A Comparison of the Reliability of Make Versus Break Testing in Measuring Palmar Abduction Strength of the Thumb

J. X. Lim · R. X. Toh · S. K. H. Chook · S. J. Sebastin · T. Karjalainen

Received: 18 April 2013 / Accepted: 2 October 2013 / Published online: 13 October 2013
© Society of the Hand & Microsurgeons of India 2013

Abstract Previous studies have established the role of quantitative measurements of palmar abduction strength of the thumb (PAST). This study compares the reliability of the ‘make’ versus the ‘break’ test in measuring PAST in healthy volunteers. In a ‘make’ test, the body part being tested is positioned at the start of its range of motion and the participant is asked to exert his/her maximal force. In a ‘break’ test, increasing force is applied to a body part after it has completed its range of motion, until the joint being tested gives way. PAST was measured in both hands in 100 healthy volunteers using a handheld device. Two examiners measured PAST using both the ‘make’ and ‘break’ test to determine inter-rater reliability. The tests were repeated in 30 volunteers 6 weeks after the initial testing to determine intra-rater reliability. Our results showed that the ‘make’ test has better inter and intra-rater reliability.

Keywords Abductor pollicis brevis · Break test · Carpal tunnel syndrome · Make test · Palmar abduction · Thumb abduction strength

The measurement of palmar abduction strength of the thumb (PAST) provides an objective assessment of thenar muscle

function. It is a useful research tool in patients with carpal tunnel syndrome (CTS). We have previously reported on normative values of PAST in our population [1]. In our previous study, we realized that PAST could be measured using a ‘make’ test or a ‘break’ test. In the ‘make’ test, the body part being tested is positioned at the start of its range of motion and the participant is asked to exert his/her maximal force against the examiner or a measuring device. In a ‘break’ test, increasing force is applied to a body part after it has completed its range of motion, until the participant’s maximal muscular efforts is overcome and the joint being tested gives way. We used the ‘break’ test in our previous study and suggested that the ‘make’ test would have better control over the amount and rate of force application.

The aim of this study is to compare the reliability of the ‘make’ versus the ‘break’ test in measuring PAST in healthy volunteers. We hypothesized that the ‘make’ test has better inter and intra-rater reliability compared to the ‘break’ test.

Subjects and Methods

Subjects A sample of a normal population was taken from the visitors, staff, and medical students at our hospital. Subjects below 21 years of age, or those with diabetes mellitus, entrapment neuropathy, degenerative joint disease, previous trauma, surgery to the upper limb, and those who were pregnant were excluded from the study. The sample population included 100 subjects and information regarding age, gender, hand dominance, height (cm), and weight (kg) was obtained. We did not collect demographic data on the subjects’ work and hobbies as these were not relevant to a study trying to compare the reliability of a test.

Fifty-eight men and 42 women aged 21–87 (mean 25; SD 11.2) were examined. The mean weight was 63 kg (SD 13.7) and the mean height 169 cm (SD 8.3). Six participants were

J. X. Lim · S. J. Sebastin (✉)

Department of Hand and Reconstructive Microsurgery, National University Health System, 1E Kent Ridge Road, NUHS Tower Block, Level 11, Singapore 119228, Singapore
e-mail: sandeepsebastin@gmail.com

R. X. Toh · S. K. H. Chook

Yong Loo Lin School of Medicine, National University of Singapore, 1E Kent Ridge Road, Singapore 119228, Singapore

T. Karjalainen

Department of Surgery, Central Hospital of Central Finland, Keskussairaalantie 19, 40620 Jyväskylä, Finland

Fig. 1 Position of the hand during the ‘make’ test

left hand dominant, two were ambidextrous, and the remainder (92 %) was right hand dominant.

