & Science in Sports*, *26*(7), 844–853. https://doi.org/10.1111/sms.12533

Brorsson, A., Sædís Jónsdóttir, U., Nygren, D., Larsson, N., & Tranberg, R. (2021). The injured limb presents lower values in foot structure measurements 6 years after an achilles tendon rupture. *Muscles, Ligaments and Tendons Journal*, *11*(4), 694–703. https://doi.org/10.32098/mltj.04.2021.12

Brorsson, A., Willy, R. W., Tranberg, R., & Grävare Silbernagel, K. (2017). Heel-rise height deficit 1 year after Achilles tendon rupture relates to changes in ankle biomechanics 6 years after injury. *The American Journal of Sports Medicine*, *45*(13), 3060–3068. https://doi.org/10.1177/0363546517717698

Byrne, C., Keene, D. J., Lamb, S. E., & Willett, K. (2017). Intrarater reliability and agreement of linear encoder derived heel-rise endurance test outcome measures in healthy adults. *Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology*, *36*, 34–39. https://doi.org/10.1016/j.jelekin.2017.07.004

Cibulka, M., Wenthe, A., Zach, B., Callier, D., Schwerdt, A., Jarman, D., & Strube, M. (2017). Variation in medial and lateral gastrocnemius muscle activity with foot position. *International Journal of Sports Physical Therapy*, *12*(2), 233–241. https://pubmed.ncbi.nlm.nih.gov/28515978

DeWolf, A., McPherson, A., Besong, K., Hiller, C., & Docherty, C. (2018). Quantitative measures utilized in determining pointe readiness in young ballet dancers. *Journal of Dance Medicine and Science: Official Publication of the International Association for Dance Medicine & Science*, *22*(4), 209–217. https://doi.org/10.12678/1089-313X.22.4.209

Fernandez, M. R., Athens, J., Balsalobre-Fernandez, C., Kubo, M., & Hébert-Losier, K. (2022). Concurrent validity and reliability of a mobile iOS application used to assess calf raise test kinematics. *Musculoskeletal Science and Practice*, *63*, 102711. https://doi.org/10.1016/j.msksp.2022.102711

Galetsi, P., Katsaliaki, K., & Kumar, S. (2022). Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: An analysis of mobile applications for health specialists. *Technovation*, *121*, 102598. https://doi.org/10.1016/j.technovation.2022.102598

Haber, M., Golan, E., Azoulay, L., Kahn, S. R., & Shrier, I. (2004). Reliability of a device measuring triceps surae muscle fatigability. *British Journal of Sports Medicine*, *38*(2), 163–167. https://doi.org/10.1136/bjsm.2002.002899

Häggmark, T. M. D., Liedberg, H. M. D., Eriksson, E. M. D., & Wredmark, T. M. D. (1986). Calf muscle atrophy and muscle function after non-operative vs operative treatment of Achilles tendon ruptures. *Orthopedics*, *9*(2), 160–164. https://doi.org/10.3928/0147-7447-19860201-06

Hamrin, S. E., Svedman, S., Svantesson, E., Danielsson, A., Krupic, F., Ackermann, P., & Westin, O. (2020). Understanding limitations in sport 1 year after an achilles tendon rupture: A multicentre analysis of 285 patients. *Knee Surgery, Sports Traumatology, Arthroscopy*, *28*(1), 233–244. https://doi.org/10.1007/s00167-019-05586-0

Hébert-Losier, K., & Balsalobre-Fernández, C. (2020). Calf raise mobile application software.

