

Review



מגזין מכון תנובה למחקר

גיליון 53 ינואר 2018

**מהפכת ה-DATA
ברפואה**

ד"ר מיכל חמו לוטם

**שיטות מחקר
בתחום הבריאות**

ד"ר עפרה קלטור ליבוביץ'

**Big Data - המהפכה
שתשנה את הדרך בה אנו
חיים, עובדים וחושבים**
ד"ר אייל בנימין

**השימוש בביג דאטה
בשירות הרפואה**
פרופ' גדעון קורן,
מאיה עדולמי
ופרופ' ורדה שלו

**מיקרוביום - ביג דאטה,
אתגר ענק, פוטנציאל עצום**
מיה אלחלל



- 3** מהפכת ה-DATA ברפואה
ד"ר מיכל חמו לוטם
- 8** שיטות מחקר בתחום הבריאות
ד"ר עפרה קלטר ליבוביץ'
- 12** Big Data - המהפכה שתשנה את הדרך בה
אנו חיים, עובדים וחושבים
ד"ר אייל בנימין
- 16** השימוש בביג דאטה בשירות הרפואה
פרופ' גדעון קורן, מאיה עדולמי
פרופ' ורדה שלו
- 20** מיקרוביום - ביג דאטה,
אתגר ענק, פוטנציאל עצום
מיה אלחלל
- 23** תיאור מקרה

עולם הרפואה והמחקר, מימיו של היפוקרטס ועד ימינו אנו, משנה את פניו בכל תקופה. במבט היסטורי אפשר למנות פריצות דרך משמעותיות גדולות, אשר שינו את פני הרפואה כליל, אם גילוי החיידקים, או גילוי האנטיביוטיקה והחיסונים, אם גילוי קרני הרנטגן או אמצעי ההרדמה, וכמובן הגנטיקה, פיתוח תרופות רבות, ועוד רבים אחרים, שהצעידו את הרפואה קדימה ושינו את פניה. לעתים עוצמתם של הגילויים מוערכת רק בראיה היסטורית, ולעתים אנו זוכים להבנים בזמן אמת.

לאן עוד יכולה להתפתח הרפואה? לו שאלנו את עצמנו לפני מספר עשורים, טרום עידן האינטרנט והתפתחות הדיגיטל, לא היינו משערים בנפשנו, ולא רואים בדמיוננו, את מה שמכונה היום המהפכה התעשייתית הרביעית, אשר עומדת לשנות את פני עולם הרפואה לבלי היכר: מהפיכת המידע.

בעידן הרפואה ה'שנה' המידע הרפואי התבסס על מספר מצומצם של מקורות: תיאור וסיפור המחלה על ידי המטופל, ממצאי הבדיקה הפיסיקלית, וממצאי בדיקות המעבדה וההדמיה. הללו שולבו עם מידע מחקרי מהספרות המדעית. השנים האחרונות מתאפיינות בפיתוח של שיטות מחשוב, שמסוגלות לנתח כמויות נתונים ענקיות (Big Data), ולהבחין בגילויים רפואיים ומדעיים שלא התגלו בניתוחי נתונים קלאסיים עד היום.

ה"ביג דאטה" ולמידת המכונה עתידים לשנות את עולם הרפואה מן היסוד, ולהביא לפריצות מדעיות בכל הנוגע לפיתוח רפואה מותאמת אישית, גילוי מוקדם, דיוק מוגבר באבחנות ומניעת מחלות.

השימוש בביג דאטה כבר נוכח בהווה אבל הפוטנציאל העתידי שלו הוא עצום.

בחרנו לתת במה לענין חשוב זה בעודו באיבו, ולהביאו לידיעת קוראנו, על הפוטנציאל העצום הגלום בו.

קריאה מהנה

טליה לביא
עורכת ראשית

פרופ' זמיר הלפרן
יועץ מדעי
מכון תנובה למחקר

עורכת ראשית: טליה לביא
עורכת משנה: רותי אבירי
יועץ מדעי: פרופ' זמיר הלפרן
מנהלת הפרויקט: נגה שורץ

כתובת למכתבים: תנובה
ת.ד. 33, רמת השרון, מיקוד 4710001

Review
מגזין מכון תנובה למחקר



כל גליונות Review לרשותכם באתר שלנו:
www.tnuva-research.co.il



מהפכת ה-DATA ברפואה

ד"ר מיכל חמו לוטם

M.D, רופאת ילדים, מומחית בחדשנות רפואית, ומייסדת אושיה - נשים מנהיגות בריאות

מהפכת הדאטה (DATA) והבינה המלאכותית משנה את כמויות ומקורות המידע (רק 10% יגיע מהמקורות הרפואיים); הופכת גם את המטופלים ליוצרי נתונים, ויוצרת מציאות בה המידע והבעלות עליו מפוצלים בין תאגידים, מערכות הבריאות והפרט הבינה המלאכותית נכנסת לסייע באינספור חזיתות, ושלל אתגרים ואופקים נפתחים.

המהפכה התעשייתית הרביעית

המדינות המפותחות הגיעו לקצה יכולתן להשקיע ברפואה, כיוון שעלויותיה עולות יותר מהעלייה בתוצר.

במקביל מתרחש ברפואה שינוי פרדיגמה, שמקורו בהפיכת העולם, והרפואה בתוכו, לדיגיטליים, ובמהפכת המידע והבינה המלאכותית (1). בסטנפורד פיתחו אלגוריתם המאבחן צילומי חזה טוב ממומחי רדיולוגיה (2), סין דיווחה על רובוט שעבר ראשון את מבחני ההסמכה ברפואה, אחרי שהזון במידע של עשרות ספרי רפואה. שינוי הפרדיגמה ברפואה בולט מאד בציר המשלב ביג דאטה ובינה מלאכותית.

השינוי הינו חלק ממציאות שקלאוס שוואב, מייסד ה-World Economic Forum - כינה ה'מהפכה תעשייתית הרביעית' (3). זו מתבטאת בקצבי שינוי הולכים וגוברים, ובכך שהטכנולוגיות מטשטשות את הקווים בין העולמות הפיסיקליים, הדיגיטליים והביולוגיים, ובעתיד - את האנושיות שלנו.

העיתונאי תומס פרידמן בספרו "תודה שאיחרת" (4) מדבר על נהרות המידע, וממליץ "להתיישב" על נהרות המייצרים מידע אמין, ולתרום להם מידע ייחודי. תיאורים אחרים ממשיכים דאטה ל'נפט החדש'.

למהפכת המידע והבינה המלאכותית השלכות רחבות, החל במקורות למידע, בדרכים לאיסוף, אכסון, עיבוד, הפיכתו לתובנות, ועד הגדרת תפקידי הרופא והאחות והכשרתם. אפשר לקוות שבינה מלאכותית תאפשר מודלים כלכליים חדשים שיוכלו לגדר את העלויות המאמירות, וכן לתמוך את הצוותים הרפואיים הנתונים בעומס גובר, עקב הזדקנות האוכלוסייה והתחלואה שבעקבותיה, ועקב הצפת המידע.

מקורות המידע

בעידן הרפואה ה'ישנה' המידע הרפואי התבסס על מספר מצומצם של מקורות: תיאור וסיפור המחלה על ידי המטופל, ממצאי הבדיקה הפיסיקלית, וממצאי בדיקות המעבדה וההדמיה. הללו שולבו עם מידע מחקרי מהספרות המדעית.

