J. Basic. Appl. Sci. Res., 2(12)13134-13138, 2012
© 2012, TextRoad Publication

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Inter-Examiner Reliability of the Manual Muscle Strength Testing

Maryam Yahyaei-Rad

University of Guilan, Iran

ABSTRACT

Standardized physical examination using manual muscle testing is a widely accepted method for evaluating strength. This study determines the inter-examiner reliability of manual test grades in assessing muscle strength in patients with a diagnosis of neuropathy or myopathy. Subjects were 11 patients, aged 10 to 40 years. 3 physical therapists participated in this study. 18 muscle groups in upper limb were tested bilaterally, using modification of the Medical Research council scale. Reliability of muscle strength grades obtained for individual muscle groups and individual muscle strength grades was analyzed using Cohen's weighted kappa. The reliability of grades for individual muscle groups was from 0.62 to 0.82, with the proximal muscles having the higher reliability values. The reliability of individual muscle strength grades from 0.52 to 0.88, with those in the gravity eliminated range scoring the highest. We find the MMT grades are reliable for assessing muscle strength in muscular weakness.
KEYWORDS: reliability, upper limb, MMT, MRC

INTRODUCTION

Manual muscle testing is the most commonly used method for documenting impairments in muscle strength (17). Manual muscle tests evaluate the ability of the nervous system to adapt the muscle to meet the changing pressure of the examiner's test (13). This requires that the examiner be trained in the anatomy, physiology, and neurology of muscle function (18). The action of the muscle being tested, as well as the role of synergistic muscles, must be understood (6). Manual muscle testing is both a science and an art. To achieve accurate results, muscle tests must be performed according to a precise testing protocol (3). The appeal of the MMT is that it can be performed simply with the patient, an examiner, and a bench or table (24). This makes it ideal for the routine clinical environment where specialized equipment is unavailable and time is short (1). Manual muscle testing is a procedure for evaluating strength and function of an individual muscle or a muscle group in which the patient voluntarily contracts the muscle against gravity load or manual resistance (10,25). It is quick, efficient, and easy to learn, however, it requires total cooperation from the patient and learned response levels by the assessor (19). MMT was developed by Lovett and described by Wright in 1912 (31). This technique has been revised, advanced and promoted so that it has resulted in a range of methods from which the investigator may select the most suitable one (24). The scale proposed by the Medical Research Council (MRC) uses the numeral grades 0–5 (23). Kendall & McCreary (20) use percentages, and Daniels & Worthingham (12) use differentiation between Normal, Good, Fair, Poor, Trace and Zero. Manual grading of muscle strength is based on palpation or observation of muscle contraction, ability to move the limb through its available ROM against or without gravity, and ability to move the limb through its ROM against manual resistance by the examiner. Manual resistance is applied by the examiner using one hand with the other hand stabilizing the joint (5). Exact locations for applying resistive force are specified and must be followed exactly to obtain accurate MMT results (10).This paper will present evidence that the MMT can be a legitimate and useful evaluation tool for the assessment of the musculoskeletal and nervous systems.

METHOD

Subjects were 11 patients, aged 10 to 40 years, with a diagnosis of neuropathy or myopathy. Muscle strength was operationally defined by the MMT grades given. The present study was designed to clarify those results and document the reliability of the evaluation procedures. All subjects were required to be able to cooperate and perform the MMT. Muscle strength grades ranged from 0 to 5. Examiners were three physical therapists with 10 to 12 years of experience. Independently of each other, 3 examiners performed identical manual physical examinations of the maximal voluntary strength in each of the 18 muscles on both sides. Examiners were blind to other testing results. Muscle strength was assessed and individual MMT grades were assigned using a modified MRC (23)grading scale, with subtlivisions of the grades 3, 4, and 5, as follows: 5, 5-, 4+, 4, 4-, 3 + , 3, 3 -, 2, 1, 0. The modifications to the MRC scale included the addition of the grading subdivisions 5-, 3 +, and 3 -.

***Corresponding Author:** Maryam Yahyaei-Rad, University of Guilan, Iran
E-mail: myahyaei56@yahoo.com

Definitions of the individual muscle testing grades are shown in Table 1. 18 upper limb muscles were selected for evaluation of individual strength (Table 2). Independently of each other, 3 examiners performed identical manual physical examinations of the maximal voluntary strength in each of the 18 muscles on both sides. The criteria proposed for optimization of muscle strength testing (Danniels and Worthingham, 2002) were met. Upper limb postures were standardized for each muscle, with the part to be tested positioned with stabilization of the part proximal or adjacent to the tested part. During testing it was aimed to maximize the length of the lever arm. Each test moment was defined from the functional anatomy of the individual muscle.

