

תוקף מערכת המיון ללימודי רפואה בחיזוי
ההצלחה בלימודי שנה א' באוניברסיטאות
בשנים תשס"ז (2006/07) – תש"ע (2009/10)

תמר קנת-כהן
אליוט טורוול
יונתן סער
כרמל אורן

דצמבר 2014



דוח מרכז 413
ISBN:978-965-502-188-2

תוכן העניינים

2	תקציר
3	מבוא
3	תוקף הניבוי של כלי המיון ללימודי רפואה – סקירת הספרות
8	מטרות המחקר
9	שיטה
9	מדגם
10	משתני המחקר
13	עיבוד הנתונים
15	תוצאות
20	סיכום ודיון
23	מקורות
26	נספח 1 – אמינות התיקון לקיצוץ תחום במצבי ברירה קיצוניים
27	נספח 2 – נוסחת התיקון לקיצוץ תחום
28	נספח 3 – סטטיסטיים נצפים

תקציר

מטרת המחקר הייתה לספק תמונה על תוקף מערכת המיון ללימודי רפואה בישראל בניבוי הישגים לימודיים. המחקר נעשה על נתוני 1,002 סטודנטים אשר למדו בחוג לרפואה במסלול שש-שנתי במחזורים תשס"ז (2006/07) עד תש"ע (2009/10) באוניברסיטה העברית, באוניברסיטת תל אביב ובטכניון ב-12 יחידות עיבוד (יחידת עיבוד הוגדרה כחוג במוסד לימודים ובמחזור).

במחקר זה נעשה לראשונה, במסגרת מחקרי תוקף הניבוי הנערכים במאלי"1, תיקון לקיצוץ תחום עבור תהליך ברירה דו-שלבי, שמאפיין את לימודי הרפואה במוסדות ובמחזורים שבמחקר. בתהליך זה מדורגים המועמדים בשלב הראשון על פי צירוף משוקלל של ממוצע הבררות וציון מכפ"ל² (להלן: סכס 1). מועמדים בעלי ציון סכס 1 שהוא גבוה מסף מסוים מוזמנים, בשלב השני, להיבחן במו"ר (מערכת מיון לרפואה) או במרק"ם (מערכת ראיונות קצרים מובנים). משתנה הברירה של השלב השני (להלן: סכס 2) מבוסס על צירוף משוקלל של סכס 1 עם הציון שהתקבל בשלב השני. במחקר זה הותאם התיקון לתהליך הברירה הייחודי של כל מוסד ומחזור. הממצאים מעידים שכלי המיון הם בעלי תוקף גבוה בניבוי ציון שנה א': תוקף ממוצע הבררות הוא 0.30, ותוקף הציון הכללי במכפ"ל הוא 0.60 בקירוב. שישה ציוני סכס 1 הוגדרו במחקר. ציונים אלה נבדלו זה מזה בציון הכללי במכפ"ל (הציון הרב-תחומי, הציון בדגש מילולי או הציון בדגש כמותי) ובמשקלות שניתנו לציון מכפ"ל ולממוצע הבררות. תוקף ציוני סכס 1 נע בין 0.60 ל-0.62. עשרים וארבעה ציוני סכס 2, המבוססים על צירוף של ציוני סכס 1 והציון הסופי במו"ר או במרק"ם, הוגדרו במחקר. ציונים אלה נבדלו זה מזה בגורמים שבהם נבדלו ציוני סכס 1, בחזאי הלא-קוגניטיבי שהם מתבססים עליו (מו"ר או מרק"ם) ובמשקלות שניתנו לסכס 1 ולחזאי זה. תוקף ציוני סכס 2 נע בין 0.56 ל-0.64.

במטרה להעריך את תוקף הניבוי של תהליך הברירה כולו, נעשתה הדמיה תחת הנחה שהמשתנים מתפלגים התפלגות רב-משתנית נורמלית, תוך שימוש בערכים שהתקבלו במחקר עבור המתאמים בין החזאים לקריטריון ובינם לבין עצמם, וביחסי ברירה דומים לאלו שקיימים בפועל. ממצאי ההדמיה הראו שציון שנה א' בקרב המתקבלים בעקבות ברירה חד-שלבית על פי סכס 1 (ציון מכפ"ל הרב-תחומי וממוצע הבררות במשקלות שווים) גבוה ב-1.07 סטיות תקן מממוצע ציון שנה א' בקרב המועמדים (לו היו כולם מתקבלים). הערך המקביל עבור ברירה דו-שלבית שבה משתמשים בסכס 1 בשלב הראשון ובמו"ר בשלב השני הוא 0.93, ועבור ברירה דו-שלבית שבה משתמשים בסכס 1 בשלב הראשון ובסכס 2, שהוגדר כצירוף של סכס 1 ומו"ר במשקלות שווים, בשלב השני הוא 1.11. ממצאים אלה מעידים שהכללת רכיב לא-קוגניטיבי בשלב השני של תהליך ברירה דו-שלבי בשקלול עם ציון דמוי סכס 1 אינה פוגעת בתוקף הניבוי ביחס לקריטריון סכולסטי – לעומת מצב שבו נעשית ברירה חד-שלבית על פי סכס 1 בלבד. שימוש ברכיב הלא-קוגניטיבי בשלב השני של תהליך הברירה במשקל של 100% מוריד רק במעט את תוקף הניבוי (כמשוער, לטובת מטרות אחרות).

לבסוף, נציין שיכולת ההשוואה של ממצאי המחקר הנוכחי עם ממצאי מחקרים קודמים על תוקף הניבוי של כלי המיון ללימודי רפואה היא מוגבלת, וזאת בשל הבדלים מתודולוגיים משמעותיים בין המחקרים, כפי שמפורט בדוח.

¹ מרכז ארצי לבחינות ולהערכה

² מבחן כניסה פסיכומטרי לאוניברסיטאות

מבוא

לימודי רפואה זוכים בהתמדה למעמד בלתי מעורער של אחד התחומים היוקרתיים והמבוקשים ביותר בישראל בקרב מועמדים ללימודים אקדמיים. מחקר זה יתמקד בתוקף הניבוי של מערכת המיון ללימודי רפואה כללית (במסלולים השש-שנתיים).

בשל האופי המיוחד שלהם, לימודי רפואה כרוכים בהשקעה כספית גדולה יחסית מצד האוניברסיטאות: היקף הקורסים הנלמדים הוא גדול בהשוואה לחוגי לימודים אחרים, ובמקביל לשיעורים התאורטיים הסטודנטים נדרשים להשתתף גם במעבדות וכן בביקורים ובעבודה מעשית בבתי חולים. ההתלמדות בבתי החולים מתבצעת בקבוצות קטנות המקבלות הדרכה צמודה מהרופא המנחה. לכן, לימודי רפואה מצריכים השקעה כספית גדולה שאינה מאפשרת לאוניברסיטאות להגדיל בצורה משמעותית את מספר המקומות בחוגים לרפואה. וכך, שלא כמו תחומי לימוד אחרים כגון משפטים או מנהל עסקים (הלשכה המרכזית לסטטיסטיקה, 2013), לא חל ב-30 השנים האחרונות שינוי דרמטי במספר הבוגרים של בתי הספר לרפואה בישראל (ורון, 2012).

השילוב של מועמדים רבים ומיעוט מקומות הפך את תנאי הקבלה לרפואה בישראל לקשים מאוד. כך, למשל, בשנת הלימודים תשע"ג (2012/13) היו 4,700 מועמדים ללימודי רפואה בארבע האוניברסיטאות המקיימות תכנית שש-שנתית של לימודי רפואה (האוניברסיטה העברית, אוניברסיטת תל אביב, אוניברסיטת בן-גוריון והטכניון) ורק 517 התקבלו ללימודים – יחס ברירה של 9:1 (ורון, 2012), שהוא באופן עקבי הנמוך מבין כל תחומי הלימוד. קבלה לרפואה מהווה, אם כן, אתגר משמעותי. מצב דומה קיים בכל מדינות העולם. להלן ייסקרו בקצרה תנאי הקבלה לרפואה בעולם ובישראל, וידווח תוקף הניבוי של כלי המיון, תוך התמקדות במרכיבים הקוגניטיביים שלהם.

תוקף הניבוי של כלי המיון ללימודי רפואה – סקירת הספרות

בעולם

בארה"ב ובקנדה מתקיימים כמעט כל לימודי הרפואה במתכונת ארבע-שנתית (graduate-entry programs – "המודל האמריקאי"), ומוגדרים כלימודים מתקדמים המיועדים לבעלי תואר ראשון. הקבלה ללימודים אלה מתבססת על צירוף של גורמים קוגניטיביים: הישגים במבחן כניסה סטנדרטי ללימודי רפואה וממוצע ציוני הלימודים לתואר הראשון, וגורמים לא-קוגניטיביים: הצהרה אישית, ריאיון, מכתבי המלצה ופעילויות מחוץ לתוכנית הלימודים. באירופה לימודי רפואה מתקיימים לרוב במתכונת שש-שנתית (undergraduate-entry programs – "המודל האירופאי"), המיועדת לבוגרי בית ספר תיכון. קריטריוני הקבלה משתנים בין מדינות (ואף בין מוסדות באותה מדינה). ברוב מדינות אירופה (כדוגמת בלגיה, פינלנד, יוון, איטליה, פולין, פורטוגל, רומניה וספרד) נעשה שימוש בלעדי בחזאים קוגניטיביים כגון ממוצע ציוני בית ספר תיכון וציון במבחן ארצי במקצועות ביולוגיה, כימיה, פיסיקה ומתמטיקה. בצרפת הקבלה ללימודי שנה א' פתוחה לכל בוגרי בית ספר תיכון והברירה מתבססת על מבחן המתקיים בסוף שנה א' (עם יחס ברירה של 9:1). גם צרפת שייכת, לפיכך, למדינות הנשענות על חזאים קוגניטיביים בלבד. מדינות אירופאיות אחרות, כדוגמת בריטניה, משתמשות בצירוף של חזאים קוגניטיביים ולא-קוגניטיביים. בגרמניה נהוג מודל המשלב חזאים קוגניטיביים ולא-קוגניטיביים

באופן הבא : 20% מהמקומות שמורים למועמדים בעלי ההישגים הגבוהים ביותר בבית הספר התיכון, 60% מהמקומות שמורים למועמדים שנבחרים על פי קריטריונים שקובע המוסד (שגם בהם ניתן בדרך כלל משקל מכריע להישגים בבית הספר התיכון) ו-20% מהמקומות ניתנים על פי זמן ההמתנה מסיום לימודי התיכון. מודל נוסף של צירוף של חזאים קוגניטיביים ולא-קוגניטיביים מתקיים בצ'כיה, הונגריה ובולגריה. במדינות אלה נעשה שימוש בחזאים קוגניטיביים בקבלה של אזרחי המדינות הללו ובצירוף של חזאים קוגניטיביים ולא-קוגניטיביים בקבלה של סטודנטים זרים (Martinho, 2012; "Medical School", 2013).

באוסטרליה מתקיימים לימודי רפואה לרוב במתכונת ארבע-שנתית. הקבלה ללימודים מתבססת על צירוף של גורמים קוגניטיביים (הישגים במבחן כניסה ובלמודים לתואר ראשון) וגורמים לא-קוגניטיביים.

אשר לתוקף הניבוי של כלי המיון, מידע רב נאסף בהתייחס לכלי המיון ללימודי הרפואה בצפון אמריקה : מבחן הכניסה ללימודי רפואה (MCAT – Medical College Admission Test) וממוצע ציוני הלימודים לתואר הראשון. ה-MCAT מורכב מארבעה חלקים : פיזיקה, ביולוגיה, חשיבה מילולית וכתיבה. שלושת החלקים הראשונים נערכים במתכונת של שאלות ברירה, וחלק הכתיבה כולל שתי מטלות כתיבה. ה-MCAT בודק, אם כן, לא רק יכולת חשיבה, אלא גם ידע ספציפי בנושאי יסוד שחיוניים ללימודי הרפואה, ונלמדים במסגרת קורסים בתואר הראשון.

במחקר שעקב אחר שני מחזורי לימוד ב-14 בתי ספר לרפואה (Julian, 2005) נמצא שתוקף

ציון MCAT בניבוי ממוצע הציונים בשנתיים הראשונות (הקדם-קליניות) הוא 0.59^3 ובניבוי ממוצע הציונים בשנה השלישית (הקלינית) 0.46 . הערכים המקבילים עבור כלי המיון השני – ממוצע ציוני הלימודים לתואר הראשון – היו 0.54 ו- 0.36 עבור השנים הקדם-קליניות והשנה הקלינית, בהתאמה. עוד נמצא שתוקף הניבוי של צירוף שני כלי המיון הוא 0.71 ו- 0.54 עבור השנים הקדם-קליניות והשנה הקלינית, בהתאמה. ניתן לסכם, אם כן, על פי הקריטריונים שהציע כהן (Cohen, 1988), שתוקף שני החזאים בניבוי הציונים בשנים הקדם-קליניות הוא גבוה, ובניבוי הציונים בשנים הקליניות בינוני, כאשר בשני המקרים התוקף של MCAT גבוה יותר. התוקף של צירוף שני החזאים עולה על תוקף הניבוי של כל אחד מהם בנפרד, כאשר התרומה השולית של MCAT גדולה מהתרומה השולית של ממוצע ציוני הלימודים לתואר הראשון. מידע אודות תוקף הניבוי של מרכיבי MCAT התקבל בניתוח-על של ממצאי 11 מחקרים (Donnon, Paolucci, & Violato, 2007). בהתאמה עם מה שתואר קודם ביחס לציון הכללי, תוקף המרכיבים היה גבוה יותר בניבוי הציונים בלימודים הקדם-קליניים מתוקפם בניבוי הציונים בלימודים הקליניים⁴. השוואה בין המרכיבים מראה שלביולוגיה יש יתרון בולט בניבוי הציונים בלימודים הקדם-קליניים (מתאם של 0.40) ולחשיבה מילולית יש יתרון מסוים בניבוי הציונים בלימודים הקליניים (מתאם של 0.18). תוקף מבחן הכתיבה היה נמוך ביחס לשני הקריטריונים.

