Arch Physioter 2025; 15: 327-336
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3722
**MASTERCLASS**

# Mastering the use of hand-held dynamometry in clinical practice

**Claire Minshull**

School of Health, Leeds Beckett University, Leeds, UK; Get Back To Sport, Nottingham, UK

## ABSTRACT

Measures of muscle performance, such as strength and the rate of force development (RFD), are important for function, rehabilitation, and successful aging. Clinicians seeking to use objective measurement methods of muscle performance in support of their assessments and rehabilitation programs have many affordable dynamometry options. However, substantive differences exist between devices on important characteristics such as sampling frequency, load capacity and determining force onset, which dramatically affect the ability to obtain accurate estimates of muscle performance. The assessment environment and setup also require careful consideration. Busy clinicians are often unaware of the extent to which methodological variability and inconsistencies in testing protocols can inflate measurement error and render tests insensitive to change. Where data inform treatment and return-to-play decisions (vs. motivational aid), ensuring validity and reliability is paramount, particularly given that clinicians typically assess individual, not group performance. This is because ascertaining change or difference in intra-individual performance demands a greater level of measurement precision compared to assessing performance between groups of people.

This evidence-informed Masterclass will exemplify some of the critical technical and methodological factors that intrude on measurement accuracy. It will provide readers with the knowledge: how to critically evaluate the utility of dynamometers, answering the question, which to buy and why? How to construct assessment protocols to improve quality data collection, and how to understand what constitutes real change in performance beyond "differences" caused by measurement error.

**Keywords:** Exercise therapy, Muscle strength, Muscle strength dynamometer, Rehabilitation

## Introduction

Do you have access to a dynamometer, and are you proficient in its use? Are you aware of its sampling frequency and technical limitations? If you are considering the acquisition of a dynamometer, what are your selection criteria, and upon what evidence or rationale is your decision based?

Measures of muscle performance, such as strength and rate of force development (RFD), are important rehabilitation outcomes; are essential for successful aging and falls avoidance, and they are important indicators of performance in sports (1-3). Knowledge of muscle performance can also help to guide clinical decisions, stratify patient treatment groups and provide prognostic markers of health (4). Clinicians seeking to obtain objective measures of muscle performance in support of their assessments and rehabilitation programming have a multitude of affordable hand-held dynamometry (HHD) options to choose from. However, the assumption that generating a numerical output equates to a valid and reliable measure is misleading.

The market for HHD has expanded rapidly, offering a wide range of affordable and portable devices, often marketed with claims of high reliability and accuracy. Yet important differences in technical specifications, such as sampling frequency and force onset detection, exist. These parameters significantly influence the fidelity of performance estimates (5,6). Consideration should also be given to load capacity, particularly in settings where high force outputs are anticipated, such as during the assessment of athletes and whole-body performance, for example, the isometric mid-thigh pull (IMTP). The absence of regulatory standards or consistent validation procedures across devices compounds the challenge, leaving clinicians to navigate a landscape of devices of variable quality and unclear clinical utility.

In addition to these technical issues, methodological inconsistencies permeate clinical practice (7). Inaccurate or poorly standardized measurements risk not only being clinically uninformative but may also lead to misinformed decisions that compromise patient outcomes and quality and direction of care. Heterogenous testing protocols, even when done well, can yield variable results, and thus the establishment of "normative values" is a futile exercise.

To add to the mix, clinicians almost exclusively measure individual performances (intra-individual), such as tracking

**Received:** August 1, 2025
**Accepted:** December 23, 2025
**Published online:** December 31, 2025

**Corresponding author:**
Claire Minshull
email: minshullc@hotmail.com

progress across sessions or measuring the difference between limbs. The precision required to detect meaningful change in individuals is substantially higher than when comparing group means (8). Without an understanding of the sources and magnitude of measurement error, tests risk being statistically and clinically insensitive to physiologic change. Thus, when "differences" between test scores are observed, the clinician cannot be sure if this represents real change or is an artefact of measurement error.

This evidence-informed Masterclass aims to equip clinicians with the conceptual and practical tools necessary to critically evaluate and apply HHD in practice. It will examine the impact of technical device characteristics, methodological choices, and operator skill on measurement quality. It will also address key questions such as: What makes a device fit-for-purpose? How can protocols be constructed to minimize error and maximize sensitivity? And, how can clinicians distinguish between true performance change and variation attributable to measurement error?

Ultimately, this manuscript aims to enhance clinicians' confidence in selecting the appropriate tool/s for their needs and provide the essential components to designing robust muscle performance assessments that can meaningfully inform patient care.

## The assessment setting

### *What are we assessing; what do we want to achieve or understand?*

A growing body of research highlights the importance of muscle performance indices, such as strength and RFD, for quality of life (9,10). Current guidelines also recommend limb symmetry indices greater than 90% for a safe return to play (RTP) (11), with evidence showing that deficits in muscle performance may increase the risk of injury (12,13). As a result, the use of dynamometry to obtain objective measures that support rehabilitation and clinical decision-making across multiple patient groups is becoming increasingly common in contemporary practice.

In clinical practice, two testing scenarios predominate: 1 - evaluation of both limbs (affected and unaffected) within a single session to calculate a Limb Symmetry Index (LSI), which is the performance of the affected/injured limb expressed as a percentage of the non-affected/uninjured limb; 2 - tracking performance metrics over time to gauge the effectiveness of rehabilitation and conditioning programmes. Each testing application (e.g., within session [intra-session] and/or between days [inter-day]) involves assessing an individual's performance and comparing results between limbs, over time, or both.

This type of intra-individual testing requires a more stringent level of measurement precision than testing a group of people (8). In empirical research, methodologies document the number of recruited participants to each experimental or control group. Calculated using data from prior and/or pilot studies, participant numbers are manipulated to achieve the required level of statistical power to detect the level of change or difference that the researchers want to see (14). These calculations take into account factors that influence measurement precision, such as variability.

In contrast, when conducting intra-individual assessments where the focus is on detecting change of difference within a single person, the measurement precision must be sufficient to distinguish true change from random variability or measurement error in that single person. Without this heightened precision, any observed differences risk being indistinguishable from normal performance fluctuations and experimental inconsistencies, undermining the clinical or experimental utility of the assessment (8).