הרפואה החדשה צריכה להתייחס למקורות חדשים, לדוגמא:

מידע אורכי/לונגיטודינלי

מידע הנאסף 24/7 מהסלולרי: פעילות ותנועה, משך השינה, כמות ואיכות התקשורת, ועוד. מידע אורכי נאסף גם מאביזרים 'לבישים' Wearables כגון שעונים חכמים, המספקים נתונים על דופק, לחץ דם, רוויון חמצן, רמת הזעה ועוד. מחקרים מראים שמידע מלבישים משפר הישרדות חולי סרטן (5). מידע בייצור עצמי אינו חדש - שעון הספורט פיטבייט הפופולרי יצא לשוק כבר בשנת 2009, ומאז ישנה התפוצצות אקספוננציאלית של 'לבישים' לניטור עצמי, הכוללים בנוסף לשעונים, גם טבעות, שרשראות, חגורות, חולצות, נעלים וכרטיסים חכמים. הדבר נתמך ותומך מגמות חברתיות כמו ה'כימות העצמי' - Quantify Self.

במקביל עברו לייצור 'בליעים' Ingestibles דוגמת אלה של חברת פרוטאוס, הכוללים חיישן ומשדר לדווח על מצב התרופה בתוך הקיבה והמע. FDA אישר בנובמבר 2017 תרופה פסיכיאטרית חדשה Abilify MyCite שתדווח על בליעתה לעד 4 אנשים לבחירת המטופל (6).

השילוב בין מידע סלולרי לבין מידע מלבישים יכול לספק מדדים דומים לאלה הנאספים ביחידות לטיפול נמרץ, ולתמוך במגמה של מעבר הרפואה הביתה.





- דם לכימיה שמיפה בעבר כ-20 מרכיבים הפך 'מטבולום' הכולל כימות אלפי מרכיבים.

מידע קולי

בקול האנושי אצור מידע בעל אינספור יישומים. לדוגמה Beyond Verbal הישראלית מבוססת על אלגוריתמיקה המבצעת ניתוח רגשות דרך אינטונציות קוליות, ומזהה 400 סוגי רגשות בזמן אמת באמצעות טון הדיבור בלבד, ללא תוכן או הבעות. האלגוריתמיקה מבוססת על ממצאי 18 שנות מחקר, ב-70,000 נחקרים וב-30 שפות, וקצת מצוייה במחקרים קליניים עם מאיו קליניק על יישומים רפואיים. דוגמה אחרת היא Healthymize, סטארט-אפ ישראלי, פיתח אלגוריתם לאבחון החמרה בקולם של חולי COPD.

מידע סביבתי

למקורות החדשים נוסף המידע הסביבתי שיגבר ככל שהבית יהפוך חכם. חיישני נפח לדוגמה יכולים לאתר תנועה, כניסה ויציאה דרך הפתחים, ולהתריע על נפילות מבלי לפגוע בפרטיות. המצלמה החכמה קינקט גם היא מקור מידע, ומאפשרת שיקום מטופלים מרחוק, כמו ב-ReAbility Online פיתוח שנולד במכון גרטנר שבתל השומר. הבית החכם יאסוף מידע גם מיישומים כאמזון אקו החמוש באלקסה, הנמכרת במיליוני עותקים, ועל התשתית שלה פותחו מספר אפליקציות בריאות. כן פותח WIFI המסוגל למדוד קצב נשימה ודופק של 10 אנשים המצויים בחדר - בלא מגע ישיר בגוף. בית חכם, ובו שלל חפצים מקושרים, יהיה ספק נוסף של מידע ושל פתרונות שיאפשרו לרפואה לעבור מבית החולים אל הבית.

בינה מלאכותית

קצב ייצור המידע עולה בצורה אקספוננציאלית וגם למקורות הנ"ל מצטרפים עוד ועוד מקורות, דוגמת מידע של ננו חלקיקים מהאוויר הננשף. ב-IBM מעריכים שרק 10% מהנתונים הנדרשים לקבלת החלטות רפואיות יגיעו מהקליניקה, והשאר מהמקורות החדשים. מנהל החדשנות של IBM אמר בכנס ביומד 2016: "90% מהדאטה בעולם נוצרו בשנתיים האחרונות. בעשור הראשון לאלף קצב הכפלת המידע היה כל כ-5 שנים, ב-2013 זמן ההכפלה התקצר לשנתיים, וב-2016 זמן ההכפלה הוא כבר פחות משנה". ב-IBM מעריכים שכל אדם ייצור במשך חייו מידע השווה ל-300 מיליון ספרים, כאן נדרשת בינה מלאכותית לחילוץ המידע הרלבנטי, והפיכתו לתובנות תומכות החלטות. IBM השקיעה סכומי עתק בהקמת חטיבת "Watson Health" המבוססת על בינה מלאכותית וננו-טכנולוגיה. ווטסון, שזכה בשנת 2011 במשחק הידיעה הכללית ג'יאופרדי, יודע שלושה דברים:

1. להבין שפה
2. להבין סיבה ותוצאה
3. "ללמוד" מתוך מידע שמזינים אותו: ווטסון מעבד 2 מיליון עמודים ב-3 שניות, מכונת "למידה עמוקה" מהירה הפועלת ללא הפסקה.

ווטסון לומד על בסיס הזנת כמויות נתונים, לדוגמה CT של איבר מסוים, עם פענוח של פתולוגיה ספציפית. לאחר הלמידה הוא מתחיל לספק פענוחים, שהולכים ומשתפרים ברמת הדיוק. כדי לאמן את ווטסון יצרה החברה שותפות עם בית החולים ממוריאל סלון קיטרינג, שמספקת לווטסון דאטה ספציפי יחד עם הפענוחים מהרדיולוגים של בית החולים, וכך הוא לומד תחום אחר תחום בהדמיה.

מעבר להדמיה ציר נוסף הוא ללמד את הבינה המלאכותית לפענח גנום, וציר שלישי הוא אבחון משטחים בפתולוגיה. אנשי IBM יצאו למסע ארוך כשחזונום הוא ליצור חיבור יכולות ניתוח משולבות: מידע

טבלה 1 - בינה מלאכותית - השוואה בין גל ראשון לגל שני

הגל ראשון	הגל שני	
סוג הבעיה	מוגדרת היטב	לא מוגדרת
סוג הפתרון	נסמך על כללים מדויקים, עצים לוגיים וכו'	למידה סטטיסטית
יכולות הבינה המלאכותית:		
הכרת העולם החיצוני Perceive	חלשה	טובה
יכולת למידה	אין	טובה
יכולת הפשטה*	אין	חלשה
סיבתיות/לוגיקה**	חזקה	חלשה
טיפול במצבי אי ודאות	אין	

חשוב להדגיש שמערכות הגל השני הן לא בעלות מוטיבציה משל עצמן, אין להן מוסר, והן לא יכולות להחליף את משרותיהם של רוב העובדים.
*נטילת ידע, הקיים בידי המערכת, ויישומו באופן שונה.
**הבנת יחסי סיבה ותוצאה

ניתוח המידע האורכי יכול להתריע מראש על מצבים אקוטיים, דבר הפותח אופקי מניעה וטיפול חדשים. במספר מקומות עובדות קבוצות מחקר ופיתוח על עור מלאכותי, ואחרות על 'שבב אדם' - יחידות אוספות מידע שתהפוכנה חלק מהגוף.

מידע מהאינטרנט והרשתות החברתיות

החיפוש הנפוץ ביותר בגוגל, אחד מכל 20 חיפושים, הוא מונחי בריאות. כך ניתן כדוגמה, לחזות מגפות ברמת האוכלוסייה, ולהבין לעומק צרכי גולשים, ברמת הפרט - שיקום בעקבות אירוע לבבי, תזונה בעקבות הריון חדש, וכו'. שכבה נוספת היא המידע ברשתות החברתיות. חברה ישראלית מצליחה לדוגמה למפות תופעות לוואי מתרופות על סמך הטקסטים הללו.

