Comparison of Rolyan and Jamar dynamometers for measuring grip strength

VIRGIL MATHIOWETZ Program in Occupational Therapy, University of Minnesota, Minneapolis, MN, USA

ABSTRACT: This study compared the Jamar and Rolyan hydraulic dynamometers to determine their concurrent validity with known weights as well as their inter-instrument reliability and concurrent validity for measuring grip strength in a clinical setting. Thirty females and 30 males were tested on these two grip strength measurement devices using a repeated measure design. Results showed that the Jamar and Rolyan dynamometers have acceptable concurrent validity with known weights (that is, correlation coefficients were $r \geq 0.9994$), excellent inter-instrument reliability (that is, intraclass correlation coefficients ranged from 0.90 to 0.97) and strong concurrent validity (that is, no significant differences between dynamometers' scores). Data indicate that Jamar and Rolyan dynamometers measure grip strength equivalently and can be used interchangeably. Thus, therapists using the Rolyan dynamometer are justified in using published normative data, which were collected with the Jamar dynamometer.

Key words: hand strength, occupational therapy evaluation, rehabilitation, reliability.

Therapists measure grip strength commonly to evaluate a client's strength relative to a normative standard (Mathiowetz, 1991), to document change in a client's clinical course (Janda et al., 1987), and to determine sincerity of effort (Niebuhr and Marion, 1987; Hildreth et al., 1989). Therapists have a choice of a number of instruments for measuring grip strength, which are available commercially. The Jamar dynamometer is recommended by the American Society of Hand Therapists (Fess, 1992) and by others (Kirkpatrick, 1956; Fess, 1987; Mathiowetz, 1990) to measure grip strength. It is a reliable and valid dynamometer when it is calibrated properly (Fess, 1987) and if standard positioning and instructions are used (Mathiowetz et al., 1984; Fess, 1986, 1992). Good to excellent test–retest reliability ($r = 0.88$ to 0.93) and excellent interrater reliability ($r = 0.99$) have been reported (Mathiowetz et al.,
1984). The clinical utility of the Jamar dynamometer is enhanced by the availability of normative data for children and adults (Mathiowetz et al., 1985, 1986). For these reasons, the Jamar dynamometer has become the 'gold standard' for the measurement of grip strength. The question remains whether other dynamometers that are available commercially measure equivalent to the Jamar dynamometer and thus could be used interchangeably.

**Review of the literature**

Nine studies have compared the inter-instrument reliability of the Jamar dynamometer with other grip strength measurement devices and the concurrent validity of the dynamometers with known weights. Participants' grip strength was tested with two or more instruments within minutes of each other to assess inter-instrument reliability. Historically, inter-instrument reliability was analysed statistically using correlation coefficients such as Pearson r and Spearman rho to assess degree of association between instruments and using t-tests to assess agreement between two instruments. In recent years, the intraclass correlation coefficient (ICC) has been recommended as a better test of reliability because it assesses both degree of association and agreement among instruments with one number (Portney and Watkins, 2000). An ICC of 0.90 or higher is considered excellent; 0.75 to 0.90 is good; 0.50 to 0.75 is moderate; and less than 0.50 is considered poor (Portney and Watkins, 2000). Assessing the agreement between two instruments can also be considered a type of concurrent or criterion-related validity (for example, a new instrument, the Rolyan dynamometer, measures grip strength in the same way at the same time as the ‘gold standard’ Jamar dynamometer). Another type of concurrent validity is evaluated when known weights are suspended from the handle of a dynamometer. Fess (1987) described this as a method to check calibration of dynamometers. The Pearson r correlation coefficient between the weights suspended and the dynamometer readout measures this type of concurrent validity. Fess (1987) suggested that $r \geq 0.9994$ is considered acceptable.

