This paper provides an overview of the two most common methods health care providers use to assess strength in both clinical and athletic settings: manual muscle testing and hand-held dynamometry. The validity and reliability of each method is discussed along with the clinical implications for their use. Considering hand-held dynamometry’s ease of use, portability, cost, and compact size, compared with isokinetic devices this instrument can be regarded as a reliable and valid tool for muscle strength assessment in a clinical or athletic setting.

Kevin Robinson, PT, DSc, OCS
Professor, Director of Motion Analysis
Belmont University, School of Physical Therapy

Christi Williams, DPT, OCS, Cert. MDT
Associate Professor
Belmont University, School of Physical Therapy
Current Method of Clinical Strength Assessment

The most common method a clinician uses to assess the strength of a patient is by way of manual examination, more specifically, a manual muscle test. A manual muscle test (MMT) is a method of rating a patient’s strength in which the examiner applies a force to the patient while the patient tries to resist that force. Strength measurements obtained by a MMT are most commonly documented on a 0-5 ordinal scale. While standardized procedures exist to decrease the subjectivity of such a manual examination, such as consistent patient positioning and location in which resistance is applied, there is still a significant amount of subjectivity that exists when establishing a grade for a MMT, which ultimately decreases the overall reliability of the exam.

When looking at the intra-rater reliability of a MMT, one study found the reliability to vary from as much as 0.65 to 0.93 when testing boys ages 5 to 15 years of age who were diagnosed with Duchenne muscular dystrophy. Beasley found that therapists could not distinguish among knee extensor forces that varied by as much as 25% with a manual muscle test examination. Beasley also found that therapists graded patients as “normal”, or a 5/5, when in actuality, the patient had lost as much as 50% of that muscle’s ability to produce tension. Frese et al looked at the inter-rater reliability of a middle trapezius test and a gluteus medius test and only found 28%-45% agreement between therapists for the same grade. They concluded that, “using a manual muscle test to make accurate clinical assessments of patient status is of questionable value.”

Manual muscle testing also lacks sensitivity. One study found that the sensitivity of a MMT ranged from 62.9 to 72.3%. In an effort to increase sensitivity to detect smaller changes in strength, a +/- scale was developed, however with a greater number of categories in which to grade a person’s strength, the subjectivity of the examination tends to increase which further leads to a decrease in overall reliability.

Other issues with manual muscle testing include the fact that the measurements are subjective in nature because they are based on the judgment of the examiner. This lack of objectivity decreases the validity of the test. The inter-rater reliability tends to decrease especially when differentiating between scores of a 4/5 and a 5/5. There are several factors which may influence the inter-rater reliability, many of which are characteristics of the examiner him/herself. Such characteristics include, but are not limited to, the strength of the examiner, as well as his/her height, weight, gender, age, and also the examiner’s overall experience in testing. When a patient is able to over-power the examiner, by definition, that patient would be graded a 5/5 because the therapist could not “break” the patient’s hold, however, if the therapist is not stronger than the patient or cannot maintain appropriate leverage for stabilization, the test will not be accurate. Mulroy et al looked at the difference between female and male therapists when assigning a grade for quadriceps strength. They found that female therapists over-graded the strength of the quads in 14 of 19 patients, while male therapists over-graded strength in 2 of 19 patients.

Another issue with manual muscle testing is the ceiling effect that occurs with a grade of 5/5. Patients may score a 5/5 on a MMT, yet still have deficits in strength that may lead to functional consequences based on the activities they perform. In fact, it is not recommended
that a MMT be used as a measure of progress when scores are greater than a grade of 3/5.² For grades over a 3/5, it is recommended that the therapist rely on instrumented assessment or functional testing to assess for side-to-side differences and to differentiate between a grade of a 4/5 and a 5/5.²

Last, but not least, there exists many variations in the manner in which therapists assess strength with a MMT. While standard positions and methods are described in textbooks, different textbooks vary in how they describe the test, and even different editions of the textbook can vary. For example, there have been modifications made to the newer editions of Daniels and Worthingham’s Muscle Testing textbook and therefore newer clinicians are performing certain tests, such as gastroc and soleus testing, differently because they learned from different editions of the same textbook.²,⁷

**Hand-held dynamometry**

Hand-held dynamometry (HHD) is an alternative and more objective method of measuring isometric force abilities of a muscle in kilograms (kg); thus creating a more quantitative system of assessing muscle force than solely relying on the subjective view of the physical therapist’s 0-5 grading scale.⁸,⁹ In a study conducted by Wadsworth on the reliability of MMT on a subjective scale versus HHD, results showed MMT on the subjective scale to be reliable at detecting only large amounts of noticeable muscular differences between sides.⁹ However, the subjective MMT scale was unable to determine small variations of muscular differences between sides, particularly between the scores of 4/5 and 5/5 on the MMT scale.⁹ In these ranges

---

*Hip abduction assessment using the Activforce system. Extremity mobilization strap used in conjunction with the Activforce device.*
Reliability and Validity of Hand Held Dynamometry

There have been several studies that have examined the reliability (both intra and inter-rater) of HHD. Most recently, Kim and Cho tested elbow flexion in 42 young adults and found that the intra-rater reliability coefficient was 0.992 and the inter-rater reliability coefficient was 0.949.

