

המהימנות והתקפות של Manual Muscle Testing - סקירת ספרות שיטתית

יגאל קצפ

פיזיותרפיסט, B.P.T, המכון לפיזיותרפיה, קופת חולים מכבי שירותי בריאות, רחובות.
סטודנט לתואר שני בפיזיותרפיה, אוניברסיטת בן גוריון.

מנחה: ד"ר תמר יעקב, אוניברסיטת אריאל בשומרון

תקציר

רקע: Manual muscle testing (MMT) היא אחת הבדיקות המקובלות להערכה של כוח השריר בקליניקה. הבדיקה היא פשוטה, זולה ואינה דורשת מכשור. למרות השימוש הנרחב בבדיקה זו, אין מידע מספק דיו לגבי התקפות והמהימנות של הבדיקה.

מטרה: לסקור את המחקרים שעסקו בהערכת המהימנות והתקפות של ה - MMT, אשר פורסמו בעשור האחרון.

שיטות: נערך חיפוש באתרים אינטרנטיים: PubMed, NARIC, CIRRIE, EBSCO, Google Scholar, EM consult, Cochrane Library, Web of Science, Medscape CME. נכללו מחקרים שעוסקים במהימנות או בתקפות של ה - MMT שהשתמשה בסקאלת ה - Medical Research Council (MRC) Scale, שפורסם בשפה האנגלית ושהיה ניתן להשיגם בגרסתם המלאה. כל המחקרים שהוכללו בעבודה זו הוערכו על-ידי "Critical appraisal tool for validity and reliability studies of objective clinical tools", ומחקרי התקפות הוערכו גם על-ידי "Levels of Evidence for Primary Research Question".

תוצאות: החיפוש הראשוני העלה יותר מ-300 מחקרים. אבל לאחר קריאת כותרתם ותקציריהם נבחרו מבניהם תשעה מאמרים בלבד שעסקו במהימנות ו\או בתקפות של MMT ושנמצאו רלוונטיים. שמונה מבין המאמרים עסקו במהימנות כלשהי: אחד עסק בתוקף מבני ועוד שניים נועדו, על פי הצהרתם, לבדוק מהימנות. האיכות של כל המחקרים הייתה נמוכה/בינונית מלבד מחקר אחד שאיכותו הייתה גבוהה.

מסקנות: על פי המחקרים שנסקרו נראה שאפשר להשתמש ב - MMT לשם הערכת רמות שונות של כוח באוכלוסיות מסוימות, אך יש צורך במחקרים נוספים בעלי איכות מתודולוגית גבוהה כדי לבסס ביסוס-יתר את המהימנות והתקפות של הכלי בהבחנה בין רמות הכוח השונות.

מילות מפתח: Manual muscle testing; Reliability; Validity; Systematic review.

הקדמה

Manual Muscle Testing (MMT) היא אחת השיטות להערכת הכוח ותפקוד השריר כמרכיב הכרחי של תנועה. זאת בדיקה פיזיקלית הנפוצה ביותר בהערכה קלינית של כוח שרירים.¹⁻⁴ השיטה פותחה בארצות-הברית בתחילת המאה ה-20 במסגרת חקר התפקוד השרירי של חולי Poliomyelitis, והוסיפה לשמש כלי קליני חשוב להערכת כוח שריר כגורם לליקויים תנועתיים.⁵

יתרונות הבדיקה נובעים מהיותה רב-תכליתית בקליניקה, זולה למדי, פשוטה לביצוע ואינה דורשת מכשור.³ את ההערכה ניתן לבצע באמצעות מבחן השבירה (Break Test), שהוא הנפוץ יותר, או באמצעות מבחן ההתנגדות האקטיבית (Active Resistance Test). ה - Break Test נעשה על-ידי הפעלת התנגדות ידנית לגפה, או לחלק גוף אחר, לאחר שהוא השלים את מלוא טווח התנועה או הונח בסוף טווח התנועה. ה - Active Resistance Test הוא חלופה למבחן השבירה ובו מיישמים התנגדות ידנית על שריר או שרירים המתכווצים אקטיבית (בכיוון ההפוך לתנועה ובניסיון לעצור את התנועה). במהלך התנועה, הבדוק מעלה את רמת ההתנגדות הידנית בהדרגתיות, עד שזו מגיעה לרמה הגבוהה ביותר שהבדק יכול להתנגד, ואז התנועה נפסקת. שיטת מבחן זו דורשת מיומנות וניסיון רב. על כן לרוב, יכולתה לזהות חולשה אמתית מוטלת בספק והיא אינה מומלצת לשימוש.¹ שיטת הדירוג הנפוצה ביותר כיום ב - MMT היא שיטת שש הרמות Medical Research Council (MRC) Scale, שהיא זהה לשיטה שהציע

בעבר Lovett & Martin להערכת חולי Poliomyelitis.⁶⁻⁸ שיטה נוספת לניקוד היא ציון 75%, 50% ו-25% מן הכוח הנורמלי אשר מייצגת מספרית את הקטגוריות "טוב", "בינוני" ו- "חלש". כמו כן, משתמשים בשיטת האחידה האויזומטרית ובשיטת השבירה האקצנטרית לשם אבחנה אם כוח השריר הוא 4 או 5.⁹

ה - MMT משמש למטרות שונות בקרב קלינאים שעוסקים בתחום התנועה.¹⁰ סקירת הספרות האחרונה שנעשתה בתחום פורסמה על-ידי Cuthbert & Goodheart.¹¹ הסקירה כללה

שיטות

הקריטריונים לבחירת מאמרים - טרם דלייתם של המאמרים, נבדקו כותרתם ותקצירם כדי לבחון באיזו מידה הם עונים על הקריטריונים להכללתם של מאמרים בסקירת ספרות. מאמרים שנראו מתאימים נקראו והרלוונטיות והאיכות שלהם הוערכה.

קריטריונים להכללה: 1. מחקרים שבדקו מהימנות או תוקף של MMT תוך כדי דירוג עוצמת כוח השרירים. 2. המאמרים שפורסמו בין השנים 2003-2013. 3. מאמרים שפורסמו באנגלית ושניתן היה להשיגם בטקסט מלא. 4. מחקרים קליניים, שנעשו בקרב בני אדם.

קריטריונים לאי-הכללה: מאמרים המתיחסים ל - MMT כחלק משיטת הבדיקה של קינזיולוגיה יישומית ואינם מתייחסים לתוצאות הבדיקה בצורה של דירוג עוצמת כוח השריר (למשל, מגדירים רק אם השריר חלש או חזק ולא מדרגים בסולם של 0-5).

אופן חיפוש המאמרים - נערך חיפוש ספרות באתרים האלה: PubMed, NARIC, CIRRIE, EBSCO, Google Scholar, EM consult, Cochrane Library, Web of Science, Medscape CME. זוהו כל המאמרים הרלוונטיים בנושא שפורסמו בין ינואר 2003 ליוני 2013.