OMICS - שם החיבה של שלושה מקורות מידע חדשים - ריצוף גנום, מיקרוביום והמטבולום

- ריצוף DNA בסיסי אנושי עולה כיום כ-1000\$ ולוקח פחות משתי יממות, המחיר ממשיך לצנוח, וצפוי להגיע לכ-100\$ בשנת 2018, ועד לעלות זניחה כשל בדיקת דם בהמשך.
- ריצוף המיקרוביום - שהשלכותיו על דיכאון, השמנה ועוד מתפרסמות במחקרים, ויכולתו לחזות תגובה אינדיבידואלית של עליה רמת סוכר בעקבות אכילת מזונות שונים (7) היא הבסיס לחברת הסטארט-אפ DayTwo.

ניתוח המידע האורכי יכול להתריע מראש על מצבים אקוטיים, דבר הפותח אופקי מניעה וטיפול חדשים.





ריצוף DNA בסיסי
אנושי עולה כיום
כ-1000\$ ולוקח
פחות משתי
יממות, המחיר
ממשיך לצנוח,
וצפוי להגיע
לכ-100\$ בשנת
2018, ועד לעלות
זניחה כשל בדיקת
דם בהמשך

בשכבות חישוביות, שבכל אחת מהן מתבצעת פעולה של עיבוד המידע, ותוצאותיה מועברות לשכבה הבאה לעיבוד נוסף. מלאכת האימון דורשת מהרשתות לחזור על ניתוח המידע עשרות-אלפים פעמים, כדי להגיע לשיפור קטן.

בינה מלאכותית, המבוססת על למידה עמוקה, מגיעה לתוצאות מרשימות בתחומים כמו עיבוד ראייה או תמלול, ותרגום דיבור אנושי. היא מאפשרת זיהוי טון הקול ואבחון רגשות, זיהוי פנים אנושיות וקיצוץ לפי שמות בסלולרי. הדוגמאות של פריצות דרך בתחומי ההדמיה רבות (10), לדוגמא חברת זברה, שהאלגוריתמים שלה מחשבים צפיפות עצם ב-CT, מסייעים בפענוח כבד שומני ב-CT, מאבחנים רמת ההסתיידות בעורקים הכליליים ב-CT עם חומר ניגוד ועוד. שותפות חדשה נוצרת בין אדם למכונה, האלגוריתמים מסייעים לרדיולוגים אחרי שאומנו באמצעות כמויות נתונים.

הוכח שבינה מלאכותית מנבאת סיכון הפרט להתקף לב או אירוע מוחי בעשור הבא טוב יותר מהכלים שפיתחו טובי הרופאים והמובילים שבארגונים. במחקר בו נוסו 4 סוגי AI (Artificial Intelligence) מהגל השני התוצאות של כולן היו טובות מאלה של הרופאים (11), עקב היכולת להתייחס למאות משתנים מורכבים. הבינה המלאכותית אינה יכולה עדיין ללמד אותנו את הפרמטרים, אך כן יכולה לחסוך חיי אלפי אנשים. כלים אחרים נגישים לציבור הרחב, כמו צילום כלל גופי בסלולר, והשוואה לאורך ציר זמן של נגעים עוריים וסימון שינויים החשודים למלנמה באפליקציית Derma-Compare של חברת אמרלד, או צילום וגילוי של צוואר הרחם באמצעות אביזר המחובר לסלולרי ב-Mobile ODT, ועוד. בניגוד למערכות הגל הראשון מערכות הגל השני אינן נסמכות על

הדמייית + מידע גנומי + מידע פתולוגי שייתן מערכת אובייקטיבית, מהירה וסקיילבילית (היכולת להתמודד עם כמויות הדרגתית של עבודה) לניתוח נתונים ותמיכה בקבלת החלטות בתחומים כגון אונקולוגיה. 'להיות קטליסט לשיפור בריאות ורפואה דרך תובנות קוגניטיביות', אומרים בחברה. ומוסיפים: "הרובוט יקבל אותך עכשיו" בפראפראזה על הספר The Patient Will See You Now (8).

סוגי בינה מלאכותית

ווטסון הוא דוגמה לבינה מלאכותית מהגל השני, כפי שמגדירה דארפ"א סוכנות הביטחון האמריקאית, שהקימה, כדוגמא, את רשת האינטרנט כדי לקשר בין מכוני מחקר לבין זרועות הביטחון בסוף שנות ה-60. בסרטון שהפיקה דארפ"א מתוארים שלבי התפתחות הבינה המלאכותית בשלושה גלים (9):

הגל הראשון של הבינה המלאכותית - ידע מתוכנת

מערכות שפותחו סביב בעיות מוגדרות לגמרי, שלפתרון קיימים כללים הגיוניים וברורים. האלגוריתמים הממוחשבים תוכנתו על בסיס הידע שקיים בידי מומחים, והוזנו בהקשרי סיבה תוצאה. מערכות בינה מלאכותית המתבססות על הגל הראשון אחראיות על מרבית המוצרים הממוחשבים כיום: חלונות של מיקרוסופט, אפליקציות בסלולר, וכו'.

הגל השני של הבינה המלאכותית: למידה סטטיסטית

הגל השני של בינה מלאכותית עוסק בבעיות פחות מוגדרות, שאין כללים מדויקים לפתרונן. במקרים הללו מהנדסים מפתחים מודלים סטטיסטיים עבור תחום מסוים של בעיה, ואז מאמנים את מערכת הבינה המלאכותית על כמויות רבות של נתונים. הן מושתתות על למידה עמוקה באמצעות 'רשתות עצבים עמוקות'. המדובר





טבלה 2 - מהפכת הדאטה ברפואה סיכום

מידע ברפואה 'ישנה'	מידע בעולם ה'ממהר'
כמויות קטנות של מידע	ביג דאטה
מידע המורכב מבדיקה גופנית, בדיקות מעבדה והדמיה	שכבות מידע חדשות, מידע «אורכי», אומיקס, מידע מרשתות ואינטרנט, מידע סביבתי, קולי וכו'
מידע אבחנתי - מה המחלה?	מידע פרדיקטיבי - מחלות עתידיות וכיצד למנוע אותן
מידע המנותח על ידי הרופא	מידע המנותח על ידי בינה מלאכותית לסוגיה וגליה
תשתיות כבדות ועלויות יקרות	מזעור וירידת עלויות
מידע המוזמן על ידי הרופא	מידע הנוצר על ידי החולה
מידע בבעלות גורמי הבריאות	מידע בבעלות מספר גורמים, לרבות החולה

- מידע הנוצר על ידי החולה
- מידע בבעלות גורמי הבריאות
- מידע בבעלות מספר גורמים, לרבות החולה

סיכונים ואתגרים

השינויים הללו מעלים אתגרים וכמה סיכונים ממשיים:

העדר אינטר-אופרביליות/קישוריות

חסם מרכזי לפריצת דרך הוא העובדה שהמידע מצוי במקטעים שונים בידי גופים עסקיים, רפואיים, והפרט עצמו, כשאין קישור בין המידעים מהמקורות השונים. פיצול המידע הוא מציאות מורכבת ובעייתית.

הבעלות על המידע

קיימת מחלוקת על הסוגיה מי צריך להחזיק את המידע, שכאמור מפוזר בין הגורמים השונים. גופי הבריאות הם בעלי המידע הרפואי ה'קלאסי', לענקיות האינטרנט בעלות על המידע שלהן, וכל אפליקציה מנסה לבנות לעצמה בעלות על מה שיותר נתונים - כבסיס למודל העסקי שלה.

השאלה היא פרקטית, וגם אתית ופילוסופית. הגנטיקאי, הקרדילוג ומוביל החדשנות הד"ר אריק טופול ערך טבלה של מעל 20 נימוקים מדוע המידע צריך להיות בידי הפרטים, בין נימוקיו - הגוף הוא של הפרט, הפרט שילם על המידע (בזמן או כסף), הפרט יוצר חלקים גדולים מהמידע, המידע ממילא נמכר או נגנב, ולתפיסתו אם הפרטיות הרפואית חשובה, הרי שהדרך היחידה להגן על המידע היא בדה-צנטרליזציה.

