Intra- and Intertester Reliability of Chatillon and MicroFet Hand-Held Dynamometers in Measuring Force Production


The purpose of this study was to investigate intratester and intertester reliability when using the Chatillon and MicroFet hand-held dynamometers (HHDs) to measure isometric force production of the wrist extensors, elbow flexors, ankle dorsiflexors, and knee extensors. Twelve subjects participated, with each joint tested four times with each HHD. Intratester and intertester intraclass correlation coefficients were measured for both devices separately and between the devices. Results indicated that the HHDs were more reliable when used by a single examiner who had been properly trained in their use. Reliability of the HHDs appears to be affected by both the magnitude of the force produced by the subject and the examiner’s ability to resist the force. There was no correlation between examiner’s stature and consistent force production values. Caution should be taken when interpreting data obtained from different testers or different HHDs. The same clinician should use the same HHD for successive tests to yield the most reliable data.

Manual muscle testing (MMT) is an inexpensive, versatile, practical (12), and widely accepted form of evaluating muscle strength (2, 4, 17, 19, 23). MMT has advanced from the spring balance (5, 29) to the cable tensiometer (20) to the hand-held dynamometer (HHD), as researchers have attempted to obtain objective and reliable assessments of muscle strength (3). Objectivity, validity, and reliability are crucial for appropriate diagnosis and treatment. HHDs were designed to improve and quantify muscle strength assessment (4, 5, 26). Data obtained from MMT are recorded in nominal, interval, or ratio levels of measurement (16) (Table 1). Despite attempts to standardize MMT, a subjective component (12) is involved for grades above a level of “fair.” Grading reliability may be influenced by the frame of reference for a “normal” grade and the strength of the examiner (26). For example, the criteria for “normal” may be different between trained versus untrained clients and stronger versus weaker examiners. However, unlike MMT, evaluation with an HHD does not require subjective judgment by the examiner.

The authors are with Temple University Graduate Athletic Training/Sports Medicine Program, Pearson Hall, Office 129, Philadelphia, PA 19122. Direct correspondence to Dawn T. Gulick.
Table 1  Manual Muscle Testing Grades and Criteria

<table>
<thead>
<tr>
<th>Manual muscle testing grades</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>0%</td>
</tr>
<tr>
<td>Trace</td>
<td>1</td>
</tr>
<tr>
<td>Poor</td>
<td>2</td>
</tr>
<tr>
<td>Fair</td>
<td>3</td>
</tr>
<tr>
<td>Good</td>
<td>4</td>
</tr>
<tr>
<td>Normal</td>
<td>5</td>
</tr>
</tbody>
</table>

Numerous investigators (6, 12, 13, 18, 22, 24–28) have evaluated and compared various HHDs and isokinetic devices. However, only four studies (3, 5, 6, 26) have compared various HHDs with each other. The purpose of this study was to determine intra- and intertester reliability of the MicroFet and the Chatillon CSD500 HHDs. In addition, we wanted to compare the data of the two units when measuring isometric force produced by two upper and two lower extremity muscle groups.

**Methodology**

Four female and 8 male NATA-certified athletic trainers (mean age 27.7 ± 1.6 years) served as examiners. Each examiner was instructed by a service representative in the proper use of the HHDs. Prior to data collection, examiners practiced isometric break tests without an HHD and then practiced with each HHD until an intraclass correlation (ICC) of at least .90 was achieved. Examiners were instructed in subject positioning, hand positioning, goniometric measurement, and limb stabilization.

Subjects were 3 females and 9 males (mean age 26.4 ± 1.3 years). All subjects completed an injury questionnaire designed to eliminate individuals who had been diagnosed with extremity pathology within 1 year prior to the study. Signed consent forms were obtained from all subjects.

The Chatillon CSD500 HHD (Chatillon Medical Products, Greensboro, NC) is an electronic device that is factory calibrated and designed to measure forces up to 500 ft-lb. The MicroFet HHD (MicroFet, Draper, UT) is a battery-operated, microprocessor-controlled device with three strain gauges to measure force vectors from three planes. The test range is 0.8 to 100 ft-lb.