מילות החיפוש היו: assess, strength, weak, force, validity, reliability, manual muscle test. מילות החיפוש הוכנסו לכל האתרים שצוינו לעיל כמילות מפתח או כטקסט חופשי כמו כן, נעשה חיפוש ברשימות המקורות של המאמרים שנדלו מן האתרים לשם מציאת מאמרים רלוונטיים נוספים.

הערכת איכותם הכללית של המחקרים - לשם הערכה כללית של איכות המחקר דורגה רמתו הכללית על-ידי רשימת שאלות מחקר ראשוניות (Levels of Evidence for Primary Research "Question").¹⁷ הכלי הזה פותח לשם הגדרה והשוואה של רמת הראיות של מחקרים לסוגיהם לשם מתן המלצה קלינית. בדירוג ישנן חמש רמות: I היא הרמה הגבוהה ביותר של מאמר ו - V היא הרמה הנמוכה ביותר.

רמה I: בדיקת קריטריונים דיאגנוסטיים שכבר פותחו בעבר על בסיס כמה מטופלים עוקבים (עם התייחסות עקבית ל - Gold standard); **רמה II:** פיתוח של קריטריונים דיאגנוסטיים על בסיס של מטופלים עוקבים (עם התייחסות עקבית ל - Gold standard); **רמה III:** חקירה של מטופלים לא עוקבים (ללא התייחסות עקבית ל - Gold standard); **רמה IV:** מחקר

מאמרים שבדקו את המהימנות (בין בודקים שונים ובין בודק לעצמו) והתקפות (תוקף מבנה ותוכן, תוקף מתכנס וזיהוי, תוקף מקביל ותוקף ניבוי) של MMT בהקשר של קינזיולוגיה יישומית (שיטה שבה משתמשים בעיקר כירופרקטים). שיטה זו משתמשת ב - MMT כבדיקה יזנית של הסטטוס שבו נמצאת המערכת העצבית והיא שונה מזו שבה משתמשים בכלי זה בבדיקה האורתופדית הקלאסית. לסקירת הספרות הזאת ישנן מגבלות רבות, כפי שצוין במאמרו של Haas וחבריו.¹² נסקרו בה מאמרים שבהם ה - Cohen's kappa coefficient היה 0.5 ומעלה בלבד, ולא נכללו בה מאמרים שהראו רמות נמוכות יותר של מהימנות ותקפות. שיטת חיפוש המאמרים הייתה חלקית, לא נבדקה התקפות החיצונית והפנימית של ה - MMT ונכללו מחקרי ניסוי קליני אקראיים אשר לא תאמו את מטרות סקירת הספרות.

אחת הדרכים שנוקטים קלינאים כדי לקבוע אם מבחן קליני הוא עקבי, אם אפשר לחזור עליו ואם אפשר להסיק ממנו מסקנה משמעותית, היא לנתח את המהימנות והתקפות של המבחן.¹³⁻¹⁶ מהימנות של מבחן נבחנת על-ידי בדיקת העקביות שלו במבחנים חוזרים, כאשר המבחן נערך בתנאים דומים. ניתן לערוך חישובים של מקדמי מהימנות שונים, לדוגמה מקדם Cohen's kappa, להערכת מהימנות של סולם קטגוריאלי ובמבחן Intraclass Correlation Coefficient (ICC) כאשר מדובר בבדיקת מהימנות של סולם מטרי.^{13,15} בהערכות של מחקרים וטכניקות בדיקה המשתמשות ב - MMT, קלינאים צריכים להכיר סוגי תוקף כמו תוקף המבנה (באיזו מידה ה - MMT כוללת את כל ההיבטים של משתנה כוח השריר), תוקף התוכן (באיזו מידה ה - MMT מצליחה לשקף ולייצג את כוח השרירים האמיתי של הנבדק), תוקף מבחין (באיזו מידה ה - MMT יכולה להבדיל בין רמות כוח שונות) ותוקף הניבוי (באיזו מידה ה - MMT יכולה לנבא משתנה אחר, כגון תמותה בטיפול נמרץ).

שימוש במבחן שאינו מהימן ואינו רגיש יכול להוביל קלינאים לקבלת החלטות שגויות לגבי כוח השריר הנבדק. בעקבות השימוש הנרחב ב - MMT בתחום הבדיקה האורתופדית והיעדר סקירת ספרות מקיפה הבודקת את מהימנות ותקפות הבדיקה, הוחלט לערוך את סקירת הספרות הנוכחית. מטרת סקירת הספרות היא לסקור מאמרים אשר בדקו את מהימנות ותקפות ה - MMT כחלק מן הבדיקה האורתופדית, ולנסות לקבוע אם ניתן להשתמש בבדיקה זו באוכלוסיות שונות של נבדקים.

297 נחקרים בסך הכול.

מחקרי תקפות

המחקר של Brookham וחבריו²⁰ נועד לבדוק את תוקף המבנה של 29 מבחני כוח שונים (אשר חולקו לשבע קבוצות תפקוד) על-ידי MMT בבדיקת ארבעת שרירי ה- Rotator cuff בהשוואה לבדיקת הפעילות השרירית שלהם באמצעות Electromyography תוך כדי הבדיקה. נוסף על שרירי ה- Rotator cuff, נבדקה במקביל גם הפעילות השרירית של עשרה שרירים נוספים בחגורת הכתף, והחישוב שנעשה הוא בדיקת הפעילות המקסימלית של השריר הנבדק ביחס לפעילות השרירית המינימלית של 13 השרירים הנותרים. תוצאות המחקר הראו ששריר ה-Infraspinatus הופעל במידה הרבה ביותר בקבוצת מבחני הרוטציה החיצונית, שריר ה-Supraspinatus הופעל במידה הרבה ביותר בקבוצת מבחני ההרחקה, מבחני הכוח הרדיאלי, מבחני הכוח הדורסאלי והרוטציה החיצונית (ללא הבדל מובהק סטטיסטית ביניהם). שריר ה- Teres Minor הופעל במידה הרבה יותר בקבוצת מבחני הרוטציה החיצונית והפנימית, הכוח הפלמרי, הכוח הדורסלי והאולנרי, ללא הבדל משמעותי סטטיסטית ביניהם. שריר ה- Subscapularis הופעל במידה הרבה ביותר בקבוצת מבחני הכוח האולנרי והרוטציה הפנימית, ללא הבדל מובהק סטטיסטית ביניהם.

מחקרי מהימנות

במחקר של Hara וחבריו²¹ נבדקה המהימנות בין בודקים שונים של MMT לגבי פשיטת בוחן במנחי זוויות שונות במפרקי הברך. נוסף על כך, נבדקה הקורלציה בין אורך שריר וזווית המפרק ובין כוח השריר. ההנחה הראשונית של החוקרים הייתה שכוח השרירים של בהונות הנבדקים תקין (כוח 5 בסולם 0-5). נעשתה השוואה בין שני בודקים שונים באמצעות מבחן Kappa. כוח השרירים נמצא תקין או מתחת לתקין (מתוך ה-80 שנבדקו). תוצאות המחקר, לפי אחוז הבהונות שנמצאו תקינים (מספר הבהונות), נעו בין K=1 לשריר MTPJ ועד K=0.28 ל IPJ with MTPJ in flexion. לא נמצא קשר מובהק סטטיסטית בין אורך השריר וזווית המפרק ובין כוח השריר. במחקר של Paternostro-Sluga וחבריה²² ניסו החוקרים לבדוק את התקפות והמהימנות (בין בודק לעצמו ובין בודקים שונים) של ה- MMT הקלאסי

מסוג Case-control או התייחסות דלה לבדיקה המקובלת; רמה V: דעה של מומחה. כל מחקר מדורג לפי שאלת המחקר, התוכן ותכנון המחקר, כאשר המחקרים הרלוונטיים שנבדקו באמצעות כלי הערכה זה הם מחקרי תקפות.