Hébert-Losier, K., & Holmberg, H. C. (2013). Biomechanics of the heel-raise test performed on an incline in two knee flexion positions. *Clinical Biomechanics*, *28*(6), 664–671. https://doi.org/10.1016/j.clinbiomech.2013.06.004

Hébert-Losier, K., Newsham-West, R. J., Schneiders, A. G., & Sullivan, S. J. (2009). Raising the standards of the calf-raise test: A systematic review. *Journal of Science and Medicine in Sport*, *12*(6), 594–602. https://doi.org/10.1016/j.jsams.2008.12.628

Hébert-Losier, K., Ngawhika, T. M., Gill, N., & Balsalobre-Fernandez, C. (2022). Validity, reliability, and normative data on calf muscle function in rugby union players from the Calf Raise application. *Sports Biomechanics*, 1–22. https://doi.org/10.1080/14763141.2022.2118158

Hébert-Losier, K., Schneiders, A. G., García, J. A., Sullivan, S. J., & Simoneau, G. G. (2012). Influence of knee flexion angle and age on triceps surae muscle fatigue during heel raises. *The Journal of Strength &*

Conditioning Research, *26*(11), 3134–3147. https://doi.org/10.1519/JSC.0b013e318243ff0e

Hébert-Losier, K., Schneiders, A. G., Sullivan, S. J., Newsham-West, R. J., García, J. A., & Simoneau, G. G. (2011). Analysis of knee flexion angles during 2 clinical versions of the heel raise test to assess soleus and gastrocnemius function. *Journal of Orthopaedic & Sports Physical Therapy*, *41*(7), 505–513. https://doi.org/10.2519/jospt.2011.3489

Holloszy, J. O. (1967). Biochemical adaptations in muscle. Effects of exercise on mitochondrial oxygen uptake and respiratory enzyme activity in skeletal muscle. *Journal of Biological Chemistry*, *242*(9), 2278–2282. https://doi.org/10.1016/s0021-9258(18)96046-1

Jakob, V., Küderle, A., Kluge, F., Klucken, J., Eskofier, B. M., Winkler, J., Winterholler, M., & Gassner, H. (2021). Validation of a sensor-based gait analysis system with a gold-standard motion capture system in patients with Parkinson's disease. *Sensors*, *21*(22), 7680. https://doi.org/10.3390/s21227680

Jan, M.-H., Chai, H.-M., Lin, Y.-F., Lin, J. C.-H., Tsai, L.-Y., Ou, Y.-C., & Lin, D.-H. (2005). Effects of age and sex on the results of an ankle plantar-flexor manual muscle test. *Physical Therapy*, *85*(10), 1078–1084. https://doi.org/10.1093/ptj/85.10.1078

Lunsford, B. R., & Perry, J. (1995). The standing heel-rise test for ankle plantar flexion: Criterion for normal. *Physical Therapy*, *75*(8), 694–698. https://doi.org/10.1093/ptj/75.8.694

Maurer, C., Finley, A., Martel, J., Ulewicz, C., & Larson, C. A. (2007). Ankle plantarflexor strength and endurance in 7-9 year old children as measured by the standing single leg heel-rise test. *Physical & Occupational Therapy in Pediatrics*, *27*(3), 37–54. https://doi.org/10.1080/j006v27n03_04

Möller, M., Lind, K., Styf, J., & Karlsson, J. (2005). The reliability of isokinetic testing of the ankle joint and a heel-raise test for endurance. *Knee Surgery, Sports Traumatology, Arthroscopy*, *13*(1), 60–71. https://doi.org/10.1007/s00167-003-0441-0

Monteiro, D. P., Britto, R. R., Lages, A. C., Basílio, M. L., de Oliveira Pires, M. C., Carvalho, M. L., Britto, Pereira, & Pereira, D. A. (2013). Heel-rise test in the assessment of individuals with peripheral arterial occlusive disease. *Vascular Health and Risk Management*, *9*, 29–35. https://doi.org/10.2147/vhrm.S39860

Nawoczenski, D., DiLiberto, F., Cantor, M., Tome, J., & DiGiovanni, B. (2016). Ankle power and endurance outcomes following isolated gastrocnemius recession for Achilles tendinopathy. *Foot & Ankle International*, *37*(7), 766–775. https://doi.org/10.1177/1071100716638128

Nordenholm, A., Nilsson, N., Senorski, E. H., Helander, K. N., Westin, O., & Olsson, N. (2022). Patients with chronic Achilles tendon rupture have persistent limitations in patient-reported function and calf muscle function one year after surgical treatment - a case series. *Journal of Experimental Orthopaedics*, *9*(1), 15. https://doi.org/10.1186/s40634-022-00451-5

