



## הדיוק של סקירת האוכלוסייה לוירוס הקורונה והערכת מספר הבדיקות הנדרש לאבחון מחלה אצל יחידים

שי לין<sup>1</sup>

### תקציר

השליטה במגפת הקורונה וגם היציאה מהסגר והיכולת להחזיר את הפעילות הכלכלית תלויים במידה רבה ביכולת לאבחן חולים על ידי מכלול בדיקות. אולם, דיוק הבדיקות איננו תמיד ידוע ולשגיאות באבחון יש משמעות אישיות וכלכליות רחבות. הנתונים הידועים הנם בדרך כלל הנתונים הטכניים של הבדיקות כלומר הרגישות והסגוליות של הבדיקות אשר יכולים רק לכמת אחוזי זיהוי נכונים באוכלוסייה מתוך החולים או הבריאים. כלומר, בדיקות הזיהוי מגלות נשאים/חולים מתוך קבוצה של נשאים/חולים שכבר אובחנו או מגלות בריאים מתוך קבוצה של אנשים שאובחנו כבריאים. בעזרת בדיקות אלו אפשר לחשב את העלויות של שימוש בבדיקות לגילוי חולים ולכן מדדי הזיהוי הטכניים חשובים להערכת מדיניות לאומית.

המאמר יראה שנתונים אלו אינם יכולים לאמוד את אחוז האבחנות הנכונות של יחידים או את אחוז הטעויות החיוביות והשליליות של האבחון. המאמר מפרט את האלגברה של חישובי דיוק האבחנות של יחידים, ומצביע על אחוז נמוך של אבחון נכון של נשאים/חולים כאשר הימצאות המחלה באוכלוסייה נמוכה וכשנדרש מספר גבוה של בדיקות כדי לאבחן חולה אחד. מאמר זה בעל חשיבות גדולה לתכנון היציאה מהסגר ולתכנון עתידי של ניהול התפרצות המגפה.

### מבוא

בזמן כתיבת הדברים (ספטמבר 2020), מגפת הקורונה הביאה לתמותה של כמעט מיליון אנשים בעולם וכ-1400 מתים במדינת ישראל, ולתחלואה של יותר מ-31 מיליון אנשים בעולם וכמעט לרבע מיליון מקרים בישראל. המשבר נמשך חודשים ארוכים שבהם האוכלוסייה בישראל הייתה בסגר כמעט מלא, ונגרם נזק כבד לכלכלה. יש האומרים שניתן היה לצמצם את הנזקים הללו אילו הייתה מוטלת חובת שימוש במסכות וריחוק חברתי מלכתחילה, בתנאי שהיה אפשר לאבחן נשאי המחלה/חולים על ידי בדיקות רבות ותכופות וכך למנוע הדבקה של אחרים. למשל, היתר לטיסות בתנאי שהיה אפשר לאבחן חולים על ידי בדיקת חום ובדיקות מעבדה. האם אבחון באמצעים אלו מדויק? כמו כל בדיקה רפואית (לקיחת היסטוריה רפואית, בדיקה פיזיקלית, בדיקה מעבדתית) הבדיקות יכולות להביא לתוצאה מוטעית. טעות שכזו, למשל קבלת תוצאות חיוביות מוטעות יכולות להביא לבידוד של אנשים שאינם נשאים או חולים ללא הצדקה. תוצאות שליליות מוטעות יכולות להביא להמשך ההדבקה על ידי נשאים/חולים שלא זוהו.

### מטרת המאמר: סקירת השיטות להערכת בדיקות באוכלוסייה לעומת בדיקות של אנשים יחידים

הנתונים הידועים לציבור על דיוק בדיקות רפואיות הנם בדרך כלל נתונים טכניים של הבדיקות: הרגישות (אחוז הזיהוי הנכון מתוך כלל הנשאים/חולים באוכלוסייה) והסגוליות (אחוז הזיהוי הנכון של בריאים

<sup>1</sup> פרופ' (אמריטוס) שי לין, בית הספר לבריאות הציבור, אוניברסיטת חיפה, [slinn@univ.haifa.ac.il](mailto:slinn@univ.haifa.ac.il)

באוכלוסייה). קל לקבל את הנתונים האלו כי אפשר למדד במדגם קטן של חולים ובריאים. יש לרגישות ולסגוליות חשיבות בהערכת היעילות של הבדיקות מנקודת מבט של עלויות ומדידת הצלחה של תוכנית לאומית לאתר נשאים או בריאים. ואולם, אלו אחוזי זיהוי באוכלוסיות אך לא מאפייני אבחון אצל יחידים. המאמר יראה שנתונים אלו אינם יכולים לתאר את אחוז האבחנות הנכונות של נשאים/חולים אצל יחידים ואת אחוז הבדיקות השליליות הנכונות אצל יחידים. המאמר יציג ויסביר את ההבדל בין הערכה נכונה לפי דגימה מתוצאות הבדיקות (כמה מאלו עם תשובה חיובית של הבדיקה הם נשאים/חולים ובדומה לגבי תשובה שלילית) לעומת הסתברויות באוכלוסיות כלומר כמה מאלו שכבר ידוע שהם נשאים/חולים נזהה (ובדומה לגבי בריאים). המאמר מסביר כיצד יהיה ניתן להשתמש בנתונים הטכניים (שאינם אינפורמטיביים לגבי יחידים) כדי להגיע למדדי האבחנה אצל יחידים.

