Physiotherapeutic treatment and clinical evaluation of shoulder disorders

Vermeulen, H.M.

9 A COMPARISON OF TWO PORTABLE DYNAMOMETERS IN THE ASSESSMENT OF SHOULDER AND ELBOW STRENGTH

H. M. Vermeulen[1], G. H. de Bock[5], H. C. van Houwelingen[2], R. L. van der Meer[6], M. C. Mol[6], B. T. Plus[6], P. M. Rozing[3], T. P.M. Vliet Vlieland[1,4],
[1]dept. of Physical Therapy
[2]dept. of Medical Statistics
[3]dept. of Orthopaedics
[4]dept. of Rheumatology Leiden University Medical Center
[5]dept. of Epidemiology University of Groningen
[6]dept. of Physical Therapy Hogeschool Leiden, University of professional education

Physiotherapy 91 (2005) 101–112

Abstract

Objectives: To compare the practical applicability and measurement properties of a hand-held dynamometer (MicroFET2®) and a fixed dynamometer (Isobex2.1®) in determining isometric strength of the shoulder and elbow.

Design: Muscle strength in four directions (glenohumeral abduction, external rotation and elevation and elbow flexion) was measured using both instruments by two examiners. The assessments were repeated by one of the examiners three days later.

Setting: Leiden University Medical Centre

Participants: 20 healthy volunteers

Main outcome measures: Time to complete a set of measurements and discomfort were recorded. To determine intra-observer and inter-observer reliability, intra-class correlation coefficients (ICCs), limits of agreement and smallest detectable difference were computed.

Results: The time to complete a set of measurements was significantly shorter for the hand-held dynamometer than for the fixed dynamometer in both examiners. The number of subjects reporting discomfort was similar with the two dynamometers. Except for glenohumeral abduction, the forces measured using the hand-held dynamometer were significantly higher than those when using the fixed dynamometer in both examiners. The intra-observer and inter-observer ICCs for the four directions ranged from 0.82 to 0.98 for both dynamometers. However, the mean differences between replications and the wide limits of agreement suggest substantial bias and variability. For example, for the measurement of shoulder abduction with the fixed dynamometer by one tester (190 N), the results suggest that on 95% of occasions the second tester's measurement would be between 158 and 275 N.

Conclusions: Although time taken and discomfort should be considered in the selection of dynamometers, due consideration should be given to the significant differences in absolute results. Neither the dynamometers nor the testers can be considered interchangeable. Both the intra-observer and inter-observer reliability of the two dynamometers were similar, yet both demonstrated systematic bias and variability in the measurements obtained.

Key messages: The intra-observer and inter-observer reliability of a hand-held (MicroFET2®) and a fixed dynamometer (Isobex2.1®) were moderate.
Practical considerations of a portable dynamometer in conjunction with reliability are important in the selection for use in clinical practice.
As absolute results obtained with various types of dynamometers or by various testers may differ significantly, neither instrument nor tester can be considered interchangeable.

Introduction

Measurement of muscle force is frequently used in physiotherapy practice, to establish anomalies, set therapy goals and evaluate the effect of intervention. The classical method of muscle testing defined be the Medical Research Council [1] utilises a 6-point nominal scale, providing a quick and global impression of muscle strength; however, its reliability is reported to be limited. [2] Dynamometers are also used to measure muscle strength, although they can be large, expensive and time-consuming to use, making them unsuitable for routine clinical practice. To address this, several portable dynamometers have been designed differing in mode of operation.

Hand-held dynamometers contain strain gauges and fit in the palm of the hand allowing the operator to provide direct resistance to movement of the extremity. A measurement of the force output is obtained electronically. Fixed dynamometers, which also incorporate strain gauges and provide an electronic output, can be attached to the wall or floor and use a cable with a fixation band against which the subject exerts a force.[3]

The hand-held MicroFET2® and the fixed dynamometer, the Isobex2.1® have been found to have good intra-rater and inter-rater reliability [4–7] and have been used in several clinical studies to measure post-operative muscle strength. [8–13]. Hayes investigated the intra-rater and inter-rater reliability of a hand-held dynamometer, a springscale dynamometer and manual muscle testing in 17 symptomatic subjects. The hand-held dynamometer was found to be the most reliable and discriminatory for assessing the strength of the rotator cuff. [14] In spite of the good reliability of portable dynamometers, their practical use in conjunction with their measurement properties have not been directly compared.

The aim of this study was to compare the practical applicability and reliability of two portable dynamometers, a hand-held dynamometer, MicroFET2® and a fixed dynamometer, Isobex2.1® in measuring muscle strength of the shoulder and elbow.

Methods

Design

The shoulder and elbow strength of the subject's right arm were measured with both the hand-held dynamometer and the fixed dynamometer by two examiners (RLM and MCM) with a 30 minute interval. The instruments were presented in a random order. After three days, one of the examiners (RLM) repeated the procedure, using a presentation sequence which balanced that of the same examiner on the first day. The study was performed at the Leiden University Medical Centre.

Subjects

Twenty healthy volunteers, who were 18 years old or over and without a history of shoulder complaints were recruited and entered the study after written informed consent had been obtained. Height (cm), weight (kg) and hand dominance of all subjects were recorded.

