

השימוש במדדי אבחנה לצורך קבלת החלטות קליניות בפיזיותרפיה

אלון רבין DPT, MS, המרכז האוניברסיטאי אריאל בשומרון, ביה"ס למדעי הבריאות, המחלקה לפיזיותרפיה, ARPT פיזיותרפיה – מנהל

תקציר

במהלך עבודתם נדרשים פיזיותרפיסטים לקבל החלטות בעלות אופי אבחנתי. הבנת תהליך האבחנה והכלים לביצועה חיונית על כן לכל פיזיותרפיסט. מאמר זה יסקור את תהליך האבחנה בכללותו תוך שימת דגש על הבנת מדדי האבחנה העיקריים: רגישות (Sensitivity), סגוליות (Specificity), ערך ניבוי חיובי/שלילי (Positive/negative predictive value) ויחס סבירות חיובי/שלילי (Positive/negative likelihood ratio). הבנת משמעות מדדים אלו תסייע לפיזיותרפיסט בבחירת הבדיקות המתאימות להערכת המטופל ובפירוש תוצאתן.

מילות מפתח: אבחנה, מהימנות, דיוק אבחנתי, רגישות, סגוליות, ערך ניבוי חיובי/שלילי, יחס סבירות חיובי/שלילי.

מבוא

למרות שטיפול פיזיותרפיה מתבצע בד"כ לאחר הפנייה מגורם רפואי ובהתאם לאבחנת אותו הגורם, הרי שתהליך הטיפול מתחיל בבדיקת המטופל והערכת צרכיו. ממצאי הבדיקה מכוונים את הפיזיותרפיסט בקביעת ההתערבות המתאימה וכן בקביעת הפרוגנוזה. פיזיותרפיסטים אינם מוסמכים או מוכשרים לבצע אבחנה רפואית. אין לנו את הידע ו/או הכלים הנדרשים על-מנת לקבוע, לדוגמא, אם מטופל המתלונן על כאבים בחגורת הכתפיים סובל מבעיית שלד-שריר, דלקת ראות או סרטן (2). יחד עם זאת, במהלך האינטראקציה עם המטופל, אנו בהחלט נדרשים להחליט מספר החלטות בעלות אופי אבחנתי. פיזיותרפיסטים נדרשים להחליט האם הבעיה ממנה סובל המטופל כלל ניתנת לטיפול ע"י פיזיותרפיסט, או האם יש להפנות את המטופל לגורם אחר (2,6). לחילופין, על הפיזיותרפיסט להחליט מהו מקור הבעיה ממנה סובל המטופל: האם מדובר בכאב הנובע מבעיית כתף או מקור הכאב כלל בעמוד השדרה הצווארי. לבסוף על הפיזיותרפיסט להחליט מהי ההתערבות בעלת הסיכוי הטוב ביותר להביא לשיפור בתלונות המטופל ותפקודו (8). אם כן, תהליך הבדיקה בפיזיותרפיה הנו בעל אופי אבחנתי, ואבחנה מקצועית מהווה חלק אינטגרלי ממקצוע הפיזיותרפיה.

בין אם מדובר בצורך להחליט אם לטפל במטופל או להפנותו לגורם אחר, ובין אם להתרכז בבדיקת הכתף או עמוד שדרה צווארי, על הפיזיותרפיסט לאסוף מספיק מידע בעל ערך אבחנתי שיאפשר מענה על אותן שאלות. את תהליך האבחנה ניתן להקביל לחקר ה"אמת". הבעיה ממנה סובל המטופל במציאות היא "האמת". האמת איננה ידועה לנו במפגש הראשון עם המטופל. במקום זאת, המטופל מציג בפנינו מגוון תסמינים וסימנים מהם הוא סובל. תסמינים וסימנים אלו מהווים "רמזים" לבעיה ממנה סובל המטופל ("אמת"). תפקידנו, כקלינאים, הנו להשתמש באותם רמזים על-מנת לחשוף את ה"אמת", או במילים אחרות, לאבחן את טיב הבעיה.

Delitto (5) מציע להפוך את שאלת האבחנה משאלה בפורמט פתוח לשאלה בפורמט סגור. לדוגמא: במקום לשאול "מה הבעיה של המטופל?", שאלה עם מגוון רחב ביותר של תשובות אפשריות, יש לשאול "האם הבעיה של המטופל נובעת ממפרק הכתף או מעמ"ש צווארי?" שאלה כזו מצריכה מיקוד בהשערות הסבירות ביותר לגורם לבעיית המטופל. לאחר שיצרנו רשימת השערות יש להעריך את מידת הסבירות של כל אחת מהן (9). מידת סבירות זו נקראת ה- Pre-test probability. Pre-test probability נקבעת עפ"י הצגת המקרה של המטופל, בסיס הידע של הבודק ו/או נתונים על שכיחות התופעה המשוערת באוכלוסייה, (8,9,11). Riddle (10) ממליץ לנקוב בערך מספרי המבטא את ה- Pre-test probability. במקרים בהם ה- Pre-test probability נמוכה מציע Riddle לתת ערך של 25%, ב- Pre-test probability בינונית 50% וב- Pre-test probability גבוהה 75%. לאחר דירוג סבירות ההשערות השונות לגבי בעיית המטופל, המשך הבדיקה יתרכז באיסוף מידע על-מנת לאמת או לשלול את ההשערות השונות.

