Validity and Reliability of a Hand-Held Dynamometer for Dynamic Muscle Strength Assessment

Lan Le-Ngoc\textsuperscript{1} and Jessica Janssen\textsuperscript{2}

\textsuperscript{1}Industrial Research Ltd, Christchurch
\textsuperscript{2}Burwood Academy of Independent Living, Christchurch
New Zealand

1. Introduction

An important component of physical therapy is to conduct assessment of a patient’s mobility including muscle strength and joint range of motion (ROM).

The purposes of this study were to investigate the possibility of measuring dynamic muscle strength using a new hand-held device and to assess its validity and reliability. If proven valid and reliable, this device will provide a practical tool for physical therapists to perform dynamic muscle assessment in a clinical setting.

The current standard clinical evaluation and diagnostic tool for muscle strength assessment is the manual muscle testing (MMT) method, using a 5-point grading scale (Clarkson (2000); Petty (2011)). Although it has been a clinically useful tool for over forty years, its accuracy and reliability remains questionable (Cuthbert & Goodheart (2007); Frese et al. (1987)).

To overcome the limitations of the MMT, isometric hand-held dynamometers (HHD) have been developed to aid therapists in clinics (Andrews (1991)). HHDs are generally small and portable, and measure strength objectively in kilograms, pounds or newtons. The clinician holds the HHD between his or her force-applying hand and the patient’s limb segment. The clinician stabilises the limb segment while encouraging the patient to exert as much force against the device as possible and the maximum force is recorded by the HHD. Such devices have been proven to have good to excellent reliability in different populations (Andrews (1991); Bohannon & Andrews (1987); Stark et al. (2011)). In a single test, however, they can assess the strength of a patient at only one joint angle, rather than through the patient’s entire ROM. Although this technique provides a crucial tool for clinical quantification of joint strength at a fixed static position (isometric), it cannot measure properties from dynamic muscle performance assessments.

Isokinetic dynamometers, such as the Cybex (USA) or the Biodex (USA), are considered as the gold standard in simultaneous strength and angle measurements for the evaluation of dynamic muscular performance (Kannus (1994); Baltzopoulos & Brodie (1989); Osternig (1986); Lund et al. (2005); Drouin et al. (2004)). Strength profiles showing instantaneous torque versus joint angle are generated and a number of properties such as dynamic peak torque, peak torque angle, angle-specific torque, power, and energy used can be determined. The dynamic strength profiles can also be used to detect weaknesses over small
regions of a specific joint’s ROM. Other advantages of the isokinetic dynamometer over the current isometric HHDs are that assessor’s strength is not an issue; the subject is stabilized consistently during testing; and the joint angle and strength are measured simultaneously during testing (Lund et al. (2005); Martin et al. (2006); Harlaar et al. (1996)). Disadvantages of these devices are their size and cost, which make them impractical for routine clinical examinations (Li et al. (2006); Mital et al. (1995)).

Recognising the needs for better clinical strength assessment tools, there have been a number of attempts to incorporate angle measurement in the strength assessment (Li et al. (2006); Roebroeck et al. (1998)). However, there have been no published results on the use of a single hand-held device to perform dynamic strength measurements on human subjects. A new device, referred to as the IRL-HHD (Fig. 1), is a single hand-held device that can measure force and angle simultaneously while the joint moves through its ROM\(^1\). The ability to measure force and angle simultaneously means that it can measure energy or power in a similar manner to an isokinetic dynamometer. In order for the IRL-HHD to capture dynamic joint strength, the assessor must provide sufficient force to resist the limb movement, but also allow the limb to move at a constant and controllable pace. This is not a trivial task and the assessor may not be able to concentrate on keeping the device in perfect alignment with the limb. The algorithm used in the IRL-HHD can measure the required joint angle accurately without having to maintain the alignment of the longitudinal axis of the device with respect to the limb. In some cases, this feature allows the joint to reach its full ROM (see Fig. 2 for an example of measuring concentric elbow flexion where the longitudinal axis of the IRL-HHD does not have to be aligned with the forearm). The IRL-HHD and the assessment techniques have been shown to be reliable and valid by measuring concentric flexion of a simulated mechanical arm, which was used to eliminate the effects of human variability (Janssen & Le-Ngoc (2009)).