During the tests, an independent observer (BTP) recorded all measurements, the total time required for instruction and for carrying out the measuring technique. After completing each set of measurements with both devices on the first day, all subjects filled in a questionnaire in which they were asked about the perception of any discomfort with the application of each device. In addition, subjects were asked which of the two devices they preferred and the reasons for their choice. The independent observer also monitored the execution of the procedure, timed the periods of contractions and rest and recorded the special features of the use of both devices.

Both examiners and subjects were unaware of the scores that were obtained. Moreover, the examiner who repeated the procedures after three days was unaware of his scores from the first round.

Devices

The MicroFET2® hand-held dynamometer (Hoggan Health Industries Inc., Draper, USA) is a compact and lightweight instrument (0.45 kg) that fits in the palm of the hand. Three independent force transducers measure changes in the applied force on the head of the instrument in orthogonal planes. The digital output is in Newtons (N).

The Isobex2.1® fixed dynamometer (Cursor AG, Bern, Switzerland) is also a lightweight (1.0 kg) and portable measurement device. It requires fixation to a smooth surface by two suction-pads. Force exerted is measured through a steel wire elongated with an adjustable strap for fixation to the patient. Force is displayed in kilograms following a three second isometric contraction. The Isobex2.1® emits an audible signal when the display should be read.

Test positions and instructions

Each participant received standardised instructions and encouragement by the examiners during the measurement sessions. Four test positions which are frequently used in daily practice to obtain an impression of muscle strength in shoulder and elbow disorders were selected. In each test position, the participant was seated on a stool. Resistance was applied perpendicular to the limb segment at the appropriate joint angle. To obtain this position the hand-held dynamometer was placed against the extremity while the fixed dynamometer was attached to the floor or to the wall with the strap around the participant's arm.

Glenohumeral abduction force was measured with the shoulder abducted to 45° and the elbow in 90° of flexion.[15] The hand-held dynamometer was placed just above the

Figure 9.1.a Testing abduction strength with hand-held dynamometer.

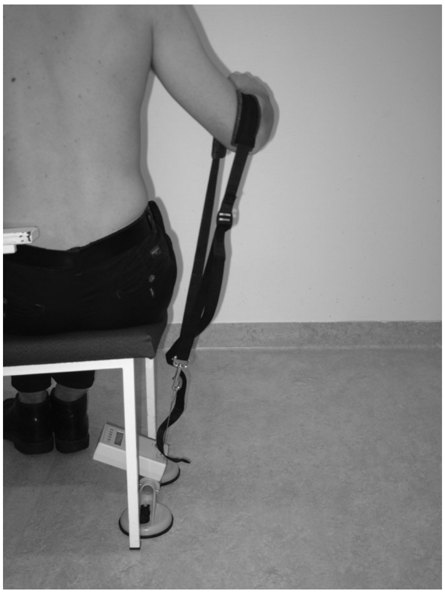

Figure 9.1.b Testing abduction strength with fixed dynamometer.

Figure 9.1.c Testing external rotation strength with hand-held dynamometer.

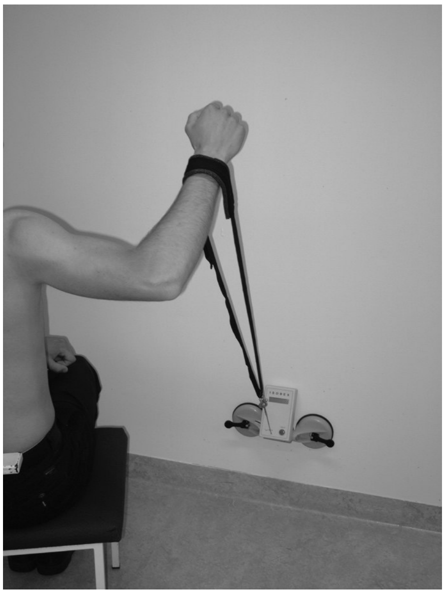

Figure 9.1.d Testing external rotation strength with fixed dynamometer.

Figure 9.1.e Testing elevation strength with hand-held dynamometer.

Figure 9.1.f Testing elevation strength with fixed dynamometer.

Figure 9.1.g Testing elbow flexion strength with hand-held dynamometer.

Figure 9.1.h Testing elbow flexion strength with fixed dynamometer.

lateral epicondyle of the humerus (Figure 9.1a). The fixed dynamometer was attached to the floor under the stool; the strap was placed over the distal end of the humerus and adjusted to take up the play in the cable (Figure 9.1b).

Glenohumeral external rotation force was measured with the humerus in 90° abduction and externally rotated to 70° and the elbow in 90° of flexion.[15] The hand-held dynamometer was placed on the distal dorsal side of the forearm while the other hand of the investigator was placed on the humerus to remind the participant to maintain arm abduction (Figure 9.1c). The fixed dynamometer was attached to the wall in front of the participant at knee height. The strap was placed proximal to the wrist (Figure 9.1d).