Methods PAST was measured with a commercially available device (Algometer™, JTech Medical, Salt Lake City, UT, USA). This device has two parts: a force transducer and a display device. Two examiners did the measurements: one positioned the subjects’ arm, gave instructions, and applied force (examiner-1) and the other recorded the values (examiner-2). Examiner-2 held the display device such that examiner-1 and the subject were unable to see the reading. The examiners and the protocol remained constant throughout the study. Subjects were seated with the shoulder abducted, elbow flexed to 90°, forearm in pronation, wrist in neutral position, and fingers in extension. This position was maintained using a custom-made frame and two straps:—one across the forearm and one across the wrist (Figs. 1 and 2). The wrist strap prevents wrist flexion.

PAST was measured using the ‘make’ and ‘break’ tests. Each test was repeated twice. The second measurement was

Fig. 2 Position of the hand during the ‘break’ test

Fig. 3 Flowchart of testing protocol

taken 20 s after the first and the average of these two measurements was used for the final analysis. The order of ‘make’ or ‘break’ testing was randomized using a restricted randomization method so as to obtain groups of equal sizes. The participants were tested in groups of ten and asked to draw lots with a 1:1 ratio between participants starting with either the ‘make’ or the ‘break’ test first. The whole procedure was repeated with examiner-1 acting as examiner-2 and vice-versa to obtain data for inter-rater reliability. Overall, eight measurements of PAST were done on each thumb. PAST was measured again in 30 subjects 6 weeks later using both the ‘make’ and ‘break’ tests to obtain data for intra-rater reliability. At this juncture, only four measurements of PAST were done on each thumb (Fig. 3).

In the ‘make’ test, the thumb was positioned at the start of its abduction arc (Fig. 1). The force transducer was held in place with an adjustable holder such that the curved pad of the device rested against the radial aspect of the interphalangeal joint of the thumb. The subject was asked to palmarly abduct his/her thumb. The maximal strength that the participant exerted was recorded. In the ‘break’ test, the thumb was placed in a position of maximal palmar abduction (Fig. 2). The curved pad of the force transducer was placed on the radial aspect of the interphalangeal joint of the thumb at 90° to the long axis of the thumb. The tester exerted a gradually increasing force until the subject could no longer maintain

Table 1 Palmar abduction strength of the thumb

Gender	Hand dominance	Testing method	Mean PAST (N) (Range)	Standard deviation	P-value
Male	Dominant	Make	32 (7–66)	10.0	0.00
		Break	59 (26–110)	15.3	
	Non-dominant	Make	30 (11–52)	8.6	0.00
		Break	57 (26–110)	15.0	
Female	Dominant	Make	19 (7–31)	5.1	0.00
		Break	38 (20–67)	9.0	
	Non-dominant	Make	19 (9–34)	5.6	0.00
		Break	36 (18–65)	9.2	

Table 2 Inter-rater reliability

Method of testing	Hand dominance	ICC (2,1)	95 % confidence interval
Make	Dominant	0.83	0.776–0.872
	Non-dominant	0.85	0.801–0.888
Break	Dominant	0.81	0.752–0.857
	Non-dominant	0.81	0.750–0.856

his/her thumb in maximal palmar abduction. The pressure exerted by the tester was released once the thumb gave way and the highest value displayed on the device was recorded.

Intraclass correlation coefficient (ICC) for single measures was used to determine inter-rater and intra-rater reliability. The ICC has a maximum value of 1.0 for perfect agreement which is when each test produces identical measurements, and a minimum value of 0. An ICC of more than 0.75 may be taken to represent excellent agreement [2].

Results

PAST was significantly higher ($p < 0.05$) in men in both dominant and non-dominant hand using both ‘break’ and ‘make’ methods (Table 1). The ‘make’ test showed a higher inter-rater reliability compared to the ‘break’ test (Table 2). The intra-rater reliability of the ‘make’ test was better in consecutive measurements. However, repeat measurements done 6 weeks later showed similar intra-rater reliability between the two methods (Table 3). The mean PAST obtained using the ‘make’

test was 53 % of the ‘break’ test. This difference was statistically significant ($p < 0.05$).