התוקף של MCAT ושל ממוצע ציוני הלימודים לתואר ראשון נבדק גם בהתייחס לחלקים

שונים של בחינת הרישוי לעסוק ברפואה בארה"ב (Donnon et al., 2007; Julian, 2005).

³ המתאמים המדווחים להלן בפרק זה מתוקנים לקיצוץ תחום.

⁴ יצוין, עם זאת, שממצאי הניתוח-על בנוגע לתוקף הציון הכללי – 0.43 ו- 0.39 בשנים הקדם-קליניות והקליניות, בהתאמה – היו נמוכים ממה שדווח לעיל (Julian, 2005).

הממצאים שנאספו מעידים שתוקף MCAT בניבוי הציונים בבחינת הרישוי עולה על תוקפו בניבוי הציונים בלימודים. תוקף ממוצע הציונים לתואר ראשון דומה ביחס לשני הקריטריונים. עוד נמצא, בהתייחס לקריטריון של הציונים בבחינת הרישוי, ש-MCAT לבדו מסביר את מרבית השונות המוסברת ע"י שני החזאים ביחד, וממוצע הציונים לתואר ראשון מוסיף מעט מאוד. קריטריון נוסף בבדיקת תוקף הניבוי של MCAT ושל ממוצע ציוני הלימודים לתואר ראשון הוא התקדמות ללא עיכובים לקראת סיום התואר (Dunleavy, Kroopnick, Dowd, Searcy, & Zhao, 2013). במחקר שכלל 119 בתי ספר לרפואה נמצא ש-MCAT היה חזאי טוב יותר מממוצע ציוני הלימודים לתואר ראשון בניבוי ההתקדמות לקראת סיום, והצירוף של שניהם היה החזאי הטוב ביותר בניבוי קריטריון זה. התמונה המסתמנת היא, אם כן, שבכל סוגי הקריטריונים שנבדקו נמצא ל-MCAT יתרון במונחי תוקף ניבוי על פני ממוצע ציוני הלימודים לתואר ראשון, ותוקף הצירוף של שניהם היה גבוה מתוקף כל אחד מהם בנפרד. מבחני כניסה ללימודי רפואה מחוץ לארה"ב הם חדשים יחסית (מבחן MCAT נהוג מאז 1928). מבחן אחד המשמש במיון מועמדים ללימודי רפואה במתכונת ארבע-שנתית באוסטרליה (החל מ-1996) ובבריטניה (החל מ-2000) הוא ה-GAMSAT (Graduate Australian Medical Schools Admission Test). ה-GAMSAT מורכב משלושה חלקים: חשיבה במדעי הרוח והחברה, תקשורת כתובה, וחשיבה במדעי הטבע. מבחן ייעודי נוסף – UMAT (Undergraduate Medical Schools Admission Test) – מיועד למיון מועמדים ללימודי רפואה במתכונת שש-שנתית באוסטרליה (נכנס לשימוש בסוף שנות התשעים של המאה העשרים). מבחן זה בודק חשיבה לוגית, הבנת בני אדם וחשיבה לא-מילולית. בבריטניה מקובל, במיון מועמדים ללימודים במתכונת השש-שנתית, להישען על הישגים בלימודי בית ספר תיכון (A level) כמרכיב הקוגניטיבי של מערכת המיון. עם זאת, עבור מוסדות מבוקשים במיוחד ללימודי רפואה, לא מגלים המבחנים הארציים שונות מספקת בקרב מועמדים בעלי יכולת גבוהה בשל אפקט תקרה, ונעשה שימוש במבחנים ייעודיים: UKCAT (UK Clinical Aptitude Test) שבודק חשיבה מילולית, חשיבה כמותית, חשיבה מופשטת וניתוח החלטות (נכנס לשימוש ב-2006) ו-BMAT (BioMedical Admissions Test) שבודק מיומנויות קוגניטיביות כלליות (פתרון בעיות, הבנת טיעונים, ניתוח נתונים, ויכולות הסקה), ידע מדעי (ביולוגיה, כימיה, פיסיקה ומתמטיקה) וכתובה (נכנס לשימוש ב-2003); החליף את קודמו – MVAT (Medical and Veterinary Admissions Test) שנכנס לשימוש ב-2000. המידע בנוגע לתוקף הניבוי של המדדים השונים המשמשים במיון ללימודי רפואה מחוץ לארה"ב אינו נרחב (Prideaux et al., 2011), ומתקבל ברובו ממחקרים שנערכו במדגמים קטנים ובמוסדות יחידים. לאחרונה דווחו ממצאים ממחקר רב-מוסדי (Edwards, Friedman, & Pearce, 2013) שבדק את תוקף הניבוי של כלי המיון בשלוש אוניברסיטאות באוסטרליה בהתבסס על נתוני 650 סטודנטים לרפואה במתכונת שש-שנתית בשנים 2006 ו-2007, המהווים כ-25% מסך הסטודנטים שהתחילו ללמוד רפואה באוסטרליה במחזוריים הנדונים. במחקר נמצאו מתאמים חיוביים בין החזאים – UMAT וההישגים בלימודים בביה"ס התיכון –

ובין ההישגים בשנים הקדם-קליניות (שנה ראשונה עד רביעית). עם זאת, נמצאו הבדלים בין מוסדות ובין שנות לימוד בגובה המתאמים וביתרון היחסי של שני החזאים הללו. הממצא המשותף לכל המסגרות היה שתוקף שני החזאים ביחד היה גבוה מתוקף כל אחד מן המרכיבים בנפרד.

אשר למרכיבים הלא-קוגניטיביים המשמשים במיון, החזאי היחיד שעבורו התקבלו אינדיקציות על תוקף ניבוי משמעותי, ובפרט על תרומה שולית לתוקף, מעבר לכלים הקוגניטיביים, הוא ה-MMI (Multiple Mini Interview), שפותח באוניברסיטת מקמסטר שבקנדה, ומורכב מסדרה של ראיונות קצרים מובנים. תוקפו של חזאי זה נבדק בסדרה של עבודות שנערכו באוניברסיטת מקמסטר. שלושת המחקרים הראשונים בדקו את הקשר שבין MMI וקריטריונים בשלבים הולכים ומתקדמים של תהליך ההכשרה: הישגים בלימודים הקדם-קליניים (Eva, Reiter, Rosenfeld, & Norman, 2004), הישגים בלימודים הקליניים ובחלק הראשון של בחינת הרישוי הארצית שהוא בעל דגש קוגניטיבי (Reiter, Eva, Rosenfeld, & Norman, 2007), והישגים בחלק השני של בחינת הרישוי הארצית (Eva et al., 2009) שיש בו דגש על ביצוע קליני ומיומנויות תקשורת. שלושת המחקרים הללו, שהתבססו על מדגמים קטנים, הראו באופן עקיף קשר בין הביצוע ב-MMI למרכיבים קליניים של הערכות הביצוע. מחקר נוסף שהתבסס על מדגמים גדולים יותר (Eva et al., 2012), מצא הבדלים בהישגים במבחן הרישוי הארצי בין מועמדים לאוניברסיטת מקמסטר (שבה תהליך הקבלה מבוסס, בין היתר, על MMI) שהתקבלו, ובין מועמדים שנדחו ולמדו במוסד אחר (שלא עושה שימוש ב-MMI). בפרט נמצא, בדומה למחקרים הקודמים, ש-MMI קשור יותר למדדי תוצאה בעלי אוריינטציה קלינית/אתית/בין-אישית ואילו ממוצע ציוני הלימודים לתואר ראשון קשור יותר למדדי תוצאה בעלי אוריינטציה של ידע. לאחרונה נערך מחקר בבריטניה (Husbands & Dowell, 2013), שסיפק אישור לממצאים הקודמים, שהביצוע ב-MMI קשור בעיקר להערכות קליניות, אך גם להערכות הטעונות בידע. יצוין ששיעור לא מבוטל מן הקריטריונים הקליניים נמדדים על ידי OSCE (Objective Structured Clinical Examination), שהיה הבסיס הרעיוני לפיתוח ה-MMI, מה שעשוי להגדיל את המתאם בין MMI לקריטריונים הקליניים בשל השיטה המשותפת. מענה מסוים לאפשרות זו מצוי במחקר שנערך בישראל (חדד, 2012), שבו נעשה שימוש בשיטה חלופית להערכת הביצוע – שאלון הערכת עמיתים. מחקר זה שנערך בקרב 99 סטודנטים לרפואה באוניברסיטה העברית בדק את התוקף של מו"ר ומרק"ם, המבוססים על ה-MMI, בניבוי הערכת עמיתים בסיום השנה הרביעית ללימודים. המתאמים (מתוקנים לקיצוץ תחום) שהתקבלו (0.39 ו-0.37 עבור מרק"ם ומו"ר, בהתאמה) סיפקו אף הם עדות תומכת לתוקף הניבוי של ה-MMI.

אשר לחזאים נוספים – ראיונות אישיים במתכונת המסורתית, מכתבי המלצה, הצהרות אישיות, מבחני אישיות או מדדים של אינטליגנציה רגשית – לא נמצאה עדות תומכת לתוקף הניבוי שלהם (Prideaux et al., 2011; Siu & Reiter, 2009). כן נמצאה עדות מסוימת לתוקף של מבחני שיפוט מצבי (מצולמים בוידאו) בניבוי קריטריונים בעלי אוריינטציה בין-אישית (Lievens & Sackett, 2006).

בישראל

לימודי רפואה בישראל נמשכים שש שנים (ועוד שנת סטאז') במסלול המלא ("המודל האירופאי"), או ארבע שנים במסלול לבעלי תואר ראשון⁵ ("המודל האמריקאי"). קבלת מועמדים למסלול השש-שנתי, שבו, כאמור, יתמקד מחקר זה, מבוססת על שילוב של חזאים קוגניטיביים ולא-קוגניטיביים. החזאים הקוגניטיביים הם ממוצע הבגרות, ציון מכפ"ל, וציון סכס ("סכס 1") המחושב כשקלול של שני החזאים הללו. החזאים הלא-קוגניטיביים הם מו"ר (נכנס לשימוש באוניברסיטת תל אביב החל משנה"ל תשס"ה ובטכניון החל משנה"ל תשס"ז), מרק"ם (נכנס לשימוש באוניברסיטה העברית החל משנה"ל תשס"ז), או ראיונות אישיים (באוניברסיטת בן-גוריון). מו"ר ומרק"ם מבוססים, כאמור, על ה-MMI הקנדי. הם כוללים שלושה מרכיבים: שאלון שיפוט והחלטה, שאלון ביוגרפי ותחנות התנהגותיות. שני המרכיבים הראשונים זהים בין מו"ר ומרק"ם, בעוד שהמרכיב השלישי הוא שונה: במו"ר כוללות התחנות סימולציה בשילוב שחקנים, ריאיון אישי ופעילות קבוצתית, ואילו מרק"ם מבוסס על ראיונות בלבד (התיאור הנ"ל מתייחס למחזורי המחקר. תיאור מפורט של המערכות ניתן למצוא אצל: Gafni, Moshinsky, Eisenberg, Zeigler, & Ziv, 2012; Ziv et al., 2008).

שילוב החזאים הקוגניטיביים והלא-קוגניטיביים נעשה באופן הבא: בשלב הראשון מדורגים המועמדים על פי חזאים קוגניטיביים בלבד ("סכס 1"). מועמדים בעלי ציוני סכס 1 הגבוהים ביותר מוזמנים, בשלב השני, לבחינת מו"ר או מרק"ם או לראיונות קבלה, המהווים את החזאים הלא-קוגניטיביים של תהליך המיון. משתנה הברירה של השלב השני ("סכס 2") מבוסס על שקלול של סכס 1 עם הציון שהתקבל בשלב השני. הן השקלול של מרכיבי סכס 1 (ממוצע הבגרות וציון מכפ"ל) והן השקלול של מרכיבי סכס 2 (ציון סכס 1 וציון במו"ר, במרק"ם או בראיונות הקבלה) משתנה בין מוסדות, ובתוך מוסד בין שנים. במחקרי התוקף שנעשו במאלי"ו עד כה הוצגו נתונים (על חזאים קוגניטיביים בלבד) לפי פקולטות או מקבצים של חוגים. במחקרים אלה נכלל החוג לרפואה כללית במקבץ החוגים "רפואה", יחד עם החוגים רפואת שיניים ורוקחות. בלוח 1 מוצגים מתאמים (מתוקנים לקיצוץ תחום, בהתבסס על הנחה שהברירה נעשתה על פי ציון סכס, המבוסס על משקלות שווים לממוצע הבגרות ולציון מכפ"ל) בין הציונים בכלי המיון ובין ממוצע ציוני שנה א' במקבץ החוגים "רפואה" (הגדרות המשתנים – ראו להלן בפרק השיטה) במחקרי התוקף האחרונים של מאלי"ו.

לוח 1: מתאמים בין הציונים בכלי המיון ובין ציון שנה א' במקבץ החוגים "רפואה" (רפואה כללית, רפואת שיניים ורוקחות) במחקרי תוקף של מאלי"ו

ציונים בתחומי מכפ"ל			מכפ"ל	בגרות	ציון הסכס	מספר		מחזורים
אנגלית	כמותי	מילולי				סטודנטים	חוגים ^a	
0.24	0.35	0.26	0.38	0.49	0.52	4,613	71	^b 1990/91-2002/03
0.36	0.40	0.34	0.44	0.35	0.47	1,593	23	^c 2002/03-2004/05
0.24	0.31	0.24	0.33	0.30	0.35	3,883	57	^d 2005/06-2009/10

^a חוג מוגדר חוג מסוים במוסד לימודים מסוים ובמחזור מסוים.

^b מקור: ברונר, 2006

^c מקור: אורן, קנת-כהן וברונר, 2007

^d מקור: Oren, Kennet-Cohen, Turvall, & Allalouf, 2014

⁵ תכנית ארבע-שנתית קיימת באוניברסיטאות תל-אביב ובר-אילן.