Glenohumeral elevation force was measured with the humerus in a position of 90° of elevation, 30° in front of the coronal plane. The humerus was internally rotated and the forearm in pronation.[14] The hand-held dynamometer was placed on the distal, ulnar side of the forearm (Figure 9.1e), while the fixed dynamometer was attached to the floor beside the participant's right foot, perpendicular to the fixation of the band at the wrist (Figure 9.1f).

Elbow flexion force was tested with the humerus by the participant's side, the elbow flexed to 90° and the forearm supinated.[15] The hand-held dynamometer was placed against the flexor aspect of the wrist (Figure 9.1g). The fixed dynamometer was attached to the floor beside the participant and the strap was placed on the flexor aspect of the wrist (Figure 9.1h).

Procedure

The subject exerted a maximal isometric muscle contraction lasting 3 seconds while the dynamometer was held stationary.[16] This was repeated 3 times with a 30-second rest period. With the hand-held dynamometer, a stopwatch with an audible signal was used by the observer to indicate that the 3 seconds were over. Both examiners issued the same instructions and encouragement to promote maximal effort in each test. As both examiners were inexperienced in the use of dynamometers, a period of ten hours testing, evaluation and improvement of the test procedure preceded the study.

Analysis

The total time to complete a set of measurements with each device in one subject (4 positions with 3 contractions including time to rest, mounting and demounting) was compared for both examiners, and between both examiners, using a paired t-test. The number of subjects reporting discomfort at the first measurement session was compared for both devices and examiners using a Chi-Square test.

Due to differences in units of measurements, the results for the fixed dynamometer (kg) were converted into Newtons (N) before analysis (1 kg=9.8 N). In all analyses the mean of three contractions in one subject for a specific test position, a specific instrument and by a specific examiner was used.

To examine the agreement between the two devices, the Bland and Altman method was used.[17,18] First, the mean difference ($d1$) with the standard deviation (sd) between the results of both instruments for all measurements done by the two observers, was computed for the 4 test positions. To test whether there was a systematic difference between the results of the two devices, the 95% confidence interval for $d1$ was calculated. It is accepted that where zero lies outside the 95% confidence interval, a systematic difference between the results of the two devices exists. The 95% limits of agreement were defined as the mean difference between the results of the two devices, $d1$, ± 1.96 sd of the difference, indicating the total error. The difference, $d1$, was than plotted against the means of the measurements obtained with the two devices. This plot shows the size, direction, and range of the differences and whether the differences are consistent across the range of measurements.

For the quantification of intra-observer and inter-observer reproducibility, both the Bland and Altman method for assessing agreement and the calculation of the intra-class correlation coefficient (ICC) for the evaluation of reliability were used. [19]

Intra-observer and inter-observer agreement were quantified by calculating the mean difference between the two observers ($d2$), and between the two measurements of one observer ($d3$), respectively, and the standard deviation (sd) for these differences. In addition, the 95% confidence intervals for $d2$ and $d3$ were calculated. The 95% limits of agreement were defined as the mean difference between the two observers or between the two measurements of one observer, *d2 or d3*, ± 1.96 sd of the differences, indicating the total error. Differences $d2$ and $d3$ were plotted against the means of the measurements of the two observers and the means of the two measurements by observer 1.

The smallest detectable difference (SDD) of the various measurements was calculated as 1.96 • S.D.diff. in which S.D.diff represents the standard deviation of the difference of the measurement values between or within the observers. Although no clear criteria for an acceptable value of intra-observer and inter-observer agreement are available, in general a difference between or within observers of 10% of the total range is considered acceptable. [20]

To examine reliability, ICCs (2,1) with their 95% confidence interval were calculated by using a two-way random effect model. [19,21] All analyses were performed in SPSS statistical software. [22]

Results

Twelve male and eight female healthy subjects participated in the study. Their mean age was 23 years (sd 4), mean height 177 cm (sd 9) and mean weight 75 kg (sd 14). All but one were right-handed.

Practical application

The time recorded to complete the measurements with both instruments is shown in Table 9.1. The results indicate that the time needed to complete a set of measurements

Table 9.1 Practical applicability of two portable dynamometers during the assessment of shoulder and elbow strength.

	Hand-held dynamometer			Fixed dynamometer			Hand-held vs. fixed	Hand-held vs. fixed
	tester 1	tester 2	*P*-value tester 1 vs. 2	tester 1	tester 2	*P*-value tester 1 vs. 2	*P*-value tester 1	*P*-value tester 2
Time needed to complete set of measurement (minutes) (sd)	9.10 (47)	8.36 (37)	0.003	12.15 (61)	11.21 (63)	0.003	<0.001	<0.001
Number of subjects reporting discomfort	4	6	0.480	7	5	0.414	0.257	0.763
Nature of discomfort								
Pressure	4	6		3	1			
Friction				4	4			

with the hand-held dynamometer was significantly shorter than with the fixed dynamometer. In addition, with both instruments, tester 1 needed significantly more time than tester 2 to complete a set of measurements.

The numbers of subjects reporting discomfort during the exertion of force was not significantly different between testers or between devices.