גורמי הבריאות ימשיכו לטעון שאי אפשר שאדם יגלה שחלה בסרטן בלא תיווך רפואי, והתאגידים יאחזו בחוזים הארוכים עליהם החתימו את המשתמשים (שלרוב לא קראו אותם). אך הקרב הזה

לוגיקה וכללים מדויקים (בדומה למוח האנושי. בני אדם יכולים להשליך כדור, ולנבא היכן הוא עומד ליפול מבלי לחשב משוואות מורכבות). לכן קשה להסיק על סיבתיות, אלא על 'מה שעובד מספיק טוב, מספיק מהפעמים'. הן גם אינן מצליחות להעביר ידע מתחום אחד למשנהו באופן יעיל.

עם הגל השני נולדות בעיות חדשות, מקרים של בוטים שיצאו משליטה - החלו לקלל, ואף תועד מקרה בו שני בוטים פיתחו שפה חדשה לצורך משימתם, יעילה מזו שהוזנה להם על ידי המפתחים. במערכות מהגל הראשון ניתן לחזות במדויק כיצד יפעלו, במערכות הגל השני בהן לא ניתן להתחזות באופן מדויק אחר סיבתיות, לא ברורה הדרך המדויקת בה קלט הופך לפלט, ומידע מתורגם להחלטה.

לפי דארפ"א מערכות הגל שלישי כבר תוכלנה לגלות את הכללים וההיגיון שמאחורי הפעילות, וכך לנמק את התשובות, להסביר, לדוגמה, מדוע הממצא בצילום חזה הינו דלקת ריאות ולא גידול. הן אפילו תוכלנה לאמן את עצמן ולשפר שוב ושוב את המודלים שבאמצעותם הן מגיעות לתובנות.

מערכות כאלה תוכלנה לעשות אינטגרציה בין מקורות המידע הרבים שנולדים, ולגבש תובנות והמלצות טיפוליות אופרטיביות. הן תבאנה לכדי מימוש את החיבור הפרקטי בין מידע מהפרט, מהבית החכם, מהמחשוב הלבניש ומהחיפושים באינטרנט, מריצופי גנום, מיקרוביום, תוך יצירת אינטגרציה בין מקורות המידע למתן תובנות ייחודיות, (בהנחה שתיתרנה הבעיות המורכבות של יצירת קישורים בין כל המידעים הללו).

אכסון, מזעור ושינויים נוספים

לצירי השינויים של היווצרות מקורות חדשים, לעליה אקספוננציאלית בכמויות המידע שנוצר, ולגלי הבינה המלאכותית נוסף ציר שינוי נוסף, הכולל מזעור וקישור העלויות. בשנת 1965 חזה גורדון מור מנכ"ל אינטל שצפיפות הטרנזיסטורים במעגלים משולבים, קרי עוצמת המחשוב ליחידות נפח, תוכפל כל שנתיים. 'חוק מור' הפך מפת דרכים לתעשייה, וגרם למוטיבציה להמשיך במזעור ובהורדת העלויות.

הירידה בעלות ריצוף DNA כבר עוברת את הצפוי לפי חוק מור! העולם התמלא בצ'פים זולים, פרובים המתחברים למכשיר הסלולרי בעלות של \$50-200 ומאפשרים פעולות כגון אקו לב, אק"ג, ושלל בדיקות דם ונוזלי גוף.

ציר שינוי נוסף הוא ביצרני ומחזיקי המידע. ברפואה הישנה המידע מוזמן, מיוצר ומוחזק בידי רופאים וארגוני הבריאות, כיום יותר ויותר מהמידע נוצר על ידי הצרכנים בסלולרי שלהם, וכמובן אצל ענקי אינטרנט.

מהפכת הדאטה ברפואה מתרחשת במספר צירים:

- מידע ברפואה 'ישנה'
- מידע בעולם ה'ממהר'
- כמויות קטנות של מידע
- ביג דאטה
- מידע המורכב מבדיקה גופנית, בדיקות מעבדה והדמיה
- שכבות מידע חדשות, מידע "אורכי", אומיקס, מידע מרשתות ואינטרנט, מידע סביבתי, קולי וכו'
- מידע אבחנתי - מה המחלה?
- מידע פרדיקטיבי - מחלות עתידיות וכיצד למנוע אותן
- מידע המנותח על ידי הרופא
- מידע המנותח על ידי בינה מלאכותית לסוגיה וגליה
- תשתיות כבדות ועלויות יקרות
- מזעור וירידת עלויות
- מידע המוזמן על ידי הרופא

נגישות המידע
בכף היד של
כל אדם הופכת
את המטופלים
לשותפים המובילים
בטיפול. ניתן למדוד
כמעט כל משתנה
פיזיולוגי, כמו
כמה שעות שינה
ממוצעות, דופק
ממוצע במנוחה,
רווייה של חמצן
בדם ועוד. כך
המטופל יודע מה
הסימפטומים, קרא
על המחלה, ומחזיק
מידע רלבנטי
מהסלולרי, השעון
החכם או כל סנסור
אחר



מבט אופטימי לעשורים הקרובים

מהפכת המידע כבר הביאה לשינוי במרקציות החולה - מעבר מהמקום היררכי, בו הרופא מחזיק בידע ובכוח, לנקודה יותר שוויונית. 'החולה במרכז' משמעו שינוי תפיסתי הן של רופאים והן של המטופלים. הטכנולוגיה היא קטליזטור לשינוי.

נגישות המידע בכף היד של כל אדם הופכת את המטופלים לשותפים המובילים בטיפול. ניתן למדוד כמעט כל משתנה פיזיולוגי, כמו כמה שעות שינה ממוצעות, דופק ממוצע במנוחה, רגישות של חמצן בדם ועוד. כך המטופל יודע מה הסימפטומים, קרא על המחלה, ומחזיק מידע רלבנטי מהסלולרי, השעון החכם או כל סנסור אחר. סטטיסטית - ככל שהמטופלים צעירים יותר, כך הם באים מצוידים ביותר מידע, באים למפגש עם שאלות רלבנטיות, והופכים את מערכת היחסים שלהם עם הרופא ליותר שוויונית.

מרכזיות הפרט משמעה גם התמקדות במניעת תחלואה ובשמירת הבריאות, מעבר לטיפול בחולי. היא גוזרת מעורבות של המטופל בתהליכי האבחון והטיפול, וכמובן טיפול רפואי מותאם אישית לפרט, העדפותיו, הגנטיקה שלו, חיידיק גופו וכך הלאה.

ניתן להעריך ש'מניעה מותאמת אישית' תוכל, בצד השיפור בבריאות הציבור, לסייע בריסון העלויות העולות של הרפואה. לכך לדוגמה מכון הסטארט אפ הישראלי DayTwo, שעוסק בריצוף גנטי של המיקרוביום במטרה להתאים תזונה אישית לכל אדם, למנוע השמנה, סוכרת ותחלואה.

רתימת מהפכת הדאטה הכרחית על מנת להמשיך לספק בריאות, ולשמור על מסגרת תקציבית מגודרת. ה - NHS, סוכנות הבריאות הלאומית בבריטניה, כבר הצהירה על אסטרטגיה המשלבת דאטה, טכנולוגיה ודיגיטל, ופרסמה 'מדריך למפת דרכים דיגיטלית מקומית', והיא גם מממנת טכנולוגיות ב'סל הבריאות' שלה.