Fig. 1. IRL Hand-held dynamometer.

This article describes the validity and reliability trials of the device to measure concentric elbow flexion and concentric knee extension on human subjects. Other possible uses of the IRL-HHD in clinical and on-field assessments are also discussed.

2. Validity and reliability of dynamic muscle strength assessment

This section describes the test protocol and the results of using the IRL-HHD to perform concentric elbow flexion and concentric knee extension assessment on human subjects.

2.1 Instrumentation

Two dynamometers, the IRL-HHD and the isokinetic dynamometer (Biodex), were used to measure maximal concentric strength for elbow flexion and knee extension. The Biodex measurements were corrected for the effect of gravity caused by the Biodex lever arm. For the IRL-HHD tests, a seat and an arm rest attached to a plinth were used to position and restrain the participants in a similar manner to the tests carried out using the Biodex (see Fig. 2 and Fig. 3).

2.2 Protocol

A registered physiotherapist conducted the tests using the IRL-HHD and another registered physiotherapist performed the Biodex tests. Both therapists were blinded from the outcome measures.

2.2.1 Participants

Fifteen able-bodied, healthy adults participated in this study, which was approved by the University of Otago (New Zealand) Ethics Committee. All participants provided informed written consent before testing.

2.2.2 Design

There were two test sessions for each participant using the IRL-HHD, and one test session using the Biodex. Each test session comprised one sub-maximal contraction, and three repeated maximal strength contractions to perform right elbow flexion and right knee extension. Each measurement was followed by a one minute rest period. The order of sessions was randomized for each participant, and within each session the order in which joints were tested was randomized. The participants were given five minutes rest between each test session to prevent fatigue.

The distances from the centre of the force pad to the rotational axis of elbow and knee were recorded for each participant and used to convert measured forces into joint torques. Peak torque, peak torque angle and total work were obtained from the torque versus joint angle curves recorded by both dynamometers.

A three-stage procedure was followed to record strength versus joint angle data using the IRL-HHD:

- Defining the zero position of the joint;
- Moving the joint to the start position, positioning the device to resist the limb motion and commencing the measurement;
- Instructing the participant to exert maximal muscular contraction while providing a resistance to control the movement of the joint, and stopping the measurement when the participant reaches the end of joint movement.

For concentric elbow flexion, the participant was seated beside the end of the plinth, and the right arm was strapped to an arm rest at 60° shoulder flexion and 30° shoulder abduction.
(Fig. 2). The zero position of the elbow was identified by placing the device lengthwise on a reference line between the acromion and the lateral epicondyle of the humerus. The device was placed with the force pad 2 cm proximal of the wrist while the arm was fully extended. It is possible to have a negative start angle, which is a measure of elbow hyperextension.

For concentric knee extension, the participant was seated using the same arrangement as on the Biodex (Fig. 3). The zero position was set against a horizontal surface. The device was placed with the force pad 10 cm proximal of the medial malleolus and the leg was moved to the starting position (110° knee flexion) before commencing the measurement.

The isokinetic mode of the Biodex was used for testing with a maximum speed of 60°/s. In this mode, the start and end ROM had to be set before starting the test. For elbow flexion, the zero elbow position was set so that the participant’s arm was supported at 60° shoulder flexion and 30° shoulder abduction (Fig. 2). Unfortunately, the Biodex strap restricted some participants from reaching end ROM, so it was not possible to provide a comparison of elbow
ROM measurements between the IRL-HHD and the Biodex. For knee extension, the Biodex chair and the fixture beneath the chair prevented participants from reaching full knee flexion. In order to provide a meaningful comparison of the peak torque angle, the starting position of the knee extension was set at the maximum possible knee flexion angle but not greater than 110°. Because of this preset starting position, it was not meaningful to report knee ROM measurements using the Biodex.