A comprehensive review of the clinimetric properties of HHD is beyond the scope of this Masterclass; however, it is important to acknowledge that numerous studies have demonstrated its strong reliability across a range of muscle groups and clinical populations. Readers are directed to systematic reviews such as Chamorro et al. (15) and empirical studies, such as Mentiplay et al. (16), for further detailed evidence on the reliability and measurement characteristics of HHD across different testing protocols and populations.

*In summary, the majority of assessments within clinical practice are to detect change in intra-individual performance over time or between limbs and to do this, precision of measurement is important.*

## Which indices of performance?

Muscle strength, experimentally termed peak force (PF), and RFD are two key indices of muscle performance, which offer critical insight into neuromuscular function and have been assessed within neuromuscular physiology settings for decades (17,18). Recent developments in HHDs have enabled the quick and easy procurement of PF and RFD data, with applications in both clinical and athletic or performance-focused environments. When appropriately obtained and reported, these data can help identify performance deficits, guide individualized rehabilitation programmes and RTP decisions, and quantify improvements resulting from therapeutic interventions. Moreover, they may also provide estimates of dynamic stabilization capabilities during joint loading (19-21), with deficits potentially indicating elevated injury risk or incomplete recovery (22).

### *Peak Force (PF)*

Peak force (PF) is the maximal voluntary contractile force of a muscle in a single contraction, typically measured in newtons (N). With a robust testing configuration and reproducible methods, PF is a relatively easy metric to obtain and, by comparison to other indices of neuromuscular function, can be relatively impervious to minor deviations in things like dynamometer technical specifications (23). Participants are instructed to produce a maximal voluntary contraction (MVC) as "hard and as fast" as they can, and the focus of the measurement is on the PF output regardless of time.

### *Rate of Force Development (RFD)*

Rate of force development (RFD) is a measure of explosive force production, and the focus of the measurement is on the rapidity of force production, often irrespective of the peak. RFD can be reported in several different ways; the

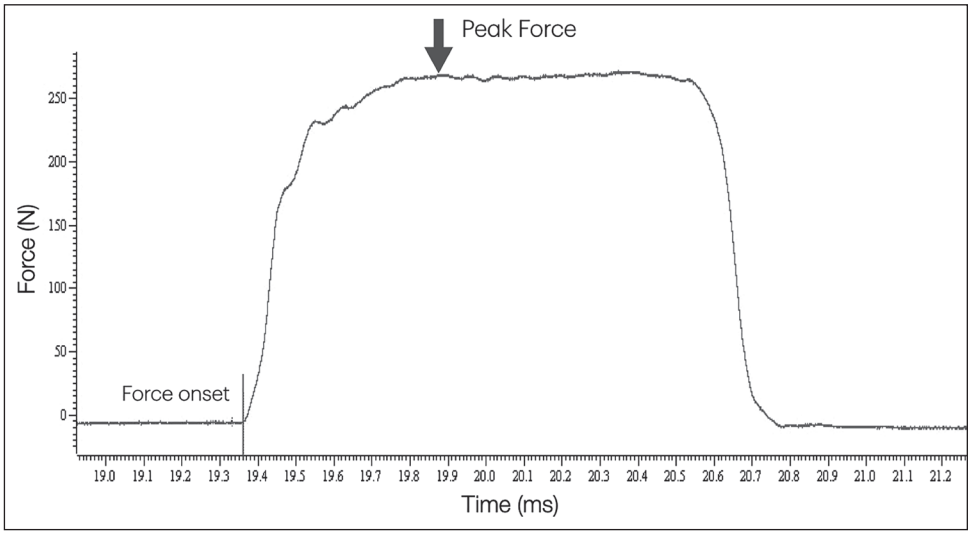


**FIGURE 1 -** Individual quadriceps force-time curve, 6 months post ACL reconstruction. Sampling Frequency: 2000 Hz. Unpublished data.

most common method involves calculating the gradient of the slope of the force-time (F-T) curve within specified time epochs from the onset of force, resulting in a newtons per second ($N \cdot s^{-1}$) value. For example, the average force produced over 0-50 ms and 0-100 ms captures early-phase RFD and 100-150 ms and 100-200 ms reports late-phase RFD (24,25). Alternative methods include "peak RFD," which is the steepest slope of the force-time signal, no matter where it occurs, and "zero to peak," which provides an average RFD across the whole force-time curve. The index of RFD involves rapid changes in force over extremely short time frames, and as such, minor deviations in methodology and differences in dynamometer technical specifications can heavily influence the accuracy of data (23,26).

It is beyond the scope to provide a detailed review of the rationale for each method of testing and RFD index selection; however, Figure 2 simply exemplifies the consequences of RFD index selection on performance interpretation.

Zero-to-peak, often reported with contemporary HHDs, can provide a general marker of RFD performance; however, it lacks temporal (time-based) precision. It may substantially underestimate RFD by averaging over a broader time interval, thus masking rapid changes in force output (27). Furthermore, zero-to-peak is heavily influenced by the threshold that determines force onset. Insensitive devices may use a large threshold, which requires participants to produce significant force before recording commences.

The zero-to-peak index is less suitable when rapid force production is the primary focus. In contrast, analyzing smaller specific segments of the F-T curve provides more detailed insights, but doing so requires accurate data collected over short sampling windows. This depends heavily on the characteristics of the hand-held dynamometer (HHD) and associated software, particularly the use of a high sampling frequency (see below).

*In summary, clinicians seeking to accurately assess muscle performance metrics, PF and RFD should recognize that both index selection and dynamometer characteristics will influence performance interpretation.*

## Dynamometers and technical specifications

### *Selecting the right tool for the job*

In many respects, the proliferation of the number of commercially available hand-held and other dynamometers is a positive development. Historically, the objective measurement of muscle force production was largely confined to research laboratories or clinical environments with access to high-cost isokinetic or custom-built dynamometry systems.