Table 1 summarizes the nine studies that assessed the reliability and validity of the Jamar dynamometer in comparison with other grip strength measurement devices. The first five studies compared the Jamar with non-pneumatic devices. Excellent inter-instrument reliability was reported between the Jamar, Dexter and Baseline dynamometers (Bellace et al., 2000; Mathiowetz et al., 2000). Because these three instruments measure grip strength in pounds and normative data is in pounds, they are very easy to use interchangeably. Moderate to excellent reliability is reported between the Jamar dynamometer, the Baltimore Therapeutic Equipment (BTE) work simulator (King and Berryhill, 1988; Beaton et al., 1995) and the BTE Primus (Shechtman et al., 2001). Because the two instruments yield different units of measurement (that is, pounds versus inch-pounds), the BTE scores must be
<table>
<thead>
<tr>
<th>Authors, year, sample</th>
<th>Grip dynamometers</th>
<th>Reliability and validity</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bellace et al., 2000, n = 62</td>
<td>Jamar and Dexter</td>
<td>Inter-instrument reliability</td>
<td>ICC = 0.99 and 0.98</td>
<td>Excellent reliability</td>
</tr>
<tr>
<td>Mathiowetz et al., 2000, n = 40 males, n = 40 females</td>
<td>Jamar and Baseline</td>
<td>Inter-instrument reliability</td>
<td>ICC = 0.95 and 0.94</td>
<td>Excellent reliability</td>
</tr>
<tr>
<td>King and Berryhill, 1988, n = 25 patients</td>
<td>Jamar and BTE Work Simulator</td>
<td>Inter-instrument reliability</td>
<td>r = 0.87 (both hands)</td>
<td>Moderate–excellent reliability</td>
</tr>
<tr>
<td>Mathiowetz et al., 2000, n = 40 males, n = 40 females</td>
<td>Jamar and Baseline</td>
<td>Concurrent validity with known weights</td>
<td>r = 0.92 (injured)</td>
<td>Moderate–excellent reliability but different units of measurement</td>
</tr>
<tr>
<td>Beaton et al., 1995, n = 14</td>
<td>Jamar with elbow at 90° and 45°, and BTE Work Simulator</td>
<td>Inter-instrument reliability</td>
<td>ICC = 0.87</td>
<td>Good–excellent reliability but different units of measurement</td>
</tr>
<tr>
<td>Shechtman et al., 2001, n = 13 patients, n = 13 controls</td>
<td>Jamar and BTE Primus, Work Simulator</td>
<td>Inter-instrument reliability</td>
<td>r = 0.63 and 0.74 patients</td>
<td>Moderate–excellent reliability but different units of measurement</td>
</tr>
<tr>
<td>Hamilton et al., 1992, n = 29 females</td>
<td>Jamar and sphygmomanometer</td>
<td>Inter-instrument reliability with known weights</td>
<td>rho = 0.75</td>
<td>Moderate reliability across three sessions (1 week apart) but different units of measurement</td>
</tr>
<tr>
<td>Lusardi and Bohannon, 1991, n = 34 females</td>
<td>Jamar and modified sphygmomanometer inflated to 20 and 30 mmHg</td>
<td>Inter-instrument reliability with known weights</td>
<td>r = 0.51 and 0.42</td>
<td>Moderate–poor reliability but different units of measurement</td>
</tr>
<tr>
<td>Fike and Rousseau, 1982, n = 486</td>
<td>Jamar and Vigorimeter</td>
<td>Inter-instrument reliability with known weights</td>
<td>r = 0.60</td>
<td>Moderate reliability but different units of measurement</td>
</tr>
<tr>
<td>Stephens et al., 1996, n = 75</td>
<td>Jamar, Modified Jamar and Tekdyne</td>
<td>Inter-instrument reliability with known weights</td>
<td>r = 0.87</td>
<td>Good reliability but different units of measurement</td>
</tr>
</tbody>
</table>

Note: Fess (1987) suggested that correlation coefficients of r ≥ 0.9994 between known weights and instrument readouts were acceptable.
converted to pounds before they can be compared with normative data. If the
BTE work simulator data is not collected following precise hand positioning
on the grip attachment, then the scores cannot be compared. If a therapist
wants to use normative data collected with the Jamar dynamometer, then the
Jamar, Dexter and Baseline dynamometers are the easiest and most reliable to
use.

The second four studies compared the Jamar dynamometer with pneumatic
devices, which have different units of measurement than the Jamar and
require special formulas to convert their data into pounds. The lowest inter-
instrument reliability scores were between the Jamar dynamometer, the
sphygmomanometer (Lusardi and Bohannon, 1991; Hamilton et al., 1992)
and the Vigorimeter (Fike and Rousseau, 1982). Consequently, these instru-
ments are not recommended to be used interchangeably with the Jamar
dynamometer. The Tekdyne dynamometer (Stephens et al., 1996) has good
reliability, as does the BTE, but has the problem of different units of measure-
ment. Therefore, it would be difficult to use interchangeably.

Three studies (Stephens et al., 1996; Bellace et al., 2000; Mathiowetz et al.,
2000) reported acceptable concurrent validity with known weights, and five
studies (King and Berryhill, 1988; Lusardi and Bohannon, 1991; Hamilton et
al., 1992; Beaton et al., 1995; Shechtman et al., 2001) reported that concur-
rent validity with known weights was checked and was acceptable but the
actual correlation values were not reported. The Fike and Rousseau (1982)
study was the only one that did not report checking concurrent validity of their
instruments with known weights. All studies that use dynamometers should
check and report evidence of concurrent validity of their instruments with
known weights, preferably before and after their studies.

