

Comparison between qualitative and quantitative measurement of strength deficit in shoulder flexors of young females: A cross-sectional study

Saadia Perwaiz¹, Muhammad Waqar Afzal², Ghulam Fatima³

Abstract

Objective: To determine the accuracy of manual muscle testing as a diagnostic tool to detect strength deficit across grades 5 and 4 in shoulder flexors at different percentage deficits of dynamometric strength.

Methods: The cross-sectional study was conducted at the University of South Asia, Lahore, in collaboration with the Association of Fatima Jinnah Old Graduates, Lahore, Pakistan, from May to June 2019. It included young healthy females aged 18-39 years without any ailment of upper extremity or generalised health problem. The strength of shoulder flexors was measured through manual muscle testing and hand-held dynamometer. Data was analysed using SPSS 25.

Results: There were 300 women with a mean age of 23.11±4.71 years. The majority 124(41.3%) were aged 20-24 years. The strength of the dominant shoulder flexors of 194(64.7%) participants and the strength of non-dominant shoulder flexors of 206(68.7%) was in grade 4. The overall diagnostic accuracy of manual muscle testing at 10% dynamometric measures was 68.67% (95% confidence interval: 63.21-73.65%) in non-dominant shoulder flexors, and 61.67% (95% confidence interval: 56.05-66.99%) in dominant shoulder flexors. The positive predictive value and specificity of manual muscle testing was better at 10% strength deficit in comparison with dynamometric measures, whereas negative predictive value and sensitivity was better at 20% strength deficit.

Conclusion: The subjective area of grades 4 and 5 showed better, though not ideal, diagnostic worth at 10% deficit of dynamometric measures in shoulder flexors.

Keywords: Diagnostic measures, Muscle strength, Dynamometry, Manual muscle testing. (JPMA 71: 2559; 2021)

DOI: <https://doi.org/10.47391/JPMA.011431>

Introduction

Manual muscle testing (MMT) and hand-held dynamometry (HHD) both have clinometric importance in measuring muscle strength and sense strength differences. Both the tools have varied approaches, as MMT grades the strength, and HHD objectively measures the strength in terms of pounds or kilograms.^{1,2} Correlation studies have found both the tools effective in the detection of strength deficit and differences from moderate to strong levels except for ankle dorsiflexors.^{3,4} The cross-sectional studies however, do not compare and establishes the diagnostic worth of both tools to detect strength differences.

The diagnostic and predictive validity of MMT grades necessitates exploration in the light of objective measures of HHD. The vast range of strength, from 50% onward, attributed to MMT grade 4 (good) and grade 5 (normal) also indicates the need of comparison of grading system with respect to quantified measures. Studies on diagnostic worth of MMT has indicated lack of accuracy to measure strength deficit across subjective range of grades 4 and 5.

The diagnostic worth was better in these studies in grade 3 and below. The overlap in strength in the context of subjectivity in grades 4 and 5 has been indicated as the reason for reluctance to predict strength differences.^{5,6}

Shoulder flexion constitutes an important segment of functional movements of the upper extremity.⁷ Grade 4 and 5 of manual muscle testing have the most subjective range of muscle strength that has been identified as an ambiguous area in many studies.⁸⁻¹⁰ The current study was planned to determine the accuracy of MMT as a diagnostic tool to detect strength deficit across grades 4 and 5 in shoulder flexors at different percentage deficits of dynamometric strength.

Subjects and Methods

The cross-sectional study was conducted at the University of South Asia, Lahore, in collaboration with the Association of Fatima Jinnah Old Graduates, Lahore, Pakistan, from May to June 2019, and comprised women without any ailment of upper extremity or generalised health problem. The subjects enrolled were aged 18-39 years as this range represents young adult females.^{11,12} The subjects were further divided into five age-based subgroups in line with earlier studies.^{13,14} Females having any ailment of upper extremity or comorbid condition were excluded.

¹Akhtar Saeed College of Rehabilitation Sciences, Akhtar Saeed Medical and Dental College, Lahore, Pakistan; ²The University of Lahore Teaching Hospital, Lahore, Pakistan; ³Life and Health Department, University of South Asia, Lahore, Pakistan.