הנתונים מעידים על הבדלים לא מבוטלים במקדמי התוקף בין שלושת המחקרים: תוקף ממוצע ציוני הברגרות גבוה במחקר הראשון ונמוך יחסית במחקרים השני והשלישי, ואילו תוקף מכפ"ל ומרכיבו גבוה במחקר השני ונמוך יחסית במחקרים הראשון והשלישי. עם זאת, שני ממצאים מרכזיים מתקיימים בשלושת המחקרים. ראשית, תוקף הניבוי של ציון הסכם גבוה מתוקף הניבוי של ממוצע ציוני הברגרות וציון מכפ"ל בנפרד, דהיינו: ציון מכפ"ל תורם לתוקף הניבוי של מערכת המיון מעבר לממוצע ציוני הברגרות ולהיפך. שנית, בהתייחס לתקפויות היחסיות של הציונים בתחומי מכפ"ל, מוצאים יתרון לציון בתחום החשיבה הכמותית על פני הציונים בשני התחומים האחרים, שהם דומים למדי זה לזה. בהתאם, במחקר השלישי, שבו נבדקו גם שני ציונים כלליים חדשים במכפ"ל – ציון מכפ"ל בדגש מילולי וציון מכפ"ל בדגש כמותי, מדווח שתוקף ציון מכפ"ל בדגש כמותי (0.35) עולה על תוקף ציון מכפ"ל בדגש מילולי (0.30) וכן על תוקף הציון הרב-תחומי במכפ"ל (0.33) (השניים הראשונים אינם מוצגים בלוח 1).

מטרות המחקר

המחקר שלפנינו מתמקד בבדיקת תוקף הניבוי של כלי המיון בקרב סטודנטים אשר למדו בשנה א' בבתי הספר לרפואה/במסלול לרפואה (יכוננו להלן: החוגים לרפואה) במסלולים השש-שנתיים במחוזות תשס"ז (2006/07) עד תש"ע (2009/10) באוניברסיטה העברית, אוניברסיטת תל אביב והטכניון. בשלושת המוסדות האלה נכללו במערכת הברירה במחוזות הנדונים מבדקי מו"ר או מרק"ם. מאחר שציוניהם של המועמדים לרפואה במו"ר ובמרק"ם ידועים לנו, ניתן לתקן לקיצוץ תחום את ערכי התוקף בהתבסס על תהליך ברירה שכולל את המרכיבים הללו. אשר לאוניברסיטת בן-גוריון, אין לנו מידע מלא על תהליך הברירה, ובפרט לא על הציונים בראיונות האישיים, שמשקלם הוא 100% בציון הקבלה בשלב השני. מאחר שבמחקר זה שאפנו להתקרב ככל האפשר למידול מציאותי של תהליך הברירה (ראו להלן), בחרנו שלא לכלול את נתוני אוניברסיטת בן-גוריון במחקר.

נחיצותו של התיקון לקיצוץ תחום מוכרת היטב בספרות המקצועית (למשל, AERA, APA,

NCME, 1999): סוגיית יכולת החיזוי נוגעת למועמדים ללימודים ולא ללומדים בפועל. אלא שכדי לאמוד את יכולת החיזוי הזאת, נאלצים בדרך כלל להתבסס על מדגמים של מועמדים שהתקבלו, שרק עבורם ידועים הביצועים בקריטריון, ולהקיש מכך על אוכלוסיית המועמדים כולה. מדגמים אלה מושפעים מברירה מוקדמת, שמקטינה את הטרוגניות המדגם, ולפיכך את מקדם התוקף הנצפה בו לעומת התוקף בקרב המועמדים. נוסחאות לתיקון לקיצוץ תחום אומדות את מקדם התוקף שהיה מתקבל לולא חל צמצום בטווח הציונים. העבודות הרבות שנעשו בהתייחס לנוסחאות אלה (ראו סקירה אצל Sackett & Yang, 2000) מעידות שהמתאמים המתוקנים מדויקים יותר מהמתאמים הלא מתוקנים, ובאופן כללי מספקים אומדן שמרני (הערכת חסר) למתאם באוכלוסייה (Bobko, Roth, & Bobko, 2001; Linn, 1983). כאשר מדובר במסגרת (תכנית לימודים, מקום עבודה, מסלול בשירות צבאי, וכיו"ב) סלקטיבית במיוחד, מובעת לעיתים טענה שיחס הברירה הקיצוני מלווה בקיצוץ תחום משמעותי כל כך שלא ניתן לשחזר מן המתאם הנצפה בקרב הלומדים את המתאם בקרב המועמדים. במענה לטענה זו ערכנו בדיקה מקדימה על איכות התיקון במדגמים עם רמת סלקטיביות דומה לזו שקיימת בלימודי הרפואה.

תוצאות הבדיקה (ראו נספח 1) מעידות שגם במצבי ברירה קיצוניים מספק התיקון לקיצוץ תחום אומדן קרוב בערכו למתאם טרם ברירה.

- חידושי המחקר, הן בהתייחס לתיקון לקיצוץ תחום והן בנושאים נוספים, הם:
1. במחזורי המחקר נכללו במערכת הברירה מבדקי מו"ר ומרק"ם. בהתאם, נעשה במחקר זה תיקון לקיצוץ תחום של ערכי התוקף בהתבסס על תהליך ברירה שכולל את המרכיבים הללו. בפרט, התיקון לקיצוץ תחום הותאם לברירה דו-שלבית המכונה גם (Sackett & Yang, 2000) ברירה סדרתית (sequential selection) להבדיל מברירה בו-זמנית (simultaneous selection).
 2. במסגרת הטיפול בתהליך ברירה מורכב כזה, בחרנו להתאים במידת האפשר את התיקון לתהליך הברירה הייחודי של כל מוסד ומחזור.
 3. במחקר נכללו (בעצה אחת עם ראשי יחידות המיון האוניברסיטאיות) רק סטודנטים שנתוני מועמדותם לא התבססו על חלופה כלשהי לממוצע הברגרות (ציון מכינה, ציון מלימודים אקדמיים קודמים, תחליף מחו"ל לממוצע הברגרות). מועמדים שהתקבלו בתנאים החלופיים עלולים לספק אינדיקציות מטעות על תוקף הניבוי של כלי המיון בקרב המועמדים הטיפוסיים.
 4. במחקר זה נבדק לראשונה תוקף הניבוי של הכלים הנכללים במבדקי מו"ר ומרק"ם, הגם שהללו מכוונים לניבוי ההצלחה בלימודים הקליניים (החל מהשנה הרביעית) ולא בלימודים הקדם-קליניים, ובפרט לא בלימודי שנה א'.

שיטה

מדגם

המחקר נעשה על נתוני 1,002 רשומות של סטודנטים אשר למדו בחוג לרפואה במסלול שש-שנתי במחזורים תשס"ז (2006/07) עד תש"ע (2009/10) באוניברסיטה העברית, באוניברסיטת תל אביב ובטכניון ב-12 יחידות עיבוד. יחידת עיבוד מוגדרת כך: חוג במוסד לימודים מסוים ובמחזור מסוים (מכאן ואילך נשתמש במונח "סטודנט" לתיאור רשומה של סטודנט ו"חוג" לתיאור יחידת עיבוד).

התנאים להיכללות במדגם היו:

1. ציון שנה א' ≤ 10 .
 2. ממוצע הברגרות וציון מכפ"ל ידועים.
 3. יש ציון מו"ר (אם למד באוניברסיטת תל אביב או בטכניון) או ציון מרק"ם (אם למד באוניברסיטה העברית) משנת המועמדות.
 4. הוזמן להיבחן במו"ר או במרק"ם על סמך ציון מכפ"ל וממוצע הברגרות בלבד, ולא על סמך חלופה כלשהי לממוצע הברגרות (ציון מכינה, ציון מלימודים אקדמיים קודמים, תחליף מחו"ל לממוצע הברגרות).
- בלוח 2 מוצגת התפלגות הסטודנטים במדגם לפי מוסד (יכונו להלן מוסד א, ב או ג) ומחזור. מאחר שהוספת תנאי 4 שלעיל היא בבחינת חידוש מוצג בלוח גם מספר הסטודנטים ללא הפעלת תנאי זה.

לוח 2: מספר הסטודנטים – במדגם (מודגש) וללא הפעלת תנאי 4 ("כל הסטודנטים") – ומספר החוגים (בסוגריים) במוסדות השונים ובמחזורים השונים

מחזור	מוסד א	מוסד ב	מוסד ג	סה"כ
2006/07	95	90	59	244
	(1)	(1)	(1)	(3)
2007/08	88	97	70	255
	(1)	(1)	(1)	(3)
2008/09	99	89	64	252
	(1)	(1)	(1)	(3)
2009/10	95	95	61	251
	(1)	(1)	(1)	(3)
סה"כ	377	371	254	1,002
	(4)	(4)	(4)	(12)

הוספת תנאי 4 הקטינה את מספר הסטודנטים במדגם בשיעור של כ-16%.

בנוסף לנתוני הסטודנטים במדגם שתואר לעיל (יכונה להלן "קובץ הלומדים") השתמשנו במחקר זה בשני קובצי נתונים נוספים:

1. מועמדים לכל החוגים תשנ"ב-תשנ"ג: מועמדים למחזורים תשנ"ב (1991/92) ותשנ"ג (1992/93) באוניברסיטאות (יכונה להלן "קובץ המועמדים").
2. מועמדים לרפואה נבחני מו"ר/מרק"ם תשס"ז-תש"ע: מועמדים לרפואה למחזורי המחקר לאוניברסיטה העברית שנבחנו במרק"ם בשנת המועמדות ומועמדים לרפואה למחזורי המחקר באוניברסיטת תל אביב או בטכניון שנבחנו במו"ר בשנת המועמדות (יכונה להלן: "קובץ נבחני מו"ר/מרק"ם").

משתני המחקר

קריטריון

ממוצע משוקלל בנקודות זכות של ציוני הקורסים שלמד הסטודנט בשנה א' (ציון שנה א')

הזאים

1. ממוצע ציוני הבגרות (B), כפי שחושב באוניברסיטאות (כולל בונוסים)
- 2-4. הציונים בתחומי מכפ"ל:
2. הציון בחשיבה מילולית (V)
3. הציון בחשיבה כמותית (Q)
4. הציון באנגלית (E)
- 5-7. הציונים הכלליים במכפ"ל⁶:
5. ציון רב-תחומי (P_{MD}): בציון זה משקלם של הציונים בתחום החשיבה המילולית ובתחום החשיבה הכמותית כפול ממשקלו של הציון באנגלית: $2V + 2Q + E$

⁶ בתקופה שבה נעשה תהליך המיון למחזורים הנכללים במחקר לא דווחו עדיין על ידי מאל"יו הציון בדגש מילולי והציון בדגש כמותי במכפ"ל. הציון הכללי היחיד שדווח היה הציון הרב-תחומי.

6. ציון בדגש מילולי (P_V): בציון זה משקלו של הציון בתחום החשיבה המילולית הוא פי

שלושה ממשקלו של כל אחד משני הציונים האחרים: $3V + Q + E$

7. ציון בדגש כמותי (P_Q): בציון זה משקלו של הציון בתחום החשיבה הכמותית הוא פי

שלושה ממשקלו של כל אחד משני הציונים האחרים: $V + 3Q + E$

13-8. שישה ציוני סכם 1, שנבדלים זה מזה במשקלות שניתנים למכפ"ל ולבגרות ובציון הכללי במכפ"ל, כמפורט להלן:

מספר	שם	משקל ^a מכפ"ל בסכם 1	הציון הכללי במכפ"ל שבו משתמשים
8	$P_{MD}50B50$	50	רב-תחומי
9	$P_{MD}70B30$	70	רב-תחומי
10	P_V50B50	50	בדגש מילולי
11	P_V70B30	70	בדגש מילולי
12	P_Q50B50	50	בדגש כמותי
13	P_Q70B30	70	בדגש כמותי

^a באחוזים

ציוני סכם 1 חושבו כך:

ראשית, חושבו ציוני סכם 1 גולמיים:

- חושבו ממוצע וסטיית תקן של מכפ"ל (כל אחד משלושת הציונים הכלליים) ובגרות בקובץ המועמדים בכל מוסד ומחזור בנפרד (על בסיס תצפית אחת לאדם במוסד ובמחזור ורק בקרב תצפיות שהיו להן הן ציוני מכפ"ל והן ציון בגרות). לאחר מכן מוצעו הסטטיסטיים הני"ל מעבר למחזורים בכל מוסד.

- על פי ממוצעי הסטטיסטיים הני"ל, תוקננו מכפ"ל ובגרות (בשלושת הקבצים: קובץ המועמדים, קובץ נבחני מו"ר/מרק"ם וקובץ הלומדים) וצורפו בשני צירופי המשקלות (50% למכפ"ל ו-50% לבגרות או 70% למכפ"ל ו-30% לבגרות), וזאת עבור כל אחד משלושת הציונים הכלליים של מכפ"ל. הצירופים הללו מהווים את ששת ציוני סכם 1 הגולמיים. שנית, חושבו ציוני סכם 1 (סופיים):

- חושבו ממוצע וסטיית תקן של ששת ציוני הסכם הגולמיים בקובץ המועמדים בכל מוסד ומחזור בנפרד. לאחר מכן מוצעו הסטטיסטיים הני"ל מעבר למחזורים בכל מוסד.

- על פי ממוצעי הסטטיסטיים תוקננו ששת ציוני סכם 1 הגולמיים (בשלושת הקבצים). לקבלת ששת ציוני סכם 1 הסופיים, הוכפלו הציונים המתוקננים ב-10 והוספו ל-50.