### **סוגי בדיקות לזיהוי מחלת הקורונה באוכלוסייה וחסרונותיהן<sup>2</sup>**

במגפת הקורונה השתמשו בכמה טכניקות של מבדקים באוכלוסייה כדי לזהות מחלה חריפה:

1. בדיקת היסטוריה רפואית והצהרה אישית
  2. בדיקת חום
  3. בדיקות מעבדה על בסיס דגימה במטוש מדרכי הנשימה ל RNA של הנגיף (בשיטת PCR (Polymerase Chain Reaction
  4. בדיקות אנטיגן (אתור חלבוני המעטפת של הנגיף)
- לאיתור אנשים שחלו בעבר, משתמשים בבדיקת דם לזיהוי נוגדנים (בדיקה סרולוגית). נדון בהטיות ואי דיוקים של בדיקות כאלו.

#### **דיוק והטיות בקבלת הצהרה ובבדיקת חום**

במדינות רבות נקטו בבדיקות חום לזיהוי חולים בוירוס הקורונה. אדם הנמצא כסובל מחום גבוה, נחשב כחולה בקורונה והושם בבידוד. כך גם לגבי אנשים שהיה לגביהם חשד שהיו במגע עם אנשים אחרים שהיו חולים בוירוס הקורונה. האם הבדיקה של חום או סיפור של חשיפה אפשרית הנם בדיקה רפואית מדויקת? האם בדיקת חום או סיפור של מגע עם אדם חולה מספיקים לקביעת אמצעים קיצוניים של בידוד?

בדיקת החום הנה בדיקה רגישה למחלה כלשהיא (אך לא לנשאות או למחלת הקורונה לכשעצמה), שכן חולה עם מחלה זיהומית יהיה גם עם חום גבוה. מאידך הרגישות לנשאות לפני המחלה או כאשר המחלה ללא תסמינים) קרובה יותר ל- 0%, או לכל היותר כ- 10%. נזכיר שהמחלה מאד מדבקת גם בשלבי טרום מחלה וגם במצב ללא תסמינים. הבדיקה הנה לא סגולית (לא ספציפית) כי יש סיבות אחרות לחום ואולי במוצק הסגוליות הנה 50% (כמו הטלת מטבע). בהמשך המאמר נצביע על כך שבאוכלוסיות עם הימצאות של נשאות/מחלה של 10%, יהיה, בנתונים אלו, קשה מאד לאבחן מחלה אצל יחידים בעזרת בדיקת חום בלבד.

#### **בדיקת PCR במעבדה לזיהוי וירוס הקורונה**

בדיקת PCR לגילוי הדבקות במחלה כלומר זיהוי RNA של הוירוס אצל נשא/חולה נחשבת טכנית כבדיקה אמינה. תוצאות הבדיקה מאפשרים לאמוד את ההיארעות של המחלה, ה incidence, כלומר הבדיקה מגלה חולים במחלה חריפה. הבדיקה לא תראה הדבקות בעבר. ואולם, הדיוק של בדיקות PCR תלוי לא רק בטיב בטכניקה והמכשור (המביאים בדרך כלל לדיוק גבוה מאד) אלא גם בגורמים חיצוניים לבדיקה הטכנית כגון:

<sup>2</sup> Watson, Whiting & Brush, 2020.

- צורת לקיחת הדגימה
  - מועד לקיחת הדגימה: בתחילת המחלה? במהלכה? לאחר ההבראה? נמצא שבדיקות PCR מראות מחלה רק לאחר 5 ימים מעת ההדבקה
  - מיקום לקיחת הדגימה: מהאף? מהפה? מהלוע? משטיפת סימפונות?
  - צורת העברת הדגימה למעבדה
- בדיקות המטושים וה-PCR של הדגימה אם כן הן בעלות רגישות נמוכה יותר: לעיתים לא יגלו הנשאות או המחלה בשל בעיות טכניות. מאידך הבדיקות הן בעלות סגוליות גבוהה: אם נמצא שיש עדות להדבקות (תשובה חיובית) יש להניח שהאדם אמנם נשא/חולה ונדבק בנגיף.

### **בדיקת אנטיגן לזיהוי וירוס הקורונה**

הבדיקה מאתרת חלבוני המעטפת של הוירוס. הבדיקות חיוביות כבר בשלבים מוקדמים לאחר ההדבקות וכך מאפשרות איתור מצב של יכולת הדבקה infectiousness בעוד שה-PCR מאתר מצב שבו אדם נדבק infectiveness. לכן בדיקות אנטיגן יעילות יותר למניעת הדבקה של אנשים אחרים ולעצירת המגיפה. הבדיקות מהירות, זולות, יכולות להתבצע בבית אבל לבדיקות רגישות נמוכה יותר יחסית לבדיקות ה-PCR.

### **בדיקות לזיהוי מצב מחלה בעבר**

בדיקת דם סרולוגית לזיהוי נוגדנים כשבועיים שלושה לאחר ההידבקות מאפשרת לזהות אנשים שחלו בעבר כיוון שרוב אלו, מי שחלו וגם אלו אשר נדבקו ולא חלו, או שהיו ללא תסמינים, יפתחו נוגדנים. לכן הבדיקות הסרולוגיות חשובות כדי לחקור את התפשטות הוירוס באוכלוסייה. לא ניתן להסתמך על בדיקות סרולוגיות כדי לזהות אנשים חולים בזמן אמת כיוון שלוקח כשבועיים עד שמתפתחים הנוגדנים. יצוין שלמרות שהבדיקה מגלה נוגדנים, לא ברור כרגע אם הנוגדנים מגינים מפני המחלה. על כן יש להתייחס לכל המספרים בזהירות ובהסתייגות. הדברים מחכים למחקר מעמיק שיתאפשר בעתיד, במשך ולאחר המגפה.