Table1. Modified medical research council scale

Grade	Grade Definition
5	Normal strength
-5	Barely detectable weakness
+4	Same as grade 4, but muscle holds the joint against moderate to maximal resistance
4	Muscle holds the joint against a combination of gravity and moderate resistance
-4	Same as grade 4, but muscle holds the joint only against minimal resistance
+3	Muscle moves the joint fully against gravity and is capable of transient resistance, but collapses abruptly
3	Muscle cannot hold the joint against resistance, but moves the joint fully against gravity
-3	Muscle moves the joint against gravity, but not through full mechanical range of motion
2	Muscle moves the joint when gravity is eliminated
1	A flicker of movement is seen or felt in the muscle
0	No movement

The aim was to optimize the position to facilitate the exertion of maximal strength by each muscle. A uniform placement of the hand for applying pressure was defined, and the pressure was applied directly and gradually, opposite to the line of pull of the muscle being tested. The intent to assess the peak strength as well as the ability of the individual to hold the force at a constant level during testing (endurance) demanded elements of isometric testing to be combined with an evaluation of eccentric dynamic resistance. Testing was performed up to three times. Assessment of strength in each muscle was based on a comparison of the intact contra lateral muscle or with bilaterally reduced strength in the particular muscle to other muscles of the individual in which the strength was assessed as intact (12). Data were analyzed using the weighted Kappa, as described Cohen (11) to determine the reliability of individual MMT grades and grades obtained for individual muscle groups. Kappa is a chance-corrected measure of agreement in which all disagreements are given equal weight. In contrast, the weighted Kappa takes into account the degree of disagreement among raters. Assessment of the reliability of individual muscle strength grades was made by using a modification of Cohen's Kappa, as described by Cicchetti and colleague (9). The total muscle score is determined by transforming individual muscle grades to a 10-point scale (5=10, 5- =9, 4+ =8, and so on), adding all converted scores, and using that sum for comparisons.

RESULTS

Table 2. Manual muscle test grades for individual muscle groups obtained using modified medical research council scale

Muscle groups	Weighted kappa
Serratus anterior	0.82
Scapula elevation	0.73
Scapula adduction	0.76
Rhomboids	0.77
Shoulder flexion	0.78
Shoulder extension	0.79
Shoulder abduction	0.78
Shoulder horizontal abduction	0.70
Shoulder horizontal adduction	0.76
Shoulder external rotation	0.80
Shoulder internal rotation	0.80
Elbow flexion	0.75
Thumb abduction	0.73
Elbow extension	0.81
Forearm supination	0.64
Forearm pronation	0.62
Wrist flexion	0.69
wrist extension	0.64

Inter-examiner reliability of MMT grades obtained with the modified MRC scale for individual muscle groups, as determined by the weighted Kappa, is shown in Table 2. Grades of proximal muscle groups were more reliable than were grades of muscle groups located distally. Inter-examiner reliability of MRC grades 0 to 5, as determined by the weighted Kappa, is shown in Table 3, along with the number of assignments within each grade. The reliability varied among individual grades, with grades in the gravity-eliminated position having the highest reliability values.

Table3. Inter-examiner reliability of individual muscle strength grades obtained using the modified medical research council scales

Grade	N	Weighted kappa
5	334	0.85
-5	54	0.56
+4	35	0.52
4	81	0.63
-4	61	0.64
+3	74	0.60
3	73	0.80
-3	17	0.69
2	179	0.86
1	246	0.88
0	34	0.88

N=number of assignments

DISCUSSION

The reliability and validity of various MMT techniques have been tested in patients with poliomyelitis (18, 22, 29) and muscular dystrophy (14, 15). Barr et al. (4) assessed the reliability of an mMRC scale that uses plus and minus sub-divisions (23). Perfect agreement was seen 35.9% of the time, consistency within one consecutive strength grade was found 66.5% of the time, and within 2 consecutive steps 84.7% of the time. The agreement for measures of proximal muscle strength (r = 0.80) was found to be more consistent than that for measures of distal muscle strength (r = 0.58) (4). Other studies addressed the reliability of a composite score, weighted by a factor that assessed muscle bulk rather than assigning grades to individual muscle groups or individual grades within a particular score (18, 22, 29). Some studies that addressed inter-rater reliability used a sum score of various muscles rather than analysing reliability for individual muscle groups (14, 21). Escolar et al. (14) determined a sum score of the mMRC scale and compared the reliability of MMT and quantitative muscle testing. MMT was not as reliable and required repeated training of evaluators to bring all groups to a correlation coefficient > 0.75 (14). Kleyweg et al. (21) registered nearly perfect inter-observer agreement of a sum score of various muscles tested with the MRC scale in patients with Guillain-Barré syndrome**.**