ערך נוסף הוא תמיכה בצוותים הרפואיים, הכורעים תחת ריבוי העבודה וכמויות המידע והפרסומים האקדמיים (42,000 בסוכרת בשנת 2016 כדוגמה). אין די שעות ביממה להתעדכן כראוי במאמרים, לעבור על נתוני החולים במטרה לעבוד בצורה מבוססת עובדות וראיות. בינה מלאכותית תוכל להיות כלי עזר מרכזי בעבודת הרופא, לרבות כריית המידע על המטופל, וכמובן המידע העדכני בספרות הרפואית. כבר היום ישנן חברות המסייעות לצוותים, דוגמת Medaware הישראלית המאגרת ומתריעה על טעויות במתן תרופות באמצעות בינה מלאכותית.

מנהלי בתי חולים חכמים כבר נמנעים מלבזבז את סכומי העתק על רכישת והטמעת מערכות מידע חדשות, וממתינים לפתרונות חדשים שיגיעו. הצורך הוא בפתרונות 'משבשים', שיתייחסו לשלל הסוגיות שהוצגו: קישוריות, רגולטוריות, טכנולוגיות ומוסריות. אין ספק כי מערכת שתצליח להחליף את שלל המערכות היקרות והחלקיות המיושמות כיום לא תהיה אוסף פתרונות צרים ואינקרימנטליים אלא תפעל באינטגרציה 'אנד טו אנד', עבור הלקוחות המשלמים והלא משלמים כאחת. ניתן להעריך שקפיצת מדרגה כזו, אם וכאשר תקרה, תחולל 'שיבוש' ומהפכה גדולה שתתווסף למהפכת הרפואה הדיגיטלית בעיצומה אנו מצויים. אפשר להעריך בזהירות שחידושים כמו בלוקצ'יין, והמשקשים בין רפואה לסייבר סקוריטי יאפשרו פריצות דרך טרנספורמטיביות בעשור או שניים הבאים, קרי מוצרים שיהיו טובים יותר פי 50-100, ויעלו הרבה פחות.

ועד שקפיצת המדרגה הזו תופיע נשאר לחזור ולהדגיש את חשיבות השמירה על אורח חיים בריא, את זה בינה מלאכותית לא יכולה לעשות עבורנו.

המאמר, כולל רשימת המקורות, מופיע באתר מכון תנובה למחקר

יזכר. כבר כיום מרבית המידע הרלבנטי לא מוחזק בידי מוסדות הבריאות, והרגולטורים מתעוררים לפעולה לאור הסיכונים בענקיות האינטרנט. במקביל למציאת פתרונות לאינטגרציה בין השכבות השונות, חייבת להיות מעוגנת זכותו של כל פרט לקבוע היכן יוחזק המידע שלו, היכן יעובד, ובמה הוא משתף את רופאיו או גופים דוגמת חברת הביטוח שלו.

אבטחת המידע

אין ספק כי ברגע בו ימצא פתרון לקישוריות יידרשו במקביל גם פתרונות שישמרו על אבטחת המידע. יתכן שהן יבואו מטכנולוגיות כמו בלוקצ'יין, שיוכלו לשמור על המידע מאובטח בצורה לא מרוכזת, או מטכנולוגיות שטרם שמענו עליהן...

ממידע לתובנות

אתגר מרכזי הוא מעבר מכמויות מידע לתובנות מעשיות שיביאו את הפרט לשינוי אורחות חיים, או יסייעו לרופא ולמטופל בקבלת החלטות רפואיות. קרי נדרש תרגום מכמויות של נתונים להשפעה על הבריאות - ברמת מניעה, טיפול או שיקום.

מעבר למצגים הקבועים של הנתונים, אותם מאפשרות לדוגמה, האפליקציות התומכות בבריאות (עם או בלי לבישים), נדרשת גם 'התאמה אישית' של הדש-בורד לכל מטופל ברמת בריאות ואורחות חיים, תוך התייחסות לגורמי הסיכון, ההעדפות האישיות, וכמובן למידע בספרות.

דבר דומה נדרש גם עבור הרופא על מנת לא להציף רופאים בכמויות מידע. המיקוד איננו בנתונים, אלא בסוג ההחלטה שנדרש לקבל עבור כל מטופל. כל החלטה צריכה לגזור את שדות המידע הנדרשים עבורה.

מידע לא מהימן

מילות חיפוש רפואיות הן מרבית החיפושים בגוגל. הנפוצים הם כאבי גב, כאבי ראש, חרדה ודיכאון. גורמי בריאות מוחים, בצדק, על כך שהמידע ברשת לא תמיד מהימן, והסיכון הוא שצריכת מידע כזה מביאה להחלטות וטיפולים שגויים. גם כאן נכנסים יזמים המפתחים אפליקציות לקבלת מידע מבוסס ראיות. הסטארט-אפ הישראלי CureFacts מנסה לספק למשתמש הקצה מידע מבוסס ראיות; החברה הישראלית Medivizor מספקת למטופלים ורופאים סקירות של המאמרים העדכניים והרלבנטיים ביותר עבורם; פיתוחים כגון אלה יצמצמו פגיעות בעקבות מידע לא מקצועי, טיפולים מיותרים ואף מזיקים. במקביל חשוב להדגיש לציבור את חשיבות מקורות המידע המהימנים (11).

הטמעת AI במערכות הבריאות

קצב ואיכות ההטמעה נגזר משני כוחות הפוכים: ההתפתחויות הטכנולוגיות מחד; וחסמי המערכות מנגד.

התנגדויות יכולות להיות של המשתמשים - רופאים, בתי חולים, מבטחים, הן נובעות משיקולי תקציב, מרצון להגן על המטופלים מסיכונים, מהצפה בפניות, ומשמרנות ושנאת סיכון. בנוסף קיימים חסמים רגולטורים של ה-FDA וגופים מקבילים. העולם הממרח מצביע אתגרי מדיניות ורגולציה גדולים. על הרגולטור להיערך במהירות לשינויים הדרמטיים, תוך שמירה על בטיחות המטופל.

אתגר מקומי בהטמעה

נובע מכך שישראל היא מעצמת חדשנות ברמת הפיתוח, אך כיוון שהשוק הישראלי קטן - מרבית היוזמות המפותחות בארץ אינן מתורגמות לטובת הציבור הישראלי, אלא יוצאות לשווקי העולם. פורום ריבוט להטמעת חדשנות רפואית בת קיימא בישראל עוסק בהתמודדות עם בעיה זו.



שיטות מחקר בתחום הבריאות

ד"ר עפרה קלטר ליבוביץ'

מנהלת היחידה לאפידמיולוגיה של מחלות לב וכלי דם,
מכון גרטנר לחקר אפידמיולוגיה ומדיניות בריאות, תל השומר

מחקרים הנשענים על שיטות מחקר תקפות מגדילים את הסיכוי שתוצאותיהם תהיינה בעלות ערך לצורך קבלת החלטות הנוגעות למניעה, לטיפול, לשיקום, או לארגון שירותי בריאות. יש חשיבות גדולה להכיר את היתרונות והחסרונות של מקורות המידע ומערכי המחקר השונים, ולבחור באלה שיתאימו באופן המיטבי למתן מענה על שאלות המחקר, וניתן יהיה ליישם במסגרת המשאבים העומדים לרשות החוקר.

המידע הנאסף במחקר איכותני הוא טקסטואלי, והוא מקובץ ומנותח בצורה שיטתית, תוך שימוש בשיטות סטנדרטיות. בתהליך ניתוח המידע אנו מזהים היגדים (themes) רלבנטיים לנושא המחקר. כאשר ההיגדים הללו חוזרים על עצמם מספר פעמים ממקורות שונים נדע שמייצגו את המדגם, ואין צורך בגיוס משתתפים נוספים. על פי רוב מספר המשתתפים במחקרים מסוג זה אינו גדול. לעיתים אנו משתמשים במחקר איכותני על מנת לייצר השערות, אשר נוכל לאשר או להפריך אותן בהמשך, במחקר כמותי. בהזדמנויות אחרות, אנו מסתייעים במחקר איכותני כשלב מקדים למחקר כמותי, לדוגמה לצורך בניית שאלונים המותאמים לנושא המחקר ולאוכלוסיית היעד. במקרים אלה מדובר במחקר המשתמש בשיטות מעורבות (mixed methods study) (4-1).