### **מתודולוגיה של הערכת בדיקות הזיהוי האוכלוסיות לעומת יכולות האבחון של המחלה ביחידים<sup>3</sup>**

הבה נבחן המדדים של דיוק בדיקות רפואיות "פשוטות" וזולות לעומת הבדיקות המדויקות יותר המשמשות קריטריון הסטנדרט - Gold Standard אשר מתאר את המצב האמיתי ולפיו נבחנות יתר הבדיקות. המצב האמיתי נקבע לעיתים רק לאחר אשפוז החולה, על סמך התסמינים הקליניים ובדיקות מורכבות כמו בדיקות CT עם תמונה אופיינית למחלה. מאידך, חשוב להחליט אם צריך אשפוז ולתכנן הטיפול בשלבים מוקדמים של המחלה. ראוי לכן לאפיין הדיוק של ההיסטוריה הרפואית, בדיקות חום ובדיקות הדם ל PCR כדי להסתמך עליהן לפני הדרדרות המחלה.

### **בדיקות דיוק אבחון של יחידים**

נפרט להלן מדדי דיוק של הבדיקות שבהם משתמשים לאבחון חולים:

**ערך האבחון החיובי *PPV, Positive Predictive Value*** של הבדיקה הנה ההסתברות שיש אמנם מחלה כאשר יש בדיקה חיובית אצל נבדקים, המבוטאת כאחוז החולים שאובחנו נכונה והם אמנם חולים, מתוך כלל אלו שיש להם בדיקה חיובית. למשל - PPV של 70% מראה שאובחנו נכון רק 70% מכלל הנבדקים עם בדיקה חיובית וישנם 30% מכלל המאובחנים שאובחנו בטעות כחולים בגלל בדיקה חיובית מוטעית (False

<sup>3</sup> Grunau & Linn, 2018a; Grunau & Linn, 2018b; Knottnerus, Buntinx & Van Weel, 2002; Linn, 2004; Linn & Grunau, 2006; Riegelman, 2000; Riffenburgh, 1999; Rothman, Greenland & Lash, 2008; Sackett, Haynes, Guyatt & Tugwell, 1991; Watson, Whiting & Brush, 2020; Weinstein & Finberg, 1980.

Positive Fraction , FPF) שמוגדר כאחוז הטעות החיובית של הבדיקה. בשל סיבות היסטוריות נהוג לקרוא למדד זה בטעות שיעור RATE, למרות שהמדובר באחוז ואולי ימצאו בספרות ובדיווחים מונחים כמו שיעור הטעות החיובית False Positive Rate , FPR.

חישוב PPV ובדיקות אחרות שייסקרו להלן כמו NPV נעשה בדרך כלל במדגם של בדיקות חיוביות ושיליות, נניח 100 בדיקות חיוביות ו-100 בדיקות שליליות. באוכלוסייה כזו ההימצאות של המחלה הנה ההימצאות הנכונה שבאוכלוסייה, שכן המדגם הנו לפי תוצאות הבדיקה (חיובית או שלילית). בניגוד לכך, נראה להלן שהמאפיינים טכניים של הבדיקות, הרגישות והסגוליות, נמדדים במדגם מלאכותי של נשאים/חולים או בריאים שבו הימצאות המחלה הנה מלאכותית (אם דגמנו 100 חולים או 100 בריאים ההימצאות הנה בהכרח 50%).

**ערך האבחון השלילי NPV, Negative Predictive Value** של הבדיקה הנה ההסתברות של העדר מחלה אצל אלו שיש להם בדיקה שלילית (אין מחלה). אחוז אלו שהם אמנם אינם נשאים/חולים שאובחנו נכונה, מתוך כלל אלו עם בדיקה שלילית. למשל- NPV של 80% מראה שמכלל אלו עם בדיקה שלילית הבדיקה אבחנה נכון ו-80% הנם אמנם לא נשאים/חולים. בנתונים כאלו יש 20% מכלל אלו עם בדיקה שלילית שהם כן נשאים/חולים שלא אובחנו בגלל בדיקה שלילית מוטעית, הטעות השלילית של הבדיקה False Negative Fraction , FNF.

**מהי תוספת המידע לאבחון לאחר ביצוע הבדיקה?** תוספת המידע הינה האחוז שבו נאבחן נכון יותר את המחלה לעומת מה שידוע לנו (או שאנחנו משערים) ממילא, משיעור ההימצאות של המחלה באוכלוסייה כלומר PPV-prevalence ותוספת המידע מאבחון בריאים NPV-(1-Prevalence) או בסי"כ לעומת 100% מידע, שהם הסתברות של 1:

$$\text{Predictive Summary Index} = \text{PSI} = \Psi =$$

$$= [\text{PPV} - \text{Prevalence}] + [\text{NPV} - (1 - \text{prevalence})] = \text{PPV} + \text{NPV} - 1$$

או באחוזים

$$\text{PSI}\% = \text{PPV}\% + \text{NPV}\% - 100\%$$

PSI זהה לאחוז התשובות הנכונות של הבדיקה ללא שגיאות שליליות או חיוביות. מתוך 100% (הסתברות של 1) של הבדיקות צריך להפחית את הטעויות החיוביות והשליליות.