However, the increased accessibility and commercialization of devices has, in some cases, been accompanied by insufficient attention to device technical specifications and methodological rigour, for example, slow sampling frequencies and variable testing methods, respectively. Consequently, clinicians may adopt these tools without critical evaluation of their measurement properties or adequate training in their use. This practice raises important concerns regarding the reliability, validity, and overall utility of the data being collected.

### *Types of hand-held dynamometers (HHDs)*

Many different types of dynamometers exist; the focus here is on HHDs, albeit many of the technical and methodological considerations apply to other forms of assessment.

Hand-held dynamometers are portable devices that measure muscle force production. Most new generations of HHDs have smart device app software and offer PF and some index of RFD as an output. Devices can be categorized according to the method of force application. Compression-type dynamometers assess force when the user pushes against the device, for example, during isometric contractions where the dynamometer is placed between the limb and a fixed surface, tester, or secured using strapping (Fig. 3a). Tension-type dynamometers measure force through pulling, often via straps or cables attached to the limb and anchored to a fixed point (Figs 3b and c).

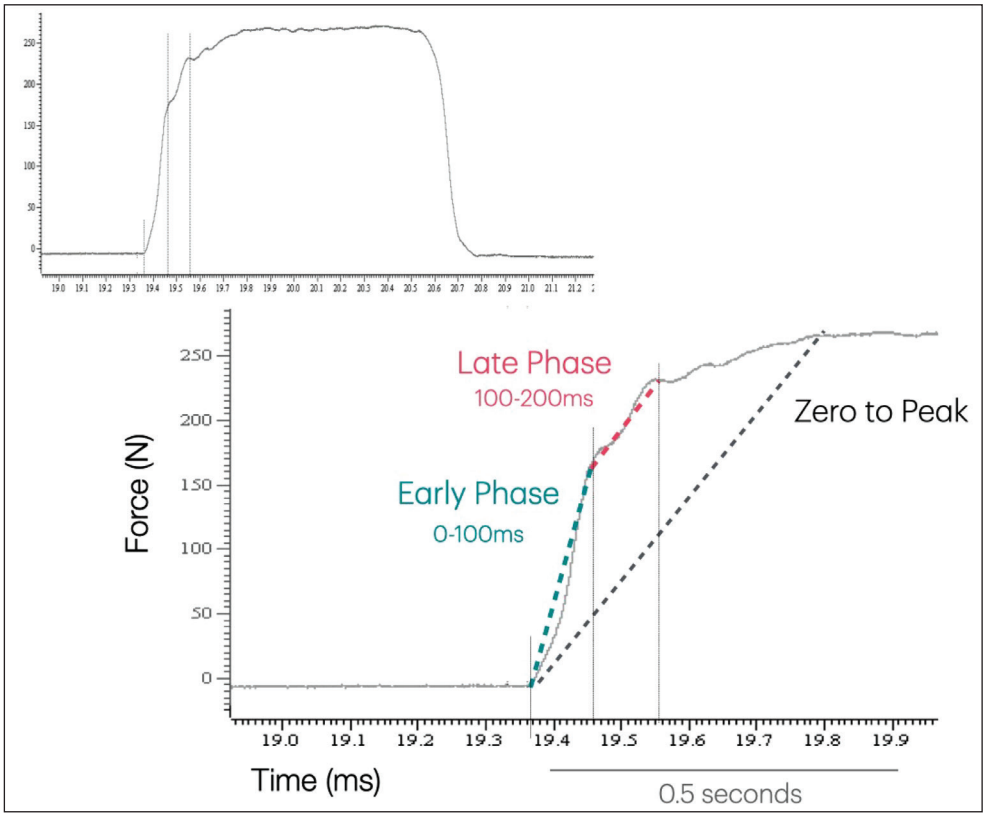


**FIGURE 2 -** Individual quadriceps force-time curve, 6 months post ACL reconstruction, illustrating three methods of reporting RFD. Sampling Frequency: 2000 Hz. Unpublished data.


**FIGURE 3 -** Assessment of quadriceps muscle force using HHD. 3a: Compression-type dynamometer secured using strapping (VALD DynaMoLite); 3b: Tension-type dynamometer secured using strapping (Kinvent Pull). Plinth-based testing is pragmatic in clinical settings; there may be some compromise in the dynamometer and strapping setup. 3c: Adaptation of the knee extensor machine to secure tension-type dynamometer (VALD DynaMoPlus).

***Important dynamometer technical specifications***

When selecting a measurement tool, financial outlay is important; however, priority should be given to the technical specifications of the selected tool to enable the accurate measurement of the data the clinician seeks to procure.

*Sampling frequency*

A critical aspect of dynamometry is sampling frequency. This refers to the number of data points of a signal collected per second, and it is measured in hertz (Hz). 1 Hz is one sample per second, 100 Hz is 100 samples per second, etc.

Sampling frequency is important in the assessment of muscle force production, especially for events that occur quickly, like RFD (5). From a rehabilitation and training perspective, different phases of RFD are driven by different physiologic events. Early phase RFD is influenced mainly by neural events, such as neural drive and late phase by muscular and morphological characteristics, such as muscle cross-sectional area (28). Should there be a need to create

specific rehabilitation and training programmes to target different parts of the F-T curve, higher sampling frequencies are required to record accurately over small sampling windows (<100 ms). Too slow and information will be lost, resulting in poor sensitivity and inaccurate representation of RFD.

To exemplify, in well-conditioned people and in non-athletes who are habituated to isometric testing, PF can occur within 300ms of force onset—one third of a second (29). In other populations, it's not dissimilar, for example, individuals with mild knee osteoarthritis (350 ms: (30)). Figure 4a shows a section of the F-T curve of a knee extensor assessment of an individual's non-injured limb prior to ACL reconstruction (time to PF 350 ms). The original data is sampled at 2000 Hz; the lines overlaid represent an estimate of what the data may look like when obtained at a sampling frequency of 10 Hz, which is the sampling frequency of the EasyForce (Meloq) HHD at the time of writing. With a sampling frequency of just 10 Hz, only three data points are obtained across a 300 ms time period, contrasted to 600 data points with a sampling frequency of 2000 Hz. This results in a different picture of the F-T curve. A low sampling frequency that captures only a few data points during rapid force production may underestimate the true change in muscle force over time and therefore distort the calculated RFD. In contrast, a high sampling frequency (e.g., 2000 Hz) provides much greater temporal resolution to more accurately represent the rapid changes in force that can define the true RFD. Users must also be aware that slow sampling frequencies, coupled with high threshold force onsets, can obscure the commencement of force production.