מחקרים כמותיים

מחקרים מסוג זה מספקים אומדן כמותי של מצבי בריאות (לדוגמה: שיעורי תחלואה, פרופורציה של אנשים הנפטרים ממחלה מסוימת), ושל קשרים הנצפים בין מאפייני הפרט, התנהגויות או טיפולים שונים, לבין תוצאי-בריאות (health-outcomes), כגון: איכות חיים, מצב נפשי או קוגניטיבי, תחלואה או מוות.

מחקרים העוסקים בבריאות מסתמכים על מידע (information), הנאסף בשיטות שונות וממקורות מגוונים, תוך שימוש במערכי מחקר שונים (study design). ראשית נפריד בין שני סוגי מחקר עיקריים: איכותני וכמותי.

מחקרים איכותניים

מחקר איכותני מתמקד בסיפור של אנשים בהקשר למצבי בריאות וחולי. המשתתפים במחקר יכולים להיות בריאים, חולים, אלו המטפלים באנשים חולים, או בעלי תפקידים שונים במערכת הבריאות. המידע הנאסף במחקר איכותני נועד לסייע לנו בהבנת הקונטקסט החברתי של מצבי בריאות וחולי, את החוויות שהמטופלים חווים במהלך מחלתם ובעקבות הטיפול בה, דפוסי התנהגות הקשורה לבריאות והמניעים להם. המידע יכול להתקבל מראיונות אישיים ("ראיונות עומק") של אנשים בודדים, משיח מונחה בו משתתפים מספר אנשים ("קבוצות מיקוד"), או ממקורות אחרים, כגון מדיה מודפסת או אלקטרונית. בסוג זה של מחקר יש חשיבות להגיע למגוון הסיפורים הרלבנטיים לנושא המחקר, אולם אין צורך שהמדגם של המשתתפים בו יהיה בהכרח מייצג.





סוג ספציפי של
מחקר עוקבה הוא
מחקר העוקבה
ההיסטורי. למרות
שהשם אינו מוכר,
מערך מחקר זה
נמצא בשימוש
רב, ההולך וגובר
בעידן ה-big data.
לעיתים מערך
מחקר זה נקרא
בטעות בשם "מחקר
רטרוספקטיבי".
הכוונה לשימוש
במידע הנאסף
ומתועד בצורה
פרוספקטיבית, אך
החוקר מנתח אותו
בשעה שהמידע על
החשיפות (גורמי
סיכון, טיפולים)
ותוצאי-הבריאות
(כגון: סיבוכי המחלה,
התקדמות/נסיגה של
הגידול) כבר נאסף

חסרונם של מחקרים מסוג זה היא הסלקטיביות שלהם: הנטייה הטבעית היא לדווח על הצלחות של טיפול חדש או גישה ניתוחית חדשנית, אך לא לדווח על כישלון.

מערך מחקר נפוץ נוסף הוא **מחקר חתך** (cross sectional study), או כפי שהוא מכונה גם מחקר הימצאות (prevalence study). בסוג זה של מחקר, מאפייני המשתתפים, כגון: גיל, מין, הרגלי תזונה או משקל, נמדדים בו-זמנית עם התוצא הבריאותי, כגון: דיכאון או מחלה ניוונית של המפרקים. היתרון בסוג זה של מחקר הוא הקלות והמהירות היחסית בביצועו. חסרונו הוא בחוסר האפשרות להסיק על קשר סיבתי; האם ההשמנה קדמה לדיכאון או שבאה בעקבותיו? מומלץ אפוא לאושש קשרים הנצפים בסוג זה של מחקר במערכי מחקר אנליטיים, הלוקחים בחשבון את הקשר בזמן שבין מועד החשיפה (גורם מחלה, טיפול) לבין התפתחות התוצא הבריאותי.

מחקר מקרה-ביקורת (case-control) הוא מחקר בו אנו צופים אחורה בציר הזמן: מהמחלה כיום, אל גורם/גורמים אפשריים למחלה שקרו בעבר. מחקר מסוג זה כולל קבוצה של "מקרים" (אנשים עם מחלה או מצב בריאותי אחר כלשהו), וקבוצת ביקורת שצריכה להיות דומה לקבוצת החולים, למטע העובדה שאין להם את המחלה. בשתי הקבוצות, אנו מודדים חשיפות שונות שחוו בעבר (כגון: טיפולים רפואיים, הרגלי עישון). אנו מעוניינים לקשור את החשיפה או החשיפות שהיו בעבר לתחלואה כיום. מערך מחקר מסוג זה כרוך בשני אתגרים עיקריים: הקושי לאתר קבוצת ביקורת מתאימה, והנטייה של אנשים חולים (המקרים) לדווח באופן שונה מאנשים בריאים (הביקורות) על גורמים שונים שנחשפו להם בעבר. נטייה זו כרויה "הטיית הזיכרון". מגבלה נוספת של מחקר מסוג זה היא שהוא בוחן קשרים עבור מחלה או מצב בריאותי יחיד (7).

מחקר עוקבה (cohort studies) הוא מחקר העוקב אחר קבוצה של אנשים לאורך זמן. כיוון ההסתכלות בסוג זה של מחקר הוא

על מנת שניתן יהיה להפיק תועלת ממחקר כמותי, יש לנקוט בשיטות התומכות בקבלת תוצאות תקפות ובלתי מוטות. שיטות אלה כוללות שימוש במערכי מחקר המאפשרים מתן תשובה לשאלת המחקר, הגדרות ברורות של משתני המחקר, תהליכים סטנדרטיים לאיסוף, טיוב, עיבוד ושמירת המידע, ושימוש בשיטות סטטיסטיות שיאפשרו לבחון את השערות המחקר.

ניתן לחלק את המחקר הכמותי בבני-אדם למחקרים תצפיתיים, בהם החוקר מודד ומתעד, אך אינו מפעיל התערבות כלשהי, ולניסויים, בהם החוקר אינו משמש כצופה בלבד, אלא מתערב במטרה לבחון את התועלת מהתערבויות שונות שנועדו להשפיע על בריאותם של אנשים: לשמור בריאות טובה, למנוע מחלה, לטפל בה, למנוע את סיבוכיה, או לשקם את מי שחלה בה (5).

מחקרים תצפיתיים

המחקרים התצפיתיים אף הם מתחלקים לשתי קבוצות: מחקרים תיאוריים ומחקרים אנליטיים.

בסוג הראשון, המטרה העיקרית היא לתאר את השכיחות של מחלה, את מהלכה, ולעיתים ביטוי שאינו שכיח של מחלה בחולה בודד, או בקבוצה של חולים בעלי מאפיינים דומים. מחקרים מסוג זה על פי רוב אינם מנסים למצוא קשרים, לדוגמה בין חשיפה לגורם מסוים לבין הסיכון לחלות במחלה ספציפית.

דוגמא למחקרים מסוג זה הם תיאור מקרה (case report) או של סדרת מקרים (case series). היתרון של מחקרים מסוג זה הוא רגישותם הרבה לחדשנות: באמצעותם לעיתים מתגלות מחלות חדשות, השפעות לא צפויות של טיפולים/חשיפות לגורמים סביבתיים, או מנגנונים ביולוגיים. יש דוגמאות רבות כיצד תיאור של סדרת מקרים הוביל למציאת מחלות חדשות, לדוגמא: תיאור של סדרת מקרים ב-1999 בארה"ב הוביל לגילוייה של מחלה חדשה - West Nile Encephalitis (6).