19-14. ציוני מערכות מו"ר ומרק"ם (לכל הציונים ממוצע 200 וסטיית תקן 20 בשנת היבחנות):

14. הציון בשאלון שיפוט והחלטה (Q_J)

15. הציון בשאלון הביוגרפי (Q_B)

16. הציון בתחנות מו"ר (ST_{MO})

17. הציון בתחנות מרק"ם (ST_{MI})

18. הציון הסופי במו"ר (MO)

19. הציון הסופי במרק"ם (MI)

43-20. עשרים וארבעה ציוני סכס 2, שנבדלים זה מזה בחזאי הלא-קוגניטיבי (מו"ר או מרק"ם), במשקלות שניתנים לסכס 1 ולחזאי הלא-קוגניטיבי, במשקלות שניתנים למכפ"ל ובגרות בחישוב סכס 1 ובציון הכללי במכפ"ל שבו משתמשים בחישוב סכס 1, כמפורט להלן:

מספר	שם	משקל ^a סכס 1 בסכס 2	החזאי הלא- קוגניטיבי	משקל ^a מכפ"ל בסכס 1	הציון הכללי במכפ"ל שבו משתמשים
20	(P _{MD} 50B50)50MO50	50	מו"ר	50	רב-תחומי
21	(P _{MD} 50B50)30MO70	30	מו"ר	50	רב-תחומי
22	(P _{MD} 70B30)50MO50	50	מו"ר	70	רב-תחומי
23	(P _{MD} 70B30)30MO70	30	מו"ר	70	רב-תחומי
24	(P _V 50B50)50MO50	50	מו"ר	50	בדגש מילולי
25	(P _V 50B50)30MO70	30	מו"ר	50	בדגש מילולי
26	(P _V 70B30)50MO50	50	מו"ר	70	בדגש מילולי
27	(P _V 70B30)30MO70	30	מו"ר	70	בדגש מילולי
28	(P _Q 50B50)50MO50	50	מו"ר	50	בדגש כמותי
29	(P _Q 50B50)30MO70	30	מו"ר	50	בדגש כמותי
30	(P _Q 70B30)50MO50	50	מו"ר	70	בדגש כמותי
31	(P _Q 70B30)30MO70	30	מו"ר	70	בדגש כמותי
32	(P _{MD} 50B50)50MI50	50	מרק"ם	50	רב-תחומי
33	(P _{MD} 50B50)30MI70	30	מרק"ם	50	רב-תחומי
34	(P _{MD} 70B30)50MI50	50	מרק"ם	70	רב-תחומי
35	(P _{MD} 70B30)30MI70	30	מרק"ם	70	רב-תחומי
36	(P _V 50B50)50MI50	50	מרק"ם	50	בדגש מילולי
37	(P _V 50B50)30MI70	30	מרק"ם	50	בדגש מילולי
38	(P _V 70B30)50MI50	50	מרק"ם	70	בדגש מילולי
39	(P _V 70B30)30MI70	30	מרק"ם	70	בדגש מילולי
40	(P _Q 50B50)50MI50	50	מרק"ם	50	בדגש כמותי
41	(P _Q 50B50)30MI70	30	מרק"ם	50	בדגש כמותי
42	(P _Q 70B30)50MI50	50	מרק"ם	70	בדגש כמותי
43	(P _Q 70B30)30MI70	30	מרק"ם	70	בדגש כמותי

^a באחוזים

ציוני סכס 2 חושבו כך:

ראשית, חושבו ציוני סכס 2 גולמיים:

- חושבו ממוצע וסטיית תקן של סכס 1 (ששת הציונים) ושל הציונים הסופיים במו"ר ובמרק"ם בקובץ נבחני מו"ר/מרק"ם בכל מוסד ומחזור בנפרד (על בסיס תצפית אחת לאדם במוסד ובמחזור ורק בקרב תצפיות שהיו להן כל הציונים הרלוונטיים).

- על פי הסטטיסטיקה הללו תוקנו – בשני קבצים: קובץ נבחני מו"ר/מרק"ם וקובץ הלומדים – ציוני סכס 1 והציונים במו"ר ובמרק"ם (סה"כ 8 משתנים). חשוב להדגיש שבמקרה זה הסטטיסטיקה ששימשו לתקנון הם ספציפיים למוסד ומחזור (וזאת להבדיל ממה שנעשה בחישוב ציוני סכס 1, ששם הסטטיסטיקה ששימשו לתקנון היו ספציפיים למוסד אך לא למחזור, כיוון שהנתונים בקובץ המועמדים לא היו מלכתחילה של מחזורי המחקר).

המשתנים המתוקנים צורפו בשני צירופי משקלות (50% לסכס 1 ו-50% למו"ר או מרק"ם או 30% לסכס 1 ו-70% למו"ר או מרק"ם). הצירופים הללו מהווים את 24 ציוני סכס 2 הגולמיים.

שנית, חושבו ציוני סכס 2 (סופיים) :

- חושבו ממוצע וסטיית תקן של 24 ציוני סכס 2 הגולמיים בקובץ נבחני מו"ר/מרק"ם בכל מוסד ומחזור בנפרד.

- על פי הסטטיסטיקה הללו תוקנו 24 ציוני סכס 2 הגולמיים (בשני הקבצים). לקבלת 24 ציוני סכס 2 (הסופיים), הוכפלו הציונים המתוקנים ב-10 והוספו ל-50.

עיבוד הנתונים

1. חישוב ממוצעים וסטיות תקן של הקריטריון והחזאים בכל אחד מהחוגים

2. חישוב מתאמים פשוטים בין החזאים לקריטריון בכל אחד מהחוגים

המתאמים תוקנו לקיצוץ תחום בהנחה של ברירה דו-שלבית, כאשר בכל שלב הברירה התבססה על משתנה יחיד : בשלב הראשון על משתנה כדוגמת סכס 1 (צירוף של בגרות ומכפ"ל)⁷, ובשלב השני על משתנה כדוגמת סכס 2 (צירוף של משתנה הברירה של השלב הראשון עם הציון במו"ר או במרק"ם). המשקלות של מרכיבי משתני הברירה בשני השלבים מוצגים בלוח 3 שלהלן. ניתן לראות שהשקלול של מרכיבי משתנה הברירה של השלב הראשון השתנה בין מוסדות, והשקלול של מרכיבי משתנה הברירה של השלב השני השתנה בין מוסדות, ובתוך מוסד בין מחזורים.

לוח 3: משקלות^a מרכיבי משתני הברירה בשלב הראשון ובשלב השני לפי מוסד ומחזור

מוסד	מחזור	השלב הראשון		השלב השני	
		מכפ"ל ^b	בגרות	משתנה הברירה בשלב הראשון	מו"ר/מרק"ם
א	2006/07	50	50	50	50
	2007/08	50	50	40	60
	2008/09	50	50	40	60
	2009/10	50	50	40	60
ב	2006/07	70	30	25	75
	2007/08	70	30	25	75
	2008/09	70	30	25	75
	2009/10	70	30	25	75
ג	2006/07	60	40	30	70
	2007/08	60	40	30	70
	2008/09	60	40	0	100
	2009/10	60	40	0	100

^a באחוזים

^b הציון הרב-תחומי

תיקון המתאמים נעשה באופן דו-שלבי : המתאמים תוקנו תחילה עבור השלב השני של הברירה (מהמתאמים הנצפים, שחושבו בקרב הלומדים, למתאמים שהיו מתקבלים בקרב המועמדים לרפואה שנבחנו במו"ר או במרק"ם לו היו כולם לומדים והיו בידינו ציוני הקריטריון שלהם), ואחר-כך עבור השלב הראשון (מהמתאמים שהיו מתקבלים בקרב המועמדים לרפואה שנבחנו במו"ר או במרק"ם למתאמים שהיו מתקבלים בקרב כל המועמדים לרפואה לו היו כולם נבחנים במו"ר או במרק"ם ולומדים).

⁷ משתנה הברירה ששימש בפועל, הן בשלב הראשון והן בשלב השני, אינו בהכרח אחד מששת משתני סכס 1 או מ-24 משתני סכס 2, בהתאמה, שהוגדרו לעיל, לכן כתבנו "כדוגמת". בחרנו להתמקד בתוקף הניבוי רק של חלק ממשתני הברירה (ואותם הגדרנו כמשתני סכס 1 וסכס 2). כך, למשל (ראו לוח 3 שלהלן), במוסד א, משתנה הברירה של השלב הראשון הוגדר כאחד מציוני סכס 1 (משתנה 8), אך משתנה הברירה של השלב השני במחזור 2007/08 ואילך לא הוגדר כאחד מציוני סכס 2.

תיקון המתאמים בכל שלב התבסס על הנחה של ברירה חד-משתנית (כאמור, בכל שלב הברירה נעשתה על פי משתנה יחיד) עבור המקרה התלת-משתני (כלומר, התיקון מתבסס על סטטיסטיים של שלושה משתנים: המשתנה שהברירה נעשתה על פיו, החזאי שאת התוקף שלו מחשבים, והקריטריון). התיקון נדון בהרחבה אצל גליקסן (Gulliksen, 1987), והנוסחה שהשתמשנו בה במחקר זה (ראו נספח 2) הוצגה אצל לין (Linn, 1983). לצורך תיקון זה צריך לדעת את שונות משתנה הברירה בקרב המועמדים (כאשר ההגדרה מי הם המועמדים תלויה כמובן בשלב שבו עוסקים). כאמור, משתנה הברירה הוא ספציפי למוסד ומחזור. לפיכך, עבור כל מוסד, מחזור ושלב בברירה הגדרנו את משתנה הברירה המתאים (בדומה לאופן שבו חושבו ציוני סכס 1 וסכס 2). שונות משתנה הברירה בקרב המועמדים הוגדרה כשונות בקרב המועמדים לחוג במוסד ובמחזור. באופן אופרטיבי, לצורך תיקון עבור השלב השני של הברירה השתמשנו, עבור כל אחת מיחידות העיבוד שנכללו במחקר, בשונותו של משתנה הברירה בקרב המועמדים לחוג לרפואה במוסד ובמחזור הנדונים שנבחנו במו"ר או במרק"ם. לצורך תיקון עבור השלב הראשון של הברירה חושבה שונותו של משתנה הברירה כמוצע – מעבר לחוגים, מוסדות ומחזורים במחזורים תשנ"ב ותשנ"ג – של שונותו בקרב מועמדים לחוג כלשהו במוסד ובמחזור. בגוף הדוח מוצגים המתאמים המתוקנים לקיצוץ תחום. המתאמים הנצפים מוצגים בנספח 3.

3. חישוב מתאמים מרובים ומקדמי רגרסיה (מתוקנים) שהתקבלו בנייתוחי רגרסיה מרובה לניבוי הקריטריון בכל אחד מהחוגים

נבדקו שישה מודלים תלת-משתניים עבור צירופים משלוש קבוצות המשתנים הבאות:

(1) בגרות

(2) ציון כללי במכפ"ל: הציון הרב-תחומי, הציון בדגש מילולי או הציון בדגש כמותי

(3) ציון סופי במו"ר או במרק"ם

כל הסטטיסטיים המדווחים בנייתוחי הרגרסיה תוקנו לקיצוץ תחום באופן שהקלט לנייתוחי הרגרסיה היה (במקום התצפיות עצמן) מטריצת המתאמים הפשוטים אשר תוקנו לקיצוץ תחום כמתואר בסעיף 2 בעיבוד הנתונים – בין כל המשתנים הנכללים בנייתוח. בגוף הדוח מוצגים הערכים המתוקנים לקיצוץ תחום. הערכים הנצפים מוצגים בנספח 3.

כאמור, הסטטיסטיים המתוארים בסעיפים 1 עד 3 חושבו בחוגים. בלוחות בפרק התוצאות יוצגו ממוצעים משוקללים (במספר הסטודנטים בחוג⁸) של הסטטיסטיים הני"ל. כמו כן, עבור הסטטיסטיים בסעיפים 2 ו-3 (מתאמים פשוטים, מתאמים מרובים ומקדמי רגרסיה), יוצג הטווח הבין-רבעוני של הסטטיסטיים, כמדד לפיזור הערכים שהתקבלו ב-12 החוגים (נזכיר שבמונח "חוג" הכוונה לחוג במוסד לימודים ובמחזור). טווח זה חושב, בדומה לסטטיסטיים עצמם, תוך התחשבות במספר הסטודנטים בחוג.

4. המחשת תוקף הניבוי של תהליך ברירה דו-שלבי כמכלול

המתאמים הפשוטים בין החזאים לקריטריון שחושבו בסעיף 2 שלעיל מתארים את תוקף הניבוי של מערכת ברירה שמתבססת על מיון חד-שלבי שבו לכל המועמדים יש ציון בחזאי הנדון והם מדורגים על פיו. כדי להעריך את תוקף הניבוי של תהליך ברירה דו-שלבי מהסוג שמתקיים בפועל

⁸ מספר התצפיות עבור משתני מו"ר ומרק"ם והמשתנים הקשורים אליהם היה נמוך מ-1,002: N=902 עבור משתנים 16, 18 ו-20, N=879 עבור משתנים 17, 19 ו-32. השקלול התבסס על מספר התצפיות שבהן חושב הסטטיסטי.

(ראו לוח 3), ובפרט בהשוואה בינו ובין תהליך ברירה חד-שלבי, נערכה הדמיה⁹ שבדקה את ממוצע ציון שנה א' בקרב המתקבלים במצבי הברירה הבאים :

א. תהליך ברירה חד-שלבי על פי $P_{MD}50B50$ (חזאי מס' 8)

ב. תהליך ברירה חד-שלבי על פי MO (חזאי מס' 18)

ג. תהליך ברירה חד-שלבי על פי $(P_{MD}50B50)50MO50$ (חזאי מס' 20)

ד. תהליך ברירה דו-שלבי שבו בשלב הראשון ממיינים על פי $P_{MD}50B50$ ובשלב השני על פי MO

ה. תהליך ברירה דו-שלבי שבו בשלב הראשון ממיינים על פי $P_{MD}50B50$ ובשלב השני על פי

$(P_{MD}50B50)50MO50$

בכל ההדמיות הוגדרו המשתנים (החזאים והקריטריון) כמתפלגים התפלגות רב-משתנית

נורמלית עם ממוצעים וסטיות תקן שוליים של 0 ו-1, בהתאמה. התפלגות זו אמורה לדמות את התפלגות המועמדים. בהתאמה להקשר של מיון לרפואה הוגדר מדגם של 1,000 מועמדים עם 100 מתקבלים. בשני המצבים של תהליך ברירה דו-שלבי (מצבים ד ו-ה שלעיל) נקבע ש-400 מתוך 1,000 המועמדים נבחרים לעבור לשלב השני. המתאמים בין החזאים ובין ציון שנה א' ובינם לבין עצמם נקבעו בהתאמה לנתונים האמפיריים¹⁰.

ההדמיה כללה 500 הרצות של כל אחד מחמשת מצבי הברירה, והערכים שמוצגים הם ממוצעים מעבר ל-500 ההרצות.