Olsson, N., Petzold, M., Brorsson, A., Karlsson, J., Eriksson, B. I., & Silbernagel, K. G. (2014). Predictors of clinical outcome after acute Achilles tendon ruptures. *The American Journal of Sports Medicine*, *42*(6), 1448–1455. https://doi.org/10.1177/0363546514527409

Österberg, U., Svantesson, U., Takahashi, H., & Grimby, G. (1998). Torque, work and EMG development in a heel-rise test. *Clinical Biomechanics*, *13*(4), 344–350. https://doi.org/10.1016/S0268-0033(98)00100-4

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, *372*, n71. https://doi.org/10.1136/bmj.n71

Pereira, R., Schettino, L., Machado, M., da Silva, P. A. V., & Neto, O. P. (2010). Task failure during standing heel raises is associated with increased power from 13 to 50 Hz in the activation of triceps surae. *European Journal of Applied Physiology*, *110*(2), 255–265. https://doi.org/10.1007/s00421-010

Peterson Silveira, R., Stergiou, P., Carpes, F. P., Castro, F. A., Katz, L., & Stefanyshyn, D. J. (2017). Validity of a portable force platform for assessing biomechanical parameters in three different tasks. *Sports Biomechanics*, *16*(2), 177–186. https://doi.org/10.1080/14763141.2016.1213875

Portney, L. G. (2020). *Foundation of clinical research: Applications to evidence-based practice* (4th ed.). F.A. Davis Company.

Ross, M. D., & Fontenot, E. G. (2000). Test–retest reliability of the standing heel-rise test. *Journal of Sport Rehabilitation*, *9*(2), 117–123. https://doi.org/10.1123/jsr.9.2.117

Sara, L., Gutsch, S., & Hunter, S. (2021). The single-leg heel raise does not predict maximal plantar flexion strength in healthy males and females. *PLoS One*, *16*(8), e0253276. https://doi.org/10.1371/journal.pone.0253276

Schurr, S. A., Marshall, A. N., Resch, J. E., & Saliba, S. A. (2017). Two-dimensional video analysis is comparable to 3D motion capture in lower extremity movement assessment. *International Journal of Sports Physical Therapy*, *12*(2), 163–172.

Shamsi, M., Mirzaei, M., & Khabiri, S. S. (2019). Universal goniometer and electro-goniometer intra-examiner reliability in measuring the knee range of motion during active knee extension test in patients with chronic low back pain with short hamstring muscle. *BMC Sports Science, Medicine & Rehabilitation*, *11*(1), 4. https://doi.org/10.1186/s13102-019-0116-x

Shaw, M. P., Satchell, L. P., Thompson, S., Harper, E. T., Balsalobre-Fernández, C., & Peart, D. J. (2021). Smartphone and tablet software apps to collect data in sport and exercise settings: Cross-sectional international survey. *JMIR mHealth and uHealth*, *9*(5), e21763. https://doi.org/10.2196/21763

Silbernagel, K. G., Brorsson, A., Olsson, N., Eriksson, B. I., Karlsson, J., & Nilsson-Helander, K. (2015). Sex differences in outcome after an acute Achilles tendon rupture. *Orthopaedic Journal of Sports Medicine*, *3*(6), 2325967115586768. https://doi.org/10.1177/2325967115586768

Silbernagel, K. G., Gustavsson, A., Thomeé, R., & Karlsson, J. (2006). Evaluation of lower leg function in patients with Achilles tendinopathy. *Knee Surgery, Sports Traumatology, Arthroscopy*, *14*(11), 1207–1217. https://doi.org/10.1007/s00167-006-0150-6

Silbernagel, K. G., Nilsson-Helander, K., Thomee, R., Eriksson, B. I., & Karlsson, J. (2010). A new measurement of heel-rise endurance with the ability to detect functional deficits in patients with Achilles tendon rupture. *Knee Surgery, Sports Traumatology, Arthroscopy*, *18*(2), 258–264. https://doi.org/10.1007/s00167-009-0889-7