**Procedures**

Standard written instructions were utilized throughout data collection. Examiner and subject testing sequences were randomly ordered. Break testing was performed on the wrist extensors, elbow flexors, ankle dorsiflexors, and knee extensors at the midpoint in the range of motion (21). The HHD was positioned between the examiner’s hand and the body segment to be tested.
All testing positions were standardized (8). When the wrist extensors were tested, the subject was seated with the forearm supported, elbow flexed to 30°, and wrist extended to 35°. Stabilization was provided at the midforearm, and resistance was applied to the distal metacarpals (Figure 1). When the elbow flexors were tested, the subject was supine with the shoulder in neutral, the elbow flexed to 75°, and the forearm fully supinated. Stabilization was provided at the distal humerus, and resistance was applied to the distal forearm (Figure 2). To test the ankle dorsiflexors, the subject was seated over the edge of the plinth with the knee flexed to 90° and the ankle in 10° of dorsiflexion. Stabilization was provided proximal to the malleoli, and resistance was applied to the distal metatarsals (Figure 3). For testing the knee extensors, the subject was seated over the edge of the plinth with a towel roll placed under the distal femur and the knee flexed to 67.5°. The subject's trunk weight was sufficient for stabilization, and resistance was applied 4 cm proximal and anterior to the lateral malleoli (Figure 4).

Each test began with the examiner counting aloud from 1 to 3 as the subject gradually increased muscle force so that the maximum force was exerted on the count of 3. This prevented quick, jerky motions that could have jeopardized transducer pad placement. All assessments of force production were performed using a break test technique in which the examiner provided sufficient counterforce to move the body segment being tested from the static position. Test–retest sequences were performed twice at each joint test site with each HHD. A minimum of 30 s rest was given prior to each retest to avoid fatigue. The force produced during each effort was recorded by a third examiner.
Figure 2 — Testing position and stabilization for elbow flexion using the MicroFet hand-held dynamometer.

Figure 3 — Testing position and stabilization for ankle dorsiflexion using the MicroFet hand-held dynamometer.
Data Analysis

Data analyses were performed using the Biomedical Statistical Software Program (10) and Excel 4.0 Statistical Software (Microsoft Corporation, Redmond, WA). Scores from the first and second test–retest sequences were averaged to obtain Test 1 and Test 2 criterion values for each of the four muscle groups tested. ICCs (2,1) were performed on the data (9) to determine intratester and intertester reliabilities for both the Chatillon CSD500 and MicroFet. An analysis (ICC) was also performed between the stature of the examiner and the ICCs of each joint tested; this was done to address the relationship between the examiner’s stature and his or her ability to produce consistent data across the various joints tested.

Results and Discussion

The ICCs for each of the test conditions are summarized in Table 2. Intratester ICCs were higher than intertester ICCs for both HHDs. ICCs were higher for the upper extremity joints than for the lower extremity joints for both HHDs. ICCs were lowest when different joints were tested by different examiners with different HHDs. ICCs for examiner stature with respect to each joint position tested were as follows: wrist .30, elbow -.11, ankle .17, and knee .38.

The results of this study revealed moderate to high intratester ICCs for both the Chatillon and MicroFet. Only two correlations for the Chatillon and five correlations for the MicroFet fell below \( r = .56 \). These results are consistent with those of previous investigators (1, 4, 24, 25) who reported HHDs to be reliable when
upper extremity and lower extremity measurements were taken by a single examiner. The high degree of intratester reliability found in the present study may be attributed to the standardization of testing protocol and the training of the examiners.