הערכת איכותם של מחקרי מהימנות ותקפות - נעשה שימוש ברשימה של 13 השאלות של Brink & Louw¹⁸ שנקראת "כלי הערכה ביקורתי למחקרי מהימנות ותקפות של כלים קליניים אובייקטיביים". כלי הערכה זה מורכב משני כלים, האחד מעריך איכות מחקרים הבודקים פרוצדורות של שיטות מבחני דיוק אבחוני ומבחני תקפות של כלי אובייקטיבי או קליני ומכונה: Quality Assessment of Diagnostic Accuracy Studies (QUADAS), השני מעריך את איכות שיטות הבדיקה במחקרים הבודקים את המהימנות של כלי אובייקטיבי או קליני ומכונה: Quality Appraisal of Diagnostic Reliability Studies (QAREL).

מתוך 13 השאלות שבכלי הערכה, חמש שאלות קשורות למחקרי מהימנות ותקפות, ארבע שאלות קשורות למחקרי תקפות בלבד וארבע שאלות קשורות למחקרי מהימנות בלבד. על השאלות ניתן לענות ב- "כן" או "לא" או "אינו ישים". מחקרים שנמנו בהם יותר מ- 60% תשובות "כן" בהערכת קריטריונים שיטתיים הרלוונטיים לסוג המחקר, נחשבו בעלי איכות גבוהה, ומחקרים שנמנו בהם פחות מ-40% תשובות "כן" נחשבו לבעלי איכות נמוכה.¹⁹

תוצאות

החיפוש באתרים המצוינים העלה יותר מ-300 מאמרים שפורסמו בשנים 2003-2013 ושהכילו את מילות החיפוש. לאחר קריאת הכותרות והתקצירים של המאמרים נותרו 22 מאמרים שסקרו את התקפות והמהימנות של MMT. שלושה עשר מחקרים הוצאו מן המחקר בגלל חוסר רלוונטיות (בדיקת כוח שריר על פי שיטה קינזיולוגית), ולבסוף נכללו בסקירה תשעה מחקרים שסקרו את המהימנות והתקפות של MMT באוכלוסיות שונות (והשתמשו בסקלה 0-5). מתוך המחקרים הללו מחקר אחד בדק תקפות של MMT בנפרד,²⁰ ושמונה בדקו מהימנות.²¹⁻²⁸ מן המחקרים שנכללו, שני מחקרים נועדו לבדוק מהימנות ותקפות יחד. באחד מהם²¹ אין כלל התייחסות לתקפות ובשני²² נראה כי בדיקת התקפות נעשתה בצורה לא רלוונטית. לפיכך, מחקרים אלו הוערכו על פי בדיקות המהימנות שלהם בלבד. תשעת המחקרים דגמו

שעה זה מזה, כאשר המטופל שוכב על הגב או בישיבה (תלוי במצבו הקליני). ה-ICC של תוצאת הסיכום יצא 0.83 (95% CI, 0.67-0.93). בקרב מטופלים שיכלו לבצע את הבדיקה, ובמיוחד בקרב מטופלים שנבדקו במחלקות בית החולים (לאחר שחרורם מן היחידה לטיפול נמרץ), לעומת ביחידה לטיפול נמרץ עצמה, נמצאה רמת הסכמה טובה בין בודקים שונים בקביעת "חולשה נרכשת עקב האשפוז בטיפול נמרץ" (בקרב כל המטופלים, רמת ההסכמה בין הבודקים עמדה על 93% עם $kappa = 0.76$ (ו-0.44-1.0 CI 95%). בחינה של רמת המהימנות בין הבודקים בהערכת שריר יחיד מראה רמה נמוכה של הסכמה בין הבודקים, במיוחד בשרירים הפרוקסימאליים, אשר נעה בין 40%-83%, עם רמת Weighted kappa -0.11-0.64 ורמת ICC 0.15-0.75.

שיטה נוספת לסיכום מספר ציוני רמת כוח השריר פותחה על-ידי האיגוד האמריקאי לפגיעות חוט השדרה (ASIA). השיטה משתמשת בציוני כוח השריר כחלק מבדיקת נפגעי חוט שדרה, אשר כוללת גם בדיקה של תחושת מגע קל ובדיקה של תחושת דקירה כדי לזהות חסרים נוירולוגיים ולקבוע את חומרת הפגיעה. מחקרים אחדים בדקו את המהימנות בין הבודקים לגבי הסטנדרטים של ASIA, אך מעטים מהם בלבד הבדילו בין מיומנות הבודק ובין ההשפעה של מיומנות זו על תוצאות הבדיקה ועל קביעת חומרת הפגיעה. לכן, Savic וחבריו²⁵ רצו לבדוק את מיומנותם של בודקים כדי לקבוע לאיזו רמת הסכמה יש לצפות על פי תוצאות הבדיקות של שני בודקים מיומנים, וכיצד ההבדלים בתוצאות משפיעים על קביעת חומרת הפגיעה בפגיעות חוט שדרה. כוח השרירים הוערך בדרגות של 0-5 ובסופו של דבר רק 43 מטופלים עברו בדיקה מוטורית ו-30 מטופלים עברו בדיקה סנסורית על-ידי שני הבודקים (וזאת בגלל אי זמינותם של הבודקים) בהפרש של חמישה ימים זה מזה. כדי לגרוע את ההשפעה של דרמטומים ומיטומים עם תפקוד נורמאלי על המהימנות, חזרו על הניתוח הסטטיסטי לאחר הוצאת השרירים שהם מעל לגובה הפגיעה, שהוגדרו כבעלי כוח 5 על-ידי שני הבודקים (הוצאו מן המחקר 21 מטופלים עם פגיעה מוטורית מלאה של עמוד השדרה הגבי). בסופו של דבר, הניתוח הסטטיסטי המוטורי השניוני נעשה לגבי מטופלים שניתן היה לגזור את רמתם המוטורית מן הבדיקה המוטורית. התוצאות של המחקר הזה הראו שיש רמת הסכמה גבוהה בבדיקת MMT של 10 השרירים שנבדקו ב-22 נבדקים (רמת מתאם $kappa = 0.649-0.993$ כאשר $P < 0.01$). ואולם, רמת ההתאמה