$$1 - \text{FPR} - \text{FNR} = 1 - (1 - \text{PPV}) - (1 - \text{NPV}) = \text{PPV} + \text{NPV} - 1$$

ה-PSI הנו מדד המציין את ההבדל בין אחוז האבחנות החיוביות הנכונות (PPV) לאחוז האבחנות השליליות השגויות (FNR)

$$\text{PPV} - \text{FNR} = \text{PPV} - (1 - \text{NPV}) = \text{PPV} + \text{NPV} - 1$$

בעזרת PSI נוכל לאמוד את המספר המינימלי של אנשים שיש לבדוק כדי לאבחן אדם אחד נשא/חולה או אדם אחד בריא ונחשב minimal number of needed to test-diagnosed, NND.

$$\text{NND} = 1 / \text{PSI}$$

לדוגמה, אם ה-PSI הנו 5% (0.05), הרי אחוז האבחנות החיוביות הנכונות (PPV) הנו ב- 5% יותר מאשר אבחנות שליליות שגויות (FNR). כדי לאבחן נכון אדם אחד שהנו נשא/חולה נצטרך לפחות 20 נבדקים 1/0.05. אם עלות כל בדיקה הנה 10 דולר, משתמע שנצטרך להשקיע בנתונים אלו 200 דולר כדי לזהות נכונה אדם אחד שאיננו חולה. באותה מידה, ה-PSI הנו מדד ההבדל בין אחוז אבחנות השליליות הנכונות (NPV) לאחוז האבחנות החיוביות השגויות (FPR) והחישוב דומה:

$$NPV - FPR = NPV - (1 - PPV) = PPV + NPV - 1$$

**טבלה 1** מתייחסת לדיוק תהליכי אבחון. הסימון  $P(S+|T+)$  מציין הסתברות למחלה  $S+$  sickness חיובית כאשר נתון  $T+$  (I) שהבדיקה = test חיובית. וכך לגבי סימונים דומים. הסימונים באותיות גדולות כדי להבדיל להלן מטבלה 2. שימו לב שאין משמעות לסיכום בטורים כי הדגימה הנה בשורות לפי תוצאות הבדיקה.

**טבלה מספר 1: דיוק תהליכי אבחון**

		מצב אמיתי		
		S+ נשא, חולה	S- לא נשא, לא חולה	
תוצאות הבדיקה	T+ חיובית	A	B	A+B
	T- שלילית	C	D	C+D

$$PPV = P(S+|T+) = \frac{A}{A+B}, \quad NPV = P(S-|T-) = \frac{D}{C+D}$$

$$\text{False Positive Rate (Fraction)} = FPR = \frac{B}{A+B} = 1 - PPV,$$

$$\text{False Negative Rate (Fraction)} = FNR = \frac{C}{C+D} = 1 - NPV$$

$$PSI = \Psi = PPV + NPV - 1$$

$$NND = \text{Minimal Number Needed to be tested to Diagnose correctly one person} = \frac{1}{PSI}$$

## בדיקות זיהוי - המדדים הטכניים של הבדיקה

נחזור כעת על הסברים דומים לגבי דגימה טכנית אשר כאמור איננה מאפשרת אבחנה אצל יחידים, אבל היא הזמינה בדרך כלל בתקשורת, ועלולה להטעות. בדיקות הזיהוי מגלות נשאים/חולים מתוך קבוצה של נשאים/חולים שכבר אובחנו או מגלות בריאים מתוך קבוצה של בריאים. בעזרת בדיקות אלו אפשר אולי לחשב העלויות או חישובי רווח/הפסד של שימוש בבדיקות לגילוי חולים ולכן מדדי הזיהוי הטכניים חשובים להערכת מדיניות ועלויותיה.

למשל, מבדק מעבדה שיש לו רגישות 90% וסגוליות 90% יזהה את מרבית הנשאים/החולים והבריאים, ויאפשר סגר אפקטיבי. אחוז הטעויות השליליות והחיוביות יהיה נמוך (10%). מבדק פחות רגיש (נניח רגישות 30%) כמו דגימת אף במטוש, יהיה שגוי ולא יזהה חולים בטעות באחוזים גבוהים, כלומר יהיו טעויות שליליות רבות. ה-FNR גבוה ויאפשר שחרור מהסגר בטעות ולכן הדבקה של רבים. מבדק פחות ספציפי (נניח סגוליות 30%), כמו בדיקת חום, יזהה בטעות חולים למרות שהם בריאים, כלומר יהיו טעויות חיוביות רבות, ה-FPR גבוה, כ-70% ויביא לסגר לא מוצדק של רבים.

מדדי הזיהוי הטכניים הנ"ל נקבעים בדרך כלל על ידי יצרני הבדיקות לאחר בדיקת התוצאות של הבדיקה אצל חולים (נניח 100 חולים) ובריאים (נניח 100 בריאים). היחס בין בריאים לחולים שונה בכל מחקר כזה, אבל תמיד היחס איננו משקף את הימצאות המחלה באוכלוסייה הכללית. למשל במחקר כלעיל הימצאות המחלה הנה 50% (100 מתוך 200 שנבדקו) בעוד שבאוכלוסייה הכללית ההימצאות יכולה להיות שונה (בישראל מעריכים ההימצאות של 10% נשאות ותחלואה). נראה להלן שלכך השפעה רחבה על דיוק הבדיקות. יודגש: מדדי הזיהוי אינן הבדיקות המעניינות אותנו לסקירת אוכלוסייה שנועדה לאבחן נשאים/חולים שעדיין איננו יודעים שהם נשאים/חולים. רק בדיקת PPV ו-NPV שהוזכרו לעיל יכולה לאבחן חולים ובריאים.