In those subjects who reported some degree of discomfort, this was described as 'pressure' on 10 occasions with the hand-held dynamometer and on 4 occasions with the fixed dynamometer. With both dynamometers, this discomfort was felt in the area of the wrist where the resistance to the force was applied. A sensation of 'friction' was reported on 8 occasions with the fixed dynamometer and was associated with skin contact with the strap. Overall, 12 of the 20 subjects preferred the hand-held dynamometer to the fixed dynamometer, the main reason given as the friction discomfort experienced between the strap and the upper arm when using the fixed dynamometer.

On three occasions, the fixed dynamometer became detached from the wall or floor when force was applied, while on one occasion when using the hand-held dynamometer, the examiner was not able to resist the abduction force applied by the subject.

The endurance of the battery of the fixed dynamometer was limited; however, connecting the device to a mains supply resolved this problem. The batteries of the hand-held dynamometer were unproblematic during the complete test period.

Table 9.2 Agreement between two portable dynamometers for assessing shoulder and elbow strength (Newtons).

Muscle strength [a]	Hand-held dynamometer mean (sd)	Fixed dynamometer mean (sd)	Mean difference (sd) Hand-held minus fixed	95% confidence interval	95% Limits of agreement
Shoulder abduction	170 (48)	176 (65)	- 6.5 (37.7)	-16.3–3.3	-80.5– 67.5
Shoulder external rotation	113 (32)	92 (36)	21.4 (21.1)	15.9–26.8	-17.6– 65.2
Shoulder elevation	99 (29)	70 (24)	29.9 (13.8)	26.3–33.5	2.3–57.5
Elbow flexion	272 (74)	222 (70)	50.5 (24.3)	44.3–56.8	1.9– 99.1

[a] mean value of three repetitions

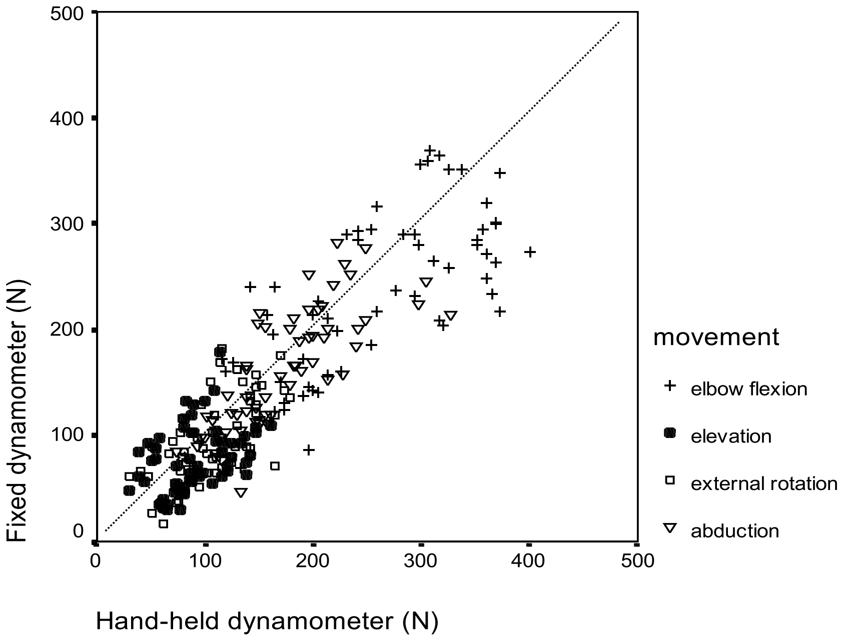

Figure 9.2. Mean results of three contractions of all subjects, on both days and by both examiners for the hand-held dynamometer (x-axis) and the fixed dynamometer (y-axis).

Measurement outcome

In Table 9.2, the results for all measurements in all four test positions for each instrument are presented. Except for the results of shoulder abduction, the mean results of the hand-held dynamometer were significantly higher than those of the fixed dyna -

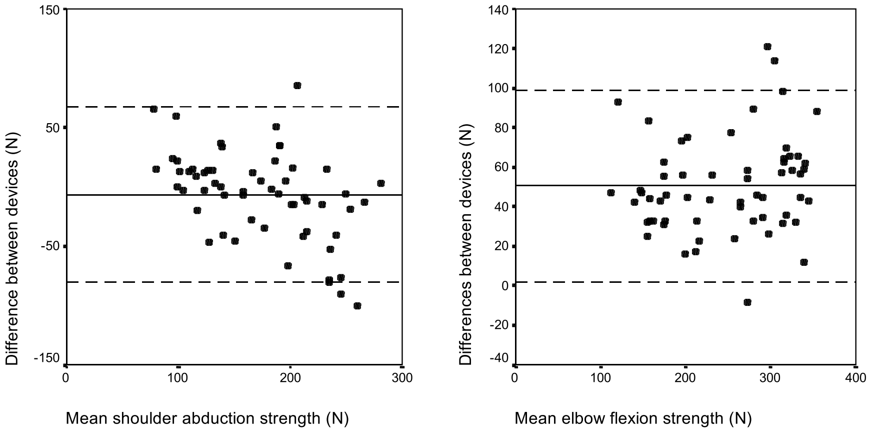

Figure 9.3a and 9.3b. The differences between the hand-held and the fixed dynamometer plotted against the mean value of both devices for shoulder abduction strength (left) and elbow flexion strength (right)(pooled data from day 1 and day 2 and tester 1 and tester 2, n=60). ——— mean differences; -------limits of agreement.