ושל ה- mMMT, עם רמות של טווח תנועה אקטיבי. בשתי השיטות שבהן העריכו את רמת כוח השריר באמצעות סקאלה של 0-5 דרגות, נעשתה הערכה לשתי הידיים, שכללה את פושטי שורש כף היד ופשטי האצבעות החיצוניים. לשם בדיקת התוקף, נעשתה הערכת כוח (Grip (Flexor digitorum superficialis, Flexor digitorum profundus, intrinsic hand muscles) באמצעות דינמומטר. המהימנות בין הבודקים נבחנה על-ידי חמישה בודקים שהעריכו 31 מטופלים. כדי לבדוק מהימנות בין בודק לעצמו בדק בודק אחד פעמיים כל אחד מ-22 מטופלים, בהפרש של שבוע בין הבדיקות. רמות ההסכמה בין בודקים שונים ב- MMT הקלאסית (ממוצע ה- Weighted pair-wise Kappa) הן: פושטי שורש כף היד - 0.78, פושטי אצבעות -0.77 וכוח Grip - 0.78, כאשר טעות התקן מוערכת ב- 0.03-0.18 בכל התוצאות. מבחינת מהימנות בין הבודק לעצמו ב- MMT הקלאסי, נמצאה הסכמה כמעט מלאה: פושטי שורש כף היד - 0.82, פושטי האצבעות - 0.86 וכוח Grip - 0.84, עם טעות תקן שלא עלתה על 0.12. בדיקת ה- MMT המותאם מראה תוצאות דומות מאוד.

בדיקות MMT נערכות באוכלוסיות שונות, אך קיים מידע מועט בלבד על המהימנות של MMT כוללת במטופלים המאושפזים ביחידה לטיפול נמרץ. במחקר של Fan וחבריו²³ ניסו להעריך את המהימנות בין בודקי MMT כללית בקרב מטופלים ששהו ביחידה לטיפול נמרץ ומטופלים מדומים שנאמר להם לךמות כוח שריר מסוים. חולשה קלינית של ממש הוגדרה כאשר הציון הכללי, הציון של הגפה העליונה או של הגפה התחתונה היה פחות מ- 80% מן הציון המרבי (למשל, ציון ממוצע פחות מ-4 בכל השרירים). Hough וחבריה²⁴ בדקו גם מאושפזים ביחידה לטיפול נמרץ. החוקרים השתמשו במחקר מסוג Cohort כדי לבחון את הישימות ואת רמת ההסכמה בין בודקים שונים לגבי סכום רמת הכוח (שנבדק באמצעות MMT) של 12 שרירים בחולים במצב קריטי ובפצועים ביחידה לטיפול נמרץ. סכום זה קובע דיאגנוזה המכונה "חולשה נרכשת של טיפול נמרץ" שזו חולשה חמורה אשר מתרחשת אצל מטופלים רבים המאושפזים ביחידה לטיפול נמרץ לאורך זמן. כל שריר הוערך על פי סקאלה שבין 0-5. סכום הניקוד של 12 השרירים שהוא נמוך מ- 48 מצביע על כך שהכוח הממוצע מוגבל לתנועה כנגד כוח המשיכה, וכי ההתנגדות היא חלקית. החולים הוערכו לאחר שהות של 48 שעות ביחידה לטיפול נמרץ ונערך אחריהם מעקב. לאחר בדיקות הסקירה נערכו בדיקות השרירים בהפרשים של חצי

בילדי CP הסובלים מחולשת גפיים בצד אחד של הגוף. אחת השיטות שנועדה לבדיקת כוח השריר הייתה MMT. השרירים שנבדקו היו שרירי הכתף (מכופפים, מרחיקים ומקרבים), המרפק (מכופפים ופושטים) ושורש כף היד (סופינציה ופרונציה, מכופפים ופושטים). כוח השרירים דורג בסקאלה של 0-5. לכל נבדק ניתן ציון כללי (0-45), ציון לכתף (0-15), ציון למרפק (0-10) וציון לשורש כף היד (0-20). הפרש הזמנים בין שתי הבדיקות של אותו הבודק היה שבועיים, והפרש הזמנים בין שני הבודקים היה שעה אחת. מתוצאות המחקר נראה שבבדיקת שריר יחיד המהימנות בין בודקים שונים (ב-30 מטופלים) הייתה בינונית עד גבוהה מאוד (ICC 0.53-0.93), כאשר ה-ICC גבוה יותר בקבוצות השרירים היותר דיסטליים (ICC 0.60 בציון המשוקלל של שרירי הכתף, 0.62 בציון המשוקלל של שרירי המרפק לעומת 0.90 בציון המשוקלל של שרירי שורש כף היד). בבדיקת המהימנות בין בודק לעצמו (ב-23 מטופלים בלבד) נמצאה מהימנות גבוהה מאוד (ICC 0.80-0.98), בבדיקת שריר יחיד, חוץ מאשר בשרירי מרחיקי הכתף שבהם נמצאה מהימנות בינונית, עם ICC 0.69. במקרים אלה, אפשר שהירידה בכוח השריר תלויה לבעיות מוסקולריות ונוירו-מוסקולריות, עם הופעה בתבניות תנועה פתולוגיות שונות. יש הסוברים שהערכה זו היא בעלת ערך מוגבל בבדיקות של הפרעות בגפיים העליונות, חוץ מאשר במצבים שבהם נמצאה רמה ניכרת של חולשת שריר. מטופלים הסובלים מבעיות בגפיים העליונות, על רקע שינויים בעמוד השדרה, נוטים להתלונן לעיתים קרובות על כאב, על חולשה ו/או על נימול/עקצוצים. לכן, במחקר של Jepsen וחבריו²⁸ ניסו לבדוק את המהימנות בין הבודקים לגבי בדיקת הכוח ב-14 שרירים של הגפיים העליונות במטופלים אלו, וכן לבדוק אם ישנו קשר בין הכוח ובין תלונות המטופל. בסך הכול השתתפו במחקר 41 מטופלים שהגיעו לקליניקה בעקבות תלונות שונות. שני בודקים ראינו את המטופלים כדי לקבוע אם הם סימפטומטיים או לא, ולאחר מכן שני בודקים אחרים (איש מהם לא ידע את הנאמר בראיונות ולא ידע את התוצאות של עמיתו) ערכו את ה-MMT ל-14 השרירים של הגפיים העליונות של שתי הידיים במקביל (בסקאלה של 0-5). מתוצאות המחקר נמצא שחציון רמת ההסכמה בין הבודקים בדבר היותו של השריר תקין או בעל כוח ירוד (נמוך מ-5) עומד על 81% (CI 72%-87%). חציון רמת ה- Kappa היה 0.54 (CI 0.25-0.72) בכל הנבדקים ו-0.57 (CI 0.32-0.82) כאשר נעשה חישוב של המטופלים בעלי

הפכה בלתי מובהקת סטטיסטית לגבי מקצת השרירים כאשר ערכו את הניתוח הסטטיסטי השניוני (שבו נגרעו התוצאות של השרירים שצוינו כבעלי כוח 5 מעל לגובה הפגיעה והשרירים שמתחת לגובה הפגיעה, בפגיעת חוט שדרה לא מלאה, שהוצגו ככוח 0 ושעליהם הייתה הסכמה בין שני הבודקים), זאת בגלל מספר מועט של נבדקים בשרירים אלו.