חשוב להבין ששלילה או אימות מוחלט של השערותינו אינו אפשרית (1). לעולם לא נגיע ל-100% ודאות לגבי נכונות השערה כלשהי (4,5). מטרותנו, אם כן, היא להגיע למידת ודאות מספקת לגבי צדקת השערה מסוימת או לחילופין לשלול במידת ודאות מספקת את ההשערות האלטרנטיביות על-מנת שנוכל להפסיק לבדוק את המטופל ונוכל להתחיל לטפל בו. מידת הודאות המינימלית המאפשרת את קביעת האבחנה ותחילת הטיפול נקראת הסף הטיפולי (1). הסף הטיפולי איננו אחיד ומשתנה בהתאם לסיטואציה הקלינית. ככל שטיפול מסוים יקר, בעל תופעות לווי משמעותיות ובעל יעילות נמוכה יותר, הסף הטיפולי גבוה יותר. טיפולים יעילים, זולים ובטוחים מצריכים סף טיפולי נמוך יותר (1,8).

על מנת לאמת או לשלול את השערותינו יש לבצע בדיקות שונות. אלו הן למעשה בדיקות אבחנתיות. תוצאת אותן בדיקות אמורה להפחית את מידת חוסר הודאות הקיימת לגבי אבחנה כלשהי עד למידה כזו שתאפשר לקבוע כי זו האבחנה בה יש לטפל (מעבר הסף הטיפולי). בדיקות אבחנתיות יכולות להיות בעלות אופי מגוון: בדיקות מעבדה (לדוגמא, רמת הסוכר בדם), בדיקות פיזיקליות (Straight leg raise) ואפילו שאלות מראיון המטופל. המשך המאמר יתרכז במידע המתקבל מביצוע בדיקות אבחנתיות שונות וכיצד ניתן להשתמש במידע זה על-מנת להעריך מחדש את מידת הסבירות של השערותינו הראשונות באשר לבעיית המטופל.

ערך אבחנתי:

כאשר משווים בין תוצאות הבדיקה הנחקרת (SLR) ותוצאות האמת (MRI), יתכנו 4 מצבים שונים (8):

- A** - תוצאת הבדיקה הנחקרת חיובית והתופעה הנבדקת אכן קיימת - True positive
- B** - תוצאת הבדיקה הנחקרת שלילית אולם התופעה הנבדקת קיימת - False negative
- C** - תוצאת הבדיקה הנחקרת חיובית אולם התופעה הנבדקת איננה קיימת - False positive
- D** - תוצאת הבדיקה הנחקרת שלילית והתופעה הנבדקת אכן איננה קיימת - True negative

מתוצאות טבלת ההשוואה הנ"ל ניתן לחשב מספר מדדים בעלי ערך אבחנתי:

רגישות (Sensitivity)

אחוז הפעמים שתוצאת הבדיקה חיובית כאשר התופעה אכן קיימת. למעשה מדובר במידת ה- True positive. רגישות מחושבת ע"י הנוסחה:

$$\text{True positive} / (\text{True positive} + \text{false negative}) \times 100$$

$$\text{או } A / (A + B) \times 100$$

סגוליות (Specificity)

אחוז הפעמים שתוצאת הבדיקה שלילית כאשר התופעה הנמדדת איננה קיימת. למעשה מדובר במידת ה- True negative. סגוליות מחושבת ע"י הנוסחה:

$$\text{True negative} / (\text{True negative} + \text{False positive}) \times 100$$

$$\text{או } D / (D + C) \times 100$$

על מנת להמשיך ולהסביר את משמעות מדדי האבחנה השונים נשתמש בדוגמא הבאה:

100 נבדקים עם כאב גב תחתון עוברים הדמיה מגנטית ובדיקת SLR. תוצאות שתי הבדיקות מסוכמות בטבלה מספר 2.

טבלה 2: רגישות וסגוליות של SLR. (שכיחות פריצת דיסק של 50%):

	+ MRI	- MRI
+ SLR	45	37
- SLR	5	13
	Sensitivity: $45 / (45 + 5) \times 100 = 90\%$	Specificity: $13 / (13 + 37) \times 100 = 26\%$

שכיחות פריצת דיסק בקרב האוכלוסייה הנבדקת הנה 50%. כלומר, ב- 50 נבדקים הראתה ההדמיה המגנטית פריצת דיסק בעמ"ש מותני וב- 50 הנותרים לא נצפו פריצות דיסק. אם נתמקד ב- 50 המטופלים עם MRI חיובי, נראה שב- 45 מהם הייתה גם תוצאת ה-SLR חיובית. זוהי מידת ה- True-positive. ב- 5 הנבדקים הנותרים עם MRI חיובי, תוצאת ה-SLR הייתה שלילית. זוהי מידת ה- False-negative