*What is an adequate sampling frequency?*

The sampling frequencies available across commercially available HHDs range from 10 Hz (EasyForce [Meloq]) to 1000 Hz (Kinvent Push/Pull [Physio Kinvent]) and recently 1200 Hz (DynamoMax [VALD Performance]). The minimal required sampling frequency will depend on the data acquisition and its importance.

The index PF is fairly resilient to sampling frequency deviations; quality data can be obtained at a sampling rate of 100 Hz with minimal differences with higher sampling rates (6). For lower sampling rates, information such as the true PF may be underrepresented during explosive contractions, and users may wish to consider maintaining the maximal contraction for 2 or 3 seconds to obtain a reliable estimate within the available sampling window.

It is widely regarded that much higher sampling frequencies are required to accurately measure RFD, especially in the early phase of the force-time curve and from explosive contractions. Several data-driven studies suggest a minimum sampling frequency of 500 Hz for PF and RFD (e.g., 6,31), with even higher rates (1000 Hz/>) recommended where high fidelity data is required. Sampling frequencies below 1000 Hz have been shown to underestimate peak and early-phase RFD due to insufficient temporal resolution to capture the steep initial rise in force (5,6). It is important to note that these findings are based on data collected using isokinetic dynamometry or wired force plates, methods that differ from most HHDs, whereby data are transmitted via Bluetooth. Wireless (Bluetooth) transmission may lead to data packet loss (32), potentially affecting the accuracy of RFD calculations. If RFD is a critical metric for clinical decision-making, practitioners would be prudent to use a device with a sampling frequency of at least 1000 Hz to ensure sufficient data quantity and quality. Figure 5 shows a simple flow chart to help guide clinicians on the required sampling frequency for their data acquisition needs.

Some manufacturers offer exportation of raw data in CSV format. For the more technically-minded clinician, this enables manual calculation of RFD across the F-T curve by using customized scripts.

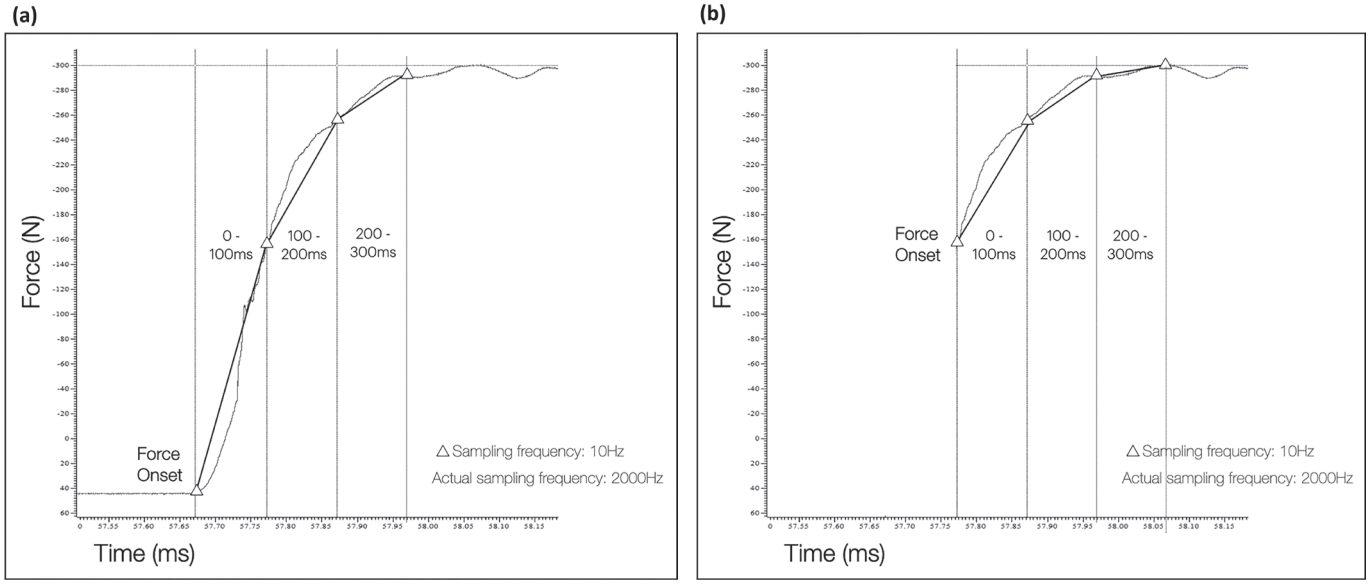


**FIGURE 4 -** a. Individual data from Minshull et al. (39). Quadriceps force-time curve, pre-ACL reconstruction, non-injured leg. Δ Example data captured at 10 Hz. b. Individual data from Minshull et al. (39). Quadriceps force-time curve, pre-ACL reconstruction, non-injured leg. Δ Example data captured at 10 Hz with a 20 N force onset threshold.