פתולוגיה נוספת שקשורה לבעיה נוירולוגית היא מחלת Charcot-Marie-Tooth (CMT) אשר ידועה גם בשם נוירופתיה מוטורית וסנסורית גנטית. הסימפטומים והסימנים הקליניים הם חולשה של השרירים המרוחקים בשרירי הגפה התחתונה. מחקרים שונים מצאו כי ישנם שרירים שונים שנחלשים, והממצאים הסותרים לגבי סדר ההיחלשות של השרירים השונים יכולים להיות מוסברים על-ידי השונות הגנטית והבדלים בחומרת המחלה. Vinci וחבריו²⁶ ערכו מחקר שמטרתו הייתה לבדוק כוח של כמה שרירים בגפה התחתונה במטופלים דומים מבחינת הליקוי הגנטי והחומרה הקלינית, בשלבי המחלה המוקדמים. במחקר נבדקו 45 חולי CMT 1A שהיו בשלבי מחלה 1 ו-2 בלבד, ומבניהם כל המטופלים נבדקו - 83 גפיים תחתונות: 26 גפיים ברמה 1 ו-57 גפיים ברמה 2 (הוצאו גפיים שהתאימו לרמה 3 או גפיים שעברו ניתוח). לפני ביצוע בדיקת השרירים הדריכו את המטופלים לכווץ בצורה נכונה את השריר הרצוי. בבדיקת הכוח נעשתה חמש פעמים ונרשמה התוצאה הטובה ביותר (להערכת הכוח השתמשו בסקאלה שבין 0-5). כל אחד משני הבודקים לא ידע על תוצאות הבדיקות של עמיתו. מתוך 1183 שרירים שנבדקו, שרירי ה-FHB וה-lumbricals נמצאו כשרירים החלשים ביותר (דורגו ברמות 0-1 ב-95.18% וב-100% מן הגפיים בהתאמה). ב-95% מן הבדיקות הייתה הסכמה מלאה כמעט בין הבודקים, כאשר $\text{kappa value} > 0.8$, וב-5% הנותרים נמצאה הסכמה ניכרת עם $\text{kappa} < 0.61$ ו- $\text{value} < 0.8$.

דוגמה נוספת היא מחלת Cerebral palsy (CP) בה קיימות אפשרויות רבות להערכת כוח השריר, אך קיימים מעט מאוד מחקרים המספקים תימוכין למהימנות ולתקפות של בדיקות אלו, במיוחד בילדים. במחקר של Klingels וחבריו²⁷ ניסו להעריך את המהימנות בין בודקים שונים ובין בודק לעצמו עבור בדיקות לליקויים סנסוריים (פרופיוספציה, זיהוי חפצים וכו') ומוטוריים (בדיקות כוח אחיזה, תנועות פסיביות וכו')

נבדקים שסבלו מ-Radial palsy, בדיקת כוח Grip נראית לכאורה בלתי מתאימה לבדיקת התקפות בקרב מטופלים אלו שכן היא מפעילה בעיקר שרירים שאינם קשורים לפגיעה הרלוונטית.

אשר למהימנות בין הבודקים, Hara וחבריו²¹ מצאו שהמהימנות הטובה ביותר קיימת בבדיקת כוח ה-Extensor hallucis longus דרך מפרק ה-MTP. לדעתם, יש לכנות בדיקה זו ברשומה הרפואית כ-"Hallux extension". עם זאת, המחקר לוקה בחסר: הבדיקות נעשו על נבדקים בריאים ולא ידוע מה תהיה רמת המהימנות בבדיקה בקרב אנשים עם פתולוגיות המשפיעות של כוח השריר הזה. קיים פירוט על בדיקת מהימנות לגבי כוח 5 בלבד, ואין נתונים על המהימנות לגבי רמות כוח אחרות; לא ברורה הסיבה לבחירת גודל המדגם.

כאשר רוצים להעריך כוח שריר, קיימות כמה שיטות כמותיות, מדויקות יותר מ-MMT, כמו שימוש בדינמומטר. עם זאת, השימוש בדינמומטר אינו מתאים במקרים שבהם השריר החלש אינו מסוגל להניע את המפרק נגד התנגדות, כפי שקורה במקרים רבים של פגיעות בעצבים פריפריים. במחקר של Paternostro-Sluga וחבריה²² נמצאה הסכמה רבה בכל השרירים באשר לשתי המהימנויות של MMT - ב-MRC לעומת mMRC. לדעת החוקרים, שתי שיטות המדידה הן בעלות מהימנות ניכרת (וכן תקפות) והן מומלצות לשימוש בקליניקה להערכת כוח שרירי האמה בפגיעות העצב הרדיאלי. הבעיות במחקר זה הן, בין השאר, פירוט לא מספק על הנחקרים, הכללת מטופלים עם פגיעה פריפריית של העצב הרדיאלי, פגיעת הקרנה של C7 או פציעות של Brachial plexus (שיכולות לפגוע גם בעצבים האולנרי והמדיאני). כל הפגיעות האלה גורמות לחולשה של שרירי האמה בצד הרדיאלי, אך אפשר שלבדיקת כל פתולוגיה יש רמת מהימנות שונה ולא ניתן להכליל את כולם באותה הקבוצה. בתוצאות המחקר צוין שלא נמצא הבדל סטטיסטי ($P > 0.05$) בין שני סולמות הבדיקה מבחינת המהימנות, אך לא צוינה רמת המובהקות הסטטיסטית של ההסכמה בכל סקאלה בנפרד. כמו כן, אחת המגבלות הגדולות של המחקר היא חוסר ההתייחסות לסמיות הבודקים.

המחקר של Jepsen וחבריו²⁸ בדק גם הוא את מהימנות ה-MMT בגפיים העליונות בלבד. ברוב השרירים שנבדקו (ב-11 מבין 14 שרירים) רמת המהימנות בין הבודקים נמצאה בינונית עד טובה. במחקר אין די פרטים על הנחקרים

סימפטומים חד-צדדיים בלבד. כאשר נעשתה חלוקה לגפיים סימפטומטיות, לעומת לא-סימפטומטיות, נראה שחציון רמת ה-Kappa יורד עוד יותר (0.32 בגפיים הסימפטומטיות). איכות המחקרים מוצגת בטבלה 1.

דיון ומסקנות

סקירת ספרות זו בוחנת את המחקרים שעסקו במהימנות ובתקפות של ה-MMT עם הערכה בסקאלה של 0-5 ואשר פורסמו בעשור האחרון. מתוך תשעת המאמרים שנכללו בסקירה נמצאו רק ארבעה בעלי איכות בינונית או גבוהה^{20,25-23} מתוכם רק אחד היה באיכות גבוהה.²⁴ כמו כן, נמצאו שלושה מחקרים בלבד שבהם אחת המטרות הייתה להעריך את תקפות ה-MMT.²²⁻²⁰

בבדיקה לעומק של המחקרים שנכללו בסקירת ספרות זו בולטות כמה מגבלות משותפות: מספר קטן של נבדקים, מחסור בנתונים על משתתפי המחקר, תיאור לא מדויק בפרוטוקול הבדיקה, ולעיתים היעדר ציון של רמות המובהקות של התוצאות.