לכל בדיקה יש ערך אבחנתי מסוים. הערך האבחנתי של הבדיקה נקבע עפ"י מידת הדיוק של הבדיקה בניבוי התופעה אותה היא בודקת. בדיקה מושלמת מנבאת את התופעה אותה היא בודקת ב-100%. כלומר, תוצאת הבדיקה תהיה חיובית בכל המקרים בהם התופעה קיימת ולחילופין, תוצאת הבדיקה תהיה שלילית בכל המקרים בהם התופעה איננה קיימת. בדיקות כאלה, לצערנו, נדירות. במציאות, הבדיקה תדייק במקרים מסוימים, כלומר תהיה חיובית כאשר התופעה הנמדדת אכן קיימת במציאות (True-positive) ושלילית כאשר התופעה הנמדדת איננה קיימת (True-negative). במקרים אחרים, הבדיקה לא תדייק ותראה תוצאה חיובית כאשר התופעה הנמדדת איננה קיימת (False-positive), או תוצאה שלילית כאשר התופעה הנמדדת אכן קיימת (False-negative) (9).

לבדיקות שונות ערך אבחנתי שונה. תוצאת בדיקה מסוימת עשויה לחזק את החשד בבעיה כלשהי ולקרב אותנו אל הסף הטיפולי יותר מתוצאת בדיקה אחרת גם אם שתי הבדיקות בודקות את אותה התופעה. לדוגמא, גם בדיקת ה-Straight leg raise וגם בדיקת ה-Crossed straight leg raise נועדו לאבחן פריצת דיסק בעמוד השדרה המותני (4,7). יחד עם זאת, תוצאה חיובית בכל אחת מהבדיקות לא תאמת פריצת דיסק באותה מידה של ודאות (4,7). בנוסף, אין זה נכון לחשוב שתוצאה חיובית בבדיקה כלשהי תחזק את החשד בבעיה באותה המידה שתוצאה שלילית תחליש אותו. במילים אחרות לבדיקות שונות ולתוצאות שונות של אותה הבדיקה יתכן ערך אבחנתי שונה. על-מנת לעשות שימוש יעיל בבדיקה אבחנתית, חשוב לדעת מראש את הערך האבחנתי של אותה הבדיקה.

מדדי ערך אבחנתי:

הערך האבחנתי של בדיקה נמדד בד"כ ע"י השוואת תוצאות הבדיקה לתוצאות מדד המשקף את המצב האמיתי בצורה המדויקת ביותר (4,8) (בדיקת Gold standard). לדוגמא, השוואת תוצאת בדיקת SLR לתוצאת הדמיה מגנטית (Magnetic resonance imaging-MRI) על-מנת לזהות פריצת דיסק בעמוד שדרה מותני. ההדמיה המגנטית מהווה את המדד הקרוב ביותר לאמת (מלבד הסתכלות ישירה בפעולה כירורגית שאיננה ישימה תמיד). את תוצאות ההשוואה ניתן לסכם בטבלת 2 X 2:

טבלה 1: תוצאות אפשריות בהשוואת תוצאות בדיקה קלינית (SLR) לתוצאות הדמיה מגנטית (MRI):

	Herniated disc (+ MRI scan)	Normal disc (- MRI scan)	
+ SLR	A True-positive	C False-positive	ערך ניבוי חיובי: $A / (A + C) \times 100$
- SLR	B False-negative	D True-negative	ערך ניבוי שלילי: $D / (D + B) \times 100$
	רגישות (Sensitivity): $A / (A + B) \times 100$	סגוליות (Specificity): $D / (D + C) \times 100$	

של ה-SLR. רגישות מבטאת את אחוז הפעמים בהם תוצאת הבדיקה חיובית מתוך כלל הפעמים שהתופעה אכן קיימת. כלומר אם פריצת דיסק קיימת ב-50 נבדקים ותוצאת ה-SLR חיובית ב-45 מאותם מקרים הרי שרגישות ה-SLR הנה: $45/(45 + 5) \times 100 = 90\%$. אם נתמקד ב-50 הנבדקים עם MRI שלילי, נראה שב-13 מתוכם תוצאת ה-SLR אכן שלילית (True-negative). ב-37 הנבדקים הנותרים תוצאת ה-SLR חיובית (False-positive). סגוליות מבטאת את אחוז הפעמים בהם תוצאת הבדיקה שלילית מתוך כלל הפעמים שהתופעה איננה קיימת. כלומר, אם פריצת דיסק איננה קיימת ב-50 נבדקים וב-13 מתוכם גם תוצאת ה-SLR שלילית, הרי שסגוליות ה-SLR הנה: $13/(13 + 37) \times 100 = 26\%$.