בעיה נוספת ספציפית שיש לתת עליה את הדעת בסוג זה של מחקר היא, שהבחירה לתת טיפול תרופתי מסוג מסוים אינה אחידה; היא נעשית על פי שיקול הדעת הקליני של הרופא בהתאם למצבו של החולה, הצפי להיענות החולה למשטר הטיפול, והחלופות האפשריות. כל אלה מכניסים סוג ספציפי של הטייה: הטיית אינדיקציה, קרי: חולים שקיבלו את הטיפול שונים מאלה שלא קיבלו אותו או קיבלו טיפול אחר, והבדלים אלה עשויים להיות קשורים בתוצאי-הבריאות ללא קשר להשפעתו הסגולית של הטיפול. אמנם יש שיטות סטטיסטיות האמורות לסייע לטפל בסוג זה של הטייה, אולם פתרון זה אינו מושלם, ויש מקרים בהם אינו ישים כלל.

ניסויים קליניים

בניסוי קליני החוקר מעוניין ללמוד על התועלת מהתערבות ספציפית, או ממספר התערבויות הניתנות בו זמנית או ברצף. ההתערבות הנבחנת יכולה להיות תרופה חדשה הניתנת לטיפול במחלה מסוימת; טיפול מונע, כגון: חיסון; תכנית שיקום; טיפול נפשי; פרוצדורה ניתוחית, שיטת אבחון חדשה לגילוי מוקדם של מחלה או סיבוכיה, או צורת ארגון חדשה של שירותי בריאות (לדוגמה: מתן של טיפולים הניתנים בדרך כלל במסגרת אשפוז בבית החולים, דווקא בביתו של המטופל).

בניסוי קליני החוקר צריך להגדיר מי היא אוכלוסיית היעד עבורה מיועדת ההתערבות, ואת מי הוא מעוניין לכלול כמשתתפים בניסוי. לרוב אין זהות מלאה בין שתי האוכלוסיות, מכיוון שיש אנשים שלא נרצה להכליל בניסוי, לדוגמה: אנשים שמראש הסיכוי שייטלו את הטיפול הניסיוני ושיבואו לביקורי המעקב מאד נמוך (כגון מכורים לסמים); או אנשים שעשויים לפתח סיבוכים בלתי צפויים מהטיפול (כגון התפתחות מומים מולדים בעובריים של נשים הרות). מכאן נובעת המגבלה הראשונה של הניסוי הקליני, והיא העובדה שאוכלוסיית המשתתפים בו היא פעמים רבות סלקטיבית. מגבלה זו עשויה לפגוע בתוקף החיצוני של המחקר, קרי: ביכולת להסיק מסקנות על התועלת מההתערבות עבור אוכלוסייה השונה במאפייניה מזו שהשתתפה בניסוי. לעיתים קרובות, לדוגמה, ניסויים קליניים בתחום הקרדיולוגיה כוללים בעיקר גברים בגיל העמידה. באיזו מידה תוצאות המחקר תקפות גם עבור נשים או קשישים החולים במחלה?

כאשר אנו כוללים משתתפים בניסוי קליני, נכנסים גורמים נוספים העשויים להשפיע על התוצאים שלהם שאינם קשורים כלל להשפעת ההתערבות שאנו בוחנים. הגורם הראשון הוא המהלך הטבעי של המחלה. מחלות רבות הן בעלות אופי גל, עם תקופות של חמירה, ותקופות בהן המחלה פחות פעילה, כגון: מחלות מפרקים, טרשת נפוצה או תסמונת המעי הרגיש. במקרים אלה, שיפור הנובע מהמהלך הטבעי של המחלה עשוי להיות מיוחס בטעות להשפעת הטיפול. הגורם השני העשוי להשפיע היא עצם ההתנהלות של המטפל והמטופל במסגרת של ניסוי קליני. מסגרת זו אינה דומה למעקב או לטיפול השגרתי. הביקורים במרפאה עשויים להיות תכופים יותר, בדיקות ההערכה מקיפות יותר, ותשומת הלב הניתנת לטיפול בכללותו, גם במחלות אחרות שיש למשתתף, עשויה להיות גדולה יותר. גם מטופל הנמצא במסגרת מחקרית עשוי להקפיד יותר על הטיפול במחלתו, ובמחלות אחרות שאינן במוקד המחקר. שינויים אלו עשויים לשפר את מצבו של המטופל ללא קשר להשפעת הטיפול הניסיוני הנבדק.

הגורם האחרון הוא אפקט הפלסבו. למטופל המשתתף בניסוי קליני יש מראש ציפייה שהטיפול יועיל לו. ציפייה זו עשויה להשפיע על תפישתו את התועלת שהוא חווה מהטיפול הניסיוני, ואת האופן בו הוא מדווח עליה למטפל. אפקט הפלסבו בולט במיוחד בהקשר לתוצאים "רכים", כגון: איכות חיים, כאב או מצב רוח.

עם ציר הזמן: מהחשיפה לגורם אחד או יותר, הנמדדת בכניסה למחקר, ולעיתים גם במהלכו, אל מצבי-בריאות המתפתחים במהלך המעקב. בכניסה למחקר, המשתתפים יהיו ללא התוצא (ים) שבמוקד המחקר. החוקר מודד בקרב משתתפי המחקר את הגורמים השונים שהוא משער שיש להם קשר לסיכון לפתח את המחלה, כגון: סיפור משפחתי של המחלה בקרובי משפחה, סוג העיסוק, היסטוריה של חשיפה לקרינה או לרעש, וכדומה. במהלך המעקב המתוכנן לתקופה המוגדרת מראש, המשתתפים עוברים הערכה נוספת אחת, או הערכות חוזרות בפרקי זמן קצובים, על מנת לבחון האם התפתחו תוצאי הבריאות שבמוקד שאלת המחקר. היתרון הגדול של מערך מחקר עוקבה הוא הוודאות הגבוהה שיש לנו, שהחשיפה לגורמים בעלי העניין קדמה להתפתחות מצבי הבריאות. זהו תנאי הכרחי להסקה על סיבתיות (causality) של הקשרים הנצפים. מערך מחקר כזה גם מאפשר בחינת קשרים בין גורמים מרובים לבין תוצאי-בריאות מרובים. החיסרון במערך מחקר כזה הוא הצורך לעיתים לחכות זמן רב עד אשר יתפתח התוצא הבריאותי בעל העניין. בזמן זה אנו עלולים לאבד קשר עם חלק ממשתתפי המחקר, מה שעלול להוביל לתוצאות מוטות עקב הטיית-בחירה (selection bias): הנושרים מהמחקר עשויים להיות שונים ממשתתפים שלא נשרו, לדוגמה בסיכון לפתח את המחלה שבמוקד המחקר. לעיתים משך הזמן העובר ממועד החשיפה ועד להתפתחות המחלה הוא כה ארוך, או שהסיכון לפתח את המחלה כה נמוך (למשל במחלות ממאירות), שהשימוש במערך המחקר הזה הופך לכמעט בלתי אפשרי. במקרים אלה, מחקר מקרה-ביקורת יהיה יותר מעשי ופחות יקר (8).