The MMT method described by Daniels & Worthingham (12) was shown to be reliable (26, 28). Brandsama et al. (7) tested the reliability of the 6-point MRC scale of intrinsic muscles of the hand. They suggested testing specific movements rather than selective muscles because it is difficult to isolate, and hence grade, most of the intrinsic muscles of the hand (7). They also introduced a mMRC scale (8), which includes the description of ROM as well as resistance into the 6-point original MRC scale. In their mMRC scale, grade 3 has to have normal ROM. In our modified scale, grade 3 has to have more than 50% of the feasible ROM and additional 3 grades (grade 2–3, 3–4 and 4–5) were included, which was assumed to represent better the clinical course of nerve regeneration.

Manual muscle testing, using the MRC scale, provides reliable grades for then assessment of strength of individual muscle groups within a sample of patients with a

Diagnosis of neuropathy or myopathy. Inter-examiner reliability ranged from 0.62 to 0.82. The weighted Kappa values for the proximal muscles were more consistent than those for the distal muscles. Weighted Kappa values for Inter-examiner reliability varied among individual muscle groups (0.52- 0.82). Inter-examiner reliability varies among individual muscle grades with those in he gravity-eliminated position (MRC 0 -2) grading most reliably (≥0.86). This finding differs from that of Frese and Colleagues (16) who found poor. Inter-examiner reliability in grades below Fair; similarly, Beasley (5) found poor differentiation in grades below Fair. These grades were the only categories that had no subdivisions (no plus or minus designations) and were strictly defined, but Frese and colleagues also stated that "compressing the scores by eliminating pluses and minuses did not appreciably change the Inter-examiner reliability coefficients (16).

The strength grading subdivisions in which gravity and resistance are factors in determining the MMT grade have come under much criticism (2, 27, 30) be cause they require judgments beyond assigning the original grade on the part of the examiner. The weighted Kappas for the MRC grades 4- to 5 ranged from 0.64 to 0.85. These grades are less reliable than those given in positions in which the factors of gravity and resistance have been eliminated but, we believe, are still acceptable for measurement in the clinical trial setting. The reliability coefficients in our MMT study ranged from 0.63 to 0.85 in the Good to Normal range (MRC 4-5). Studies relating to the sensitivity of these methods of measurement of muscle strength are needed to determine the most useful method for documentation, as both methods appear reliable in the assessment. The lowest reliability coefficient for an individual muscle grade in this study was 0.60, with a grade of 3+. The definition of this grading subdivision in our study required considerable judgment by the examiner because this grade indicates that the muscle "is capable of transient resistance, but collapses abruptly."

Controversy exists in the literature over the use of the MMT as a measurement tool in the documentation of muscle strength. Our study suggests that MMT grades obtained with the MRC scale are reliable. Some authors have suggested that MMT grades below Fair (MRC <3) are not reliable (16, 27) and that grades of Good (4) and Normal (5) are subjective." Our study suggests that we need to be most cautious of the grades Fair minus (3-), Fair plus (3+), Good minus (4-), Good (4), Good plus (4+) and Normal minus (5-) and that the most reliable grades are those made with the factor of gravity eliminated. The MMT is shown to yield reliable grades within individual muscle groups, but reliability varied proximal to distal within an extremity. The best agreement was shown when MMT grades for individual muscle groups were combined into a total muscle score. This finding suggests the most stable measure for documenting muscle strength in systemic diseases or with systemic interventions is a composite score.