טבלה 3: Positive predictive value (PPV) ו-Negative predictive value (NPV), רגישות וסגוליות של SLR בשכיחות פריצת דיסק של 50%-20%:

שכיחות פריצה 50%	+ MRI	+ MRI	Positive Value
+ SLR	45	37	Positive: $45/(45 + 37) \times 100$ 55%
- SLR	5	13	Negative: $13/(13 + 5) \times 100$ 72%
	Sensitivity: $45/(45 + 5) \times 100$ 90%	Specificity: $13/(13 + 37) \times 100$ 26%	
שכיחות פריצה 20%	+ MRI	+ MRI	Predictive Value
+ SLR	18	59	Positive: $18/(18 + 59) \times 100$ 23%
- SLR	2	21	Negative: $21/(21 + 2) \times 100$ 91%
	Sensitivity: $18/(18 + 2) \times 100$ 90%	Specificity: $21/(21 + 59) \times 100$ 26%	

כאשר שכיחות התופעה הנחקרת יורדת, ערך PPV יורד ואילו ערך NPV עולה. על כן, נוכל להתחשב בערכים אלה רק אם שכיחות התופעה הנחקרת, בסיטואציה בה אנו נמצאים, דומה לשכיחות התופעה בקבוצת המחקר ממנה הופקו ערכים אלו. ערכי הרגישות והסגוליות לעומת זאת, אינם משתנים עם השינוי בשכיחות התופעה הנחקרת. ערכים אלה אינם תלוי שכיחות היות והם מתייחסים לתוצאות האמת ולא לתוצאות הבדיקה.

הרגישות והסגוליות קשורות בכל זאת ל-PPV ו-NPV (10): כאשר הרגישות גבוהה, גם ה-NPV גבוה וכאשר הסגוליות גבוהה, גם ה-PPV גבוה. לכן, ניתן לומר, שכאשר הרגישות גבוהה, תוצאה שלילית בבדיקה בד"כ תסמן שהתופעה הנחקרת איננה קיימת (האבחנה נשללת). כאשר הסגוליות גבוהה, תוצאה חיובית בבדיקה בד"כ תסמן שהתופעה הנחקרת קיימת (10) (האבחנה מאומתת). ניתן לזכור את הקשר הנ"ל עפ"י ראשי התיבות: SPPIN ו-SNNOUT (8).

SNNOUT - Sensitive test, Negative result, rule **OUT**
SPPIN - Specific test, Positive result, rule **In**

רגישות מבטאת את אחוז הפעמים בהם תוצאת הבדיקה חיובית מתוך כלל הפעמים שהתופעה אכן קיימת. כלומר אם פריצת דיסק קיימת ב-50 נבדקים ותוצאת ה-SLR חיובית ב-45 מאותם מקרים הרי שרגישות ה-SLR הנה: $45/(45 + 5) \times 100 = 90\%$. אם נתמקד ב-50 הנבדקים עם MRI שלילי, נראה שב-13 מתוכם תוצאת ה-SLR אכן שלילית (True-negative). ב-37 הנבדקים הנותרים תוצאת ה-SLR חיובית (False-positive). סגוליות מבטאת את אחוז הפעמים בהם תוצאת הבדיקה שלילית מתוך כלל הפעמים שהתופעה איננה קיימת. כלומר, אם פריצת דיסק איננה קיימת ב-50 נבדקים וב-13 מתוכם גם תוצאת ה-SLR שלילית, הרי שסגוליות ה-SLR הנה: $13/(13 + 37) \times 100 = 26\%$.

רגישות וסגוליות מבטאות למעשה את יכולת הבדיקה לזהות את תוצאות האמת (תוצאות ה-gold standard). מדדים אלה נותנים, למעשה, את המידע ההפוך למידע הנדרש לנו כקלינאים (8). בסיטואציה הקלינית איננו יודעים את תוצאות האמת. אנו יודעים אך ורק את תוצאת הבדיקה האבחנתית. לכן, בסיטואציה הקלינית, אנו רוצים לדעת מה ההסתברות שהאבחנה אכן קיימת במציאות כאשר תוצאת הבדיקה חיובית, או לחילופין, מהי ההסתברות שהאבחנה איננה קיימת במציאות כאשר תוצאת הבדיקה שלילית. את המידע הזה לא ניתן, כאמור, לקבל מערכי הרגישות ו/או הסגוליות.

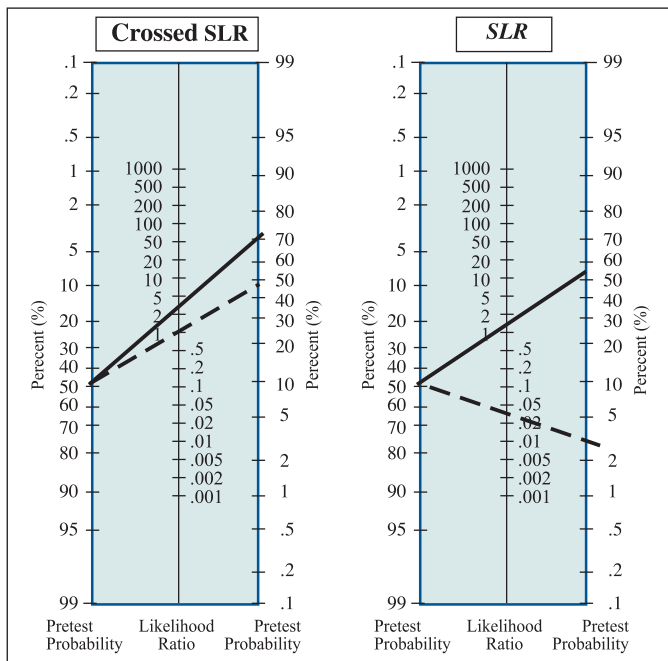
ערך ניבוי (Predictive value):