תוצאות

1. ממוצעים וסטיות תקן של משתני המחקר

בלוחות 4א עד 4ד מוצגים הממוצעים וסטיות התקן של משתני המחקר. מתחת לשמו של כל חזאי בלוח מופיע מספרו, כדי לסייע באיתורו בתיאור החזאים בפרק השיטה.

לוח 4א: ממוצעים וסטיות תקן (בסוגריים) של ציון שנה א', ממוצע ציוני הבגרות, הציונים בתחומי מכפ"ל והציונים הכלליים במכפ"ל

ציון שנה א'	B 1	V 2	Q 3	E 4	P_{MD} 5	P_V 6	P_Q 7
86.3 (5.6)	112.8 (2.7)	143 (5)	144 (5)	142 (7)	749 (20)	748 (22)	750 (21)

לוח 4ב: ממוצעים וסטיות תקן (בסוגריים) של ציוני סכם 1

$P_{MD}50B50$ 8	$P_{MD}70B30$ 9	P_V50B50 10	P_V70B30 11	P_Q50B50 12	P_Q70B30 13
73.0 (1.7)	71.7 (1.8)	73.2 (1.9)	71.7 (2.1)	72.6 (1.8)	71.2 (2.0)

לוח 4ג: ממוצעים וסטיות תקן (בסוגריים) של הציונים במערכות מו"ר ומרק"ם

Q_J 14	Q_B 15	ST_MO 16	ST_MI 17	MO 18	MI 19
208 (17)	207 (17)	209 (15)	206 (16)	210 (14)	207 (14)

⁹ הדמיה דומה ניתן למצוא אצל Sackett & Roth, 1996, ותוצאותיה תואמות לפתרון אנליטי שהוצע לסוגיה הנדונה (De Corte, Lievens, & Sackett, 2006).

¹⁰ המתאמים הרלוונטיים בין החזאים ובין ציון שנה א' מוצגים בלוחות 4ב עד 4ד שלהלן; המתאמים בין חזאי השלב הראשון לחזאי השלב השני הם 0.19 במצב ד ו-0.91 במצב ה.

לוח 24: ממוצעים וסטיות תקן (בסוגריים) של ציוני סכמ 2

(P _V 50B50)30 MO70 25	(P _V 50B50)50 MO50 24	(P _{MD} 70B30)30 MO70 23	(P _{MD} 70B30)50 MO50 22	(P _{MD} 50B50)30 MO70 21	(P _{MD} 50B50)50 MO50 20
54.8 (7.0)	54.6 (7.2)	54.6 (6.9)	54.5 (7.1)	54.9 (7.0)	54.7 (7.1)

לוח 24 (המשך): ממוצעים וסטיות תקן (בסוגריים) של ציוני סכמ 2

(P _Q 70B30)30 MO70 31	(P _Q 70B30)50 MO50 30	(P _Q 50B50)30 MO70 29	(P _Q 50B50)50 MO50 28	(P _V 70B30)30 MO70 27	(P _V 70B30)50 MO50 26
54.6 (7.0)	54.4 (7.3)	54.8 (7.0)	54.6 (7.2)	54.6 (7.0)	54.4 (7.1)

לוח 24 (המשך): ממוצעים וסטיות תקן (בסוגריים) של ציוני סכמ 2

(P _V 50B50)30 MI70 37	(P _V 50B50)50 MI50 36	(P _{MD} 70B30)30 MI70 35	(P _{MD} 70B30)50 MI50 34	(P _{MD} 50B50)30 MI70 33	(P _{MD} 50B50)50 MI50 32
53.8 (7.8)	53.5 (8.2)	53.6 (7.5)	53.4 (7.9)	53.8 (7.7)	53.5 (8.1)

לוח 24 (המשך): ממוצעים וסטיות תקן (בסוגריים) של ציוני סכמ 2

(P _Q 70B30)30 MI70 43	(P _Q 70B30)50 MI50 42	(P _Q 50B50)30 MI70 41	(P _Q 50B50)50 MI50 40	(P _V 70B30)30 MI70 39	(P _V 70B30)50 MI50 38
53.6 (7.5)	53.3 (7.9)	53.7 (7.7)	53.4 (8.1)	53.6 (7.6)	53.4 (8.0)

2. מתאמים פשוטים (מתוקנים לקיצוץ תחום) בין החזאים לקריטריון

בלוחות 5א עד 5ד מוצגים המתאמים בין החזאים לקריטריון. מתחת לכל מתאם מופיע (בסוגריים) הטווח הבין-רבעוני שלו.

לוח 5א: מתאמים בין ציון שנה א' ובין

ממוצע ציוני הבגרות, הציונים בתחומי מכפ"ל והציונים הכלליים במכפ"ל והטווח הבין-רבעוני של המתאמים (בסוגריים)

P _Q 7	P _V 6	P _{MD} 5	E 4	Q 3	V 2	B 1
0.61 (0.51-0.83)	0.59 (0.47-0.84)	0.61 (0.49-0.84)	0.54 (0.48-0.79)	0.57 (0.42-0.73)	0.55 (0.41-0.78)	0.30 (0.14-0.56)

בלוח 5א ניתן לראות שתוקף הבגרות בינוני ושתוקף הציונים הכלליים במכפ"ל גבוה. תקפויות שלושת הציונים הכלליים במכפ"ל דומות זו לזו, כאשר תוקף הציון בדגש מילולי מעט יותר נמוך.

אשר לתחומי מכפ"ל, תקפויותיהם גבוהות ודומות זו לזו, עם יתרון קל לציון בתחום הכמותי. תוקף כל אחד מהציונים הכלליים במכפ"ל גבוה מתוקף מרכיביו (תחומי מכפ"ל) בנפרד.

לוח 5ב: מתאמים בין ציון שנה א' ובין ציוני סכמ 1

והטווח הבין-רבעוני של המתאמים (בסוגריים)

P _Q 70B30 13	P _Q 50B50 12	P _V 70B30 11	P _V 50B50 10	P _{MD} 70B30 9	P _{MD} 50B50 8
0.62 (0.51-0.83)	0.61 (0.49-0.80)	0.60 (0.47-0.84)	0.60 (0.47-0.82)	0.61 (0.49-0.84)	0.61 (0.49-0.82)

בלוח 5ב ניתן לראות שתקפויות ציוני סכמ 1 גבוהות ודומות זו לזו (בין 0.60 ל-0.62).

השוואת הממדים השונים שבהם נבדלים ששת ציוני סכס 1 מלמדת שתוקף ציוני הסכס המבוססים על ציון מכפ"ל בדגש מילולי הוא מעט יותר נמוך. תוקף ציוני הסכס המבוססים על ציון מכפ"ל בדגש כמותי גבוה או דומה לתוקף ציוני הסכס המבוססים על ציון מכפ"ל הרב-תחומי. ההבדלים במשקלות היחסיים של מכפ"ל ובגרות אינם משפיעים באופן עקבי על התוקף. הציון התקף ביותר הוא זה המבוסס על ציון מכפ"ל בדגש כמותי במשקל של 70% ועל הבגרות במשקל של 30%. השוואה בין תוקף ציון סכס 1, על חלופותיו השונות, לתוקף מרכיביו, מלמדת שתרומת הבגרות לתוקף ציון הסכס היא קטנה.

לוח ג5: מתאמים בין ציון שנה א' ובין הציונים במערכות מו"ר ומרק"ם והטווח הבין-רבעוני של המתאמים (בסוגריים)

MI 19	MO 18	ST_MI 17	ST_MO 16	Q_B 15	Q_J 14
0.37 (0.13-0.53)	0.35 (0.04-0.70)	0.27 (0.06-0.42)	0.30 (0.05-0.47)	0.33 (0.20-0.56)	0.31 (0.14-0.44)

לוח ג5 ניתן לראות שתקפויות הציונים הסופיים במו"ר ובמרק"ם בינוניות ודומות זו לזו. חשוב להזכיר שהנתונים עבור מערכות מו"ר ומרק"ם לא התבססו על אותן תצפיות. תוקף הציונים בקרב בעלי ציונים בשתי המערכות (N=779) הוא 0.36 ו-0.29 עבור הציונים הסופיים במו"ר ובמרק"ם, בהתאמה.

תקפויות המרכיבים בתוך כל מערכת הן בינוניות, עם יתרון קל לשאלונים על פני התחנות. בשתי המערכות תוקף הציון הכללי גבוה מתוקף המרכיבים.

לוח ד5: מתאמים בין ציון שנה א' ובין ציוני סכס 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

(P _V 50B50)30 MO70 25	(P _V 50B50)50 MO50 24	(P _{MD} 70B30)30 MO70 23	(P _{MD} 70B30)50 MO50 22	(P _{MD} 50B50)30 MO70 21	(P _{MD} 50B50)50 MO50 20
0.58 (0.39-0.84)	0.62 (0.51-0.85)	0.58 (0.36-0.85)	0.63 (0.52-0.86)	0.59 (0.41-0.85)	0.63 (0.53-0.86)

לוח ד5 (המשך): מתאמים בין ציון שנה א' ובין ציוני סכס 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

(P _Q 70B30)30 MO70 31	(P _Q 70B30)50 MO50 30	(P _Q 50B50)30 MO70 29	(P _Q 50B50)50 MO50 28	(P _V 70B30)30 MO70 27	(P _V 70B30)50 MO50 26
0.57 (0.34-0.85)	0.64 (0.54-0.86)	0.59 (0.36-0.85)	0.64 (0.54-0.86)	0.56 (0.31-0.85)	0.62 (0.51-0.86)

לוח ד5 (המשך): מתאמים בין ציון שנה א' ובין ציוני סכס 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

(P _V 50B50)30 MI70 37	(P _V 50B50)50 MI50 36	(P _{MD} 70B30)30 MI70 35	(P _{MD} 70B30)50 MI50 34	(P _{MD} 50B50)30 MI70 33	(P _{MD} 50B50)50 MI50 32
0.60 (0.45-0.84)	0.63 (0.49-0.86)	0.60 (0.46-0.85)	0.64 (0.50-0.87)	0.61 (0.45-0.84)	0.64 (0.51-0.86)

לוח 25 (המשך): מתאמים בין ציון שנה א' ובין ציוני סכם 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

(P _Q 70B30)30 MI70 43	(P _Q 70B30)50 MI50 42	(P _Q 50B50)30 MI70 41	(P _Q 50B50)50 MI50 40	(P _V 70B30)30 MI70 39	(P _V 70B30)50 MI50 38
0.60 (0.45-0.85)	0.64 (0.51-0.87)	0.60 (0.44-0.85)	0.64 (0.52-0.86)	0.59 (0.46-0.84)	0.63 (0.48-0.86)

בלוח 25 ניתן לראות שתקפויות ציוני סכם 2 גבוהות ודומות זו לזו (בין 0.56 ל-0.64).
 השוואת הממדים השונים שבהם נבדלים 24 ציוני סכם 2 מלמדת שההשפעה העיקרית על התוקף קשורה למשקלו של סכם 1 בסכם 2: התוקף גבוה יותר (בכ-4 נקודות בממוצע) כאשר משקל סכם 1 בסכם 2 גבוה יותר (50% לעומת 30%). לשניים מן הממדים האחרים יש השפעה קטנה (כנקודה אחת בממוצע) על התוקף, שכיוונה הוא כדלקמן: התוקף גבוה יותר כאשר הציון הכללי במכפ"ל שבו משתמשים הוא הציון הרב-תחומי או הציון בדגש כמותי וכאשר החזאי הלא-קוגניטיבי הוא מרק"ם. למשקלות של מכפ"ל ובגרות בסכם 1 יש השפעה זניחה על התוקף.
 הציונים התקפים ביותר, לפיכך, הם אלה שבהם משקל סכם 1 בסכם 2 הוא 50% ושהציון הכללי במכפ"ל הוא הציון בדגש כמותי, אם משתמשים במו"ר, ואלה שבהם משקל סכם 1 בסכם 2 הוא 50% ושהציון הכללי במכפ"ל הוא הציון הרב-תחומי או הציון בדגש כמותי, אם משתמשים במרק"ם.

השוואה בין תוקף ציוני סכם 2 לתוקף מרכיביו מלמדת שבאופן כללי תוקפים דומה לתוקף ציוני סכם 1. צירוף ציוני מו"ר או מרק"ם לציוני סכם 1 אינו משנה באופן משמעותי את התוקף.

3. תוצאות ניתוחי רגרסיה מרובה

למרות שמקסימיזציה של התוקף בניבוי ההישגים בלימודים, ובפרט בניבוי ציוני שנה א', אינה המטרה היחידה של תהליכי הברירה לחוגים לרפואה, בחרנו לבחון את הרכבו של תהליך ברירה שעושה אופטימיזציה (או קירוב לאופטימיזציה) של תוקף הניבוי. בלוחות 6א ו-6ב מוצגים המתאמים המרובים ומקדמי הרגרסיה שהתקבלו משישה מודלים של רגרסיה לניבוי ציון שנה א'. מתחת לכל סטטיסטי מופיע (בסוגריים) הטווח הבין-רבעוני שלו.