נפרט להלן מדדי דיוק של הבדיקות שבהם משתמשים לזיהוי חולים ובריאים:

**רגישות SENSITIVITY** של הבדיקה הנה ההסתברות של בדיקה חיובית (יש מחלה) לזיהוי הנשאים/חולים בקורונה באוכלוסייה, המבוטאת כאחוז הנשאים/חולים שזוהו נכונה והם אמנם נשאים/חולים, מתוך כלל הנשאים/חולים באוכלוסייה. למשל, רגישות של 90% מראה שהבדיקה זיהתה נכון 90% מכלל הנשאים/חולים. בהקשר למגפת הקורונה המשמעות הנה שרק 90% מהנשאים/חולים אמנם זוהו וטופלו (או הושמו בהסגר).

**סגוליות SPECIFICITY** של הבדיקה הנה ההסתברות של בדיקה שלילית (אין מחלה) לזיהוי אלו שאינם נשאים/חולים בקורונה באוכלוסייה, המבוטאת כאחוז שאינם נשאים/חולים שזוהו נכונה והם אמנם אינם נשאים/חולים, מתוך כלל אלו שאינם הנשאים/חולים באוכלוסייה. למשל, סגוליות של 95% מראה שהבדיקה זיהתה נכון 95% מכלל הלא נשאים/חולים. בהקשר למגפת הקורונה המשמעות הנה שרק 95% מהלא הנשאים/חולים אמנם זוהו נכון ושחררו מבידוד.

**תוספת המידע לזיהוי מביצוע הבדיקה נמדד על ידי YOUDEN INDEX (1950):** בכמה נזהה נכון יותר את החולים לעומת מה שידוע לנו (או שאנחנו משערים) מאחוז הבדיקות החיוביות P(T+) או השליליות P(T-). כלומר P(T+) - prevalence - sensitivity = specificity - P(T-).  
P(T+)=specificity - P(T-)

Youden Index = J = sensitivity + specificity - 1

J זהה לאחוז התשובות הנכונות של הזיהוי פחות הטעויות. היינו רוצים 100% כלומר הסתברות של 1 של תשובות נכונות, אבל למעשה צריך להפחית את הטעויות החיוביות והשליליות.

$$1-fpr-fnr=1-(1-specificity)-(1-sensitivity)=sensitivity+specificity-1$$

ה-J הנו מדד ההבדל בין אחוז זיהוי מחלה נכון (sensitivity) לאחוז הזיהוי החיובי השגוי (fpr)

$$sensitivity-fpr=sensitivity-(1-specificity) = sensitivity+specificity-1$$

בדומה לדיון הקודם על ה-NND נוכל ליישם את המושג גם למדד ההפרש J כדי לאמוד את המספר המינימלי של אנשים שיש לסקור בתוכנית גלוי מחלה כדי לגלות אדם אחד נשא/חולה או אדם אחד בריא, ונחשב minimal number of needed to screen, nns. לדוגמה, אם ה-J הנו 5% (0.05) הרי אחוז זיהוי חולים נכון, ה-sensitivity הנו ב-5% יותר מאשר אבחנות חיוביות שגויות, ה-FPR. כדי לזהות נכון אדם אחד שהנו נשא/חולה נצטרך לפחות 20 נבדקים  $20=1/0.05=nns$ .

אם עלות כל בדיקה הנה 10 דולר, משתמע שנצטרך להשקיע בנתונים אלו 200 דולר כדי לזהות נכונה אדם אחד שאיננו חולה. באותה מידה, ה-J הנו מדד ההבדל בין אחוז הזיהוי הנכון של בריאים (specificity) לאחוז האבחנות השליליות השגויות (FNR) והחישוב דומה:

$$specificity-fnr=specificity-(1-sensitivity) = sensitivity+specificity-1$$

**טבלה 2** מתייחסת למדדים טכניים של דיוק תהליכי זיהוי.

הסימון  $P(T+ | S+)$  מציין הסתברות לבדיקה חיובית כשנתון שיש מחלה  $S+ = sickness$  וכך לגבי סימונים דומים. הסימונים באותיות קטנות מטרם להבדיל מטבלה 2. שימו לב שאין משמעות לסיכום בשורות כי הדגימה הנה בטורים לפי מצב המחלה (בריא/חולה).