ההסתברות שאבחנה כלשהי אכן קיימת כאשר תוצאת הבדיקה חיובית מבטאת ע"י ערך הניבוי החיובי של הבדיקה (Positive predictive value - PPV). ה-PPV הוא אחוז הפעמים שהאבחנה קיימת כאשר תוצאת הבדיקה חיובית. PPV מחושב ע"י הנוסחה: $\text{True positive} / (\text{True positive} + \text{False positive}) \times 100$. ככל ש-PPV גבוה יותר, הרי שהסתברות שהאבחנה אכן קיימת במציאות, כאשר תוצאת הבדיקה חיובית, גבוהה יותר. אם נשתמש בנתונים מטבלה 2, הרי ש-PPV של ה-SLR באבחנת פריצת דיסק הנו: $45/(45 + 37) \times 100 = 55\%$. משמעות נתון זה היא שב-55% מהמקרים בהם בדיקת SLR חיובית אכן קיימת פריצת דיסק. זהו אחוז נמוך יחסית ולכן, בהתחשב בנתונים אלו, תוצאה חיובית בבדיקת SLR לא תביא למידת בטחון גבוהה במיוחד באשר לקיום פריצת דיסק. אחוז הפעמים שהאבחנה איננה קיימת כאשר תוצאת הבדיקה שלילית נקרא ערך הניבוי השלילי של הבדיקה (NPV-Negative predictive value). NPV מחושב ע"י הנוסחה:

$$\text{NPV} = \text{True negative} / (\text{True negative} + \text{False negative}) \times 100$$

ככל ש-PPV גבוה יותר, הרי שהסתברות שהאבחנה איננה קיימת במציאות, כאשר תוצאת הבדיקה שלילית, גבוהה יותר. אם נשתמש בנתונים מטבלה 2, הרי ש-PPV של ה-SLR באבחנת פריצת דיסק הנו: $13/(13 + 5) \times 100 = 72\%$. משמעות נתון זה היא שב-72% מהמקרים בהם תוצאת ה-SLR שלילית, אין פריצת דיסק. ה-PPV של ה-SLR גבוה יותר מה-PPV שלו ולכן הבדיקה תתרום לנו יותר מבחינה אבחנתית כאשר תוצאתה שלילית מאשר כאשר תוצאתה חיובית. במילים אחרות, עפ"י הנתונים הנ"ל, SLR שלילי יעזור לשלול פריצת דיסק במידת בטחון בינונית. לעומת זאת, SLR חיובי לא יאמת פריצת דיסק בביטחון רב.

איור 1: נומוגרמה לציון השינוי בסבירות לפריצת דיסק כאשר תוצאת בדיקת SLR ו-Crossed SLR חיובית (קו מלא) או שלילית (קו שביר):



הטור השמאלי בנומוגרמה מסמל את ה-pre-test probability (אותה קבענו בתחילת תהליך האבחנה). הטור המרכזי מסמל את ה-LR. הטור הימני מסמל את ה-post-test probability.

כדי להעריך את ה-post-test probability לפריצת דיסק כאשר תוצאת ה-SLR חיובית נמתח קו מערך ה-pre-test probability המשוער, נאמר 50%, דרך ערך ה-LR + של ה-SLR ועד לטור ה-post-test probability. הנקודה בה קו זה חוצה את טור ה-SLR חיובית. במקרה זה, הסבירות לדיסק פרוץ כאשר ה-SLR חיובי עלתה רק במעט ועומדת על כ-55% (איור 1- גרף ימין). במילים אחרות SLR חיובי מגדיל את הסבירות לדיסק פרוץ רק ב-5% מעבר ל-pre-test probability. כדי להעריך את ה-post-test probability לקיום דיסק פרוץ כאשר תוצאת ה-SLR שלילית נמתח קו מערך ה-pre-test probability, דרך ערך ה-LR - של ה-SLR ועד לטור ה-post-test probability. במקרה זה, ה-post-test probability תעמוד על כ-3% (איור 1- גרף ימין). כלומר SLR שלילי מפחית את הסבירות לדיסק פרוץ, מ-50% ל-3%. זהו שינוי משמעותי בסבירות לדיסק פרוץ לעומת הערכתנו הראשונית (pre-test probability). מנתונים אלה מצטייר שוב כי בדיקת SLR אינפורמטיבית יותר כאשר תוצאתה שלילית מאשר חיובית.

לעומת ה-SLR, הרי שלבדיקת ה-Crossed SLR ערכים אבחנתיים שונים לגמרי (4,7):

רגישות: 29%, סגוליות: 88%, LR +: 2.41, ו-LR -: 0.81. אם נשתמש בנתונים אלה על-מנת להעריך את ה-post-test probability לפריצת דיסק לאחר ביצוע ה-Crossed SLR כאשר ה-pre-test probability לפריצת דיסק מוערכת ב-50% נגלה שכאשר ה-Crossed SLR חיובי ה-post-test probability תעלה לכ-70% (איור 1- גרף שמאל). כאשר ה-Crossed SLR שלילי ה-post-test probability תרד לכ-48%, שינוי מינימלי (איור 1 - גרף שמאל).