Correspondence: Saadia Perwaiz. e-mail: s_perwaiz@yahoo.com

After taking approval from the ethics review committee of the University of Lahore, the sample size was calculated using the formula,

$$\frac{Z_{1-\frac{\alpha}{2}}^2 \cdot S_n(1-S_n)}{L_2 \times \text{Prevalence}}$$

which is used to calculate the diagnostic worth of any tool.^{5,15} The sample was raised using non-probability convenience sampling technique.

After taking informed consent from all the participants, data was collected using HHD and MMT. Baseline 250 hydraulic push-pull HHD (Baseline Corporation, Irvington, New York) was used for quantification of strength, which measures the strength in both pounds and kilogrammes. The strength of shoulder flexors of both extremities of the participants was measured by MMT and HHD by two female therapists.

MMT was performed using the standardised technique.² The position of the participant was short sitting, arm flexed 90 degrees, elbow slightly flexed. The therapist was standing at the side of the subject, commanding the patient to raise the arm upward while applying maximal resistance to distal humerus with one hand and stabilising with the other hand. The subject was advised to hold against maximal resistance and then, depending upon the effort of the subject, the strength was graded 4 or 5.

For quantification of strength through HHD, the protocol of isometric make test were followed. The position of participants was short sitting. The shoulder was flexed to 90 degrees, the elbow was straight, palm facing inward, the dynamometer was held stationary above the elbow by the therapist and the participant was instructed to exert maximum force against the dynamometer.¹⁷ Second reading involved break of a few seconds and in case more than 10% difference was found between the two readings, a third reading was taken to reduce the risk of potential bias.

The weakness detected by MMT was compared with 10%, 15% and 20% of dynamometric means with respect to

specific age groups.

Data was analysed using SPSS 25. To compare the prediction of accuracy of detection of strength deficit by grades, a 2x2 contingency table was used.^{17,18} Sensitivity, specificity, positive predictive value (PPV) and negative protocol value (NPV) were calculated at different percentages. OpenEpi¹⁹ was used to calculate 95% confidence intervals.

Results

There were 300 women with a mean age of 23.11±4.71 years. The majority 124 (41.3%) were aged 20-24 years (Table 1).

The strength of the dominant shoulder flexors of 194(64.7%) participants and the strength of non-dominant shoulder flexors of 206(68.7%) was in grade 4. In 106(35.3%) participants the strength of dominant shoulder flexors was grade 5 and for non-dominant shoulder flexors 94(31.3%) were in grade 5. Mean values in MMT grade 5 subjects was in used to calculate cut-off reference point with regard to strength deficit 10%, 15% and 20%, Table 2.

Accuracy of MMT for detecting strength deficit in shoulder

Table-1: Mean age of participants in different age groups.

Age groups (years)	n	Mean±SD
18-19	109	19.26±0.69
20-24	124	22.35±1.16
25-29	42	28.50±1.40
30-34	16	33.06±1.73
35-39	9	37.33±1.00

SD: Standard Deviation.

Table-2: Reference dynamometric strength of shoulder flexors with respect to age groups.

Age groups (years)	Dominant (Kgs.)	Non- Dominant (Kgs.)
18-19	10.88±2.63	10.70±2.80
20-24	8.75±2.17	8.56±1.76
25-29	14.20±3.01	12.76±2.71
30-34	12.68±1.28	12.19±1.70
35-39	14.06±6.09	17.23

Mean value in Grade 5 provides as cut-off reference point with regard to which strength deficit at 10%, 15% and 20% was calculated.

Table-3: Accuracy of manual muscle testing for detecting strength deficit in shoulder flexors of young females at different percentages of deficits identified by dynamometry.