לוח 6א: תוצאות ניתוחי רגרסיה מרובה לניבוי ציון שנה א' ע"י בגרות, מכפ"ל ומו"ר והטווח הבין-רבעוני של הסטטיסטים (בסוגריים)

בגרות, מכפ"ל כמותי ומו"ר				בגרות, מכפ"ל מילולי ומו"ר				בגרות, מכפ"ל ר"ת ומו"ר			
מקדמי רגרסיה			מתאם מרובה	מקדמי רגרסיה			מתאם מרובה	מקדמי רגרסיה			מתאם מרובה
MO	P _Q	B		MO	P _V	B		MO	P _{MD}	B	
0.32	0.48	0.23	0.67	0.30	0.44	0.24	0.66	0.30	0.48	0.22	0.67
(0.20-)	(0.33-)	(0.07-)	(0.54-)	(0.19-)	(0.27-)	(0.06-)	(0.52-)	(0.17-)	(0.31-)	(0.06-)	(0.53-)
(0.37)	(0.65)	(0.34)	(0.87)	(0.32)	(0.59)	(0.34)	(0.87)	(0.33)	(0.64)	(0.34)	(0.87)

לוח 6ב: תוצאות ניתוחי רגרסיה מרובה לניבוי ציון שנה א' ע"י בגרות, מכפ"ל ומרק"ם והטווח הבין-רבעוני של הסטטיסטים (בסוגריים)

בגרות, מכפ"ל כמותי ומרק"ם				בגרות, מכפ"ל מילולי ומרק"ם				בגרות, מכפ"ל ר"ת ומרק"ם			
מקדמי רגרסיה			מתאם מרובה	מקדמי רגרסיה			מתאם מרובה	מקדמי רגרסיה			מתאם מרובה
MI	P _Q	B		MI	P _V	B		MI	P _{MD}	B	
0.21	0.51	0.16	0.67	0.20	0.48	0.17	0.65	0.20	0.51	0.16	0.66
(0.12-)	(0.39-)	(0.07-)	(0.52-)	(0.09-)	(0.36-)	(0.08-)	(0.48-)	(0.09-)	(0.39-)	(0.07-)	(0.50-)
(0.28)	(0.70)	(0.24)	(0.87)	(0.22)	(0.66)	(0.24)	(0.87)	(0.23)	(0.70)	(0.23)	(0.87)

בלוחות 6א ו-6ב ניתן לראות שהמתאמים המרובים שהתקבלו מניתוחי הרגרסיה גבוהים מאוד, בפרט כשמדובר במודל תלת-משתני ולא באופטימיזציה מלאה של משקלות מרכיבי מערכת המיון¹¹.

המתאמים המרובים מששת ניתוחי הרגרסיה השונים דומים מאוד זה לזה (בין 0.65 ל-0.67). אשר למשקלות המרכיבים, בשקלול האופטימלי (לצורכי מקסימיזציה של התוקף) ניתנת, בקירוב, כמחצית מן המשקל לציון הכללי במכפ"ל, והמחצית הנותרת נחלקת שווה בשווה, בין הבגרות למו"ר או למרק"ם.

4. המחשת תוקף הניבוי של תהליך ברירה דו-שלבי

הסטטיסטיים שדווחו עד כה בסעיפים 2 ו-3 שלעיל התייחסו לתוקף הניבוי של חזאים שונים תחת המצב השכיח שבו הם משמשים, בנפרד או בצירופים שלהם, בתהליך ברירה חד-שלבי. כאמור, בחוגים שנכללים במחקר זה נעשה שימוש באותם חזאים, אך בתהליך ברירה דו-שלבי. נשאלת השאלה מה ניתן להסיק מן הממצאים שדווחו עד כה, על תהליך הברירה בפועל. בלוח 7 מוצגות תוצאות של שתי הדמיות שנערכו במטרה להמחיש את תוקף הניבוי של תהליך ברירה דו-שלבי, וזאת בהשוואה לתוקף הניבוי של תהליך ברירה שמשמש באותם חזאים באופן חד-שלבי. שתי ההדמיות נועדו להציג שני טיפוסים ברירה דו-שלבית. בשתיהן החזאי המשמש בשלב הראשון הוא $P_{MD}50B50$ (שהוא אב-טיפוס של ציון סכס 1). אשר לחזאי המשמש בשלב השני, בהדמיה הראשונה זהו המשתנה MO (הציון הסופי במו"ר). בהדמיה השנייה זהו המשתנה $(P_{MD}50B50)50MO50$ (צירוף במשקלות שווים של ציון סכס 1 של השלב הראשון ושל מו"ר). בשתי ההדמיות ציון שנה א' בקרב המועמדים הוא משתנה נורמלי סטנדרטי (ממוצע 0 וס"ת 1). לוח 7 מציג את ממוצע ציוני שנה א' בקרב המתקבלים בחלופות השונות של שתי ההדמיות.

לוח 7: ממוצע ציון שנה א' בקרב המתקבלים בחלופות שונות של תהליך הברירה

ציון שנה א' בקרב המתקבלים	מתאם עם ציון שנה א'			
1.07	0.61	$P_{MD}50B50$	ברירה חד-שלבית עפ"י	הדמיה ראשונה: ברירה עפ"י $P_{MD}50B50$ ו/או MO
0.61	0.35	MO		
0.93	?	ברירה דו-שלבית		
1.07	0.61	$P_{MD}50B50$	ברירה חד-שלבית עפ"י	הדמיה שנייה: ברירה עפ"י $P_{MD}50B50$ ו/או $(P_{MD}50B50)50MO50$
1.11	0.63	$(P_{MD}50B50)50MO50$		
1.11	?	ברירה דו-שלבית		

ההדמיה הראשונה מלמדת שבהשוואה לברירה חד-שלבית על פי סכס 1 ($P_{MD}50B50$) בלבד, מעבר לברירה דו-שלבית שבה השלב השני מתבסס על החזאי MO לבדו מוריד אמנם את ציון שנה א' בקרב המתקבלים, אך עדיין משמר פער של קרוב לסטיית תקן שלמה בין ממוצע ציון שנה א' בקרב המועמדים (0) לממוצע ציון שנה א' בקרב המתקבלים (0.93). להמחשת המשמעות של ממצא זה במונחי מתאם, נציין שפער כזה בין המועמדים למתקבלים יתקבל מברירה חד-שלבית עם חזאי שתוקף הניבוי שלו הוא 0.53, תחת תנאי הברירה (ההדמיה) הנוכחיים. במילים אחרות, תוקף הניבוי של הברירה הדו-שלבית בהדמיה זו הוא 0.53, שהוא נמוך מתוקף הניבוי של ברירה חד-

¹¹ אופטימיזציה מלאה, שמתבססת על הבגרות, הציונים בשלושת תחומי מכפ"ל ומרכיבי מו"ר או מרק"ם (סה"כ 7 חזאים) מניבה, בממוצע, מתאם מרובה של 0.70.

שלבית על פי סכס 1 (P_{MD}50B50) בלבד (0.61), אך עדיין תורם משמעותית לאיכות המתקבלים, כפי שהודגם לעיל.

ההדמיה השנייה מעידה שכאשר השלב השני של הברירה מבוסס על צירוף של MO עם סכס 1, אין כל פגיעה בתוקף הניבוי של מערכת המיון בהשוואה לברירה חד-שלבית ע"י אחד משני המרכיבים. בפרט, תוקף הניבוי של ברירה דו-שלבית (0.63) בהדמיה זו עולה על תוקף הניבוי של ברירה חד-שלבית על פי סכס 1 (P_{MD}50B50) בלבד (0.61).

סיכום ודיון

מטרתו העיקרית של מחקר זה היא לספק תמונה על תוקף מערכת המיון לחוגים לרפואה במסלולים השש-שנתיים באוניברסיטה העברית, אוניברסיטת תל-אביב והטכניון – בחיזור ההצלחה בלימודי שנה א' במחזורים תשס"ז (2006/07) עד תש"ע (2009/10). המחקר כלל רק סטודנטים שהחלטת הקבלה שלהם התבססה על תעודת הבגרות (וציון מכפ"ל) ולא על חלופות לתעודת הבגרות.

להלן עיקרי הממצאים:

1. תוקף החזאים המרכיבים את מערכת המיון

הבגרות: תוקף הבגרות הוא בינוני (0.30).

ציוני מכפ"ל: תקפיות הציונים הכלליים במכפ"ל גבוהות ודומות זו לזו, עם יתרון קל לציון הרב-תחומי ולציון בדגש כמותי (0.61) על פני הציון בדגש מילולי (0.59). תוקף כל אחד מהציונים הכלליים במכפ"ל גבוה מתוקף מרכיביו בנפרד.

תקפיות שלושת תחומי מכפ"ל גבוהות ודומות זו לזו, עם יתרון קל לציון בחשיבה כמותית (0.57) על פני הציון בחשיבה מילולית (0.55) ובאנגלית (0.54).

ציוני סכס 1: שישה ציוני סכס 1, המבוססים על צירוף של מכפ"ל ובגרות, הוגדרו במחקר. ציונים אלה נבדלו זה מזה במשקלות שניתנו למכפ"ל ולבגרות (50% למכפ"ל ו-50% לבגרות או 70% למכפ"ל ו-30% לבגרות) וכן בציון הכללי במכפ"ל שבו נעשה שימוש (הציון הרב-תחומי, הציון בדגש כמותי או הציון בדגש מילולי). נמצא שתקפיות כל ציוני סכס 1 גבוהות ודומות זו לזו (בין 0.60 ל-0.62). השוואת הממדים השונים שבהם נבדלים ששת ציוני סכס 1 מלמדת שתוקף ציוני סכס 1 המבוססים על ציון מכפ"ל בדגש מילולי מעט יותר נמוך. ההבדלים במשקלות היחסיים של מכפ"ל ובגרות אינם משפיעים באופן עקבי על התוקף. השוואה בין תוקף ציון סכס 1, על חלופותיו השונות, לתוקף מרכיביו, מלמדת שתרומת הבגרות לתוקף ציון הסכס היא קטנה.

מערכות מו"ר ומרק"ם: חשוב לציון שמערכות מו"ר ומרק"ם לא נועדו לנבא את ההישגים בלימודים הקדם-קליניים, ובפרט לא בלימודי שנה א'. עם זאת, נמצא שלציונים הסופיים במו"ר ובמרק"ם תוקף בינוני (0.35 ו-0.37, בהתאמה) בניבוי הישגים אלה.

ציוני סכס 2: עשרים וארבעה ציוני סכס 2, המבוססים על צירוף של ציוני סכס 1 והציון הסופי במו"ר או במרק"ם, הוגדרו במחקר. ציונים אלה נבדלו זה מזה בממדים שבהם נבדלו ציוני סכס 1, בחזאי הלא-קוגניטיבי שעליו הם מתבססים (מו"ר או מרק"ם) ובמשקלות שניתנו לסכס 1 ולחזאי זה (50% לסכס 1 ו-50% לחזאי הלא-קוגניטיבי או 30% לסכס 1 ו-70% לחזאי הלא-קוגניטיבי).

נמצא שתקפיות כל ציוני סכס 2 גבוהות ודומות זו לזו (בין 0.56 ל-0.64). השוואת הממדים השונים שבהם נבדלים 24 ציוני סכס 2 מלמדת שהשפעה העיקרית על התוקף קשורה למשקלו של

סכס 1 בסכס 2 : התוקף גבוה יותר כאשר משקל סכס 1 בסכס 2 גבוה יותר (50% לעומת 30%).
השוואה בין תוקף ציון סכס 2, על חלופותיו השונות, לתוקף מרכיביו, מלמדת שבאופן כללי תוקף ציון סכס 2 דומה לתוקף ציון סכס 1. צירוף ציוני מו"ר או מרק"ם לציון סכס 1 אינו משנה באופן משמעותי את תוקף הניבוי ביחס לקריטריון הסכולסטי שנבדק.

2. תוצאות ניתוחי רגרסיה מרובה

סוגיית האופטימיזציה של משקלות החזאים במערכת הברירה לא קיבלה מקום נרחב במחקר זה, וזאת בעיקר על רקע העובדה שמקסימיזציה של התוקף בניבוי ההישגים בלימודים הקדם-קליניים, ובפרט בניבוי ציוני שנה א', אינה המטרה הבלעדית של תהליכי הברירה הקיימים כיום לחוגים לרפואה. לסוגיית תוקף הניבוי בעיצוב תהליך הברירה, בחרנו להתייחס על ידי בדיקת תוקף הניבוי של צירופים מסוימים של מרכיבי תהליך הברירה שהוגדרו אפריורית (ציוני סכס 1 וסכס 2).

לצד זאת, הצגנו ממצאים מניתוח רגרסיה מרובה שעושה אופטימיזציה חלקית של משקלות החזאים¹², מה שמאפשר לבחון את המחיר, במונחי תוקף ניבוי, של שימוש בצירופים הקיימים על פני צירופים אופטימליים יותר. ששת ניתוחי הרגרסיה שנבדקו כללו כחזאים את הבררות, את אחד משלושת הציונים הכלליים במכפ"ל, ואת הציון הכללי במו"ר או במרק"ם. נמצא שהמתאמים המרובים מששת ניתוחי הרגרסיה גבוהים ודומים זה לזה (בין 0.65 ל-0.67). את גובהם של המתאמים ניתן להשוות לתוקף הניבוי של החזאים הבודדים או של הצירופים שהוגדרו מראש (ציוני סכס 1 וציוני סכס 2), שמכסים בקירוב את משתני הברירה שמשמשים בפועל (ראו לוח 3), כדי להעריך מה גודל ההפסד – במונחים של תוקף ניבוי – של ויתור על אופטימליות. אשר למשקלות המרכיבים, בשקלול האופטימלי (לצורכי מקסימיזציה של התוקף) ניתנת בקירוב כמחצית מן המשקל לציון הכללי במכפ"ל, והמחצית הנותרת נחלקת שווה בשווה, בין הבררות למו"ר או למרק"ם.

3. תוקף הניבוי של תהליך ברירה דו-שלבי

הממצאים שתוארו עד כה : מקדמי מתאם פשוטים או מרובים, התייחסו לתוקף הניבוי של חזאים שונים תחת המצב הפשוט שבו הם משמשים – בנפרד או בצירופים שלהם – בתהליך ברירה חד-שלבי. כאמור, בחוגים שנכללים במחקר זה נעשה שימוש באותם חזאים, אך בתהליך ברירה דו-שלבי. במטרה להעריך את תוקף הניבוי של תהליך הברירה בפועל, נערכה הדמיה שהתבססה על הסטטיסטיים (מקדמי מתאם בין החזאים לקריטריון ובין החזאים לבין עצמם) ועל אפיוני תהליך הברירה (יחסי הברירה בשלב הראשון ובשלב השני) שתואמים לנתוני המחקר. ממצאי ההדמיה הראו ש :

- ציון שנה א' בקרב המתקבלים בעקבות ברירה חד-שלבית על פי ציון סכס 1 (בחרנו באב-טיפוס : ציון מכפ"ל ר"ת והבררות במשקלות שווים) גבוה ב-1.07 סטיות תקן מממוצע ציון שנה א' בקרב המועמדים.