Discussion

There have been no previous studies that have compared ‘make’ versus ‘break’ tests in measurement of PAST. We could find only two previous articles that have directly compared ‘make’ and ‘break’ testing [3, 4]. Both of them collected data on elbow flexion strength with a single examiner and thus were reporting intra-rater reliability. Bohannon compared ‘make’ and ‘break’ tests using hand-held dynamometry and concluded that one testing method cannot be interpreted as easier to perform or preferable to the other as the reliability of the tests differed by less than 1.5 %. On the other hand, Stratford found that there was a higher reliability coefficient for the ‘make’ test (ICC 0.95) compared to the ‘break’ test (ICC 0.87). However, both of these studies were done on elbow flexors, and only on a small number of female participants (27 and 32 participants respectively). The angular velocity of thumb abduction introduces variability in ‘break’ testing [5]. In addition, factors like the angle at which the handheld device is maintained, the position of the limb during testing, as well as the point to which the device is applied will affect the results. It may be possible to standardize these factors in a research study. However in a clinical situation, it is difficult to control these factors and the results may not be reliable and comparable across the different examiners.

There were seven previous studies that measured thumb abduction strength in normal subjects [1, 6–11] and we have summarized six of them in a previous publication [1]. There was great variation in the results of PAST obtained between

Table 3 Intra-rater reliability

Method of testing	Examiner	Hand dominance	ICC (2,1)	95 % confidence interval
At week 0				
Make	1	Dominant	0.96	0.947–0.976
		Non-dominant	0.95	0.928–0.967
	2	Dominant	0.96	0.943–0.974
		Non-dominant	0.94	0.912–0.959
Break	1	Dominant	0.91	0.875–0.942
		Non-dominant	0.92	0.881–0.944
	2	Dominant	0.92	0.878–0.943
		Non-dominant	0.87	0.816–0.912
Comparing PAST taken at week 0 and week 6				
Make	1	Dominant	0.84	0.748–0.913
		Non-dominant	0.79	0.677–0.884
Break		Dominant	0.86	0.778–0.925
		Non-dominant	0.82	0.715–0.900

the different studies even when a similar method of testing was used, ranging from 1.50 kg for the right hand (without specification of hand dominance or gender) [6] to 6.16 kg in the dominant hand of male participants [1]. This highlighted the importance of standardizing the position to produce reliable and comparable results. Three of them used the ‘break’ test method [1, 7, 10] whereas two used the ‘make’ test method [6, 9]. In addition, four studies measured PAST with the hand in the supine position [1, 6, 8, 10], one measured PAST with the hand in the prone position [7], while Chao et al. [9] did not specify the position of the hand during testing. Rozmaryn et al. [11] measured isometric thumb strength with the thumb abducted to 45° which is neither a ‘make’ nor a ‘break’ test. In addition, the forearm was held in a mid-prone position. These studies also utilized a variety of complicated custom-made devices and gauges to measure thumb abduction strength. While these custom-made devices are useful, data generated from them cannot be used as a reference by the rest of the medical field since they positioned the hand and thumb differently from each other and the devices are not commercially available for use in our clinical work. A hand-held dynamometer like the one used in this study is easily available and was also used by Liu et al. [10].

Only one previous study has assessed the inter- or intra-rater reliability of their test protocol. Liu et al. [10] repeated the measurements on 12 participants to establish inter- and intra-rater reliability. They found an inter-rater reliability of 0.89 and intra-rater reliability of 0.92 with the ‘break’ test, which corresponds well with our results. However, our results suggest that the inter-rater reliability can be further improved by using the ‘make’ test. In the ‘make’ test, the position of the thumb as well as the device can be controlled better and this leads to improved inter-rater reliability. We also noted that PAST obtained using the ‘break’ method was about twice as much as that of the ‘make’ test (Table 1). Therefore, only one method can be used in the same study or in follow-up of patients.