Dynamometer Deficit %	Cell counts				Sensitivity % (Lower & Upper 95% CI)	Specificity % (Lower & Upper 95% CI)	PPV % (Lower & Upper 95% CI)	NPV % (Lower & Upper 95% CI)	Diagnostic Accuracy % (Lower & Upper 95% CI)
	a	b	c	d					
Dominant 10%	114	80	35	71	76.51 (69.09, 82.6)	47.02 (39.23, 54.96)	58.76 (51.73, 65.45)	66.98 (57.57, 75.2)	61.67 (56.05, 66.99)
Non-Dominant 10%	164	42	52	42	75.93 (69.80, 81.14)	50.00 (39.54, 60.46)	79.61 (73.59, 84.55)	44.68 (35.04, 54.74)	68.67 (63.21, 73.65)
Dominant 15%	95	99	24	82	79.83 (71.74, 86.06)	45.30 (38.22, 52.58)	48.96 (42.02, 55.95)	77.36 (68.52, 84.29)	59.00 (53.35, 64.42)
Non-Dominant 15%	109	97	24	70	81.95 (74.55, 87.56)	41.92 (34.70, 49.5)	52.91 (46.11, 59.61)	74.47 (64.81, 82.20)	59.67 (54.03, 65.06)
Dominant 20%	79	115	18	88	81.44 (72.56, 87.93)	43.35 (36.72, 50.23)	40.72 (34.05, 47.75)	83.02 (74.75, 88.98)	55.67 (50.01, 61.18)
Non-Dominant 20%	81	125	17	77	82.65 (73.96, 88.88)	38.12 (31.70, 44.98)	39.32 (32.90, 46.13)	81.91 (72.93, 88.39)	52.67 (47.02, 58.25)

PPV: Positive predictive value, NPV: Negative predictive value, CI: Confidence interval, a: True positive, b: False positive, c: False negative, d: True negative. The cut-off point was 10%, 15% and 20% of mean dynamometric strength of dominant and non-dominant shoulder flexors in grade 5 according to age groups.

Table-4: Accuracy of manual muscle testing for detecting strength deficit in shoulder flexors of young females in different age groups.

Dynamometric deficit %	Diagnostic Accuracy %	18-19 (years)	20-24 (years)	25-29 (years)	30-34 (years)	35-39 (years)
Dominant 10 %	Sn. (%)	85.94	52.27	89.65	80.00	85.71
	Sp. (%)	51.11	43.75	61.53	36.36	50.00
	Accuracy (%)	71.55	46.77	80.95	50.00	77.77
Non- Dominant 10%	Sn. (%)	83.95	60.76	83.33	83.33	100
	Sp. (%)	46.42	53.33	33.33	50.00	100
	Accuracy (%)	74.31	58.06	80.95	75.00	100.00
Dominant 15 %	Sn. (%)	84.31	61.76	91.67	100	85.71
	Sp. (%)	41.37	47.78	50.00	38.46	50.00
	Accuracy (%)	61.46	51.61	73.80	50.00	77.77
Non- Dominant 15%	Sn. (%)	81.48	69.76	95.45	100	100.00
	Sp. (%)	29.09	51.85	35.00	40.00	100.00
	Accuracy (%)	55.04	58.06	66.66	62.50	100.00
Dominant 20 %	Sn. (%)	84.61	61.53	95.65	100	83.33
	Sp. (%)	37.14	46.93	52.63	38.46	33.33
	Accuracy (%)	54.12	50.00	76.19	50.00	66.66
Non- Dominant 20%	Sn. (%)	83.33	68.75	100.00	100.00	100.00
	Sp. (%)	28.36	48.91	26.67	30.77	100.00
	Accuracy (%)	49.54	54.03	47.61	43.75	100.00

Sn: Sensitivity, Sp: Specificity.

flexors at different percentages of deficits identified by HHD were also calculated, Table 3.

Age group analysis of MMT at different percentage deficits of dynamometric measures is elaborated in Table 4. Age groups 25-30 years and 35-39 years showed better diagnostic worth.