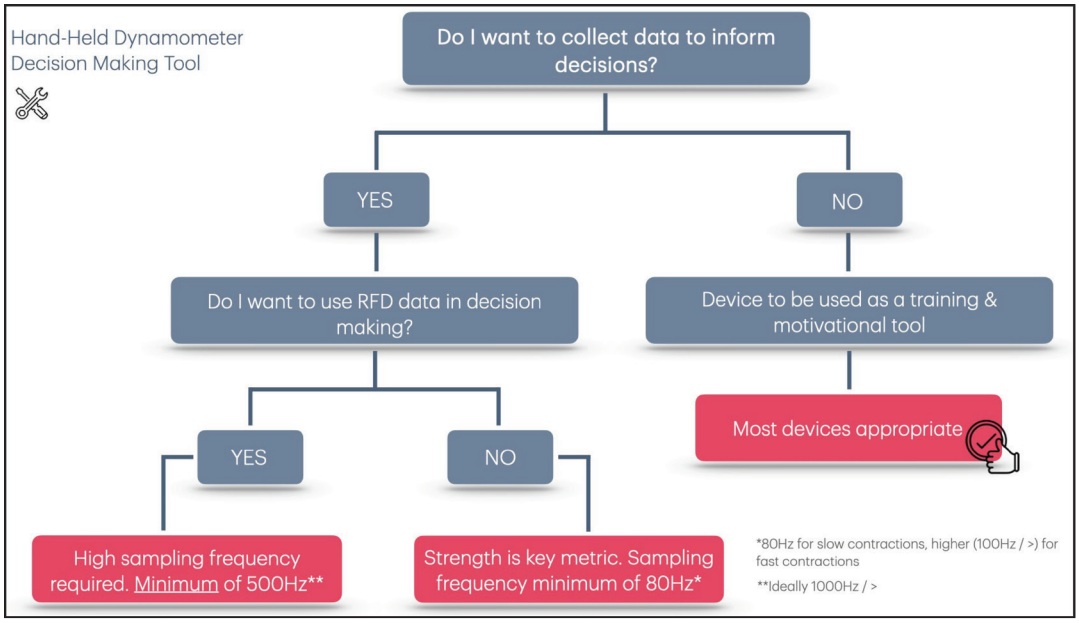


**FIGURE 5 -** HHD decision-making tool, reproduced with permission, Get Back To Sport 2025.

***Reliability does not necessarily confer validity***

A recent study (33) reported good to excellent inter- and intra-session reliability of the Microfet2 (Hoggan Scientific) in measuring RFD, supporting their recommendation of its use for assessing knee extensor RFD. The Microfet2 samples at 100 Hz, and thus the recommendation appears at odds with the literature cited above. However, reliability does not confer validity (i.e., the extent to which the test measures what it claims to measure). The slow sampling frequency captures only a small number of data points during the critical early phase of contraction. Under-sampling blurs the force-time curve, masking rapid changes in muscle force and reducing the device's ability to detect small but meaningful differences between trials (5), thereby compromising its validity. Consequently, between-trial differences may appear smaller than they actually are, leading to inflated within-trial consistency and potentially misleading conclusions about reliability.

***Force Onset***

The threshold for determining force-onset with HHD software isn't always adjustable by end-users; however, it is important to be aware of how this is determined. High absolute values signalling force onset may compromise accurate evaluation of RFD, particularly in smaller muscles. A 20 N onset threshold for registration of muscle force production, with a shoulder internal rotation strength of, for example, 100 N, represents 20% of the total force production. In this instance, much of the early phase of the F-T curve could be missed, precluding the accurate determination of RFD.

Some manufacturers recommend a pre-tensioning setup, whereby the participant produces a low-level contraction to remove the slack from the securing straps prior to MVC. Compliance, or slack within a measurement setting, should be removed to improve assessment accuracy; however, pretensioning the muscle prevents an accurate picture of early-phase RFD as the contraction does not start from a resting level. Furthermore, depending on the strength of contraction, the prior contraction may potentiate the preceding contraction, thereby enhancing force output compared to that elicited from a relaxed state (34).

*In summary, the clinician must determine which data they want to obtain to aid their decision-making in advance of the procurement of a device. Secondarily, evaluation of the technical specifications of the tools available will determine whether or not they are fit for purpose. Readers are referred to Online for further HHD comparison resources.*

## Measurement error

Testing performance is important, but not as important as testing performance accurately.

***Understanding Measurement Error***

The greater the error, or variability of a measurement, the less confident we can be about the score representing the person's true value (35), and the less we can rely on it to help make data-informed decisions, for example, do we keep going with the rehabilitation or has that person's strength increased by the target amount?

To illustrate, Figure 6 shows a hypothetical scenario in which quadriceps muscle strength was assessed using an HHD across multiple visits during the course of a rehabilitation programme. The clinician is interested in monitoring the patient's progress over time following the implementation of a targeted muscle strengthening intervention. A central question arises: at what point does a meaningful improvement in muscle strength occur?

Although a calculated change of +8% in force output relative to baseline suggests an improvement in strength at week 3, this observation cannot be confirmed with confidence

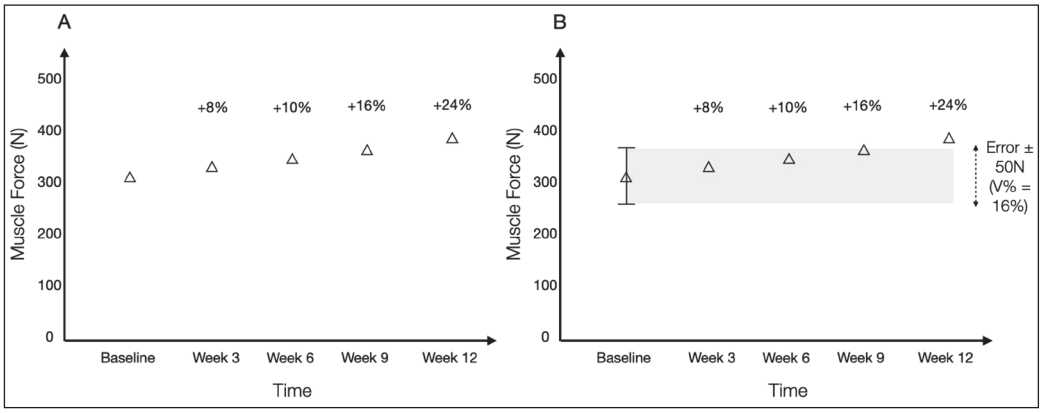


**FIGURE 6 -** Hypothetical quadriceps PF data collected at baseline and four additional times across a 12-week rehabilitation program. A: showing percentage change relative to baseline value; B: greyed area showing measurement error: coefficient of variation [V%] (multiplied by 1.96 to represent 95% of the variance) relative to baseline value.

without knowledge of the measurement error. For instance, if testing procedures were inconsistent, involving non-standardized participant positioning, inadequate stabilization of the dynamometer, or a lack of formal protocol, then measurement error is likely to intrude considerably upon the precision of measurement and the inference of change.