אם כן, בדיקות בעלות רגישות גבוהה שימושיות יותר על-מנת לשלול את התופעות אותן הן בודקות. לעומת זאת, בדיקות בעלות סגוליות גבוהה שימושיות יותר לאימות התופעות אותן הן בודקות. בדוגמה הקודמת, היות ובדיקת SLR נמצאה כבעלת מידת רגישות גבוהה ומידת סגוליות נמוכה (7), הרי שכאשר תוצאת הבדיקה שלילית נוכל להפיק ממנה יותר מידע אבחנתי (אין פריצת דיסק) מאשר כאשר תוצאת הבדיקה חיובית (לאו דווקא קיימת פריצת דיסק).

יחס סבירות (Likelihood Ratio):

מדד אבחנה נוסף הוא יחס הסבירות (Likelihood ratio). ה-LR (Likelihood ratio) משלב בין ערכי הרגישות והסגוליות על-מנת לתאר את מידת השינוי בסבירות שהאבחנה קיימת כאשר תוצאת הבדיקה חיובית או שלילית (8). ה-LR מבוסס בפורמט עשרוני ולא כאחוז.

Positive LR (+LR) מבטא את מידת השינוי בסבירות שהאבחנה קיימת כאשר תוצאת הבדיקה חיובית. ככל שערך זה גדול יותר כך גם גדלה מידת הסבירות שהאבחנה קיימת כאשר תוצאת הבדיקה חיובית. הנוסחה לחישוב ה-LR + היא: Sensitivity/(1-Specificity). אם נשתמש בנתוני הדוגמה הקודמת על בדיקת ה-SLR נראה כי ה-LR + של ה-SLR הנו: $0.9/(1-0.26) = 1.21$. משמעות נתון זה היא שבנבדקים בהם תוצאת הבדיקה חיובית (SLR חיובי), הסבירות לפריצת דיסק גדולה פי 1.21 מאשר הסבירות לדיסק נורמלי.

Negative LR (-LR) מבטא את מידת השינוי בסבירות שהתופעה הנחקרת קיימת כאשר תוצאת הבדיקה שלילית. ככל שערך זה קטן יותר כך גם תקטן הסבירות שהאבחנה קיימת כאשר תוצאת הבדיקה שלילית. הנוסחה לחישוב ה-LR - היא: (1-Sensitivity)/Specificity.

מנתוני הדוגמה הקודמת הרי שה-LR - של ה-SLR הנו: $(1-0.9)/0.26 = 0.35$. משמעות נתון זה היא שבנבדקים בהם תוצאת הבדיקה שלילית (SLR שלילי), הסבירות לפריצת דיסק היא כשליש ממידת הסבירות לדיסק תקין.

באופן עקרוני ערך LR + משמעותי יותר ככל שהנו גדול יותר, וערך LR - משמעותי יותר ככל שהנו קטן יותר. LR +/- של 1 לא ישנה את מידת הסבירות לקיום התופעה הנבדקת בכלל. כלי עזר לשימוש בערכי Positive/negative LR מפורט בטבלה מספר 5 (8):

טבלה 5: משמעות ערכי LR +/-

משמעות	- LR	+ LR
שינוי גדול ולעיתים קרובות מכריע במידת הסבירות	<.01	>10
שינוי בינוני במידת הסבירות	0.1-0.2	5-10
שינוי קטן אך לעיתים בעל משמעות במידת הסבירות	0.2-0.5	2-5
שינוי קטן ולעיתים נדירות משמעותי במידת הסבירות	0.5-1	1-2

כדי לקבוע את הסבירות לקיום האבחנה לאחר ביצוע הבדיקה (Post-test probability) בעזרת ה-LR נשתמש בנומוגרמה (4,8) (איור 1):

הסבירות שהסימפטומים בכל זאת נובעים מ-Sub-acromial impingement היא 8% בלבד. סבירות זו נמוכה ביותר ולכן תאפשר לנו להתרכז בעמוד השדרה הצווארי. בשלב זה, נרצה לאמת את האבחנה של Cervical radiculopathy. כפי שראינו קודם, בדיקות בעלות רמת סגוליות גבוהה, או לחילופין, +LR גבוה, עוזרות לאמת אבחנה כאשר תוצאתן חיובית. עפ"י Wainner et. al. בדיקת ההחזר הגידי של ה-Biceps היא בעלת ערך סגוליות של 0.95 ו-LR+ של 4.9 (12). נוכל להשתמש בנומוגרמה על-מנת לגלות את ה-Post-test probability ל-Cervical radiculopathy. במקרה זה ה-Post-test probability test תעמוד על כ-85%. כלומר, אם נבצע את בדיקת ההחזר הגידי של ה-Biceps ונקבל תוצאה חיובית (החזר ירוד/לא קיים), הסבירות ל-Cervical radiculopathy תעלה מ-50% Pre-test probability (test probability) ל-85%. זהו אחוז ודאות גבוה יחסית המאפשר לנו להתחיל לטפל במטופל שלנו בהנחה שהיא סובלת מ-Cervical radiculopathy.

הדוגמא הנ"ל ממחישה את הצורך להכיר את הערך האבחנתי של הבדיקות אותן אנו מבצעים באופן שגרתי. הערך האבחנתי של בדיקה מכתוב כאמור את האינדקס לביצוע שלה (שלילה או אימות אבחנה) ואת מידת השימושיות שלה (מידת השינוי בסבירות לקיום או שלילת האבחנה בהתחשב בתוצאת הבדיקה).