Recently, the Rolyan hydraulic hand dynamometer became available com-
mercially. Although it looks similar to the Jamar dynamometer, there are no
studies documenting that it measures equivalently.

**Purposes**

The purpose of this study was to compare the Jamar and Rolyan hydraulic
dynamometers to determine their concurrent validity with known weights as
well as their inter-instrument reliability and concurrent validity for measuring
grip strength in a clinical setting. The Rolyan dynamometer was designed to
measure equivalent to the Jamar dynamometer. Thus, it was expected to show
acceptable concurrent validity with known weights, and excellent inter-
instrument reliability and concurrent validity with the Jamar dynamometer.
An ICC of 0.90 or better would be considered excellent inter-instrument reli-
ability. Non-significant differences between the mean scores of the two
dynamometers and differences of 5.6% or less of the average mean would be
considered evidence of strong concurrent validity (Mathiowetz et al., 2000). If
these types of reliability and validity can be shown, published normative data
collected with the Jamar dynamometer could be used when testing with Rolyan dynamometers.

**Method**

**Design**

This study of reliability and validity used a repeated measures design to control for individual differences of participants. Right and left grip strength of odd-numbered female and male participants were measured with the Rolyan dynamometer first and the Jamar dynamometer second. Even-numbered female and male participants were measured in the opposite sequence. The alternate sequence controlled for potential fatigue effects. The design was consistent with the Mathiowetz et al. (2000) study, which compared the Jamar and Baseline dynamometers.

**Participants**

A convenience sample of 30 females and 30 males, 20 to 50 years of age, right- or left-hand dominant, were recruited. All participants were healthy and free from any neuromuscular or orthopaedic dysfunction affecting hand strength. The actual sample included 30 females from 30 to 49 years of age with a mean age of 38.4 years and 30 males from 23 to 48 years of age with a mean age of 37.8 years.

**Instruments**

The Jamar® hydraulic hand dynamometer (J. A. Preston Corporation, Clifton, NJ) and Rolyan® hydraulic hand dynamometer (Smith & Nephew, Inc., Germantown, WI) were used for measuring grip strength of each participant. Both instruments were set at the second handle position from the inside for all testing.

**Procedures**

Elaine Fess used her standard procedures (Fess, 1987) to evaluate the concurrent validity of the Jamar and Rolyan dynamometers with known weights (that is, the relationship between the instrument readout and applied force) before and after the study. The standard procedures recommended by the American Society of Hand Therapists (Fess, 1992) and used by Mathiowetz et al. (1984) were followed to collect grip strength data. Participants were seated with their shoulders adducted and neutrally rotated, elbow flexed at 90°, forearm in neutral position and wrist between 0 and 30° of flexion and between 0 and 15° of ulnar deviation. After participants were positioned appropriately
and were instructed to squeeze the dynamometer, the standard verbal reinforcement was, ‘Harder! ... Harder! ... Relax’ (Mathiowetz et al., 1984: 224). Three successive measurements were taken for the right hand and then the left hand. The time between trials was about 15 seconds, which was the time needed to read and record each score. The mean of the three trials was used for data analysis. Mathiowetz et al. (1984) achieved excellent inter-rater reliability ($r = 0.99$) and good to excellent test–retest reliability ($r = 0.88$ and $0.93$) using these standard procedures.

Following the measurement of the right hand with the first dynamometer, there was a timed 5-minute interval until the right hand was tested again with the second dynamometer. The left hand was tested during this 5-minute interval. Based on a previous study (Mathiowetz, 1990), the time between trials and between dynamometers was an adequate amount of time to reduce the risk of fatigue. One trained evaluator collected all of the data.

**Data analysis**

Inter-instrument reliability was tested using an ICC (3,1). According to Portney and Watkins (2000), ICC should exceed 0.90 for most clinical measurements. Concurrent validity was evaluated with t-tests for paired data. The $p < 0.05$ level of significance was chosen. To help determine the practical significance of any differences between the means, the percentage of the difference (that is, difference between the means divided by the average mean) was calculated. A difference between the means of greater than 5.6% was considered to be of practical significance, based on a previous study (Mathiowetz et al., 1985), which reported that the standard error of measurement ranged from 2% to 5.6% of the mean scores. In other words, a practical difference between the dynamometers would need to be a difference larger than the standard error of measurement reported in a previous study.

**Results**

**Concurrent validity with known weights**

Concurrent validity with known weights for the Rolyan dynamometer was $r = 0.9994$ and 0.9997 (before and after the study) and for the Jamar dynamometer was $r = 0.9998$ and 0.9998 (before and after the study). Thus, concurrent validity with known weights was acceptable (that is, correlation coefficients were $r \geq 0.9994$) (Fess, 1987) for both dynamometers before and after the study.