¹² השימוש במונח "חלקי" מבטא את העובדה שהשתמשנו במודל של שלושה חזאים (בגרות, ציון כללי במכפ"ל וציון כללי במו"ר או במרק"ם), שהשניים האחרונים מתוכם הם צירוף, במשקלות שנקבעו מראש, של חזאים.

- ציון שנה א' בקרב המתקבלים בעקבות ברירה דו-שלבית שבה משתמשים בציון סכס 1 בשלב הראשון ובציון הכללי במו"ר בשלב השני גבוה ב-0.93 סטיות תקן מממוצע ציון שנה א' בקרב המועמדים.

- ציון שנה א' בקרב המתקבלים בעקבות ברירה דו-שלבית שבה משתמשים בציון סכס 1 בשלב הראשון ובסכס 2, שהוגדר כצירוף של סכס 1 ומו"ר במשקלות שווים בשלב השני, גבוה ב-1.11 סטיות תקן מממוצע ציון שנה א' בקרב המועמדים.

הממצאים הללו מעידים שהכללת רכיב לא-קוגניטיבי בשלב השני של תהליך ברירה דו-שלבי בשקלול עם ציון דמוי סכס 1 אינה פוגעת בתוקף הניבוי (ואולי אף מעלה אותו מעט) לעומת מצב שבו נעשית ברירה חד-שלבית על פי ציון סכס 1 לבדו. שימוש ברכיב הלא-קוגניטיבי בשלב השני של תהליך ברירה דו-שלבי במשקל של 100% מוריד מעט את תוקף הניבוי (כמשוער, לטובת מטרות אחרות) אך עדיין מותיר רמה גבוהה של תוקף ניבוי.

ההדמיה שתוארה לעיל השתמשה, כאמור, בהגדרות מסוימות של ציון סכס 1 ושל ציון סכס 2. עם זאת, ניתן להתייחס לתמונה שהיא מייצרת כבת הכללה מעבר לחלופות שנבדקו ביחס להגדרת ציון סכס 1 וציון סכס 2, בהינתן הדמיון בתקפויות הניבוי של החלופות השונות.

יכולת ההשוואה של ממצאי המחקר הנוכחי עם ממצאים על תוקף הניבוי של כלי המיון ללימודי רפואה בעולם היא מוגבלת: מרבית המידע נצבר מלימודי הרפואה בצפון אמריקה, שמתקיימים במתכונת ארבע-שנתית ושכלי המיון המשמשים בהם הם מבחן הכניסה ללימודי רפואה (MCAT) וממוצע ציוני הלימודים לתואר הראשון. התמונה העולה ממחקרי תוקף הניבוי האלה (למשל, Julian, 2005) היא שתוקף שני החזאים בניבוי הציונים בשנים הקדם-קליניות הוא גבוה (0.59 ו-0.54, בהתאמה), ושתוקף הניבוי של צירוף שני החזאים (0.71) עולה על תוקף הניבוי של כל אחד מהם בנפרד, כאשר התרומה השולית של MCAT גדולה מהתרומה השולית של ממוצע ציוני הלימודים לתואר הראשון. ממצאים אלה מתיישבים עם ממצאי המחקר הנוכחי, להוציא פער מסוים בין תוקף הבגרות במחקר הנוכחי לתוקף ממוצע ציוני הלימודים לתואר הראשון. אשר למדדים הלא-קוגניטיביים המשמשים בעולם במיון (ה-MMI), תוקף הניבוי שלהם נמצא (למשל, Eva et al., 2012), כצפוי, קשור יותר למדדי תוצאה בעלי אוריינטציה קלינית, שלא נבדקו במחקר הנוכחי. לאחרונה נערך מחקר בבריטניה (Husbands & Dowell, 2013), שסיפק אישוש לממצאים הקודמים, שהביצוע ב-MMI קשור בעיקר להערכות קליניות, אך גם להערכות הטעונות בידע. קיימת הסכמה בדבר קיומם של רכיבים קוגניטיביים במדדים כדוגמת ה-MMI (Eva et al., 2009). ביצוע מיטבי במדדים אלה מחייב הבנה של מצבים חברתיים מורכבים, יכולת ביטוי וכיוצא בזה. ממצאי המחקר הנוכחי, שמעידים על קשר בינוני בין ציוני מו"ר ומרק"ם ובין ציוני שנה א' מתיישבים, אם כן, עם מסקנה זו.

יכולת ההשוואה של ממצאי המחקר הנוכחי עם ממצאים קודמים על תוקף הניבוי של כלי המיון ללימודי רפואה בישראל היא מוגבלת, וזאת בשל הבדלים מתודולוגיים משמעותיים בין המחקרים, כגון ההחלטה לכלול במחקר הנוכחי רק סטודנטים שהחלטת הקבלה שלהם התבססה על תעודת הבגרות (וציון מכפ"ל) ולא על חלופות לתעודת הבגרות, הרכב החוגים במחקרים השונים, שיטת התיקון לקיצוץ תחום ועוד. כך או כך, כפי שצוין בסקירת הספרות, גם במחקרי

העבר שנסקרו נמצאו הבדלים לא מבוטלים בין המחקרים השונים במקדמי התוקף של החזאים. כדי להבין יותר את מקור ההבדלים בממצאים בין מחקרי העבר ובינם ובין המחקר הנוכחי, צריך לחזור ולנתח את הנתונים של מחקרי העבר ושל המחקר הנוכחי בשיטה אחידה. לבסוף, חשוב לדון במידת הדיוק וביכולת ההכללה של ממצאי המחקר הנוכחי. ראשית, יש להזכיר שהתיקון לקיצוץ תחום נשען על הנחות (לינאריות, הומוסקדסטיות) שלא בהכרח מתקיימות. המחקרים שנעשו על הפרה של הנחות התיקון לקיצוץ תחום (למשל, Greener & Osburn, 1979, 1980) מצאו שהאומדנים המתוקנים הינם עמידים למדי בפני הפרות של הנחת ההומוסקדסטיות ורגישים יותר (בשני הכיוונים) לחוסר לינאריות, אך בכל מקרה הם מדויקים יותר מאשר המתאמים הנצפים (Bobko, Roth, & Bobko, 2001; Linn, 1983). ההסתייגויות הללו הן כלליות במובן שהן חלות על כל מחקר תוקף שמחייב תיקון לקיצוץ תחום. הן מועצמות במחקר הנוכחי שהתמודד עם תהליך ברירה סלקטיבי ומורכב במיוחד. ממצאי הבדיקה של אמינות התיקון במצבי ברירה קיצוניים, שתוארה בנספח 1, מתיישבים עם הספרות שהוזכרה לעיל, בנוגע להיותם של המתאמים המתוקנים אומדן טוב יותר של המתאם באוכלוסייה מאשר המתאמים הנצפים. אפשר, אם כן, שחלק מהערכים המתוקנים שהוצגו במחקר זה מאופיינים בנטייה קלה לאומדן יתר, אך הם בוודאי קרובים יותר לאמת מאשר הערכים הלא מתוקנים. בנוסף לסוגיית ההפרות האפשריות של הנחות התיקון, נזכיר שהערך המייצג את שונות משתנה הברירה בקרב המועמדים לצורך התיקון לקיצוץ תחום (עבור השלב הראשון של הברירה) מחושב כשונות הממוצעת בקרב המועמדים לכל אחד מחוגי הלימוד (במחזורים תשנ"ב ותשנ"ג). ערך זה הוא משותף לכל מחקרי התוקף ומאפשר השוואתיות של התוקף בתחומי לימוד שונים. הוא אינו מייצג במדויק את שונות משתנה הברירה בקרב המועמדים לרפואה במחזורי המחקר, ובפרט לא את השונות בקרב המועמדים שנתוני מועמדותם לא התבססו על חלופה לתעודת הבגרות. האמור לעיל אינו מיועד לערער את המסקנות מממצאי המחקר אלא להציג את מידת הדיוק של האומדנים (של מקדמי התוקף ובפרט של ההשוואות ביניהם) בזהירות הדרושה.

שנית, באופן דומה, חשוב להציג בזהירות את מידת הדיוק של ממצאי ההדמיה שנועדה להמחיש את תוקף הניבוי של תהליך ברירה דו-שלבי. הדמיה זו התבססה על הנחה של התפלגות רב-משתנית נורמלית שלא בהכרח מתארת במדויק את ההתפלגות המשותפת של המשתנים בקרב המועמדים לרפואה במוסד ובמחזור. נציין, עם זאת, שהדמיה שהתבססה על הנחה של התפלגות רב-משתנית אחידה הפיקה תוצאות דומות. משמעות הדבר היא שהתוצאות עמידות ביחס לחלופות התפלגותיות מסוימות, אך הן עדיין מספקות ערכים קרובים ולא אומדנים מדויקים. לבסוף, יש מקום להתייחס לשונות בין מוסדות ומחזורים בממצאי המחקר. ערכי הטווח הבין-רבעוני שהוצגו עבור הסטטיסטיים השונים מלמדים שהתמונה שהתקבלה ביחידות העיבוד השונות אינה אחידה. עובדה זו תורמת אף היא להמלצתנו לסייג את מידת הביטחון בדיוק וביכולת ההכללה של הממצאים.

מקורות

אורן, כ', קנת-כהן, ת' וברונר, ש' (2007). **נתונים מקובצים על תוקף מערכת המיון לאוניברסיטאות בחיזוי הצלחה בלימודי שנה א' (מחזורים תשס"ג - תשס"ה) (דוח מס' 342)**. ירושלים: מרכז ארצי לבחינות ולהערכה.

ברונר, ש' (2006). **תוקף מכפ"ל ובגרות בניבוי ציון שנה ראשונה באוניברסיטאות מחזורים תשנ"א**

- תשס"ג (2002/3 - 1990/1) (דוח טכני פנימי מס' 162). ירושלים: מרכז ארצי לבחינות ולהערכה.
- הלשכה המרכזית לסטטיסטיקה (2013). **שנתון סטטיסטי לישראל 2013** (מס' 64). ירושלים: המחבר.
- ורון, ת' (2012, 27 בנובמבר). מישהו מטפל בך: מהו החוג הטוב ביותר בישראל ללימודי רפואה? **פורבס ישראל**. נשלף ב-15.10.2013, מ-
<http://www.forbes.co.il/news/new.aspx?0r9VQ=FIGM>
- חדד, א' (2012). **הערכת עמיתים בקרב סטודנטים לרפואה, ככלי לבחינת תוקף ניבוי מבחני הקבלה הלא-קוגניטיביים (מרק"ם – מערכת ראיונות קצרים מובנים) ללימודי רפואה באוניברסיטה העברית**. עבודת מוסמך, האוניברסיטה העברית בירושלים.
- AERA (American Educational Research Association), APA (American Psychological Association), & NCME (National Council on Measurement in Education). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bobko, P., Roth, P. L., & Bobko, C. (2001). Correcting the effect size of d for range restriction and unreliability. *Organizational Research Methods, 4*, 46-61.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- De Corte, W., Lievens, F., & Sackett, P. R. (2006). Predicting adverse impact and mean criterion performance in multistage selection. *Journal of Applied Psychology, 91*, 523-537.
- Donnon, T., Paolucci, E. O., & Violato, C. (2007). The predictive validity of the MCAT for medical school performance and medical board licensing examinations: A meta-analysis of the published research. *Academic Medicine, 82*, 100-106.
- Dunleavy, D. M., Kroopnick, M. H., Dowd, K. W., Searcy, C. A., & Zhao, X. (2013). The predictive validity of the MCAT exam in relation to academic performance through medical school: A national cohort study of 2001-2004 matriculants. *Academic Medicine, 88*, 666-671.
- Edwards, D., Friedman, T., & Pearce, J. (2013). Same admissions tools, different outcomes: A critical perspective on predictive validity in three undergraduate medical schools. *BMC Medical Education, 13*, 173.
- Eva, K. W., Reiter, H. I., Rosenfeld, J., & Norman, G. R. (2004). The ability of the multiple mini-interview to predict preclerkship performance in medical school. *Academic Medicine, 79*, S40-S42.
- Eva, K. W., Reiter, H. I., Trinh, K., Wasi, P., Rosenfeld, J., & Norman, G. R. (2009). Predictive validity of the multiple mini-interview for selecting medical trainees. *Medical Education, 43*, 767-775.
- Eva, K. W., Reiter, H. I., Rosenfeld, J., Trinh, K., Wood, T. J., & Norman, G. R. (2012). Association between a medical school admission process using the multiple mini-interview and national licensing examination scores. *Journal of American Medical Association, 308*, 2233-2240.
- Gafni, N., Moshinsky, A., Eisenberg, O., Zeigler, D., & Ziv, A. (2012). Reliability estimates: Behavioural stations and questionnaires in medical school admissions. *Medical Education, 46*, 277-288.
- Greener, J. M., & Osburn, H. G. (1979). An empirical study of the accuracy of

- corrections for restriction in range due to explicit selection. *Applied Psychological Measurement*, 3, 31-41.
- Greener, J. M., & Osburn, H. G. (1980). Accuracy of corrections for restriction in range due to explicit selection in heteroscedastic and nonlinear distributions. *Educational and Psychological Measurement*, 40, 337-346.
- Gulliksen, H. (1987). *Theory of mental tests*. Hillsdale, NJ: Erlbaum. (Original work published 1950)
- Husbands, A., & Dowell, J. (2013). Predictive validity of the Dundee multiple mini-interview. *Medical Education*, 47, 717-725.
- Julian, E. R. (2005). Validity of the Medical College Admission Test for predicting medical school performance. *Academic Medicine*, 80, 910-917.
- Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgement tests: A comparison in terms of predictive validity. *Journal of Applied Psychology*, 91, 1181-1188.
- Linn, R. (1983). Pearson selection formulas: Implications for studies of predictive bias and estimates of educational effects in selected samples. *Journal of Educational Measurement*, 20, 1-15.
- Martinho, A. M. (2012). Becoming a doctor in Europe: Objective selection systems [Electronic version]. *Virtual Mentor*, 14, 984-988.
- Medical School. (n.d.). In Wikipedia. Retrieved October, 28, 2013, from http://en.wikipedia.org/wiki/Medical_school.
- Oren, C., Kennet-Cohen, T., Turvall, E., & Allalouf, A. (2014). Demonstrating the validity of three general scores of PET in predicting higher education achievement in Israel. *Psicothema*, 26, 117-126.
- Prideaux, D., Roberts, C., Eva, K., Centeno, A., Mccrorie, P., Mcmanus, C. et al. (2011). Assessment for selection for the health care professions and specialty training: Consensus statement and recommendations from the Ottawa 2010 conference. *Medical Teacher*, 33, 215-223.
- Reiter, H. I., Eva, K. W., Rosenfeld, J., & Norman, G. R. (2007). Multiple mini-interviews predict clerkship and licensing examination performance. *Medical Education*, 41, 378-384.
- Sackett, P. R., & Roth, L. (1996). Multistage selection strategies: A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology*, 49, 549-572.
- Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85, 112-118.
- Siu, E., & Reiter, H. I. (2009). Overview: What's worked and what hasn't as a guide towards predictive admissions tool development. *Advances in Health Science Education*, 14, 759-775.
- Ziv, A., Rubin, O., Moshinsky, A., Gafni, N., Kotler, M., Dagan, Y. et al. (2008). MOR: A simulation-based assessment center for evaluating the personal and interpersonal qualities of medical school candidates. *Medical Education*, 42, 991-998.