מהימנות

תכונה נוספת בעלת חשיבות בכל הנוגע לבדיקות אבחנתיות הנה מהימנות. על-מנת שנוכל לקבל החלטות קליניות על סמך תוצאות בדיקות אבחנה, על אותן הבדיקות להיות מהימנות (4). מהימנות היא המידה בה תוצאת בדיקה נשארת קבועה במדידות חוזרות כאשר לא התרחש שינוי בתופעה הנמדדת. לדוגמא: בדיקת SLR תראה תוצאה חיובית בבדיקות חוזרות במטופל הסובל מפריצת דיסק בעמוד שדרה מותני. ללא מידה מספקת של מהימנות, לא נדע אם תוצאת הבדיקה משקפת את המצב האמיתי או משקפת טעות הקשורה בביצוע הבדיקה (4). מהימנות מושפעת מגורמים שונים: גורמים הקשורים לכלי המדידה, גורמים הקשורים למבצע המדידה וגורמים הקשורים לתופעה הנמדדת. כל אחד מהגורמים הללו עלול להטות את תוצאות הבדיקה ולגרום לטעות מדידה (4).

בד"כ נמדדת מהימנות ביצוע הבדיקה ע"י אותו המודד בהזדמנויות שונות (intra-tester reliability) ומהימנות של ביצוע הבדיקה ע"י מודדים שונים (inter-tester reliability) (4).

את מידת המהימנות נהוג לציין ע"י מדדים שונים. בבדיקות שמשתמשות בסולם מדידה מסוג רווח או סולם יחס (לדוגמא מדידת טווח תנועה) נהוג להשתמש במדד קורלציה (Pearson או intraclass correlation coefficient). בבדיקות שנמדדות בסולם שמי (לדוגמא SLR חיובי או שלילי) נהוג להשתמש במדד Kappa המציין את מידת ההסכמה בין מדידות שונות מעבר למידת ההסכמה הצפויה מיד המקרה (4,8,9).

נהוג לדרג את מידת המהימנות, בין אם מבוטאת ע"י Kappa ובין אם ע"י Correlation coefficient עפ"י הסולם הבא (4): $0.5 < \text{Correlation coefficient}$ - מהימנות נמוכה, $0.5-0.75$ - מהימנות בינונית, $0.75 >$ - מהימנות גבוהה.

סיכום

הצורך היומיומי של הפיזיותרפיסט בקבלת החלטות בעלות אופי אבחנתי מחייב הבנה של תהליך האבחנה. השימוש בבדיקות אבחנתיות מצריך, מעבר למיומנות הפסיכומטריות בביצוען, גם הבנה של הערך האבחנתי של אותן הבדיקות. הערך האבחנתי של בדיקה מבוטא ע"י מידת הרגישות, סגוליות, positive/negative predictive value ו-likelihood ratio. לבדיקות שונות ערך אבחנתי שונה שיכתוב את האינדקס לביצוע הבדיקה ואת מידת השימושיות שלה.

היות וערכי מדדי האבחנה של ה-Crossed SLR שונים מאלו של ה-SLR ובפרט, הרגישות הנמוכה והסגוליות הגבוהה, הרי שבדיקה זו אינפורמטיבית יותר כאשר תוצאתה חיובית מאשר כאשר תוצאתה שלילית (ההפך מבדיקת ה-SLR).

ראינו אם כן, שלמרות ששתי הבדיקות (SLR ו-Crossed SLR) מתימרות לאבחן פריצת דיסק, הרי שבשל ההבדל בערכים האבחנתיים של שתי הבדיקות, גם המידע שנפיק מביצוען שונה. בדיקת SLR תעזור לנו יותר לשלול פריצת דיסק כאשר תוצאתה שלילית מאשר לאמת פריצת דיסק כאשר תוצאתה חיובית. לעומתה, בדיקת ה-Crossed SLR תעזור לנו יותר לאמת פריצת דיסק כאשר תוצאתה חיובית מאשר לשלול פריצת דיסק כאשר תוצאתה שלילית.

Positive/negative LR מאפשרים לנו להעריך את מידת השינוי בסבירות לקיום אבחנה כלשהי בהתחשב במידת הסבירות של אותה אבחנה עוד לפני ביצוע הבדיקה. זהו למעשה הצורך הקיים בסיטואציה הקלינית. בנוסף, היות ו-LR מורכבים מערכי רגישות וסגוליות, הרי שהם אינם תלויים בשכיחות התופעה הנבדקת (8) (להבדיל מ-Positive/negative predictive value). תכונות אלה הופכות את ה-LR למדד האבחנה השימושי ביותר.