**טבלה מספר 2: מדדים טכניים של דיוק תהליכי זיהוי**

מצב אמיתי

		S+	S-
		נשא, חולה	לא נשא, לא חולה
תוצאות הבדיקה	T+ בדיקה חיובית	a	b
	T- בדיקה שלילית	c	D
		a + c	b+d

$$sensitivity = P(T+ | S+) = \frac{a}{a+c}, specificity = P(T- | S-) = \frac{d}{b+d}$$

$$false\ positive\ rate = fpr = \frac{b}{b+d} = 1 - specificity,$$

$$false\ negative\ rate = fnr = \frac{c}{a+c} = 1 - sensitivity$$

$$Youden\ Index = J = sensitivity + specificity - 1$$

$$NNS = minimal\ number\ of\ needed\ to\ be\ screened\ to\ detect\ correctly\ one\ person = \frac{1}{J}$$

### שימוש במדדים הטכניים לזיהוי כדי לחשב מדדי האבחון

קשה יחסית לחשב המדדים הנחוצים לסקירת אוכלוסייה (המדדים לאבחון) ה-PPV וה-NPV (והמדדים הנגזרים) באוכלוסייה האמיתית: אפשר למשל לעקוב אחרי כל הנבדקים ולראות אם חלו או לא: לעקוב אחר אלו עם תוצאה חיובית (המצביעה על הנשאות) לצורך חישוב ה-PPV ואפשר לעקוב אחרי כל אלו שבבדיקה שלהם לא היתה עדות לנשאות, כלומר כל אלו עם הבדיקה השלילית. אבל בדיקה כזו יקרה אם ההימצאות של המחלה נמוכה: אם רק 1% מהאוכלוסייה נשא/חולה, הרי נצטרך לעקוב ולבדוק 100 אנשים כדי לאבחן נשא/חולה אחד.

בדרך כלל נתוני האבחון (טבלה 1) לא זמינים כיוון שהמעקב יקר וגם משום שאנשים עם תוצאה שלילית אינם תמיד חוזרים למערכת הרפואית. מאידך נתוני הזיהוי הטכניים של הבדיקות מסופקות בדרך כלל על ידי היצרנים או מחושבות במחקרים מצומצמים וזולים של חולים ובריאים (שאבחנתם ידועה) כפי שהוסבר לעיל. משום כך נאלץ להשתמש במדדים טכניים של הזיהוי ומתוכם לחשב מדדי אבחון וסריקה שמעניינים אותנו.

### חישוב PPV ו-NPV מנתוני הרגישות והסגוליות

נניח לדוגמא שידוע (מהיצרן) שלבדיקת PCR יש רגישות של 30% וסגוליות של 90%. כאשר ההימצאות הנה 10%, היחס בין חולים ובריאים הנו 1:9. מתוך 1000 אנשים יהיו 100 חולים ו-900 בריאים. ההימצאות מכתיבה את היחס בין חולים ובריאים ואת כל המספרים והמדדים בטבלה 3.

אם הרגישות הנה 30% הרי רק 30 חולים מתוך 100 חולים יזוהו נכון, ו-70 חולים יזוהו כבריאים בטעות. בדומה, אם הסגוליות הנה 90% הרי מתוך 900 הבריאים 810 יזוהו נכון כבריאים אבל 90 בריאים יזוהו בטעות כחולים.

המספרים בטבלה, המוכתבים על ידי ההימצאות ויחס של 1:9 בין החולים לבריאים, תהיה כבטבלה 3. אותה בדיקה, עם 30% רגישות ו-90% סגוליות תאבחן נכון (ותשגה) במספר שונה לחלוטין של אנשים אם ההימצאות תהיה שונה כפי שייראה להלן:



טבלה מספר 3: מצב בו נתון שהרגישות 30% והסגוליות 90%. נתון שההימצאות הנה 10%:

מצב אמיתי

		S+ נשא, חולה	S- לא נשא, לא חולה	
תוצאות הבדיקה	T+ חיובית	30	90	120
	T- שלילית	70	810	880
		100	900	1000

$$PPV = 30 / 120 = 25\% , NPV = 810 / 880 = 92.1$$

$$FPF = 7.9\% , FNF = 75\%$$

$$PSI = \Psi = PPV + NPV - 1 = 17.1$$

$$NND = \text{Minimal Number Needed to be tested to Diagnose correctly one person} = \frac{1}{PSI} = 5.85 \sim 6$$

#### שימוש במשפט בייס

אפשרות נוספת לחשב את מדדי האבחון PPV ו-NPV היא בעזרת נתוני הזיהוי, הרגישות והסגוליות, אם ידוע (נתון ממקור כלשהוא) אחוז ההימצאות  $P(S+)$  של המחלה באוכלוסייה. החישוב בעזרת משפט בייס (פיתוח המשפט והחישובים בספרות המצורפת).

$$PPV = P(S+ | T+) = \frac{P(T+ | S+) P(S+)}{P(T+ | S+) P(S+) + P(T+ | S-) P(S-)} =$$

$$\frac{\text{sensitivity} * \text{prevalence}}{\text{sensitivity} * \text{prevalence} + (1 - \text{specificity})(1 - \text{prevalence})} =$$

$$= \frac{\text{sensitivity} * \text{prevalence}}{\text{sensitivity} * \text{prevalence} + f_{pf} * (1 - \text{prevalence})}$$

וכך גם לגבי ה-NPV נחשב בעזרת נתוני הזיהוי, הרגישות והסגוליות, אם נתון אחוז ההימצאות של העדר מחלה (ההימצאות של הבריאים)  $1 - \text{prevalence} = P(S-)$  של המחלה באוכלוסייה:

$$NPV = P(S- | T-) = \frac{P(T- | S-) P(S-)}{P(T- | S-) P(S-) + P(T- | S+) P(S+)} =$$

$$= \frac{specificity * (1 - prevalence)}{specificity * (1 - prevalence) + (1 - sensitivity) * prevalence} =$$

$$= \frac{specificity * (1 - prevalence)}{specificity * (1 - prevalence) + fnr * prevalence}$$

השפעת ההימצאות על דיוק בדיקה לצורך אבחון וטיפול של יחידים בהנחה של רגישות וסגוליות של 90% בבדיקה כזו מדד יודג J קבוע ועומד על 80%. לכן מספר מינימלי של נסקרים שנחוצים כדי לזהות נכון חולה או בריא הנו כ- 2 בני אדם  $nns = 1/0.8 = 1.6$ . אם נסקור בבדיקה טובה כזו לפחות 2 בני אדם נזהה נכון לפחות אדם אחד שכבר אובחן כחולה או אדם אחד שכבר אובחן כבריא. דיוק סריקה כדי לאבחן החולים ולטפל בהם מושפעת מההימצאות.