The intertester ICCs were more variable and generally lower than the intratester ICCs for both HHDs. This was particularly noted when larger muscle groups were tested. The results displayed in Table 2 reveal lower ICCs for the knee extensors and ankle dorsiflexors than for the elbow flexors and wrist extensors. Numerous researchers (1, 7, 14, 22, 28) have suggested that the stature of the examiner may influence reliability. In the present study, correlations between examiner stature and each joint tested ranged from $-.11$ to $.38$. These values indicate that stature and the ability to produce consistent muscle force data are poorly correlated. However, it appears logical that examiners who are stronger, not larger, should be able to produce more consistent results. Wadsworth et al. (27) studied the effects of gender, body weight, and grip strength on strength testing with an HHD. Their results indicated that male and female examiners demonstrated good interrater reliabilities when testing a female subject. Male examiners also demonstrated good interrater reliabilities when testing a male subject; however, female examiners demonstrated poor interrater reliabilities when testing elbow flexors and knee extensors. Results of this study indicated that grip strength was a significant factor and correlated with force values in larger, harder to stabilize muscle groups. Low forces produced by smaller muscle groups revealed nonsignificant correlations between body weight and grip strength.

Wikholm and Bohannon (28) reported Pearson Product Moment correlations ranging from $.226$ to $.932$. The correlation for weaker muscle groups such as the shoulder external rotators was $r = .932$, while the correlation for the elbow was $r = .799$ and for the knee $r = .226$. The authors concluded that as force production increased, stabilization was compromised. Hayes and Falconer (13) reported intertester ICCs of $.89$ to $.98$ and an intratester ICC of $.92$ for knee extension with patients who had osteoarthritis. However, force production never exceeded 60 lb. In the present study of healthy college students, force production ranged from 50

### Table 2: Mean ICC\(^a\) (9) for Wrist Extension, Elbow Flexion, Ankle Dorsiflexion, and Knee Extension Between Hand-Held Dynamometers and Among Examiners

<table>
<thead>
<tr>
<th></th>
<th>Wrist extensors</th>
<th>Elbow flexors</th>
<th>Ankle dorsiflexors</th>
<th>Knee extensors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intratester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chatillon</td>
<td>.86</td>
<td>.88</td>
<td>.79</td>
<td>.72</td>
</tr>
<tr>
<td>Microfet</td>
<td>.81</td>
<td>.79</td>
<td>.73</td>
<td>.68</td>
</tr>
<tr>
<td>Chatillon vs. Microfet</td>
<td>.65</td>
<td>.69</td>
<td>.45</td>
<td>.37</td>
</tr>
<tr>
<td><strong>Intertester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chatillon</td>
<td>.55</td>
<td>.56</td>
<td>.43</td>
<td>.32</td>
</tr>
<tr>
<td>Microfet</td>
<td>.52</td>
<td>.68</td>
<td>.34</td>
<td>.41</td>
</tr>
<tr>
<td>Chatillon vs. Microfet</td>
<td>.33</td>
<td>.31</td>
<td>.14</td>
<td>.45</td>
</tr>
</tbody>
</table>

\(^a\)ICC = intraclass correlation coefficient.
to 148 lb. Hosting et al. (14) concluded that reliable force measurements using an HHD were limited to pathologically weakened, naturally weak, or smaller muscle groups. Since the reliability of HHD measurements is affected by both the magnitude of the forces produced by the subject (22) and the examiner’s ability to resist the subject’s force, Brinkmann (6) recommended that the use of HHDs be limited to muscle groups producing forces of less than 15 kg. A study by Hyde, Goddard, and Scott (15) revealed that few examiners could consistently apply force greater than 30 kg. Practitioners need to realize that an isometric break test cannot be achieved if the examiner is unable to exceed the force production of the subject. If the subject’s force overcomes that of the examiner, then the force measured is the maximum force of the examiner and not of the subject.

When the two HHDs were tested across various examiners, the force values recorded were more variable than any of the other comparisons. These results are consistent with previous investigations (5, 11) which concluded that testing devices should not be used interchangeably. Data obtained via one device should not be used as a reference for a subject measured with another device.

Conclusion

Both the Chatillon and MicroFet HHDs were found to be reliable devices for MMT when used by the same examiner. The devices were found to be more reliable for upper extremity testing than for lower extremity testing. However, using two different HHDs interchangeably on the same subject did not yield reliable data. Therefore, examiners should consistently perform manual muscle tests with the same HHD in order to gather reliable data.

References


Acknowledgments

We wish to extend our appreciation to Chatillon for the loan of the dynamometer.