באשר למחקרי התקפות, Brookham וחבריו²⁰ מצאו שקשה לבדוד לגמרי את שרירי ה-Rotator cuff בעת שימוש ב-MMT, עם זאת, מבחני השרירים הקיימים מבדדים יחסית את השריר העיקרי לפעולה ויכולים להצביע על שינויים גדולים בפעילות השריר. כמו כן, אף על פי שבמבחני ה-MMT, המבוססים על הפעולה העיקרית של השריר, הצליחו למדי לבדוד את שרירי ה-Rotator cuff, היה ניתן לראות בידוד דומה למדי גם במבחנים אחרים. לדעת החוקרים, יש צורך לערוך מחקר על מדגם גדול יותר לשם הכללת הממצאים, וכן לא ידוע אם התגובה של מטופלים עם כאבי כתפיים תהיה דומה בבדיקה. אפשר שמטופל שחש כאב בכתף ינסה לגייס שרירים שונים כדי להתנגד, ולא את השרירים העיקריים לפעולה שדווקא הם מפיקים אצלו כאב. לכן, יכול להיות הבדל בתוצאות בין אנשים בריאים לאנשים הסובלים מכאבי כתפיים. שני מחקרים נוספים רצו לבדוק את נושא התקפות של בדיקת ה-MMT: Hara וחבריו²¹ טענו שאחת המטרות שלהם היא להעריך, באמצעות MMT, את התקפות של בדיקת כוח פשיטת הבהון, ואולם במחקר נבדקה למעשה המהימנות לעומת מחקר זה, Paternostro-Sluga וחבריה²² ניסו להעריך את תקפות ה-MMT על-ידי השוואה בין בדיקה זו לבין כוח Grip הנמדד על-ידי דינמומטר. כיוון שבמחקר הזה השתתפו

פירוט של רמות ההסכמה אך ללא ציון רמת המובהקות הסטטיסטית של התוצאות. לפיכך, קשה לקבוע את מהימנות ה-MMT במטופלים המאושפזים ביחידה לטיפול נמרץ. בדיקת מהימנות ה-MMT נבדקה גם בקרב נפגעי חוט שדרה. Savic וחבריו²⁵ מצאו רמת מהימנות גבוהה מאוד בין בודקים שונים בכל המרכיבים של הבדיקה הניורולוגית (כולל בדיקה מוטורית שהוערכה באמצעות MMT) במטופלים נפגעי חוט שדרה.

פגיעה עצבית מרכזית נוספת היא CP, ומהימנות ה-MMT בפגיעה זו נבדקה על-ידי Klingels וחבריו.²⁷ לעומת המחקר הקודם הם מצאו שהמהימנות בין בודקים שונים היא נמוכה בהרבה לעומת המהימנות בין הבודק לעצמו. לדעת החוקרים, ניתן להשתמש ב-MMT לשם הערכת כוח שרירים של הגפה העליונה בילדים הסובלים מ-CP, אך יש צורך לערוך את הבדיקות המעקב על-ידי אותו בודק. מסקנה נוספת הייתה שיש צורך לערוך את הבדיקות על פי פרוטוקול מדויק, ולשמור על סטנדרטיזציה ועל דיוק מבחינת ההוראות בכל בדיקת שריר לשם שיפור המהימנות בין בודקים שונים. גם מחקרים אלו אינם חפים מבעיות, כגון היעדר תיאור מספק של הנבדקים שהוכנסו למחקר, היעדר פירוט של היקף הניסיון של הבודקים בשימוש ב-MMT והיעדר נימוק למספר הנבדקים במחקר. במחקר של Savic וחבריו²⁵ נכללו נבדקים עם טווח רחב מאוד של זמן הפגיעה (3 חודשים עד 43 שנים) ושל טווח הגילאים (18-72 שנים), עובדה שעלולה להטות לחיוב את התוצאות של ההסכמה בין הבודקים. במחקר של Klingels וחבריו²⁷ אין הפרדה בין הרמות השונות של CP וייתכן שלרמות שונות של CP נמצא רמות מהימנות שונות, אין פירוט על פרוטוקול הבדיקה לשם שחזורו. בעיה נוספת ועיקרית במחקר זה היא העובדה שאין התייחסות לסמיות הבודקים, מה שעלול להשפיע מאוד על המהימנות. בגלל עובדות אלו קשה לקבוע את רמת המהימנות של ה-MMT בפתולוגיות אלו.

פתולוגיה נוספת שגורמת לחולשה במספר שרירים היא מחלת CMT. את המהימנות בין בודקים של ה-MMT בחולים במחלה זו בדקו Vinci וחבריו,²⁶ ונמצאה רמת מהימנות טובה מאוד. לפיכך, החוקרים הסיקו שניתן להשתמש בבדיקה זו להערכת כוח שריר במטופלים הסובלים ממחלת CMT. הבעיות העיקריות במחקר זה הן פירוט לא מספק של הנבדקים שנכללו במחקר, שיקולים לא ברורים בבחירת המדגם, לא צוינה רמת הניסיון שיש לבודקים בשימוש בבדיקות ה-MMT בחולי

וחלוקת המטופלים הסימפטומטיים לקטגוריות נעשתה לא על פי הפתולוגיה שלהם. במקום זאת, קיבצו אותם לקבוצת הנבדקים הסימפטומטיים, מה שעלול להוות בעיה משום שייתכן שבכל מחלה לכל שריר רמת מהימנות שונה. נוסף על כך, במחקר זה, ההגדרה שניתנה לרמת כוח השריר היא אם השריר הוא בעל כוח נורמלי או חלש בלבד, אך יש צורך לבדוק גם את הדירוגים של רמות כוח השרירים הנמוכות יותר. הסיבה לכך היא שייתכן שרמות כוח שונות יראו רמות שונות של מהימנות. אחת הבעיות העיקריות של המחקר היא היעדר דיווח על המובהקות הסטטיסטית של רמות ההסכמה בין הבודקים.

רמת המהימנות של ה-MMT להערכת כוח שריר נבדקה גם במטופלים המאושפזים ביחידה לטיפול נמרץ. אוכלוסייה זו חווה במקרים רבים חוסר תנועתיות לאורך זמן, מה שיכול לגרום לחולשה כללית. לפיכך חשוב לעשות הערכה כללית של כוח שרירים במטרה להעריך את רמת החולשה ואת השפעתה של ההתערבות הטיפולית. כמו כן, חשובה היכולת להעריך את כוח הגפיים התחתונות והעליונות בנפרד לשם הערכה פרוגנוסטית של יכולת המטופל לבצע משימות תפקודיות ספציפיות (כגון הליכה, אכילה עצמאית).²³ מחקרים שעסקו באוכלוסייה זו בוצעו על-ידי Fan וחבריו²³ ו-Hough וחבריה.²⁴ במחקר של Fan וחבריו²³ הסיקו שבקרב צוות המחקר המיומן, MMT היא שיטה מהימנה לבדיקת כוח השרירים של מאושפזים ביחידה לטיפול נמרץ, בהערכה כוללת ובהערכת גפיים עליונות ותחתונות בנפרד. כמו כן, במחקר זה ניתן לראות מהימנות טובה בין הבודקים גם כאשר מעריכים שרירים יחידים.