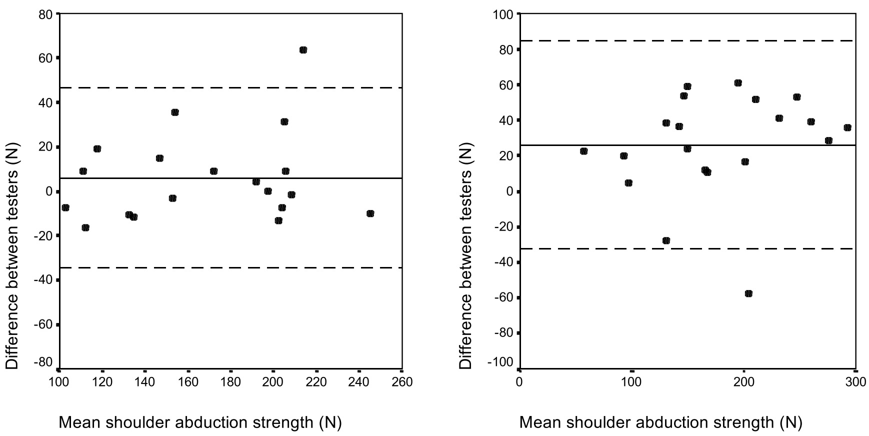

Figure 9.4a and 9.4b. The differences between testers plotted against the mean value of both testers for shoulder abduction strength for the hand-held dynamometer (left) and the fixed dynamometer (right) (n=20). ——— mean difference; -------limits of agreement.

mometer. As zero lies outside the 95% confidence interval of the mean difference between the devices for shoulder external rotation, shoulder elevation and elbow flexion, it was concluded that a systematic difference existed between the results for these devices and test positions.

Figure 9.2 also illustrates that the scores of the hand-held dynamometer were in general slightly higher than those of the fixed dynamometer, irrespective of the examiner who performed the assessment or the day the assessment was done. Figures

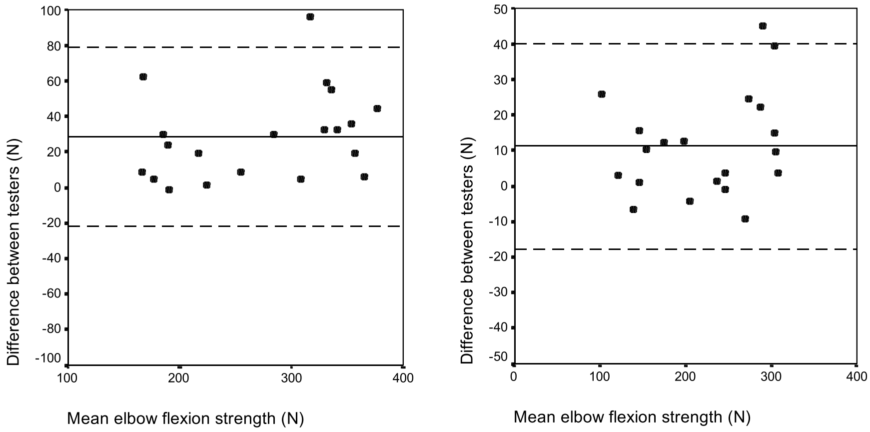

Figure 9.5a and 9.5b. The differences between testers plotted against the mean value of both testers for elbow flexion strength for the hand-held dynamometer (left) and the fixed dyna - mometer (right) (n=20). —— mean differences; -------limits of agreement.

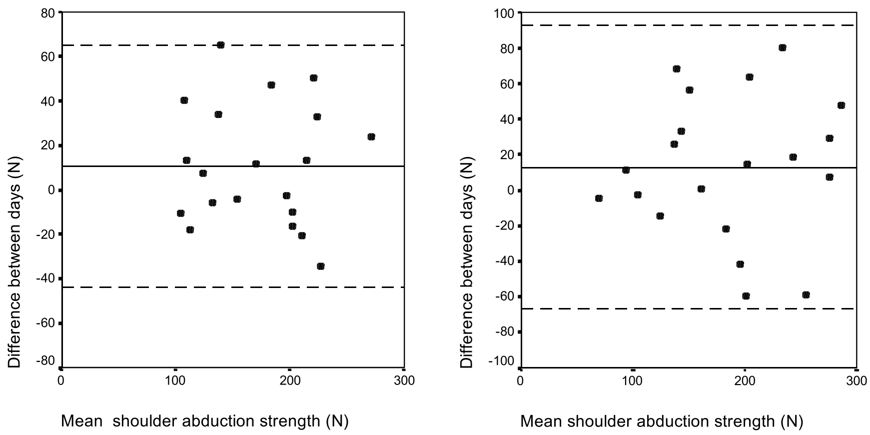

Figure 9.6a and 9.6b. The differences between day 1 and day 2 for tester 1 plotted against the mean value of both days for shoulder abduction strength for the hand-held dynamometer (left) and the fixed dynamometer (right) (n=20). —— mean differences; -------limits of agreement.