**טבלה מספר 4: אבחון כתלות בהמצאות המחלה בהנחה שהרגישות 90% והסגוליות 90%**

אחוז ההימצאות של נשאות/מחלה Prevalence	דוגמא	אחוז אבחון נכון של חולים PPV	טעות חיובית באבחון FPF	אחוז נכון של אבחון בריאים NPV	טעות שלילית באבחון FNF	תוספת המידע PSI	מספר בדיקות כדי לקבל אבחנה נכונה אחת NDD=1/PSI
1%	כך באיסלנד	8.3%	91.7%	99.9%	0.1%	8.2%	12.2
5%	כך בספרד	33.1%	67.9%	99.4%	0.6%	31.6%	3.2
10.0%	כך בישראל	50.0%	50.0%	98.8%	1.2%	48.8%	2.1
30.0%	ניו יורק אפריל 2020	79.4%	20.6%	95.5%	4.5%	74.9%	1.3
50.0%	לא ידוע, או כמעט חסינות עדר	90.0%	10.0%	90.0%	10.0%	80.0%	1.3
70.0%	חסינות עדר	95.5%	4.5%	79.4%	20.6%	74.9%	1.3
90.0%	שעורים בסוף המגיפה	98.8%	1.2%	50.0%	50.0%	48.8%	2.1

## סיכום

מאמר זה סקר מספר שיטות להערכת דיוק בדיקות לקורונה. התלות של דיוק האבחנות בהימצאות חשובה לבדיקות בכל השיטות השונות שפורטו לעיל, וגם לבדיקות רפואיות אחרות: ההימצאות של המחלה קובעת את דיוק האבחון של המחלה יחד עם המדדים הטכניים של הבדיקה עצמה. מדדי הזיהוי הטכניים של המחלה, הרגישות והסגוליות, מגלים נשאים/חולים מתוך קבוצה של נשאים/חולים **שכבר אובחנו** או מגלות בריאים מתוך קבוצה של בריאים. בעזרת בדיקות אלו אפשר לחשב את העלויות או חישובי רווח/הפסד של שימוש בבדיקות לגילוי נשאים/חולים מתוך החולים באוכלוסייה ולכן מדדי הזיהוי הטכניים חשובים להערכת מדיניות ועלויותיה. ואולם, מדדים טכניים טובים אינם מאפשרים בהכרח אבחנות נכונות. הראינו שדיוק האבחנות נמוך כאשר אחוז ההימצאות של המחלה נמוך, גם כאשר לבדיקה מאפיינים טכניים גבוהים של דיוק. אחוז ההימצאות בישראל נאמד כיום ב- 10%.

הממצאים מראים שיהיה קשה לסמוך על בדיקות חום או בדיקות ליחה בכניסה למפעלים וקניונים או לפני עליה למטוס. משום כך יהיה ראוי להקפיד על אמצעי ריחוק חברתי ומסכות עד למציאת שיטות אמינות לאבחון חולים או עד למציאת דרכים להדברת המגפה על ידי חיסון וטיפול מונע. סביר שיהיה צורך בבדיקות רבות כדי לאבחן אחוז גבוה של חולים ולו רק בשל המתמטיקה של חישוב אחוז האבחון הנכון (PPV ו-NPV לעיל). יתרה מכך, יהיה צורך לבדוק כל נבדק כמה פעמים: אבחון המחלה אפשרי רק לאחר 5 ימים ויותר מעת ההדבקה (Kucirka et al., 2020; Woloshin et al., 2020). משום כך צפוי שבדיקות רבות תהינה שליליות למרות שהאדם נדבק במחלה, ואף יכול להדביק את סביבתו בסבירות עולה במשך 5 ימים שבהם עדיין לא אובחן. אלו הן כאמור הבדיקות השליליות המוטעות.

ארגונים עסקיים, מוסדות ציבוריים וחברות תעופה המתכוונים לפתוח את שעריהם לציבור חייבים להיות מודעים לסכנות ולמגבלות הבדיקות בכניסה למתחמן. ראוי יהיה להקפיד בכל מוסד וארגון על אמצעים למניעת הדבקה כגון ריחוק פיזי וחובת לבישת מסכות מאיכות טובה. היה טוב אם מסכות, כפפות ומתקנים לחיטוי ידיים יהיו זמינים במתקנים ציבוריים מתוך הנחה שרבים מהחולים לא מאובחנים ויכולים להוות מקור הדבקה. תשומת לב לאמצעי זהירות יכולה למנוע הדרדרות ופגיעה בעסקים וארגונים. מאידך, הימנעות מאמצעי הזהירות עשויה להביא להטלת אחריות על ארגונים במקרה של הדבקות בתחומם. הנהלות ארגונים חייבות לתכנן את הסיכונים והתגובה לסכנות אלו, לפתח תהליכים מתאימים ואמצעים טכנולוגיים למניעת הדבקה ולהקדיש לכך כוח אדם ותקציבים.