Measurement error can be calculated in a number of ways, and readers are directed to the excellent paper by Atkinson and Nevill (8) for a comprehensive discussion of the various methods within the context of assessment of the individual, including, standard error of the measurement (SEM), Bland-Altman, coefficient of variation (V%), and intra-class correlation (ICC). It is unwise to rely on a sole index to describe measurement error; however, the simply calculable index: the coefficient of variation (V%), has been used here solely to illustrate the practical impact of measurement error in this hypothetical example.

The V% is an estimate of measurement error. Calculated as the standard deviation of a dataset as a percentage of its mean, it represents a method by which to express the variability of a single measurement, regardless of the unit of measurement (35).

$$V\% = \left( \frac{\text{Standard Deviation (pooled)}}{\text{Mean}} \right) \times 100$$

Here measurement error was calculated by the experimenters to be approximately ±50 N, or ±16.6% of the baseline strength value. A sample of >40-100 measures is generally recommended to provide stable error estimates (36). Only 68% of the error is represented with standard V% calculations. To capture 95% of the variance, the sample standard deviation should be multiplied first by 1.96 before being expressed as the coefficient of variation (V%) (8).

Consequently, only changes exceeding this threshold can be interpreted as genuine improvements. Based on this criterion, meaningful gains in muscle strength are not suggested until approximately week 9, and most likely week 12 into the intervention.

This example highlights how measurement error can undermine the utility of strength assessments. It also underscores the importance of understanding and, where possible, minimizing the sources of error, or at the very least being aware of the specific error margins associated with a given test. These considerations are equally relevant for:

- Inter-limb comparisons and when aiming to achieve, for example, an LSI of 90%. In such cases, it is essential to determine whether the error margins are adequate to detect a 10% inter-limb difference and, if not, efforts should be directed toward reducing measurement error (see next section).
- Determining minimal clinical important difference (MCID). For example, the MCID of grip strength in patients treated by volar locking plate fixation for a distal radius fracture has been reported as a decrease of 6.6 kg (63.7 N), or 19.5% (37). In situations of excessive measurement error, MCID may be undetectable in the individual patient.
- Comparing individual scores to normative values, where such data are available for the relevant population. Normative values, if representative of the individual tested and obtained using comparable procedures, can be useful for categorizing individuals as above or below average, setting rehabilitation goals, and, for example, identifying those at risk of sarcopenia (38). However, confident categorization of individual performance and subsequent clinical decision-making based on such data requires an understanding of the measurement error to ascertain whether observed deviations from population norms reflect true differences in muscle performance rather than normal variability or measurement noise.

### *Minimizing Measurement Error*

Measurement errors are typically categorized into systematic and random. Systematic errors are consistent, repeatable inaccuracies often caused by flaws in measurement instruments or procedures that cause all measurements to deviate in the same direction from the true value. Random errors are unpredictable fluctuations that arise from uncontrollable variables, such as natural biological variation or inconsistent effort, leading to measurements that scatter around the true value without a consistent pattern (8). To facilitate pragmatic improvements in measurement accuracy, it may be more constructive to understand the underlying sources from which these errors originate. As such, errors will be categorized into technical and biological sources.

### *Technical error*

Technical error encompasses variability in equipment, testing protocol, and setup procedures.

*Equipment*

If a force of, for example, exactly 10 kg (98.1 N) is applied repeatedly to an HHD under controlled conditions, a reliable device should yield nearly identical readings across trials. While minor deviations are expected, the degree of variability should be minimal. Most commercially available HHDs are designed to meet this level of precision, and the calibration data to verify this should be accessible from the manufacturer.

*Protocol*

Assuming a reliable, sensitive and calibrated device, most likely the largest source of technical error will arise from the protocol and test setup, that is to say, being able to replicate exactly the same procedures each time. This includes fundamental things like:

1. minimizing extraneous movement (e.g., ensuring the person's limb is secured)
2. standardizing instructions (and motivational prompts)
3. standardizing joint and HHD position
4. ensuring maximal effort is understood and performed
5. standardizing warm-up procedures
6. removing the "slack"/compliance from the measurement system if using strapping to secure the HHD

These might appear basic; however, each can influence the variability of force output substantially and confound the utility of a test. To exemplify, variation of knee flexion angle of just 10° resulted in >20% difference in isometric quadriceps strength (39). Clinicians can minimize changes in joint positioning within the test by paying strict attention to points 3 and 6 on each assessment occasion.

Other measures, such as RFD, can vary substantially with subtle differences in verbal cuing. Instructing someone to contract as "*hard and fast*" vs., "*as fast*" as possible can underestimate explosive performance by over 30% and may double measurement error (26). Clinicians can minimize performance variation by being focused on points 2 and 4 and providing sufficient opportunity for practice. In situations where RFD and PF data are required, but separation of assessments is not possible (40), then adequate data can be procured from instructions cueing "*hard and fast*" contractions, as long as participants are appropriately practised (5,41).

### *Biological error*

Biological error refers to the variability of 'the person' and their responses, not to be confused with actual change caused by training or deconditioning.

*The assessor*

Within a hand-held setup, a significant source of error is the tester themselves (42,43). Assessor strength, fatigue and experience can all contribute unnecessary and additional amounts of error into assessments. Tethering the HHD securely to something immovable (see Figs 3a and b) is recommended to eliminate this effect.

On repeated tests, using the exact same setup, even when the device is tethered, strength scores will still vary a little; people are naturally variable. However, good methodological design should minimize this.

*Learning*

Learning effects refer to when people become better at the test with practice. Here, the assessor will observe continually increasing scores on repeated tests.

To minimize learning effects:

– Ensure the person/patient is familiarised with the test. Ensure they understand what to do by clearly and concisely explaining the test and the requirements. Check their understanding.
– Habituate the person to the test. Allow a few practice attempts before you start the test so that the person is able to constantly deliver the type of muscle contractions required. To save time, this can be incorporated into the end stages of a warm-up.