9.3a–9.3b show the Bland and Altman plots of agreement between the two devices for shoulder abduction and elbow flexion strength. These figures illustrate that the mean difference for shoulder abduction strength (Figure 9.3a) is around zero, whereas the mean difference and the limits of agreement for elbow flexion strength (Figure 9.3b) are above zero, indicating a systematic difference between the two devices.

Tables 9.3 and 9.4 summarize the inter-observer and intra-observer agreement for both instruments in all four test positions.

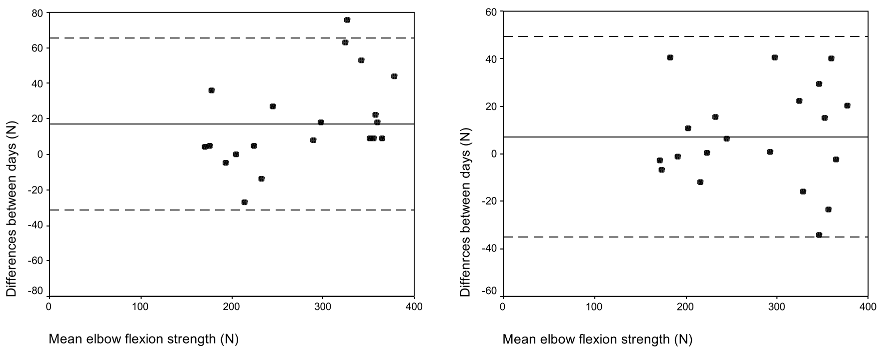

Figure 9.7a and 9.7b. The differences between day 1 and day 2 for tester 1 plotted against the mean value of both days for elbow flexion strength for the hand-held dynamometer (left) and the fixed dynamometer (right) (n=20). —— mean differences; -------limits of agreement.

The data in Table 9.3 illustrates the significant bias between the two testers when using the hand-held dynamometer to measure elbow flexion force and with the fixed dynamometer when measuring shoulder abduction, shoulder elevation and elbow flexion force (zero lying outside the 95% confidence interval for *d2)*.

The inter-observer ICCs were similar for the two instruments, ranging from 0.86 to 0.98. With the exception of shoulder abduction and external rotation with the fixed dynamometer (12% and 15% for tester 2, respectively), all smallest detectable differences were larger than the predefined 10% difference of the total range of measurement. Figures 9.4a–9.5b and 9.5a–9.5b show the differences between the testers plotted against the mean value of both testers for shoulder abduction and elbow flexion strength.

For intra-observer agreement, a significant difference between the results on the first and the second occasion was seen with the hand-held dynamometer for elbow flexion (zero lying outside the 95% CI for *d3)*. The intra-observer ICCs suggest good agreement and were similar for the two instruments (0.82 to 0.96). All SDDs were larger than the predefined 10% difference of the total range of measurement.

Figures 9.6a–9.6b and 9.7a–9.7b show the differences between the same tester on two occasions, plotted against the mean value of the two occasions for shoulder abduction and elbow flexion strength.

Discussion

A comparison of the practical applicability and measurement properties of two portable dynamometers has shown that both instruments cause some discomfort during the testing procedure. The use of the hand-held dynamometer was the least time-consuming. The intra-observer and inter-observer variation of the instruments were similar, but as the absolute results of the two instruments, measured in four different movement directions in the same subjects, varied, the two instruments should not be

Table 9.3 Inter-observer agreement between tester 1 and tester 2 for two portable dynamometers on the same day (Newtons)

Muscle strength [a]*	Tester 1 mean (sd)	Range	Tester 2 mean (sd)	Range	Mean difference (sd) Testers 1and 2	95% confidence interval	95% Limits of agreement	ICC (95% confidence interval)	SDD
Hand-held dynamometer									
Shoulder abduction	172 (46)	99–283	166 (42)[b]	106–250	6.0 (20.2)	-3.8–15.8	-34.4–46.4	0.89 (0.74–0.96)	39.6
Shoulder external rotation	112 (32)	63–166	116 (32)	49–170	- 4.9 (15.7)	-12.3– 2.5	-36.3–26.5	0.88 (0.71–0.95)	30.8
Shoulder elevation	102 (29)	59 to155	96 (29)	56–146	6.4 (15.1)	-6.4–13.4	-23.8–36.6	0.86 (0.68–0.94)	29.6
Elbow flexion	287 (81)	170–398	259 (73)	136–362	28.6 (25.1)	16.8–40.3	-21.6–78.8	0.95 (0.87–0.98)	49.2
Fixed dynamometer									
Shoulder abduction	190 (68)	68–311	164 (62)	20–170	26.1 (29.2)	12.4–39.7	-32.3–84.5	0.90 (0.77–0.96)	57.2
Shoulder external rotation	90 (33)	26–152	87 (37)	31–105	2.7 (18.7)	- 6.0–11.5	-34.7–40.1	0.86 (0.68–0.94)	36.6
Shoulder elevation	72 (24)	36–113	65 (23)	89–306	6.8 (10.6)	1.8–11.8	-14.4–28.0	0.90 (0.76–0.96)	20.8
Elbow flexion	228 (73)	115–323	217 (69)	72–283	11.2 (14.5)	4.4–17.9	-17.8–40.2	0.98 (0.95–0.99)	28.4