Silbernagel, K. G., Steele, R., & Manal, K. (2012). Deficits in heel-rise height and Achilles tendon elongation occur in patients recovering from an Achilles tendon rupture. *The American Journal of Sports Medicine*, *40*(7), 1564–1571. https://doi.org/10.1177/0363546512447926

Sman, A. D., Hiller, C. E., Imer, A., Ocsing, A., Burns, J., & Refshauge, K. M. (2014). Design and reliability of a novel heel rise test measuring device for plantarflexion endurance. *BioMed Research International*, *2014*, 391646–391647. https://doi.org/10.1155/2014/391646

Svantesson, U., Osterberg, U., Grimby, G., & Sunnerhagen, K. S. (1998). The standing heel-rise test in patients with upper motor neuron lesion due to stroke. *Scandinavian Journal of Rehabilitation Medicine*, *30*(2), 73–80. https://doi.org/10.1080/003655098444165

Svantesson, U., Osterberg, U., Thomeé, R., & Grimby, G. (1998). Muscle fatigue in a standing heel-rise test. *Scandinavian Journal of Rehabilitation Medicine*, *30*(2), 67–72. https://doi.org/10.1080/003655098444156

Svensson, R., Couppe, C., Agergaard, A., Josefsen, C., Jensen, M., Barfod, K., Nybing, J. D., Hansen, P., Krogsgaard, M., & Magnusson, S. (2019). Persistent functional loss following ruptured Achilles tendon is associated with reduced gastrocnemius muscle fascicle length, elongated gastrocnemius and soleus tendon, and reduced muscle cross-sectional area. *Translational Sports Medicine*, *2*(6), 316–324. https://doi.org/10.1002/tsm2.103

Tengman, T., Coleman, S., Silbernagel, K. G., Karlsson, J., & Riad, J. (2015). Muscle fatigue after achilles tendon rupture: A limited heel-rise test with electromyography reveals decreased endurance. *European Journal of Physiotherapy*, *17*(4), 200–207. https://doi.org/10.3109/21679169.2015.1078842

van der Linden, M. L., Andreopoulou, G., Scopes, J., Hooper, J. E., & Mercer, T. H. (2018). Ankle kinematics and temporal Gait characteristics over the duration of a 6-minute walk test in people with multiple sclerosis who experience foot drop. *Rehabilitation Research and Practice*, 1260852. https://doi.org/10.1155/2018/1260852

Van Cant, J., Pitance, L., & Feipel, V. (2017). Hip abductor, trunk extensor and ankle plantar flexor endurance in females with and without patellofemoral pain. *Journal of Back and Musculoskeletal Rehabilitation*, *30*(2), 299–307. https://doi.org/10.3233/BMR-150505

Westin, O., Nilsson Helander, K., Gravare Silbernagel, K., Samuelsson, K., Brorsson, A., & Karlsson, J. (2018). Patients with an Achilles tendon re-rupture have long-term functional deficits and worse patient-reported outcome than primary ruptures. *Knee Surgery, Sports Traumatology, Arthroscopy*, *26*(10), 3063–3072. https://doi.org/10.1007/s00167-018-4952-0

Zellers, J. A., Carmont, M. R., & Silbernagel, K. G. (2018). Achilles tendon resting angle relates to tendon length and function. *Foot & Ankle International*, *39*(3), 343–348. https://doi.org/10.1177/1071100717742372

Zellers, J. A., Pohlig, R. T., Cortes, D. H., & Grävare Silbernagel, K. (2020). Achilles tendon cross-sectional area at 12 weeks post-rupture relates to 1-year heel-rise height. *Knee Surgery, Sports Traumatology, Arthroscopy*, *28*(1), 245–252. https://doi.org/10.1007/s00167-019-05608-x

Zellers, J. A., van Ostrand, K., & Silbernagel, K. G. (2017). Calf endurance and Achilles tendon structure in classical ballet dancers. *Journal of Dance Medicine and Science*, *21*(2), 64–69. https://doi.org/10.12678/1089-313x.21.2.64

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Fernandez, M. R., & Hébert-Losier, K. (2023). Devices to measure calf raise test outcomes: A narrative review. *Physiotherapy Research International*, e2039. https://doi.org/10.1002/pri.2039