*Fatigue*

The effects of acute muscle fatigue will be visible by an observed decline in performance over repeated efforts. The assessor may see reduced PF and/or RFD scores on sequential efforts associated with insufficient recovery.

To minimize acute muscle fatigue

– Provide 10-30 s rest between maximal contractions to enable recovery. Observe the data in real time and provide more rest if the force profile is declining.
– Allow approximately 2 minutes of rest after every 5 maximal contractions. Multiple maximal efforts may be required to obtain scores of true maximal performance in individuals new to testing or those being assessed following recovery from injury. Split the efforts into sets and provide adequate time for recovery in between.

## Summary

### *Refining Testing Protocols*

Designing an effective protocol for assessment of muscle performance using HHD begins with clarity on assessment and clinical intent. Protocols must be aligned with specific goals, whether it is identifying muscular asymmetries, monitoring rehabilitation progress, or evaluating readiness to return to activity. The purpose will determine the schedule of testing and which performance indices are most relevant.

Each index of performance requires distinct methodological considerations. For example, accurate assessment of RFD, particularly in the early phase, demands high sampling rates and attention to force onset, while PF may allow for slightly more flexibility in setup but still requires consistent stabilization and instruction.

From a resource standpoint, selecting the appropriate device is crucial. Clinicians should prioritize dynamometers with sufficient force capacity, adequate sampling frequency, and the ability to export raw data where needed. Devices should also allow for consistent tester positioning and secure fixation.

Minimizing error is central to robust protocol design. Simple but influential methodical issues such as strap compliance, inconsistent participant positioning and variable participant effort can compromise data integrity. Strategies such as standardized patient, HHD and limb positioning, use of external fixation, habituation trials, and consistent verbal cues can help mitigate these sources of variability. Where early phase RFD is important and where software settings like force-onset thresholds are not user-adjustable, clinicians should be aware of how these values are defined, as they can influence force measurements.

In summary, effective protocol design for HHD involves aligning the method with clinical goals, selecting appropriate performance indices, choosing reliable equipment, and controlling for known error sources. Clinicians are encouraged to pilot their own protocols and estimate within-session reliability (e.g., using the coefficient of variation with ≥10 participants), not as a definitive evaluation of measurement quality, but as a practical introduction to how error may infiltrate personal testing procedures. This process is intended to promote critical reflection on protocol design and to inform subsequent refinement or more comprehensive reliability assessment, balancing clinical feasibility with measurement integrity.

## Disclosures

**Conflict of interest:** The author declares that they have no financial relationships or affiliations with any companies or organizations that could be perceived as a potential conflict of interest in relation to this work. However, the author has received products free of charge from multiple companies for research and/or evaluation purposes. These companies had no role in the design, execution, analysis, or writing of this manuscript.

**Financial support:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

1. Reid KF, Fielding RA. Skeletal muscle power: a critical determinant of physical functioning in older adults. Exerc Sport Sci Rev. 2012;40(1):4-12. CrossRef PubMed
2. Lee J. Associations between handgrip strength and disease-specific mortality including cancer, cardiovascular, and respiratory diseases in older adults: a meta-analysis. J Aging Phys Act. 2020;28(2):320-331. CrossRef PubMed
3. Turpeinen JT, Freitas TT, Rubio-Arias JÁ, et al. Contractile rate of force development after anterior cruciate ligament reconstruction—a comprehensive review and meta-analysis. Scand J Med Sci Sports. 2020;30(9):1572-1585. CrossRef PubMed
4. Gil S, Jacob Filho W, Shinjo SK, et al. HCFMUSP COVID-19 Study Group. Muscle strength and muscle mass as predictors of hospital length of stay in patients with moderate to severe COVID-19: a prospective observational study. J Cachexia Sarcopenia Muscle. 2021;12(6):1871-1878. CrossRef PubMed
5. Maffiuletti NA, Aagaard P, Blazevich AJ, et al. Rate of force development: physiological and methodological considerations. Eur J Appl Physiol. 2016;116(6):1091-1116. CrossRef PubMed
6. Thompson BJ. Influence of signal filtering and sample rate on isometric torque – time parameters using a traditional isokinetic dynamometer. J Biomech. 2019;83:235-242. CrossRef PubMed
7. Stone CA, Nolan B, Lawlor PG, et al. Hand-held dynamometry: tester strength is paramount, even in frail populations. J Rehabil Med. 2011;43(9):808-811. CrossRef PubMed
8. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-238. CrossRef PubMed
9. Wang DXM, Yao J, Zirek Y, et al. Muscle mass, strength, and physical performance predicting activities of daily living: a meta-analysis. J Cachexia Sarcopenia Muscle. 2020;11(1):3-25. CrossRef PubMed
10. Khodadad Kashi S, Mirzazadeh ZS, Saatchian V. A systematic review and meta-analysis of resistance training on quality of life, depression, muscle strength, and functional exercise capacity in older adults aged 60 years or more. Biol Res Nurs. 2023;25(1):88-106. CrossRef PubMed
11. Lynch AD, Logerstedt DS, Grindem H, et al. Consensus criteria for defining 'successful outcome' after ACL injury and reconstruction: a Delaware-Oslo ACL cohort investigation. Br J Sports Med. 2015;49(5):335-342. CrossRef PubMed
12. Achenbach L, Laver L, Walter SS, et al. Decreased external rotation strength is a risk factor for overuse shoulder injury in youth elite handball athletes. Knee Surg Sports Traumatol Arthrosc. 2020;28(4):1202-1211. CrossRef PubMed
13. Namazi P, Zarei M, Hovanloo F, et al. The association between the isokinetic muscle strength and lower extremity injuries in young male football players. Phys Ther Sport. 2019;39:76-81. CrossRef PubMed
14. Serdar CC, Cihan M, Yücel D, et al. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med (Zagreb). 2021;31(1):010502. CrossRef PubMed
15. Chamorro C, Arancibia M, Trigo B, et al. Absolute reliability and concurrent validity of hand-held dynamometry in shoulder rotator strength assessment: systematic review and meta-analysis. Int J Environ Res Public Health. 2021;3;18(17):9293. CrossRef
16. Mentiplay BF, Perraton LG, Bower KJ, et al. Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: a reliability and validity study. PLoS One. 2015; 28;10(10):e0140822. CrossRef
17. Norman RW, Komi PV. Electromechanical delay in skeletal muscle under normal movement conditions. Acta Physiol Scand. 1979;106(3):241-248. CrossRef PubMed
18. Ismail HM, Ranatunga KW. Isometric contractions of normal and spastic human skeletal muscle. Muscle Nerve. 1981;4(3):214-218. CrossRef PubMed
19. Buckthorpe M, Roi GS. The time has come to incorporate a greater focus on rate of force development training in the sports injury rehabilitation process. Muscles Ligaments Tendons J. 2018;7(3):435-441. CrossRef PubMed
20. Domire ZJ, Boros RL, Hashemi J. An examination of possible quadriceps force at the time of anterior cruciate ligament injury during landing: a simulation study. J Biomech. 2011;44(8):1630-1632. CrossRef PubMed
21. Minshull C, Eston R, Bailey A, et al. Repeated exercise stress impairs volitional but not magnetically evoked electromechanical delay of the knee flexors. J Sports Sci. 2012;30(2):217-225. CrossRef PubMed
22. Grindem H, Snyder-Mackler L, Moksnes H, et al. Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: the Delaware-Oslo ACL cohort study. Br J Sports Med. 2016;50(13):804-808. CrossRef PubMed