[a] Mean value of three repetitions [b] n=19

Table 9.4 Intra-observer agreement for tester 1 for two portable dynamometers on two occasions (Newtons)

Muscle strength [a]	Day 1 mean (sd)	Range	Day 2 mean (sd)	Range	Mean difference (sd) Day 1 and 2	95% confidence interval	95% limits of agreement	ICC (95% confidence interval)	SDD
Hand-held dynamometer									
Shoulder abduction	177 (51)	99–283	167 (52)	87–259	10.8 (27.3)	-2.0–23.6	-43.9–65.3	0.86 (0.68 - 0.94)	53.3
Shoulder external rotation	112 (32)	63–166	111 (33)	56–171	0.5 (13.5)	-5.8–6.9	-26.5–27.5	0.91 (0.80 - 0.97)	26.5
Shoulder elevation	102 (29)	59–155	101 (32)	48–172	1.5 (16.1)	-6.1–9.0	-28.7–31.7	0.86 (0.68 - 0.94)	31.5
Elbow flexion	287 (81)	170–398	271 (69)	167–357	17.1 (24.2)	5.8–28.4	-31.3–65.5	0.95 (0.87 - 0.98)	47.4
Fixed dynamometer									
Shoulder abduction	190 (68)	68–311	177 (66)	72–283	12.6 (39.9)	-6.0–31.3	-67.2–92.4	0.82 (0.61 - 0.93)	78.2
Shoulder external rotation	90 (33)	26–152	97 (40)	28–170	- 7.4 (16.1)	-14.9– 0.1	-39.6–24.8	0.90 (0.77 - 0.96)	31.5
Shoulder elevation	72 (24)	36–113	67 (25)	30–110	5.0 (12.4)	-0.8–10.8	-19.8–29.8	0.88 (0.71 - 0.95)	24.3
Elbow flexion	228 (73)	115–323	221 (73)	75–333	7.1 (21.1)	-2.7–17.0	-35.1–49.3	0.96 (0.90 - 0.98)	41.3

[a] mean value of three repetitions

used interchangeably. Moreover, a systematic difference between the two testers using the same instrument was seen with both devices.

Practical applicability is not often examined in studies on dynamometers, although this may be one of the main factors influencing clinical use. [6] Discomfort was reported several times with both instruments but may be resolved by using a wider and softer strap for the fixed dynamometer or a larger pressure pad for the hand-held dynamometer.

Use of the hand-held dynamometer was less time-consuming than the fixed dynamometer, primarily as a result of the need to fix the device to a suitable surface on the floor or wall and adjust the strap appropriately. The loosening of the fixed dynamometer from the surface could be harmful to patients and safety should be given due consideration; however, this happened in only 3 out of 720 measurements.

On one occasion, the examiner was not able to resist the maximal force of glenohumeral abduction with the hand-held dynamometer. This can be limiting when testing normal subjects or non-affected limbs. [6,23]

The recruitment of healthy subjects without shoulder disorders means that the results of this study cannot be generalised to patients with shoulder disorders. In patients with partial rotator cuff ruptures or impingement syndrome, strength can be reduced to between 37 and 70% of the non-involved side. [24,25] In symptomatic shoulders, the presence of pain may heighten discomfort during testing, and in patients with severe loss of motion, for example, in frozen shoulders, it may be difficult to achieve the test positions. However, loosening or insufficient tester strength will be less problematic when testing patients with decreased muscle strength. The therapist's satisfaction with each portable dynamometer was not investigated and may be a valuable addition to future research.

The absolute results of the hand-held and the fixed dynamometers varied. Apart from glenohumeral abduction, the forces measured for the hand-held dynamometer were higher than those of the fixed dynamometer. The large differences indicate that these devices are not interchangeable. The ICCs calculated (all greater than 0.8) for both intra-observer and inter-observer variation of the portable dynamometers, suggest that both instruments are sufficiently reliable. The results are equal to those of dynamometers used in the studies of Bohannon and Hayes. [14,26] In spite of this, systematic differences between the two testers were seen. This observation indicates that both in daily practice, as in a research setting, any change of assessors may bias the results.

Most importantly, the inclusion of the Bland and Altman method of analysis illustrates the substantial bias and variability between sets of results. For example, for shoulder abduction, when one tester repeats a measurement on a subsequent day, on 95% of occasions the second measurement will be between 67 N less and 92 N more than that measured on the first occasion. This suggests that to be reasonably sure that a real difference in the subject's strength has occurred, an increase of more than 92 N would be required. When combined with the smallest detectable difference of 78 N, a substantial force, this perhaps illustrates the misleading interpretation of ICCs (in this instance 0.82) in reliability studies.