על-מנת להבין טוב יותר את השפעת הערך האבחנתי של בדיקות שונות על מידת השימושיות שלהן נשתמש בדוגמא הבאה:

לפניך מטופלת בת 45 המתלוננת על כאב צוואר, שכמה זרוע ימין. לאחר סיום הראיון, את/ה מעריך כי שתי האפשרויות העיקריות למקור הסימפטומים הן:

1. Cervical radiculopathy (pre-test probability 50%)
2. Sub-acromial impingement (pre-test probability 25%)

כדי שתוכל/י להמשיך ולהתרכז בבדיקת עמוד השדרה הצווארי, עליך לשלול קודם לכן את ה-Sub-acromial impingement כמקור אפשרי לסימפטומים. אחת האפשרויות לשלילת הכתף כמקור לסימפטומים היא ביצוע בדיקה מהירה (Quick test) לכתף כמו תנועה אקטיבית של המטופל והערכת הסימפטומים בזמן הבדיקה. אם הסימפטומים עליהם מתלוננת המטופלת לא יופקו בביצוע ה-Quick test הרי שהדבר שולל את הכתף כמקור לסימפטומים. למרות שלבדיקה זו יש בהחלט היגיון, הרי שישנה בעיה מרכזית בביצוע כזה: האם לא יתכנו מקרים בהם אכן קיים Sub-acromial impingement, אך תנועה אקטיבית כלשהי, לא דווקא תפיק את הכאב? במונחים אבחנתיים, אם הבדיקה האבחנתית היא תנועה אקטיבית, תוצאה חיובית בבדיקה היא הפקת כאב ותוצאה שלילית היא חוסר הפקת כאב, האם לא תתכן תוצאה שלילית כאשר הבעיה אכן בכתף? במילים אחרות האם לא תתכן תוצאת False-negative? אם לא נדע מהו אחוז ה-False-negative לא נדע עד כמה נוכל להיות בטוחים לאחר קבלת תוצאה שלילית ב-Quick test, שהבעיה אכן איננה נובעת מ-Sub-acromial impingement.

לחילופין, אם נכיר בדיקה אבחנתית ל-Sub-acromial impingement שכאשר נבצע אותה ונקבל תוצאה שלילית נוכל לשלול במידת וודאות גדולה את הכתף כמקור לסימפטומים, נוכל להמשיך ולהתרכז בבדיקת הצוואר. כפי שראינו קודם, בדיקות בעלות רמת רגישות גבוהה, או לחילופין, LR - גבוה, יסייעו לנו לשלול אבחנה כלשהי כאשר תוצאתן שלילית.

בדיקת Hawkins ל-Sub-acromial impingement הנה בעלת מידת רגישות של 92% ו-LR של 0.32 (3) אם תחילה הערכנו את הסבירות ל-Sub-acromial impingement ב-25% ואנו יודעים ש: LR - הוא של 0.32, נוכל להשתמש בנומוגרמה על-מנת לגלות את ה-post-test probability ל-Sub-acromial impingement. במקרה זה ה-Post-test probability תעמוד על כ-8%. כלומר, לאחר ביצוע בדיקת Hawkins וקבלת תוצאה שלילית (לא הופק כאב),

מקורות

1. Bernstein J. Decision Analysis. *J Bone Joint Surg Am.* 1997; 79:1404-1414.
2. Boissonault W., Goodman C. Physical Therapist as Diagnosticians: Drawing the Line on Diagnosing Pathology. *J Orthop. Sports Phys. Ther.* 2006; 6:351-353.
3. Calis M., Akgun K., Birtane M., Karacan I., Calis H., Tuzun F. Diagnostic Values of Clinical Diagnostic Tests in Subacromial Impingement Syndrome. *Ann Rheum Dis.* 2000; 59: 44-47.
4. Cleland J. Orthopedic Clinical Examination: An Evidence-Based Approach for Physical Therapists (1st Ed). Carlstadt, NJ: *Icon Learning Systems* 2005.
5. Delitto A., Snyder-Mackler L. The Diagnostic Process: Examples in Orthopedic Physical Therapy. *Phys. Ther.* 1995; 75: 203-211.
6. Delitto A., Erhard R.E., Bowling R.W. A Treatment-based Classification Approach to Low Back Syndrome: Identifying and Staging Patients for Conservative Treatment. *Phys. Ther.* 1995; 75: 470-489.
7. Deville W, van der Windt D, Dzaferagic A, Bezemer P, Bouter L. The Test of Lasegue. Systematic Review of the Accuracy in Diagnosing Herniated Discs. *Spine* 2000; 25: 1140-1147.
8. Fritz JM, Wainner RS. Examining Diagnostic Tests: An Evidence-Based Perspective. *Phys. Ther.* 2001; 81: 1546-1564.
9. Mayer D. Essential Evidence-Based Medicine (1st Ed.). Cambridge, UK: *Cambridge University Press* 2004.
10. Riddle D.L., Stratford P.W. Interpreting Validity Indexes for Diagnostic Tests: An Illustration Using the Berg Balance Test. *Phys. Ther.* 1999; 79: 939-948.
11. Stratford P.W., Binkley J. A Review of the McMurray Test: Definition, Interpretation and Clinical Usefulness. *J Orthop Sports Phys. Ther.* 1995; 22:116-120.
12. Wainner R.S., Fritz J.M., Irrgang J.J., Boninger M.L., Delitto A., Allison S. Reliability and Diagnostic Accuracy of the Clinical Examination and Patient Self-Reported Measures for Cervical Radiculopathy. *Spine* 2003; 28: 52-62.