נספח 1

אמינות התיקון לקיצוץ תחום במצבי ברירה קיצוניים – הערה מתודולוגית

בבדיקת תוקף הניבוי במסגרות לימוד סלקטיביות, כדוגמת החוג לרפואה, מושמעת לעיתים קרובות הסתייגות אפריורית מערכם הפוטנציאלי של הממצאים, בהתבסס על הטענה שיחס הברירה הקיצוני מלווה בקיצוץ תחום משמעותי כל כך שלא ניתן לשחזר מן המתאם הנצפה (שיהיה קרוב לוודאי נמוך מאוד ואפילו אפס) את המתאם באוכלוסיית המועמדים טרם ברירה. בדקנו טענה זו בשני הקשרים: 1. המתאם בין ציון מכפ"ל לציון בחינת מתא"ם (בחינת קבלה לתואר מתקדם בפסיכולוגיה) בקרב נבחני מתא"ם בשנים 2002-2013 (N=9,000), ו-2. המתאם בין ציון מכפ"ל וממוצע הבגרות בקרב נבחני עברית במכפ"ל שהחלו את לימודיהם באוניברסיטה בין השנים 2001-2010 (N=304,427)¹³. בשני ההקשרים נבדק המתאם בין שני המשתנים באוכלוסייה כולה וכן בתת-קבוצה של תצפיות שלהן ציון מכפ"ל מעל 720. האוכלוסייה כולה מייצגת את אוכלוסיית המועמדים, שהמתאם שמחושב בקרבה הוא שמעניין אותנו, ואילו תת-הקבוצה של תצפיות עם ציון מכפ"ל מעל 720 מייצגת את מדגם המתקבלים בעקבות תהליך ברירה שמתבסס על ציון מכפ"ל עם נקודת חתך של 720. נקודת החתך של 720 נבחרה כך שממוצע מכפ"ל במדגם המתקבלים (כ-740) יהיה דומה למה שמוצאים בקרב סטודנטים לרפואה. יחס הברירה שהוכתב על ידי בחירה זו היה בסדר גודל של מה שמוצאים במיון לרפואה (8:1 בהקשר הראשון ו-11:1 בהקשר השני). המתאם המחושב ב"מדגם המתקבלים" ("המתאם הנצפה") בשני ההקשרים מושפע, אם כן, מקיצוץ תחום ניכר (סטיית התקן של החוזאי הממין בקרב ה"מועמדים" גדולה פי 4-5 מסטיית התקן בקרב ה"מתקבלים"), ונשאלת השאלה באיזו מידה התיקון לקיצוץ תחום משחזר את המתאם בקרב ה"מועמדים" (שבמקרה זה אנו יכולים לחשב אותו גם ישירות). הסטטיסטיים הרלוונטיים מוצגים בלוח 8.

לוח 8: המחשה של איכות התיקון לקיצוץ במצבי ברירה קיצוניים בשני הקשרים

מתאם מתוקן	מתאם נצפה	מתא"ם / בגרות		מכפ"ל		מספר תצפיות	
		ס"ת	ממוצע	ס"ת	ממוצע		
הקשר 1: מתאם בין מכפ"ל למתא"ם							
-	0.71	21	101	67	652	9,000	"מועמדים"
0.77	0.28	13	122	16	740	1,166	"מתקבלים"
הקשר 2: מתאם בין מכפ"ל לבגרות							
-	0.64	10	95	87	614	304,427	"מועמדים"
0.70	0.18	7	104	17	740	27,178	"מתקבלים"

ניתן להיווכח שבשני ההקשרים גם במצב ברירה קיצוני, כדוגמת זה המתקיים ברפואה, המתאם הנצפה אינו אפסי, ובפרט, שבאמצעות תיקון לקיצוץ תחום, ניתן לקבל אומדן קרוב בערכו למתאם טרם ברירה.

¹³ בשני ההקשרים נלקחה תצפית אחת לאדם, כאשר בהקשר הראשון במצב של ריבוי תצפיות לאדם נלקח ציון מכפ"ל מרבי (על מנת לשחזר את הסיטואציה בקרב מועמדים לרפואה) וציון מתא"ם ראשון (שהוא תקף יותר מאשר הציון המרבי), ובהקשר השני במצב של ריבוי תצפיות לאדם, סוננו תצפיות שבהן, כאשר היה ריבוי תצפיות לאדם, היו הבדלים ביניהן בממוצע הבגרות או בציון מכפ"ל.

נספח 2

נוסחת התיקון לקיצוץ תחום¹⁴

אנו מבקשים לאמוד את המתאם המתקיים באוכלוסייה (טרם ברירה) בין זוג משתנים: X ו- Y , כאשר ידוע לנו ("נצפה") המתאם ביניהם במדגם (שעבר ברירה).

U מציין את המשתנה שעל-פיו בוצעה הברירה.

המשתנה U עבר ברירה ישירה והמשתנים Y ו- X עברו ברירה עקיפה, כיוון שהם מתואמים עם U . S ו- s יצינו סטיות תקן ו- R ו- r יצינו מתאמים, כאשר אותיות קטנות מתייחסות לסטטיסטיים במדגם (שעבר ברירה) ואותיות גדולות מתייחסות לאומדנים של פרמטרים באוכלוסייה (טרם ברירה).

נתונים ברמת האוכלוסייה (S) ידועים רק לגבי המשתנה שעבר ברירה ישירה. אם מתקיימות ההנחות הבאות:

1. הרגרסיות של Y על U ושל X על U הן לינאריות

2. השונויות (והשונויות המשותפות) המותנות של Y ושל X אינן תלויות ב- U (הומוסקדסטיות) אזי:

$$R_{xy} = \frac{r_{xy} + w_u r_{ux} r_{uy}}{\sqrt{(1 + w_u r_{ux}^2)(1 + w_u r_{uy}^2)}}$$
$$\cdot w_u = \left(\frac{S_u^2}{s_u^2} - 1 \right) \text{ כאשר}$$

¹⁴הניסוח המוצג כאן של נוסחאות התיקון לקיצוץ תחום לקוח מ-Linn, 1983.

נספח 3

סטטיסטיים נצפים

1. מתאמים פשוטים בין החזאים לקריטריון

לוח 9א: מתאמים נצפים בין ציון שנה א' ובין ממוצע ציוני הבגרות, הציונים בתחומי מכפ"ל והציונים הכלליים במכפ"ל והטווח הבין-רבעוני של המתאמים (בסוגריים)

P_Q 7	P_V 6	P_{MD} 5	E 4	Q 3	V 2	B 1
0.18 (0.09-0.28)	0.10 (0.03-0.20)	0.15 (0.07-0.27)	0.14 (0.13-0.18)	0.14 (0.07-0.24)	0.01 (-0.10-0.11)	0.05 (-0.12-0.25)

לוח 9ב: מתאמים נצפים בין ציון שנה א' ובין ציוני סכם 1 והטווח הבין-רבעוני של המתאמים (בסוגריים)

P_{Q70B30} 13	P_{Q50B50} 12	P_{V70B30} 11	P_{V50B50} 10	$P_{MD70B30}$ 9	$P_{MD50B50}$ 8
0.22 (0.11-0.28)	0.20 (0.07-0.34)	0.14 (0.00-0.28)	0.14 (-0.03-0.27)	0.20 (0.06-0.33)	0.18 (-0.02-0.33)

לוח 9ג: מתאמים נצפים בין ציון שנה א' ובין הציונים במערכות מו"ר ומרק"ם והטווח הבין-רבעוני של המתאמים (בסוגריים)

MI 19	MO 18	ST_MI 17	ST_MO 16	Q_B 15	Q_J 14
0.09 (0.03-0.17)	0.17 (0.15-0.26)	0.05 (-0.01-0.13)	0.14 (0.09-0.24)	0.10 (-0.02-0.26)	0.09 (0.01-0.20)

לוח 19: מתאמים נצפים בין ציון שנה א' ובין ציוני סכם 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

$(P_{V50B50})_{30}$ MO70 25	$(P_{V50B50})_{50}$ MO50 24	$(P_{MD70B30})_{30}$ MO70 23	$(P_{MD70B30})_{50}$ MO50 22	$(P_{MD50B50})_{30}$ MO70 21	$(P_{MD50B50})_{50}$ MO50 20
0.23 (0.15-0.36)	0.24 (0.09-0.31)	0.25 (0.20-0.33)	0.27 (0.20-0.35)	0.26 (0.19-0.38)	0.28 (0.17-0.36)

לוח 19 (המשך): מתאמים נצפים בין ציון שנה א' ובין ציוני סכם 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

$(P_{Q70B30})_{30}$ MO70 31	$(P_{Q70B30})_{50}$ MO50 30	$(P_{Q50B50})_{30}$ MO70 29	$(P_{Q50B50})_{50}$ MO50 28	$(P_{V70B30})_{30}$ MO70 27	$(P_{V70B30})_{50}$ MO50 26
0.26 (0.23-0.33)	0.29 (0.24-0.34)	0.27 (0.22-0.39)	0.30 (0.21-0.38)	0.22 (0.17-0.33)	0.22 (0.15-0.33)

לוח 19 (המשך): מתאמים נצפים בין ציון שנה א' ובין ציוני סכם 2 והטווח הבין-רבעוני של המתאמים (בסוגריים)

$(P_{V50B50})_{30}$ MI70 37	$(P_{V50B50})_{50}$ MI50 36	$(P_{MD70B30})_{30}$ MI70 35	$(P_{MD70B30})_{50}$ MI50 34	$(P_{MD50B50})_{30}$ MI70 33	$(P_{MD50B50})_{50}$ MI50 32
0.18 (0.08-0.32)	0.21 (0.07-0.36)	0.18 (0.07-0.31)	0.22 (0.13-0.35)	0.20 (0.11-0.32)	0.24 (0.12-0.38)

**לוח 19 (המשך): מתאמים נצפים בין ציון שנה א' ובין ציוני סכס 2
והטווח הבין-רבעוני של המתאמים (בסוגריים)**

(P _Q 70B30)30 MI70 43	(P _Q 70B30)50 MI50 42	(P _Q 50B50)30 MI70 41	(P _Q 50B50)50 MI50 40	(P _V 70B30)30 MI70 39	(P _V 70B30)50 MI50 38
0.19 (0.09-0.29)	0.24 (0.14-0.31)	0.20 (0.12-0.32)	0.25 (0.16-0.39)	0.16 (0.06-0.28)	0.18 (0.08-0.33)

2. תוצאות ניתוחי רגרסיה מרובה

**לוח 10א: תוצאות ניתוחי רגרסיה מרובה לניבוי ציון שנה א' ע"י בגרות, מכפ"ל ומו"ר
והטווח הבין-רבעוני של הסטטיסטים (בסוגריים) – ערכים נצפים**

בגרות, מכפ"ל כמותי ומו"ר			בגרות, מכפ"ל מילולי ומו"ר				בגרות, מכפ"ל ר"ת ומו"ר				
מקדמי רגרסיה			מקדמי רגרסיה				מקדמי רגרסיה				
מתאם	מקדמי רגרסיה			מתאם	מקדמי רגרסיה			מתאם	מקדמי רגרסיה		
מרובה	MO	P _Q	B	מרובה	MO	P _V	B	מרובה	MO	P _{MD}	B
0.35	0.23	0.25	0.23	0.31	0.20	0.14	0.17	0.34	0.22	0.22	0.22
(0.26-)	(0.15-)	(0.16-)	(0.09-)	(0.20-)	(0.10-)	(0.03-)	(0.08-)	(0.22-)	(0.12-)	(0.10-)	(0.11-)
(0.42)	(0.28)	(0.32)	(0.40)	(0.37)	(0.28)	(0.27)	(0.35)	(0.40)	(0.27)	(0.33)	(0.39)

**לוח 10ב: תוצאות ניתוחי רגרסיה מרובה לניבוי ציון שנה א' ע"י בגרות, מכפ"ל ומרק"ם
והטווח הבין-רבעוני של הסטטיסטים (בסוגריים) – ערכים נצפים**

בגרות, מכפ"ל כמותי ומרק"ם			בגרות, מכפ"ל מילולי ומרק"ם				בגרות, מכפ"ל ר"ת ומרק"ם				
מקדמי רגרסיה			מקדמי רגרסיה				מקדמי רגרסיה				
מתאם	מקדמי רגרסיה			מתאם	מקדמי רגרסיה			מתאם	מקדמי רגרסיה		
מרובה	MI	P _Q	B	מרובה	MI	P _V	B	מרובה	MI	P _{MD}	B
0.31	0.14	0.23	0.19	0.28	0.12	0.13	0.15	0.30	0.13	0.20	0.18
(0.22-)	(0.08-)	(0.15-)	(-0.01-)	(0.17-)	(0.05-)	(-0.01-)	(-0.04-)	(0.17-)	(0.07-)	(0.07-)	(-0.01-)
(0.45)	(0.22)	(0.29)	(0.42)	(0.41)	(0.21)	(0.27)	(0.36)	(0.42)	(0.21)	(0.31)	(0.41)