In conclusion, we have shown that both the hand-held and a fixed dynamometer demonstrate a large range of variation and bias when measuring shoulder and elbow strength even in healthy subjects. However, some movements demonstrated less variability than others and the limits of agreement and mean differences for each movement should be considered individually. Both instruments showed some practical disadvantages although the hand-held dynamometer was reasonably quick to use. Future research using a patient cohort may demonstrate greater reliability but should also address therapist's satisfaction in the practical applicability of different measurement devices.

Ethical approval

Medical Ethical Committee of the Leiden University Medical Center

References

1. Medical Research Council. Aids to the investigation of the peripheral nerve injuries. 2nd edn, Her Majesty's Stationary Office, London;1943.
2. Dvir Z. Grade 4 in manual muscle testing: the problem with submaximal strength assessment. *Clin Rehab* 1997;11: 36–41.
3. Brinkmann JR. Comparison of a hand-held and fixed dynamometer in measuring strength of patients with neuromuscular disease. *J Orthop Sports Phys Ther* 1994; 19(2): 100–104.
4. Gerber C, Arneberg O. Measurement of abductor strength using an electronical device (Isobex). *J Shoulder and Elbow Surg* 1993; 2: S6.
5. Janssen AJWM, Elvers JWH, Oostendorp,RAB.Hand Held Dynamometry: intrabeoordelaarsbetrouwbaarheid bij kinderen met neuromusculaire aandoeningen. *Ned T Fysioth* 2001; 111 (6): 159–165.[in dutch]
6. Leggin BG, Neuman RM, Iannotti JP, Williams GR, Thompson EC. Intrarater and interrater reliability of three isometric dynamometers in assessing shoulder strength. *J Shoulder Elbow Surg* 1996;. 5 (1): 18–24.
7. Munneke M, Bekkering WP, Lamers FW. Betrouwbaarheid van het meten van spierkracht met de MicroFET2 bij gezonde, zesjarige kinderen . *Ned T Fysioth* 1998; 108, (4): 88–94. [in dutch]
8. Agteresch HJ, Dagnelie PC, Gaast van der A, Stijnen T, Wilson JHP. Randomized clinical trial of adenosine 5'-triphosphat in patients with advanced non-small-cell lung cancer. *J Natl Cancer Inst* 2000; 92 (4): 321–8.
9. Chammas M, Meyer zu Reckendorf G, Allieu Y. L'arthrodèse d'épaule pour paralysie post-traumatic du plexus brachial. *Revue de Chirurgie Orthopédique* 1996;. 82:386–395.[in french]
10. Galatz LM, Ball MC, Teefey SA, Middleton WD, Yamaguchi K. The outcome and repair integrity of completely arthroscopically repaired large and massive Rotator Cuff tears. *JBJS* 2004; 86-A (2):219–224.
11. Jost B, Pfirrmann CWA, Gerber C. Clinical outcome after structural failure of Rotator Cuff repairs. *JBJS* 2000; 82-A (3):304–14.

12. Jost B, Puskas G, Gerber C. Outcome of Pectoralis Major transfer for the treatment of irrepairable Subscapularis tears. *JBJS* 2003; 85-A (10):1944–51.
13. Lang CE, Bastian AJ. Cerebellar subjects show impaired adaptation of anticipatory EMG during catching. *Am Physiol Soc* 1999:2108–19.
14. Hayes K, Walton JR, Szomor ZL, Murrell GC. Reliability of 3 methods for assessing shoulder strength. *J Shoulder Elbow Surg* 2002; 11 (1): 33–9.
15. Kendall FP, Kendall – McCreary E, Geise Provance P. Muscles, testing and function., 3th ed. Houten The Netherlands 2000.
16. Bohannon RW. Make tests and brake tests of elbow flexor muscle strength. *Phys Ther*1988; 68:193–4.
17. Altman DG, Bland JM. Measurement in medicine. *Statistician* 1983; 32:307–17.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet* 1986;8:307–10.
19. Rankin G, Stokes M. Reliabilty of assessment tools in rehabilitation: an illustration of appropriate statistical analysis. *Clin Rehabil* 1998;12 (3):187–99.
20. Smidt N, Windt van der DAWM, Assendelft WJ, Mourits AJ, Devillé WL, Winter de F, Bouter LM. Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. Arch Phys Med Rehabil 2002;83:1145–50. *Erratum in Arch Phys Med Rehabil* 2003;84:938.
21. Shrout PE, Fleiss JL. Intraclass correlations: use in assessing rater reliability. *Psychol Bull* 1979;86(2):420–8.
22. SPSS [9.0.0]. Chicago: SPSS inc.;1998.
23. Wikholm JB, Bohannon RW. Hand-held dynamometer measurements: tester strength makes a difference. *J Orthop Sports Phys Ther.* 1991;13 (4):191–7.
24. Ben-Yishay A, Zuckerman JD, Gallagher M, Cuomo F. Pain inhibition of shoulder strength in patients with impingement syndrome. *Orthopedics* 1994; 17:685–8.
25. Itoi E, Minagawa H, Sato T, Sato K, Tabata S. Isokinetic strength after tears of the supraspinatus tendon. *JBJS* 1997; 79-B(1):77–82.
26. Bohannon RW. Comparability of force measurements obtained with different strain gauge hand-held dynamometers. *J Orthop Sports Phys Ther* 1993;18 (4):564–7.