Discussion

The current study showed that MMT at 10% strength deficit of dynamometric measures had better diagnostic worth. The PPV showing probability of having weakness was highest at 10% deficit. It was 79.61% (95% CI: 73.59-84.55%) in non-dominant shoulder flexors and 58.76% (95% CI: 51.73-65.45%) in dominant shoulder flexors. The NPV was maximum at 20% strength deficit in contrast with the dynamometric mean. It was 83.02% (95% CI: 74.75-88.98%) for dominant shoulder flexors at 20% strength deficit and 81.91% (95% CI: 72.93-88.39%) for non-dominant shoulder flexors. The sensitivity and NPV were directly linked at different percentage deficits, while PPV was directly associated with specificity. The age group of 35-39 years showed deviation from the usual trend, showing 100% diagnostic worth.

Previously, Bohannon compared both the tools through various diagnostic measures to determine the accuracy of MMT with respect to HHD. In studies^{5,6} on knee and grip strengths, he concluded that the accuracy never met the ideal criteria²⁰ to determine diagnostic accuracy of a screening test. In knee extensors, when weakness detected by MMT was compared with 15%, 20%, 25% and 30%

strength deficit of dynamometric measures, it was observed that the diagnostic accuracy was maximum at 30% strength deficit.⁶ It was 73.6% in both dominant and non-dominant extremities. Sensitivity and NPV were directly linked and they also showed better results at 30% deficit. The specificity and PPV were always >80% for different percentage deficits. Similarly, in a study on grip strength, the sensitivity and specificity were always <80% when weakness, in manually-assessed grip strength, was compared with 20% deficit of dynamometric grip strength.⁷ The lack of sensitivity of MMT in measuring strength deficit across grades 4 and 5 have been pointed out as an area requiring further investigations in both the studies.^{6,7} The present study was designed to focus grades 4 and 5 in shoulder flexors. The comparison of MMT with HHD at different percentage deficits through various diagnostic measures revealed insufficient diagnostic accuracy.

Sensitivity, specificity and diagnostic worth never met the criteria for screening tests.²⁰ The 10% strength deficit showed diagnostic accuracy of 61.67% (95% CI: 56.05-66.99%) in dominant shoulder flexors and 68.67% (95% CI: 63.21-73.65%) in non-dominant shoulder flexors.

The interesting finding in the current study was the higher frequency of false positive (FP) cases at different percentage deficits of dynamometric mean. At 20% strength deficit, although sensitivity was satisfactory, but a higher FP rate has reduced the probability of detecting the weakness indicated through PPV. The 10% deficit of dynamometric measures came out as a better approach to detect weakness in collaboration with MMT. In studies on diagnostic accuracies on knee and grip, FP frequencies were lower than false negatives (FN). The sensitivity of MMT grades 4 and 5 to detect strength deficit in shoulder flexors of young females at different percentage deficit of dynamometric measures was always better than specificity, which is in contrast with the finding of Bohannon on knee extensors.⁶

On the basis of the findings, the current study suggests that criteria for formulating reference data needs to be standardised owing to a broad range of strength and overlap in strength in grades 4 and 5. Dynamometer is sensitive enough to detect strength deficits as a prognostic indicator, but it is debatable if it can label a muscle 'normal' while studying strength in young healthy population. It needs to be explored. Most of the reference data compiled earlier considers a normal healthy population as the source

of compilation of data.^{14,21-23} In the light of the current study, deliberation is needed in that process.

MMT, according to the criteria of Hislop and Montgomery,² has suggested 50% or over strength falling in the range of grades 4 and 5. This range, when actualised in clinical practice, gives dynamometer the advantage of being a sensitive instrument to measure strength deficit as a prognostic indicator. Diversity in muscle strength needs to be addressed by devising a combined protocol using both MMT and HHD as strength-measuring tools.

The current study has limitations as it only comprised the female population within a limited age group. Future studies should include the male population and diverse age groups. Further standardisation of dynamometry by considering strength variables also needs to be incorporated to systematise the process of strength measurement. The usual deviation of age group 35-39 years indicates a need for further exploration with a larger number of participants, as this age group lacks ample participants. Moreover, different muscle groups are needed to be explored in the context of objectification of strength and predicting accuracy of strength differences.

Conclusion

The subjective area of grades 4 and 5 showed better, though not ideal, diagnostic worth at 10% deficit of dynamometric measures in shoulder flexors.