בדומה למחקר של Fan וחבריו,²³ במחקרם של Hough וחבריה²⁴ הגיעו למסקנה שבקרב מאושפזים ביחידה לטיפול נמרץ, שיכלו לעבור את הבדיקה, ישנה רמת הסכמה גבוהה בין הבודקים בקביעת "חולשה נרכשת של טיפול נמרץ" (בהערכה כוללת של השרירים), אך בניגוד למחקר הקודם, נמצאה רמה נמוכה של מהימנות בהערכת שריר יחיד, במיוחד בשרירים הפרוקסימליים.

לשני המחקרים האחרונים ישנן כמה מגבלות, כגון מספר קטן של נבדקים (מה שיכול לפגוע בהכללה של תוצאות המחקר) היעדר הסבר לכמות הנבדקים במחקר, היעדר פירוט על המטופלים שנכללו, היעדר הסבר מדוע לא היה ניתן לבדוק אצל חלק מן המטופלים את כל השרירים, אין תיאור של פרוטוקול בבדיקת השרירים כדי שיהיה ניתן לשחזרו, ישנו

במחקרים נוספים לשם הערכת מדדים אלו של שיטה זו.

מקורות

1. Nicholas JA, Sapega A, Kraus H, et al. Factors influencing manual muscle tests in physical therapy. J Bone Joint Surg Am 1978;60(2):186-90.
2. Bohannon RW. Manual muscle testing: does it meet the standards of an adequate screening test? Clin Rehabil 2005;19:662.
3. Wadsworth CT, Krishnan R, Sear M, et al. Intrarater reliability of manual muscle testing and hand-held dynamometric muscle testing. Phys Ther 1987;67:1342-47.
4. Perossa DR, Dziak M, Vernon HT, et al. The intra-examiner reliability of manual muscle testing of the hip and shoulder with a modified sphygmomanometer: a preliminary study of normal subjects. J of the Canadian Chiropractic Association 1998;42(2):73.
5. Frese E, Brown M, Norton BJ. Clinical reliability of manual muscle testing: Middle Trapezius and Gluteus Medius muscles. Phys Ther 1987;67:1072-6.
6. Lovett RW, Martin EG. Certain aspect of infantile paralysis: With a description of a method of muscle testing. J Am Med Assn 1916;66:729-33.
7. Brandsma JW, Schreuders TA. Sensible manual muscle strength testing to evaluate and monitor strength of the intrinsic muscles of the hand: a commentary. J of Hand Therapy 2001; 14(4): ProQuest Central pg. 273.
8. Hoppenfeld S. Physical Examination of the Spinal and Extremities. New York, Appleton-Century-Crofts 1976; 26.
9. Sapega AA. Current Concepts Review: Muscle Performance Evaluation in Orthopedic Practice. J Bone Joint Surg 1990;72-A(10):1562-74.
10. Conable KM, Rosner AL. A narrative review of manual muscle testing and implications for muscle testing research. J of Chiro Med 2011;10:157-65.

CMT מסוג A1. במחקר מוצג אחוז הגפיים שהייתה הסכמה לגביו בלבד ולא רמת ההסכמה לגבי כל שריר בנפרד. כמו כן, אחד החסרונות הגדולים במחקר זה, בדומה למחקרים רבים קודמים, הוא חוסר פירוט של רמת המובהקות הסטטיסטית של אחוז ההסכמה. בגלל בעיות אלו ולמרות מסקנות החוקרים, לדעת עורך הסקירה לא ניתן לקבוע בצורה ברורה אם בדיקת ה-MMT מהימנה באוכלוסייה זו ויש צורך במחקרי המשך לביצוע בדיקות כוח כמותיות ובדיקות של התקדמות הירידה בכוח של השרירים הדיסטליים במחלה זו.

לסיכום, המחקרים שבוצעו בעשור האחרון ערכו את בדיקות ה-MMT באוכלוסיות שונות, ורובם בדעה שניתן להשתמש בשיטה כדי להעריך את כוח השרירים במטופלים שנבדקו. למרות מסקנות אלו, במחקרים רבים היו ליקויים בשיטות המחקר, מה שהוביל להגדרתם כבעלי איכות נמוכה או בינונית. בעקבות סקירת הספרות הנוכחית, אפשר להסיק שיש צורך במחקרים נוספים, בעלי איכות מתודולוגית גבוהה, כדי לקבוע אם בדיקה זו היא מהימנה ו\או תקפה לאוכלוסייה ספציפית. כיום, בגלל היעדר מחקרים המאששים או השוללים בצורה מובהקת את מהימנות ותקפות השיטה, אין אלא להוסיף להשתמש בקליניקה בבדיקה זו כדי להעריך את כוח השרירים משום שהיא קלה, נוחה ואינה דורשת אביזרים כלשהם.

מגבלות סקירת הספרות

- כותב סקירת הספרות הנוכחית היה היחיד שחיפש והעריך את המאמרים שנסקרו.
- אף על פי שנעשה חיפוש מקיף של הנושא הרלוונטי בספרות, ייתכן שמאמרים רלוונטיים הוחמצו בעקבות מילות חיפוש לא מדויקות או שמחקרים הוצאו בטעות מן הסקירה על סמך הכותרת והתקציר שלהם.
- בסקירה זו נכללו רק מחקרים שגרסתם המלאה הייתה זמינה (תשעה בסך הכול).

נקודות עיקריות העולות מן הסקירה

- על אף רמת המחקרים שנסקרו, תוצאותיהם מראות שניתן להשתמש בשיטת בדיקה זו כדי להעריך כוח שריר באוכלוסיות מסוימות.
- בעשר השנים האחרונות לא נעשו מספיק מחקרים באיכות גבוהה בנושא מהימנות ותקפות של ה-MMT, ולכן יש צורך