23. White MS, Graham MC, Janatova T, et al. Effect of sampling rate, filtering, and torque onset detection on quadriceps rate of torque development and torque steadiness. Sensors (Basel). 2024;24(13):4250. CrossRef PubMed
24. Andersen LL, Andersen JL, Zebis MK, et al. Early and late rate of force development: differential adaptive responses to resistance training? Scand J Med Sci Sports. 2010 Feb;20(1):e162-9. CrossRef
25. D'Emanuele S, Maffiuletti NA, Tarperi C, et al. Rate of force development as an indicator of neuromuscular fatigue: a scoping review. Front Hum Neurosci. 2021;15:701916. CrossRef PubMed
26. Jaafar H, Lajili H. The influence of verbal instruction on measurement reliability and explosive neuromuscular performance of the knee extensors. J Hum Kinet. 2018;65(1):21-34. CrossRef PubMed
27. Andersen LL, Aagaard P. Influence of maximal muscle strength and intrinsic muscle contractile properties on contractile rate of force development. Eur J Appl Physiol. 2006;96(1):46-52. CrossRef PubMed
28. Andersen LL, Andersen JL, Zebis MK, et al. Early and late rate of force development: differential adaptive responses to resistance training? Scand J Med Sci Sports. 2010;20(1):e162-e169. CrossRef PubMed
29. Aagaard P, Simonsen EB, Andersen JL, et al. Increased rate of force development and neural drive of human skeletal muscle following resistance training. J Appl Physiol. 2002;93(4):1318-1326. CrossRef PubMed
30. Suzuki Y, Iijima H, Nakamura M, et al. Rate of force development in the quadriceps of individuals with severe knee osteoarthritis: a preliminary cross-sectional study. PLoS One. 2022;17(1):e0262508. CrossRef PubMed
31. Dos'Santos T, Jones PA, Kelly J, et al. Effect of sampling frequency on isometric midthigh-pull kinetics. Int J Sports Physiol Perform. 2019;14(4):525-530. CrossRef PubMed
32. Tipparaju VV, Mallires KR, Wang D, et al. Mitigation of data packet loss in Bluetooth low energy-based wearable healthcare ecosystem. Biosensors (Basel). 2021;11(10):350. CrossRef PubMed
33. Yona T, Fischer AG. Assessing the knee extensor's rate of force development using a fixed hand-held dynamometer: an inter and intrasession reliability study. J Sport Rehabil. 2024;34(4):443-448. CrossRef PubMed
34. Sale DG. Postactivation potentiation: role in human performance. Exerc Sport Sci Rev. 2002;30(3):138-143. CrossRef PubMed
35. Atkinson G. What is this thing called measurement error? In: Kinanthropometry VIII: Proceedings of the 8th International Conference of the International Society for the Advancement of Kinanthropometry (ISAK). London: Taylor & Francis; 2003. P.3–13. CrossRef
36. Mokkink LB, Eekhout I, Boers M, et al. Studies on reliability and measurement error of measurements in medicine – from design to statistics explained for medical researchers. Patient Relat Outcome Meas. 2023;14(14):193-212. CrossRef PubMed
37. Kim JK, Park MG, Shin SJ. What is the minimum clinically important difference in grip strength? Clin Orthop Relat Res. 2014;472(8):2536-2541. CrossRef PubMed
38. Landi F, Calvani R, Martone AM, et al. Normative values of muscle strength across ages in a 'real world' population: results from the longevity check-up 7+ project. J Cachexia Sarcopenia Muscle. 2020;11(6):1562-1569. CrossRef PubMed
39. Minshull C, Rees D, Gleeson NP. Joint angle affects volitional and magnetically-evoked neuromuscular performance differentially. J Electromyogr Kinesiol. 2011;21(4):672-677. CrossRef PubMed
40. Hannah R, Minshull C, Buckthorpe MW, et al. Explosive neuromuscular performance of males versus females. Exp Physiol. 2012;97(5):618-629. CrossRef PubMed
41. Minshull C, Gallacher P, Roberts S, et al. Contralateral strength training attenuates muscle performance loss following anterior cruciate ligament (ACL) reconstruction: a randomized-controlled trial. Eur J Appl Physiol. 2021;121(12):3551-3559. CrossRef PubMed
42. Wadsworth C, Nielsen DH, Corcoran DS, et al. Interrater reliability of hand-held dynamometry: effects of rater gender, body weight, and grip strength. J Orthop Sports Phys Ther. 1992;16(2):74-81. CrossRef
43. Beshay N, Lamb P, Murrell G. The reliability of fixed versus hand-held dynamometer for shoulder strength testing. J Sci Med Sport. 2010;13:e34-e35. CrossRef