כדי למנוע טעויות חיוביות (הגדרה של יחיד כחולה למרות שאיננו כזה) אפשר לבצע שתי בדיקות או יותר, ולקבוע אבחון של אנשים כחולים/נשאים רק אם שתי הבדיקות מראות שהנבדקים הנם נשאים או חולים. וגם ההפך: כדי למנוע טעויות שליליות (הגדרה של יחיד כבריא למרות שהיחיד נשא/חולה) אפשר לקבוע אבחנה של העדר נשאות/מחלה כחולה רק אם שתי הבדיקות מאבחנות כך. אם שתי הדרישות תופעלנה בו זמנית כדי להגיע לדיוק מרבי, יהיו נבדקים שלגביהם לא תהיה אבחנה חד משמעית ויידרשו לבדיקה שלישית או רביעית. ברור שהצורך בכמה בדיקות מקשה על קטיעת מעגל ההדבקה והדברת המגפה.

הצורך בבדיקות כדי לאתר נשאים/חולים ילווה את המאבק בהתפשטות המגפה ובוודאי שיימצאו בדיקות מדויקות ומהירות יותר בעתיד (Larremore et al., 2020). יתכן שבדיקות סרולוגיות יוכלו בעתיד לאבחן יחידים אשר פיתחו נוגדים לנגיף והנם מחוסנים מפני המחלה. כיום עדיין לא ידועים הנתונים הטכניים של הבדיקות הסרולוגיות, ועל כן קשה לצפות את היכולות של בדיקות אלו לאבחן יחידים כבעלי נוגדנים. לא ברור אם הנוגדנים אכן מקנים חסינות, ולכמה זמן יש חסינות לאחר מחלה לאלו שבדמם יש נוגדנים.

ההתקדמות המהירה במחקר ובפיתוח הבדיקות תשנה את האומדנים לאחוזי אבחון יחידים ואת הגישה לסקירת אוכלוסיות.

המתודולוגיה שתוארה במאמר כבר פורטה בחוזרים של המשנה למנכ"ל משרד הבריאות, וצריכה להוות בסיס לקבלת החלטות לגבי בדיקות הקורונה (גרוטו, 2020). בניהול ארגונים הנדרשים להסתמך על בדיקות קורונה, ראוי לשים לב לאפשרות שהבדיקות אינן מדויקות, שחלק מהתוצאות הן טעויות חיוביות או שליליות ולהיערך לבדיקות מרובות ומהירות יותר כדי לאפשר חזרה לשגרה (ברנע, וייס ושמר, 2020).

## רשימת מקורות

ברנע, ר., וייס, י., ושמר, י. (2020). לחיות עם משבר מתמשך-תובנות מנגיף הקורונה. *הרפואה- עיתון ההסתדרות הרפואית בישראל*, 9(159), 630-635.

גרוטו, א. (2020). המלצת גורמי הרפואה במשרד הבריאות בנוגע לבדיקות קורונה באנשים אסימפטומטיים. חוזר המשנה למנכ"ל 241294420.

Grunau, G., & Linn, S. (2018a). Detection and diagnostic overall accuracy measures of medical tests. *Rambam Maimonides medical journal*, 9(4), e0027.

Grunau, G., & Linn, S. (2018b). Commentary: sensitivity, specificity, and predictive values: foundations, pliabilitys, and pitfalls in research and practice. *Frontiers in public health*, 6, 256.

Knottnerus, J. A., Buntinx, F., & Van Weel, C., (2002). General introduction: evaluation of diagnostic procedures. In J.A. Knottnerus (Ed.), *The evidence base of clinical diagnosis* (pp. 1-18). London: BMJ Books.

Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., & Lessler, J. (2020). Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. *Annals of Internal Medicine*, M20-1495.

Larremore, D. B., Wilder, B., Lester, E., Shehata, S., Burke, J. M., Hay, J. A., ... & Parker, R. (2020). Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. *MedRxiv*.

Linn, S. (2004). A new conceptual approach to teaching the interpretation of clinical tests. *Journal of Statistics Education*, 12(3), 1-9.

Linn, S., & Grunau, P. D. (2006). New patient-oriented summary measure of net total gain in certainty for dichotomous diagnostic tests. *Epidemiologic Perspectives & Innovations*, 3(1), 11.

Riegelman, R. K. (2000). *Studying a study and testing a test: how to read the medical evidence* (4th ed.). Philadelphia: Lippincott Williams & Wilkins.

Riffenburgh, R.H. (1999). *Statistics in medicine*. San Diego, CA: Academic Press.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed). Lippincott Williams & Wilkins.

Sackett, D. L., Haynes, R. B., Guyatt, G. H., & Tugwell, P. (1991). *Helping patients follow the treatments you prescribe. Clinical Epidemiology: A Basic Science for Clinical Medicine* (2nd ed.). Boston, Mass: Little Brown & Co Inc.

Watson, J., Whiting, P.F., Brush, J.E. (2020). Interpreting a covid-19 test result. *British Medical Journal*, 369. doi: 10.1136/bmj.m1808

Weinstein, M.C., & Finberg H.V. (1980). *Clinical decision analysis*. Philadelphia: WB Saunders.

Woloshin, S., Patel, N., & Kesselheim, A. S. (2020). False Negative Tests for SARS-CoV-2 Infection Challenges and Implications. *New England Journal of Medicine*, 383(6), e38.

[חזרה לתוכן העניינים](#)