11. Cuthbert SC, Goodheart G. On the reliability and validity of manual muscle testing: A literature review. *Chiropractic & Osteopathy* 2007;15:4. DOI: 10.1186/1746-1340-15-4.
12. Haas M, Cooperstein R, Peterson D. Disentangling manual muscle testing and Applied Kinesiology: Critique and reinterpretation of a literature review. *Chiropractic & Osteopathy* 2007;15:11. DOI: 10.1186/1746-1340-15-11.
13. Fleiss, J.L. *The Design and Analysis of Clinical Experiments*. Wiley, New York, Chichester-Brislane-Toronto-Singapore 1986.
14. Eliasziw M, Young SL. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: Using goniometric measurements as an example. *Phys Ther* 1994;74:777-88.
15. Gwet KL. Intrarater reliability. In: Wiley & Sons, eds. *Wiley Encyclopedia of Clinical Trials* 2008. DOI:10.1002/9780471462422.
16. Payton O.D. *Research: The Validation of Clinical Experience*. Philadelphia, FA Davis 1994.
17. Wright JW, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am* 2003; 85: 1-3.
18. Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract* 2012;18(6):1126-32.
19. Bellet RN, Adams L, Morris NR. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness—a systematic review. *Physiotherapy* 2012;98(4):277-86.
20. Brookham RL, McLean L, Dickerson CR. Construct validity of muscle force tests of the rotator cuff muscles: an electromyographic investigation. *Phys Ther* 2010;90(4):572-80.
21. Hara Y, Matsudaira K, Hara N, et al. A comparison of muscle strength testing for great toe extension. *J Orthop Sci* 2011;16:765-7.
22. Paternostro-Sluga T, Grim-Stieger M, Posch M, et al. Reliability and validity of the Medical Research Council (MRC) scale and a modified scale for testing muscle strength in patients with radial palsy. *J Rehabil Med* 2008;40 :665-71.
23. Fan E, Ciesla ND, Truong AD, et al. Inter-rater reliability of manual muscle strength testing in ICU survivors and simulated patients. *Intensive Care Med* 2010;36:1038-43. DOI 10.1007/s00134-010-1796-6.
24. Hough CL, Lieu BK, Caldwell ES. Manual muscle strength testing of critically ill patients: feasibility and interobserver agreement. *Critical Care* 2011; 15: R43. DOI: 10.1186/cc10005.
25. Savic G, Bergstrom EMK, Frankel HL, et al. Inter-rater reliability of motor and sensory examinations performed according to American Spinal Injury Association standards. *Spinal Cord* 2007;45:444-51.
26. Vinci P, Serrao M, Pierelli F, et al. Lower limb manual muscle testing in the early stages of Charcot-Marie-Tooth disease type 1A. *Functional Neurology* 2006;21(3):159-63.
27. Klingels K, De Cock P, Molenaers G, et al. Upper limb motor and sensory impairments in children with hemiplegic cerebral palsy. Can they be measured reliably? *Disability and Rehabilitation* 2010;32(5):409-16.
28. Jepsen JR, Laursen LH, Larsen AI, et al. Manual strength testing in 14 upper limb muscles: A study of inter-rater reliability. *Acta Orthop Scand* 2004;75(4):442-8.
29. Geere J, Chester R, Kale S, et al. Power grip, pinch grip, manual muscle testing or thenar atrophy – which should be assessed as a motor outcome after carpal tunnel decompression? A systematic review. *BMS Musculoskeletal Disorder* 2007;20(8):114.
30. Michener LA, Boardman ND, Pidcoe PE, et al. Scapular Muscle Tests in Subjects With Shoulder Pain and Functional Loss: Reliability and Construct Validity. *Phys Ther* 2005;85(11):1128-38.

טבלה 1: רמת המחקרים ואיכותם

רמת המחקר ואיכותו	סוג המחקר	כותב המאמר ושנה
רמה I* (מבחינת Levels of Evidence (for Primary Research Question), איכות בינונית**).	בדיקת התקפות של MMT***	Brookham et al. 2010 ²⁰
איכות נמוכה** (לא הוערכה רמתו מבחינת התקפות כי אין מספיק פרטים)	בדיקת המהימנות והתקפות של MMT	Hara et al., 2011 ²¹
איכות נמוכה** (לא הוערכה רמתו מבחינת התקפות כי שיטת הבדיקה אינה מתאימה)	בדיקת התקפות והמהימנות - בין בודקים שונים ובין בודק לעצמו - של MMT, קלאסי ומותאם, בשרירי האמה בפגיעה בעצב הרדיאלי	Paternostro-Sluga et al. 2008 ²²
איכות בינונית**	בדיקת המהימנות בין בודקים שונים של MMT במטופלים לאחר אשפוז ביחידה לטיפול נמרץ.	Fan et al. 2010 ²³
איכות גבוהה**	בדיקת המהימנות בין בודקים שונים של MMT בחולים קריטיים ביחידה לטיפול נמרץ.	Hough et al. 2011 ²⁴
איכות בינונית**	בדיקת המהימנות, בין בודקים שונים, של ASIA ⁺ מוטורי ב- 10 שרירי מפתח שונים בגפיים (הכולל MMT) וסנסורי בפגועי חוט שדרה.	Savic et al. 2007 ²⁵
איכות נמוכה**	בדיקת המהימנות בין בודקים שונים של MMT בשרירים שונים בגפה התחתונה בחולי Charcot-Marie-Tooth מסוג A1	Vinci et al. 2006 ²⁶
איכות נמוכה**	בדיקת המהימנות - בין הבודק לעצמו ובין בודקים שונים - של MMT בילדים חולי CP**.	Klingels et al. 2010 ²⁷
איכות נמוכה**	בדיקת המהימנות - בין בודקים שונים - של MMT בשרירי הגפיים העליונות	Jepsen et al. 2004 ²⁸

* מחקרי התקפות דורגו על פי "Levels of Evidence for Primary Research Question" בין רמות I-V.¹⁹
 ** מחקרי המהימנות דורגו על פי "Critical appraisal tool for validity and reliability studies of objective clinical tools" כאשר מבין 13 השאלות בטופס, מחקר שהיה לו יותר מ- 60% תשובות "כן" נחשב לבעל איכות גבוהה, מחקר בעל פחות מ-40% תשובות "כן" נחשב לבעל איכות נמוכה, ובטווח שביניהם - נחשב לבעל איכות בינונית.²⁰
 *** Manual Muscle Testing=MMT
 + American Spinal Injury Association=ASIA
 ++ Cerebral Palsy=CP

Reliability and Validity of Manual Muscle Testing: A Systematic Literature Review

Yigal Ktzap, BPT

Physical therapist, Maccabi Health Services,
Rehovot, Israel.

Masters' Degree Student, Ben-Gurion University

Abstract

Background: Manual muscle testing (MMT) is one of the accepted tests for assessing muscle strength in the clinic. It is a simple and inexpensive test, and does not require the use of equipment. Despite the widespread use of this test, there is not enough information about its validity and reliability.

Objective: To review the studies published in the past ten years that assessed the reliability and validity of the MMT.

Methods: PubMed, NARIC, CIRRIE, EBSCO, Google Scholar, EM consult, Cochrane Library, Web of Science & Medscape CME were searched. Inclusion criteria for the study were articles that deal with the reliability and/or validity of the MMT, used the Medical Research Council (MRC) Scale, were published in English and could be obtained in full text. All studies included in the study were assessed using the "Critical appraisal tool for validity and reliability studies of objective clinical tools". The validity studies were evaluated using also the "Levels of Evidence for Primary Research Question".

Results: The initial search found over 300 studies; after reading their titles and abstracts, only nine articles that dealt with the reliability and/or validity of the MMT seemed relevant. Of these, eight dealt with reliability; one dealt with structural validity; and two dealt with both reliability and validity. Most of the articles seemed to be of low

or moderate quality, while only one study was of high quality.

Discussion and Conclusions: The assessment of the studies reviewed shows that it is possible to use the method of MMT to evaluate levels of muscle strength in certain populations; however, additional studies employing a high quality methodology are needed to establish the reliability and validity of the MMT method for distinguishing different levels of muscle strength.

Keywords: muscle testing; manual muscle testing; reliability; validity; systematic review.