Limitations Firstly, all muscle testing relies on the subject’s cooperation and willingness to exert maximum effort. This may not be significant when normal subjects are evaluated, but can become challenging when patients are evaluated. The mean of two trials rather than a single test was used to increase the consistency of results. In addition, we use only two trials instead of three for each set of test to decrease the chance of muscle fatigue. Next, it is difficult to ensure that the subject is not recruiting other muscles during ‘break’ and ‘make’ testing. We attempted to eliminate the contribution of the wrist flexors by positioning the forearm in pronation with the palm facing downwards, with the wrist straps preventing wrist flexion. Thirdly, it is difficult to keep the direction, amount, and rate of force application constant. We utilized the side of the frame

as a guide for direction of thumb movement for ‘break’ testing and fixed the position of the force transducer for ‘make’ testing. The subject and the recording examiner were blinded to the result, which was noted by the recorder. In addition, the examiners remained constant for all subjects. This was done in order to reduce errors in amount and rate of force application.

It is important to understand the limitations of the measurements we use in clinical practice. PAST testing is a classic example of a poorly precise measurement, because the test set up is hard to standardise. Any variance in the thumb position or angle of the measurement device as well as the velocity of the applied force produces variability in the results. In conclusion, our study has shown that the ‘make’ test is a more reliable method to measure PAST compared to the ‘break’ test. It is easy to standardise and has superior inter-rater reliability when compared to the ‘break’ test. It also has high intra-rater reliability. We have also shown that there is a significant difference between measurements using the ‘make’ test compared to the ‘break’ test. Thus, the results of these tests cannot be compared directly without standardizing and specifying the testing position and method.

Acknowledgments None

Declaration of Conflicting Interests None

Funding None

Ethical Approval Approved by National Healthcare Group (NHG), Singapore Domain Specific Review Board Ref: 2012/00108

Informed Consent Written informed consent were obtained

References

1. Nallakurupan V, Tung EY, Sebastin SJ et al (2012) The effect of blocking radial abduction on palmar abduction strength of the thumb. *J Hand Surg Eur* 37:269–274
2. Fleiss JL, Levin B, Paik MC (2003) The measurement of interrater agreement. In: Fleiss JL, Levin B, Paik MC (eds) *Statistical methods for rates and proportions*. Wiley-Interscience, New York, pp 598–626
3. Bohannon RW (1988) Make tests and break tests of elbow flexor muscle strength. *Phys Ther* 68:193–194
4. Stratford PW, Balsor BE (1994) A comparison of make and break tests using a hand-held dynamometer and the Kin-Com. *J Orthop Sports Phys Ther* 19:28–31
5. Burns SP, Spanier DE (2005) Break-technique handheld dynamometry: relation between angular velocity and strength measurement. *Arch Phys Med Rehabil* 86:1420–1426
6. Agabegi SS, Freiberg RA, Plunkett JM et al (2007) Thumb abduction strength measurement in carpal tunnel syndrome. *J Hand Surg [Am]* 32:859–866

7. Bigelow E, Bell M (1998) Carpal tunnel syndrome: a new objective evaluation technique. *Can J Plast Surg* 6:99–103
8. Boatright JR, Kiebzak GM, O'Neil DM et al (1997) Measurement of thumb abduction strength: normative data and a comparison with grip and pinch strength. *J Hand Surg [Am]* 22:843–848
9. Chao EYS, An KN, Cooney WP, Linscheid RL (1989) Hand functional strength assessment and its clinical application. In: Chao EYS, An KN, Cooney WP, Linscheid RL (eds) *Biomechanics of the hand*. World Scientific, Singapore, pp 97–129
10. Liu F, Carlson L, Watson HK (2000) Quantitative abductor pollicis brevis strength testing: reliability and normative values. *J Hand Surg [Am]* 25: 752–759
11. Rozmaryn LM, Bartko JJ, Isler MLD (2007) The ab-adductometer: a new device for measuring the muscle strength and function of the thumb. *J Hand Ther* 20:311–324