During testing, the physiotherapist manually recorded any unusual events, such as loss of control, or excessive movement of the IRL-HHD. These tests were discarded from the data set, which was justified on the basis that it would be standard clinical practice to ignore erroneous tests at the time of testing.

The ability of the therapist to maintain the control of the dynamic measurement is discussed in Section 4.1.

Fig. 3. Concentric knee extension measurements with the IRL-HHD and the Biodex.
2.3 Statistical analysis

2.3.1 Descriptive statistics

Descriptive statistics of muscle torques, joint angles and muscular work are presented in Nm, degrees (°) and J respectively. Torque is calculated from the measured peak force times the length from the centre of the force pad to the rotational axis of the elbow or knee. Work is defined as output of mechanical energy, that is, externally applied force multiplied by the distance through which it is applied. In the concentric measurements, work can be found by calculating the area under the torque versus angular displacement curve. Mean and standard deviations (SDs) are reported. All analyses were performed using the Matlab software package (USA).

2.3.2 Intratester reliability

The degree of correlation between six repetitions of all the maximal strength tests using the IRL-HHD is calculated with the intraclass correlation co-efficient (ICC1,1) defined by Schrout and Fleiss (1979). The same test was performed on the three repetitions of the Biodex. The most critical reliability assessment is the ICC1,1, which assumes that every individual measurement is independent and the error of measurement is assumed to be normally distributed. Other authors have used ICC2,1 for their reliability measurement, which tends to give more optimistic values than ICC1,1. In this article all ICC1,1 results are almost equal to the ICC2,1 values. According to Fleiss (1986), the reliability of an ICC over 0.75 is considered to be excellent, and between 0.4-0.75 as fair to good.

2.3.3 Validity

The agreement between the two devices can be quantified using the Bland-Altman 95% limits of agreement (LOA) method (Bland & Altman (1986)). The LOA method is based on the mean and SD of the differences between the measurements by the two devices. For repeated measurements, a one-way ANOVA is performed for each device separately. Outcomes of the one-way ANOVA are then used to calculate the lower and upper LOA (mean ± 1.96 times SD)(Bland & Altman (2007)).

3. Results

3.1 Descriptive Statistics

Five men and ten women participated in this research. The participants’ ages ranged from 23 to 45 years (mean±SD, 32.6±7.2y). Fig. 4 shows typical strength profile plots between the IRL-HHD and the Biodex for one participant. Although the shape of torque versus angle graphs were not the same for the IRL-HHD and the Biodex, both methods show consistency in repeated measurements.

Fig. 5 shows the speed of all measurements obtained with the IRL-HHD. It shows that the physiotherapist was able to control the speed of each measurement very well for the elbow flexion. Only two participants generated speeds more than 100°/s while nine generated speeds less than 80°/s. It was more difficult for the physiotherapist to control the speed for the knee extension and five participants generated speed greater than 100°/s. The range of speeds for those participants was also greater, suggesting that the physiotherapist was not in control of all the tests. Since the speed is controlled entirely from the perception of the assessor, an error of ±20°/s is considered to be reasonable in this study.
Fig. 4. Strength profiles of one participant for concentric elbow flexion and knee extension obtained from the IRL-HHD (a, c) and the Biodex (b, d).

Fig. 6 shows scatter graphs of the mean peak torque and mean work between the Biodex (x-axis) and the IRL-HHD (y-axis) for elbow flexion, and Fig. 7 shows the corresponding data for knee extension. The error bars show the individual SDs for the Biodex and the IRL-HHD. It is interesting to note that the error bars for the Biodex are generally larger than those for the IRL-HHD, indicating that variability of the tested participants is a significant factor in strength measurements. For elbow flexion, fourteen out of fifteen participants generated peak torques less than 50 Nm. For knee extension, the physiotherapist was unable to resist any torque greater than 100 Nm, whereas five participants generated more than 100 Nm on the Biodex.