Whether using traditional MMT or HHD, the wide range of physical abilities, clinical backgrounds, and experience with MMT or HHD of the clinicians can be a factor in the strength assessments. Kelln et al examined the reliability of HHD when testing the lower extremity musculature in healthy, physically active, young adults by three different testers and found that the intra-tester ICC values ranged from 0.77 to 0.97. Additionally, the authors concluded “that regardless of experience or physical ability, HHD is a very reliable measurement tool.” The results of this study also noted that as long as the testers had a mechanical advantage and the necessary strength to isometrically resist a subject’s maximal contraction, reliable measures can be expected. The amount of force that the tester has to provide has also been examined and shown that when measuring forces above 200 N (45 lbs), an underestimation of muscle strength occurs when the tester holds the hand-held dynamometer versus when the patient applies force to the dynamometer against a fixated device.

In order to perform strength assessments on patients that the tester is unable to adequately isometrically resist, a stabilization belt should be used. Thorborg et al looked at inter-tester reliability concerning strength assessments of the hip and knee, using a HHD and belt fixation. In this study, hip and knee strength in an athletic population was shown to have ICC values that ranged from 0.76 to 0.95. The testers in this study were two physical therapy students (one female and the other one male) and each received one hour of instruction on how to perform the strength measurements. The authors concluded that HHD with belt fixation is an excellent option for evaluating and monitoring of athletes.

Make or Break Testing

When performing a standard MMT, the examiner performs a break test, which means the patient is asked to hold their limb in a specific position and not allow the therapist to “break” the hold with manual resistance. When performing HHD, there are two methods commonly described in the literature: a make test and a break test. As mentioned earlier, during a break test, the patient attempts to maintain an isometric force while the therapist applies resistance. The opposite occurs during a make test. With a make test, the therapist or some other stable device is the isometric force and the patient attempts to push against this isometric force to the best of their ability. While both methods have been reported in the literature, Stratford et al found the reliability of the make test to be higher than that of a break test.

In a study looking at the reliability of HHD in
relation to MMT, the authors found that a MMT tended to overestimate strength in comparison to HHD measures. They also noted that the force generated by a MMT utilizing the break test method of force application tended to exceed the force generated during a make test performed with hand-held dynamometry. When examining the strength of patients with knee osteoarthritis, Hayes et al identified measurable weakness in the knee extensor muscles when using HHD, which was not detected by a MMT. This can obviously lead to functional consequences for the patient. The authors concluded that HHD measurements are less subjective than MMT grades, especially in stronger patients.

When comparing strength values between various individuals, it has been found that there are several factors that can influence an individual’s strength in otherwise healthy individuals. Some of these factors include gender, age, extremity dominance and body weight of the individual. To compare normative strength values across various individuals, these values need to be broken down by age range of the individual, gender and body weight, which can be expressed as a percentage of body weight. Bohannon et al developed reference values for several muscles in the upper and lower extremities of individuals aged 20-79 years old, and have published these values expressed as a percentage of body weight (%BW), categorized by each decade of age and different scales exist for males and females. Other studies have published relative normative values for different muscle groups and serve as a great reference tool, however one must be careful to pay close attention to the method in which the tests were performed (position of the patient, where the resistance was applied, whether a make test or break test was used, and whether or not external fixation was used for stronger muscle groups).

Conclusion

Hand-held dynamometry has been shown to be a more objective assessment of strength when compared to MMT. This allows the health care provider to set goals tailored to the patient (expressed in %BW) and show objective changes over time with treatment. Hayes and Falconer compared MMT to HHD
with patients with osteoarthritis and found that significant weakness in the knee extensors with HHD where the MMT scores for the same subjects indicated good strength. MMT tends to over-estimate a patient’s strength, especially if the examiner is not able to provide sufficient stabilization or able to generate enough force to adequately challenge the muscle being tested. Additionally, Hislop et al does not advise using MMT for a measure of progress with grades over a 3/5, it is recommended that the therapist rely on instrumented assessment or functional testing to assess for side-to-side differences and to differentiate between a grade of a 4/5 and a 5/5.

Hand-held dynamometry can also be used to screen athletes for their level of risk of an injury. Nadler found for Division 1 female athletes during pre-season screening, those athletes with more than a 10 percent difference between the right and left hip extensor strength was predictive of them having low back pain (LBP) during the ensuing season. Similar findings have been seen with gluteus medius weakness in individuals in the performing arts, such as ballet, as well as with hip adductor weakness in Rugby. Additionally, converting the force output to a percentage of body weight allows for development of normalized data by age, gender and muscle group. This would enable a clinician to determine minimal strength values for return to work or sport. Considering hand-held dynamometry’s ease of use, portability, cost, and compact size, compared with isokinetic devices this instrument can be regarded as a reliable and valid instrument for muscle strength assessment in a clinical or athletic setting.

Shoulder abduction assessment using the Activforce system. Extremity mobilization strap used in conjunction with the Activforce device.
Bibliography