**Concurrent validity and inter-instrument reliability**

Descriptive data on grip strength scores for male and female participants are presented in Table 2. The paired-data t-tests for male and female participants
indicate no significant difference between the two dynamometers for both the right and left hands. In addition, the differences between the Jamar and Rolyan dynamometers are much less than 5.6% of the means as expected and are smaller than the 3.1 and 4.8 pounds differences reported by Blair (1989) between three Jamar dynamometers. Therefore, there is no practical significant difference between the means. The lack of statistical or practical differences between the means provides strong evidence of concurrent validity between the Jamar and Rolyan dynamometers. ICCs between the two dynamometers ranged from 0.90 to 0.97 (Table 2), which suggests excellent inter-instrument reliability.

Discussion

The Rolyan dynamometer shows acceptable concurrent validity with known weights, strong concurrent validity, and excellent inter-instrument reliability, supporting the original assumptions of the study. Therefore, it is concluded that the Jamar and Rolyan dynamometers measure equivalently for practical purposes and can be used interchangeably. As a result, individuals using the Rolyan dynamometer are justified in using the normative data, which were collected with the Jamar dynamometer (Mathiowetz et al., 1985, 1986).

When the results of this study are compared with those of previous studies (Table 1), the Rolyan dynamometer has comparable evidence of reliability and validity with the Jamar dynamometer as do the Dexter and Baseline dynamometers with the Jamar dynamometer. Because these four instruments have the same unit of measurement, it is easy to use them interchangeably.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Hand</th>
<th>Jamar M</th>
<th>Rolyan M</th>
<th>d</th>
<th>%a</th>
<th>t (29)</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>R</td>
<td>109.7</td>
<td>110.7</td>
<td>–1.0</td>
<td>(0.9)</td>
<td>–0.92 NS</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>25.0</td>
<td>24.5</td>
<td>6.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>L</td>
<td>108.7</td>
<td>108.6</td>
<td>0.1</td>
<td>(0.1)</td>
<td>0.06 NS</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>19.7</td>
<td>19.1</td>
<td>8.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>R</td>
<td>71.1</td>
<td>69.5</td>
<td>1.6</td>
<td>(2.3)</td>
<td>1.47 NS</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>12.9</td>
<td>14.2</td>
<td>6.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>L</td>
<td>62.9</td>
<td>63.1</td>
<td>–0.2</td>
<td>(0.3)</td>
<td>–0.15 NS</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>12.4</td>
<td>12.4</td>
<td>4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a% = d divided by the average of the two means
NS = No significant differences (p < 0.05, two-tailed) between the means of the two dynamometers
contrast, the Rolyan dynamometer has better evidence of reliability with the Jamar than the other instruments (that is, the BTE work simulator, the BTE Primus, the sphygmomanometer, the Vigorimeter, and the Tekdyne) and has the added advantage of having the same unit of measurement, which these other instruments do not have.

**Limitations**

The ability to interchange grip dynamometers should never be assumed. This study used one Jamar and one Rolyan dynamometer. It cannot be assumed that all Jamar and all Rolyan dynamometers measure equivalently unless their concurrent validity with known weights is acceptable. Similarly, because two reports (Flood-Joy and Mathiowetz, 1987) have provided evidence that different versions of the Jamar do not always measure equivalently, it cannot be assumed that all Rolyan dynamometers read equivalently. Thus, therapists are encouraged to always use the same dynamometer when pre- and post-testing clients (Flood-Joy and Mathiowetz, 1987). Finally, the results of this study should not be generalized to children and older adults because they were not represented in this study. These issues need to be addressed in future studies.

**Acknowledgement**

Smith & Nephew, which manufactures the Rolyan dynamometer, funded this study. However, the author retained the rights to publish the results of the study in a refereed journal without interference of the company. The author has no financial or proprietary interest in the Rolyan dynamometer.

**References**


Address correspondence to Virgil Mathiowetz, PhD, OTR, FAOTA, Associate Professor, Department of Physical Medicine & Rehabilitation, Program in Occupational Therapy, University of Minnesota, MMC 388, Minneapolis, MN 55455. Email: mathi003@umn.edu

---

Comparison of Rolyan and Jamar dynamometers

---


Address correspondence to Virgil Mathiowetz, PhD, OTR, FAOTA, Associate Professor, Department of Physical Medicine & Rehabilitation, Program in Occupational Therapy, University of Minnesota, MMC 388, Minneapolis, MN 55455. Email: mathi003@umn.edu