3.2 Intratester reliability

Six repeated measurements with the IRL-HHD and three with the Biodex were used to calculate the ICCs and their 95% confidence intervals. The results are shown in Table 1. The ICC1,1 values of both devices indicates excellent intratester reliability in the peak torque and work for both elbow flexion and knee extension. Repeatability of the peak torque angle of both tests by both devices is rated fair to good. However the confidence intervals indicates that only the knee peak torque angle obtained from the Biodex can be considered as fair to good, while all other peak torque angle measurements are poor. To determine if the mean of three measurements is a more reliable measure of the peak torque angle, the ICCs1,3 of
Fig. 5. Average speeds obtained with the IRL-HHD.

The peak torque angles in the first session of the IRL-HHD tests were found to be 0.80 (0.50, 0.93) for elbow flexion and 0.75 (0.38, 0.91) for knee extension which are within the range of excellent.

### 3.3 Validity

The overall mean differences and their SDs between the two devices, and all lower and upper LOA values are shown in Table 2. The differences were calculated by subtracting the Biodex values from the corresponding IRL-HHD values, hence a negative value indicates that the IRL-HHD measurement is smaller than the Biodex measurement. The table also shows the LOA for screened data as will be discussed in Section 4.3.
4. Discussion

4.1 Descriptive statistics

The graphs of torque versus angle in Fig. 4 suggest that the standardized methods of measurement using the IRL-HHD provided reliable concentric measurement. Speed variation during a single test using the IRL-HHD may be a factor in producing different shapes of the torque-angle curves between the IRL-HHD and the Biodex.
Table 1. ICC$_{1,1}$ for six repetitions with the IRL-HHD and three repetitions with the Biodex, and ICC$_{1,3}$ for the first session with the IRL-HHD.

<table>
<thead>
<tr>
<th>Joint Movement</th>
<th>Measurements</th>
<th>IRL-HHD</th>
<th>Biodex</th>
<th>IRL-HHD Session 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elbow flexion</td>
<td>PT</td>
<td>0.95(0.91 to 0.98)</td>
<td>0.96(0.90 to 0.98)</td>
<td>0.80(0.50 to 0.93)</td>
</tr>
<tr>
<td></td>
<td>PT Angle</td>
<td>0.41(0.20 to 0.68)</td>
<td>0.56(0.25 to 0.81)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Work</td>
<td>0.97(0.93 to 0.99)</td>
<td>0.95(0.88 to 0.98)</td>
<td></td>
</tr>
<tr>
<td>Knee extension</td>
<td>PT</td>
<td>0.99(0.94 to 0.99)</td>
<td>0.97(0.93 to 0.99)</td>
<td>0.75(0.38 to 0.91)</td>
</tr>
<tr>
<td></td>
<td>PT Angle</td>
<td>0.46(0.24 to 0.71)</td>
<td>0.67(0.41 to 0.86)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Work</td>
<td>0.86(0.73 to 0.94)</td>
<td>0.98(0.96 to 0.99)</td>
<td></td>
</tr>
</tbody>
</table>

NOTE. 95% confidence intervals shown in parenthesis
Abbreviation: PT, peak torque.

Table 2. Agreement between the IRL-HHD and the Biodex for assessing elbow flexion and knee extension.

<table>
<thead>
<tr>
<th></th>
<th>Elbow flexion</th>
<th>Knee extension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PT</td>
<td>PT angle</td>
</tr>
<tr>
<td>Mean difference</td>
<td>1.0(6.4)</td>
<td>24(23)</td>
</tr>
<tr>
<td>(SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% LOA for all data</td>
<td>-11.6, 13.5</td>
<td>-21, 19</td>
</tr>
<tr>
<td>95% LOA for screened data</td>
<td>-7.0, 9.9</td>
<td>-15, 53</td>
</tr>
</tbody>
</table>

An angular speed measurement greater than 100°/s indicates that the physiotherapist is overpowered by the participant and that the result is likely to be invalid. Fig. 5 shows that it is possible for a trained assessor to control the concentric assessment speed to within ±20°/s from a target speed of 60°/s, provided that the force generated by the subject is less than the strength limit of the assessor.