Disclaimer: The text is based on an M. Phil thesis.

Conflict of Interest: None.

Source of Funding: None.

References

1. Bohannon RW. Hand-held dynamometry: adoption 1900–2005. *Percept Motor Skills*. 2006; 103:3-4.
2. Hislop HJ, Montgomery J. Daniels and Worthingham's Muscle Testing: Techniques of Manual Examination. New York: Saunders Elsevier, 2007.
3. Bohannon RW. Measuring knee extensor muscle strength. *Am J Phys Med Rehabil*. 2001; 80:13-8.
4. Pfister PB, de Bruin ED, Sterkele I, Maurer B, de Bie RA, Knols RH. Manual muscle testing and hand-held dynamometry in people with inflammatory myopathy: An intra-and interrater reliability and validity study. *PLoS One*. 2018; 13:e0194531.
5. Bohannon RW. Manual muscle testing: does it meet the standards of an adequate screening test? *Clin Rehabil*. 2005; 19:662-7.
6. Bohannon RW. Grip strength measured by manual muscle testing lacks diagnostic accuracy. *Isokinet Exerc Sci*. 2018; 26:1-4.
7. Oatis CA. *Kinesiology: The Mechanics and Pathomechanics of Human Movement*: London: Lippincott Williams & Wilkins, 2004; pp-946.
8. Knepler C, Bohannon RW. Subjectivity of forces associated with manual-muscle test grades of 3+, 4-, and 4. *Percept Motor Skills*. 1998; 87:1123-8.
9. Dvir Z. Grade 4 in manual muscle testing: the problem with submaximal strength assessment. *Clin Rehabil*. 1997; 11:36-41.
10. Sisto SA, Hudson TD. Dynamometry testing in spinal cord injury. *J Rehabil Res Dev*. 2007; 44:123-36.
11. Scharf G, Heineke J. Finding good biomarkers for sarcopenia. *J Cachexia Sarcopenia Muscle*. 2012; 3:145-8.
12. Horng WB, Lee CP, Chen CW. Engineering. Classification of age groups based on facial features. *Tamkang J Sci Engine*. 2001; 4:183-92.
13. Van Harlinger W, Blalock L, Merritt JL. Upper limb strength: study providing normative data for a clinical handheld dynamometer. *PM R*. 2015; 7:135-40.
14. Bäckman E, Johansson V, Häger B, Sjöblom P, Henriksson KJ. Isometric muscle strength and muscular endurance in normal persons aged between 17 and 70 years. *Scand J Rehabil Med*. 1995; 27:109-17.
15. Zaidi M, Waseem H, Ansari MFA, Irfan M, Fahim S. Sample size estimation of diagnostic test studies in health sciences. Karachi: 14th International Conference, 2016; pp - 239-46.
16. Bohannon R. Testing isometric limb muscle strength with dynamometers. *Crit Rev Phys Med Rehabil*. 1990; 2:75-86.
17. Davidson M. The interpretation of diagnostic tests: *Aust J Physiother*. 2002; 48:227-32.
18. Fetterers L, Tilton J. *Evidence Based Physical Therapy*: FA. Davis Company, 2012.
19. Dean AG, Sullivan KM, Soe MM. *OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version.[Online]* [Cited on 11. September 2021]. Available from: URL: www.OpenEpi.com
20. Lincoln N, Nicholl C, Flannaghan T, Leonard M, Van der Gucht E. The validity of questionnaire measures for assessing depression after stroke. *Clin Rehabil*. 2003; 17:840-6.
21. Andrews AW, Thomas MW, Bohannon RW. Normative values for isometric muscle force measurements obtained with hand-held dynamometers. *Phys Ther*. 1996; 76:248-59.
22. Bohannon RW. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Arch Phys Med Rehabil*. 1997; 78:26-32.
23. Bohannon RW, Peolsson A, Massy-Westropp N, Desrosiers J, Bear-Lehman J. Reference values for adult grip strength measured with a Jamar dynamometer: a descriptive meta-analysis. *Physiother*. 2006; 92:11-5.