REFERENCES

1. Amundsen L. Muscle Strength Testing: Instrumented and Non-Instrumented Systems. New York: Churchill Livingstone; 1990.
2. Andres PL, Skerry LM, Munsat TL. Measurement of strength in neuromuscular diseases. In: Munsat TI., ed. Quantification of Neurologic Deficit. Boston, Mass: Butterworth- Heinemann; 1989:87-100.
3. Barbano RL: Handbook of Manual Muscle Testing. Neurology 2000, 54(5):1211.
4. Barr AE, Diamond BE, Wade CK, Harashima T, Pecorella WA, Potts CC, et al. Reliability of testing measures in Duchenne or Becker muscular dystrophy. Arch Phys Med Rehabil 1991; 72: 315–319.
5. Beasley WC. Quantitative muscle testing: principles and applications to research and clinical services. Arch Phys Med Rehabil. 1961; 42: 398-435.
6. Bohannon RW: Manual muscle testing: does it meet the standards of an adequate screening test? Clin Rehabil 2005, 19(6):662-7.
7. Brandsama JW, Schreuders TA, Birke JA, Piefer A, Oostendorp R. Manual Muscle Strength testing: intraobserver and interobserver reliabilities for the intrinsic muscles of the hand. J Hand Ther 1995; 8: 185–190.
8. Brandsama JW, Schreuders TA. Sensible manual muscle strength testing to evaluate and monitor strength of the intrinsic muscles of the hand: a commentary. J Hand Ther 2001; 14: 273–278.
9. Cicchetti DV, Lee C, Fontana AF, Dowde A. A computer program for assessing specific category rater agreement for qualitative data. Educational and Psychological Measurement. 1978; 38:805-813.
10. Clarkson HM. Musculoskeletal Assessment: Joint Range of Motion and Manual Muscle Strength. Philadelphia: Lippincott Williams & Wilkins; 2000.
11. Cohen J. Weighted Kappa: normal scale agreement with provision for scaled disagreement on partial credit. Psycho1 Bull. 1968; 70:213-220.

12. Daniels L, Worthingham K: Muscle Testing – Techniques of Manual Examination 7th edition. Philadelphia, PA: W.B. Saunders Co; 2002.24.
13. Dvir Z. Isokinetics: Muscle Testing, Interpretation and Clinical Applications. New York; Churchill Livingstone; 1995.
14. Escolar DM, Henricson EK, Mayhew J, Florence J, Leshner R, Patel KM, Clemens PR. Clinical evaluator reliability for quantitative and manual muscle testing measures of strength in children. Muscle Nerve 2001; 24: 787–793.
15. Florence JM, Pandya S, King WM, Robison JD, Signore LC, Wentzell M, Province MA. Clinical trials in Duchenne dystrophy. Standardization and reliability of evaluation procedures. Phys Ther 1984; 64: 41–45.
16. Frese E, Brown M, Norton BJ. Clinical reliability of manual muscle testing: middle trapezius and gluteus medius muscles. Phys Ther. 1987;67:1072-1076.
17. Hough CL, Lieu BK, Caldwell ES. Manual muscle strength testing of critically ill patients: feasibility and interobserver agreement. Critical care 2011; 15:254-263.
18. Iddings DM, Smith LK, Spencer WA. Muscle testing. 2. Reliability in clinical use. Phys Ther Rev 1961; 41: 249–256.
19. Kendall F P. Manual muscle testing: there is no substitute. J Hand Ther 1991; 4: 159 61.
20. Kendall FP, McCreary EK, Provance PG: Muscles: Testing and Function Baltimore, MD: Williams & Wilkins; 1993.
21. Kleyweg RP, van der Meché FG, Schmitz PI. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré Syndrome. Muscle Nerve 1991; 14: 1103–1109
22. Lilienfeld AM, Jacobs M, Willis M. A study of the reproducibility of muscle testing and certain other aspects of muscle scoring. Phys Ther Rev 1954; 34: 279–289.
23. Medical Research Council (MRC). Aids to the investigation of peripheral nerve injuries. London: Her Majesty's Stationery Office; 1976.
24. Mendell JR, Florence JM. Manual muscle testing. Muscle Nerve. 1990; 13:S16-S20.
25. Neistadt M, Crepeau E. Willard and Spackman's Occupational Therapy. Philadelphia: Lippincott Williams & Wilkins; 1998.
26. Nitz JC, Burns YR, Jackson RV. A longitudinal physical profile assessment of skeletal muscle manifestation in myotonic dystrophy. Clin Rehabil 1999; 13: 64–73.
27. Rothstein JM. Measurement and clinical practice: theory and application. In: Rothstein JM, ed. Measurement in Physical Therapy. New York, NY: Churchill Livingstone Inc; 1985: 1-46.
28. Silver M, McElroy A, Morrow L, Heafner BK. Further standardization of manual muscle test for clinical study: applied in chronic renal disease. Phys Ther. 1970; 50:1456-1466.
29. Smith LK, Iddings DM, Spencer WA, Harrington PR. Muscle testing, part I: description of a numerical index for clinical research. Phys Ther Rev 1961; 41: 99–105.
30. Stuberg WA, Metcalf WK. Reliability of quantitative muscle testing in healthy children and in children with Duchenne muscular dystrophy using a hand-held dynamometer. Phys Ther. 1988; 68:977-982.
31. Wright W. Muscle training in the treatment of infantile paralysis. Boston Med Surg J 1912; 167: 567.