Further examination shows that most of the variability in the Biodex arises from the first test in a series of three repeats being sub-maximal. It is recommended for future study that the warm-up phase should consist of more than one sub-maximal concentric movement.

For knee extension, the physiotherapist was unable to resist any torque greater than 100 Nm, whereas five participants generated more than 100 Nm on the Biodex. From the elbow tests, the assessor was overpowered by one participant, who generated 52 Nm peak torque, but was able to perform tests satisfactorily at 43 Nm peak torque, suggesting that the strength limit of this assessor is between 43 Nm and 52 Nm for elbow flexion (approximately 200 N to 250 N in force). Several authors have specified minimum upper limits of assessor’s strength necessary for performing isometric measurements using an HHD (Wikholm & Bohannon (1991)). A conservative value is 12 kg of resistive force (Edwards & McDonnell (1974)) while others have suggested a value of 30 kg force (Hyde et al. (1983)). van der Ploeg et al. (1984) stated that an HHD range beyond 220 N is not useful due to stabilization and strength issues. The upper limit of the assessor’s strength in this study is in agreement with the published results for isometric measurements. The torque limit of this assessor is expected to be between 52 Nm and 65 Nm for knee extension. Only four of the fifteen participants (27%) generated less than 52 Nm for knee extension, hence it may be concluded that the IRL-HHD and the test protocol described in this article is not feasible for general use in measuring knee extension of healthy...
adults. Most participants generated peak elbow flexion torques less than 50 Nm, suggesting that the IRL-HHD can be used to measure concentric strength of upper extremities or minor muscle groups in general healthy population. It may also be possible to use the IRL-HHD to assess children’s concentric strength and subjects with strength deficiency resulting from conditions such as stroke or spinal cord injury.

4.2 Intratester reliability

Intratester reliabilities of peak torque and work are excellent in both elbow and knee measurements with the IRL-HHD and with the Biodex, while the intratester reliabilities of peak torque angle are poor for the IRL-HHD. The intratester reliabilities of the peak torque angle using the Biodex are slightly better than those obtained with the IRL-HHD. The ICCs1,3 of the first sessions using the IRL-HHD indicates a significant improvement in the reliability of measuring the peak torque angle. These values suggest that peak torque angle should be measured by taking the mean of three repeated tests. The ICCs1,3 for peak torque angle are within the range of excellent for both elbow flexion and knee extension.

4.3 Validity

The conventional method of assessing and grading muscle strength is the manual muscle test. In this study, all of the participants would be rated with a score of 5 as they were all healthy. Quantitative assessments of concentric strength are mostly associated with research or specialized assessments of top athletes, and have not been used in clinical settings. As far we are aware this is the first study using an HHD to perform concentric measurement, so it is not possible to define clinical agreement values to assess the LOA calculated in this paper. Instead, the LOA have been calculated to provide useful benchmarks for future research and a subjective analysis of the LOA is provided.

For elbow flexion, the LOA for peak torque are -11.6 and 13.5 Nm and for work are -24.1 and 26.0 J. The LOA of the peak torque angle are -21 and 69° which is unacceptable as a valid measurement of peak torque angle.

Eliminating participants who generate torque greater than 50 Nm, any tests with speed greater than 100°/s, and the first run of all the Biodex results improves the LOA of all the parameters. They are: -7.0 and 9.9Nm for peak torques, -15 and 53° for peak torque angle, and -12.4 and 16.2 J for work in elbow flexion.

Considering that the maximum peak torque is approximately 50 Nm, the LOA are approximately ±20% of the range of measurement, therefore we suggest that the use of IRL-HHD in muscle strength assessment provides the clinician with at least 5 additional scales above the MRC score of 5, assuming that the Biodex measurements are the accepted peak torques of the participants.

For knee extension, the LOA in all measurements show unacceptably large ranges. There is an obvious trend between mean strength and difference between the two devices, showing that the stronger the participant, the bigger the difference between the IRL-HHD and the Biodex measurements in peak torque. The LOA calculation with the proposed reduced dataset as discussed for elbow flexion are -23.1 and 28.7 Nm for peak torques, -16 and 41° for peak torque angle, and -13.1 and 41.0 J for total work. This means that for the knee extension test, the LOA of peak torque are approximately ±50% of the range of measurement, which is not a significant improvement over the conventional method.
Other factors that may affect the IRL-HHD assessments include: discomfort over the anterior tibial region because of the hard padding of the IRL-HHD force plate; the participants might be trying to control the speed; or they might think that the physiotherapist would not be able to control a maximal effort were they to exert it.

5. Conclusions

The new IRL-HHD has excellent intratester reliability, when used by an experienced user on healthy adults, for measuring peak torque and total work for elbow flexion and knee extension. Therefore, the device and the associated test protocols described in this paper can be used to measure these two physical attributes. The device is only reliable for determining peak torque angle if the mean of at least three repeated measurements is taken.

The LOA between the IRL-HHD and the isokinetic dynamometer are only reasonable for measuring elbow flexion peak torque and work. There were no agreements for peak torque and work of knee extension and peak torque angles of both elbow flexion and knee extension. Therefore, the IRL-HHD cannot be used on large muscle groups, such as the quadriceps, of healthy adults. The LOA also imply that the strength of the assessor using the IRL-HHD constrains the maximum forces that may be exerted by the subject, similar to the constraints reported for other hand-held isometric dynamometers.

The results obtained with the IRL-HHD cannot be compared with those obtained with an isokinetic dynamometer. However, since it has excellent intratester reliability, it can be used to compare strengths of different subjects or of one subject at different times, if used by the same assessor with the same test protocol.

6. Potential usage and future work

Recently, a study has been published on the reliability of shoulder assessment in patients with shoulder pain using the IRL-HHD (Cadogan et al. (2011)). These results show a good to excellent reliability of the IRL-HHD in practice.

The ability of measuring simultaneously the orientation of the device and the force imposed on the force plate may lead to many other potential usages. Other applications in which the IRL-HHD could be used include:

- In an isometric setting, the device can provide additional feedback on the tested angle. An audible angle warning feature can help the therapist to keep the joint within a pre-defined range, making the assessment more reliable (Sole et al. (2010), Hanna et al. (2010), Fulcher et al. (2010)).

- In the above described study of shoulder assessment (Cadogan et al. (2011)), a standardized shoulder lateral abduction active end range measurement was introduced. Since the end range of the shoulder is dependent on the amount of force the clinician exerts, it is impossible to compare measurements made by different assessors. However, with the IRL-HHD, a pre-set force can be entered into the IRL-HHD and when the force exerted on the force pad reaches the pre-set level, the IRL-HHD gives an audible warning sound so that the clinician knows when to click a button on the IRL-HHD to record the angle measurement. This should alleviate the assessors’ variable strength issue.

- For measuring joint stiffness. Stiffness is defined as the rate of change of force with respect to the rate of change of displacement. Since the IRL-HHD can measure force and angle simultaneously, it is ideal for measuring stiffness.
• Measurement of children's dynamic strengths, as children are generally weaker than clinicians. Children are often too small to fit the isokinetic machines, and it would be difficult to strap a young child to the Biodex machine. Children may not be as patient as adults and so a rapid assessment using the IRL-HHD could offer some advantages.
• In people with disability, where transferring patients in and out of the isokinetic dynamometer is difficult.
• In cases when it is impossible to restrain the patient to the machine e.g. in patients with spasticity.

Future work should concentrate on developing and carrying out clinical trials for measuring the dynamic strength of people with injury or disability, small muscle groups in adult population or all muscle groups in children. For large muscle group assessments, additional fixtures to provide mechanical advantages for the assessors may be a solution for low-cost functional dynamic strength assessment tools.

7. Acknowledgements

This work was supported by the Foundation for Research Science and Technology, New Zealand (C08X0816). We thank Burwood Academy of Independent Living and the School of Physiotherapy, Otago University for their assistance in this research.

8. References