סוג ספציפי של מחקר עוקבה הוא מחקר העוקבה ההיסטורי. למרות שהשם אינו מוכר, מערך מחקר זה נמצא בשימוש רב, ההולך וגובר בעידן ה-big data. לעיתים מערך מחקר זה נקרא בטעות בשם "מחקר רטרוספקטיבי". הכוונה לשימוש במידע הנאסף ומתועד בצורה פרוספקטיבית, אך החוקר מנתח אותו בשעה שהמידע על החשיפות (גורמי סיכון, טיפולים) ותוצאי-הבריאות (כגון: סיבוכי המחלה, התקדמות/נסיגה של הגידול) כבר נאסף. כך, לדוגמה, חוקר יכול לאסוף מידע מתוך התיקים האישיים של עובדי הקריה למחקר גרעיני על מידת החשיפה שלהם לקרינה לאורך זמן, ולהצליב את המידע המזוהה שאסף עם מידע מזוהה הנמצא ברשם הסרטן הלאומי, על מנת לבחון את הקשר בין העוצמה ומשך החשיפה לקרינה מייננת לסיכון לפתח סוגים שונים של סרטן. כיוון ההסתכלות בציר הזמן זהה למחקר עוקבה רגיל: מחשיפה בקרב אנשים ללא מחלה ממארת, אל התפתחות הסרטן במהלך המעקב. השוני הוא רק בכך שכל המידע קיים בעת ביצוע המחקר.

בסוג זה של מחקר יש יתרון גדול; הוא מנצל מידע זמין שנאסף בשגרה למטרות אחרות, ולכן חוסך זמן ומשאבים. החיסרון של מערך מידע זה נובע אף הוא מהעובדה שהמידע לא נאסף למטרות המחקר אלא למטרות אחרות, קליניות או אדמיניסטרטיביות. אם המידע לא נאסף על פי פרוטוקול סטנדרטי ובאופן אחיד בכל משתתפי המחקר, עלולה להיווצר בעיה.

לעיתים קרובות, חוקרים משתמשים במידע הקיים ברשומה הרפואית על מנת להסיק מסקנות על התועלת הרפואית מטיפולים שונים. מידע שנאסף ברשומה נוצר בעקבות מפגשים בין מטופלים למטפלים שלהם, ומהערכה נוספת (כגון בדיקות מעבדה והדמיה) שהמטפלים שולחים אותם לבצע בהמשך. תכיפות המפגשים, הבדיקות הנוספות שהמטופל נשלח לבצע ותיעוד המידע בתיק החולה מושפעים מגורמים נוספים פרט למצבו של המטופל, כגון: היענות החולה לביקורי המעקב, התנהלות הרופא ואילוצי המערכת. כל אלו עשויים להשפיע על התוצאים הנמדדים ללא קשר להשפעת הטיפול. יש לתת את הדעת שהמידע על הגורמים הנוספים הללו עשוי להיות חלקי או חסר ברשומה הרפואית, ולכן לא נוכל לקחת אותו בחשבון בעת ניתוח הנתונים.

**מכאן נובעת
המגבלה הראשונה
של הניסוי הקליני,
وهיא העובדה
שאוכלוסיית
המשתתפים בו
היא פעמים רבות
סלקטיבית. מגבלה
זו עשויה לפגוע
בתוקף החיצוני
של המחקר, קרי:
ביכולת להסיק
מסקנות על התועלת
מההתערבות עבור
אוכלוסייה השונה
במאפייניה מזו
שהשתתפה בניסוי**



חוזק הראיות המתקבלות מסקירה שיטתית תלוי בחוזקם של המחקרים המקוריים מהם היא נוצרה. אם נכללו בה מחקרים שבוצעו במערכי מחקר החשופים פחות להטיה, ושנשענו על מתודולוגיה קפדנית, היא תהיה בעלת ערך רב (10, 11).

באיזה מערך ושיטות מחקר לבחור?

הבחירה בתבנית המחקר מושפעת משאלת המחקר, ומהמשאבים (תקציב, כוח אדם, ציוד, מיומנות של הצוות, נגישות למשתתפים פוטנציאליים) העומדים לרשות החוקר.

ההערכה הכי פחות מוטת של התועלת מהתערבות טיפולית תתקבל מניסויים קליניים מבוקרים אקראיים. מחקרי מקרה-ביקורת יהיו מערך המחקר המועדף על מנת לבחון השפעה של חשיפות שונות על הסיכון למחלות בעלות תקופה לטנטית ארוכה.

בלי קשר למערך המחקר, יש חשיבות רבה להגדרה ברורה וסטנדרטית של משתני המחקר, והאופן בו יימדדו, לדוגמה: באיזו שיטת מעבדה נשתמש למדידת רמת הכולסטרול בדם? (מה התוקף והמהימנות שלה?); באילו שאלונים נשתמש על מנת להעריך תסמיני דיכאון (האם תוקפו באוכלוסייה הדומה לאוכלוסייה של המחקר הנוכחי?); האם נבדק תוקף התרגום שלהם לשפת הדיבור של אוכלוסיית המחקר?); איזו הכשרה יש לתת לצוות שמבצע את בדיקת האולטרה-סאונד לצורך הערכת התכווצות שריר הלב, על מנת שיבצעו את הבדיקה בצורה מדויקת, ולהקטין ככל האפשר הבדלים בין הבודקים השונים?

ובהקשר לזה, חשוב מאד שנכיר בתועלת הפוטנציאלית אך גם במגבלות של שימוש משני במאגרי המידע הגדולים המצויים במערכת הבריאות, לצורכי מחקר. חלק ממידע זה נוצר לצרכים מנהליים, כמו התחשבות עם ספקי שירות-בריאות (למשל בתי החולים מול קופות החולים). אם מידע מסוג זה אינו עובר תהליכי בקרה וטיוב לפני שמשמשים בו לצרכי מחקר, הוא עלול להיות שגוי או חסר.

דוגמא נוספת היא אבחנות שניתנות ע"י רופאים; בין ארגוני הבריאות השונים, ולעיתים אף באותו ארגון, ניתן למצוא שיטות שונות לקידוד של אותו סוג של מידע. עם השנים גם חלו שינויים בשיטות ובנורמות של ערכי בדיקות המעבדה בארגוני הבריאות. ארגוני הבריאות פועלים לעיתים קרובות בתנאי מחסור מתמשך של מקורות תקציב וכוח-אדם; במצב זה, טיוב מאגרי המידע עבור צרכים שהם מעבר למטרות המקוריות שלשמן הם נוצרו אינו נמצא בראש סדר העדיפויות. על מנת שנוכל להפיק את המיטב ממאגרים אלה ולהישמר מטעויות, נחוץ לנו שיתוף פעולה הדוק עם איש מידע בעל הבנה מעמיקה והתמצאות במאגרי המידע של הארגון, והרבה מאד רצון טוב.

לסיכום

מחקרים הנשענים על שיטות מחקר תקפות מגדילים את הסיכוי שתוצאותיהם תהיינה בעלות ערך לצורך קבלת החלטות הנוגעות למניעה, לטיפול, לשיקום, או לארגון שירותי בריאות.
יש חשיבות גדולה להכיר את היתרונות והחסרונות של מקורות המידע ומערכי המחקר השונים, ולבחור באלה שיתאימו באופן המיטבי למתן מענה על שאלות המחקר, וניתן יהיה ליישם במסגרת המשאבים העומדים לרשות החוקר.

המאמר, כולל רשימת המקורות, מופיע באתר מכון תנובה למחקר

על מנת שנוכל לבחון את ההשפעה נטו של ההתערבות הנחקרת מעבר לכל ההשפעות לעיל, אנו זקוקים לקבוצת ביקורת שתהיה דומה מכל הבחינות לקבוצת המחקר, למעט העובדה שאינה מקבלת את ההתערבות הנבדקת. באופן אידאלי, קבוצת הביקורת תקבל טיפול בתרופת-דמה (פלסבו). כאשר מדובר במחלה שיש עבורה כבר טיפול מקובל, מתן פלסבו לקבוצת הביקורת עשוי להיות לא אתי, ואז הבדיקה של השפעת הטיפול החדשני תיעשה מול קבוצת ביקורת המטופלת בטיפול המקובל.

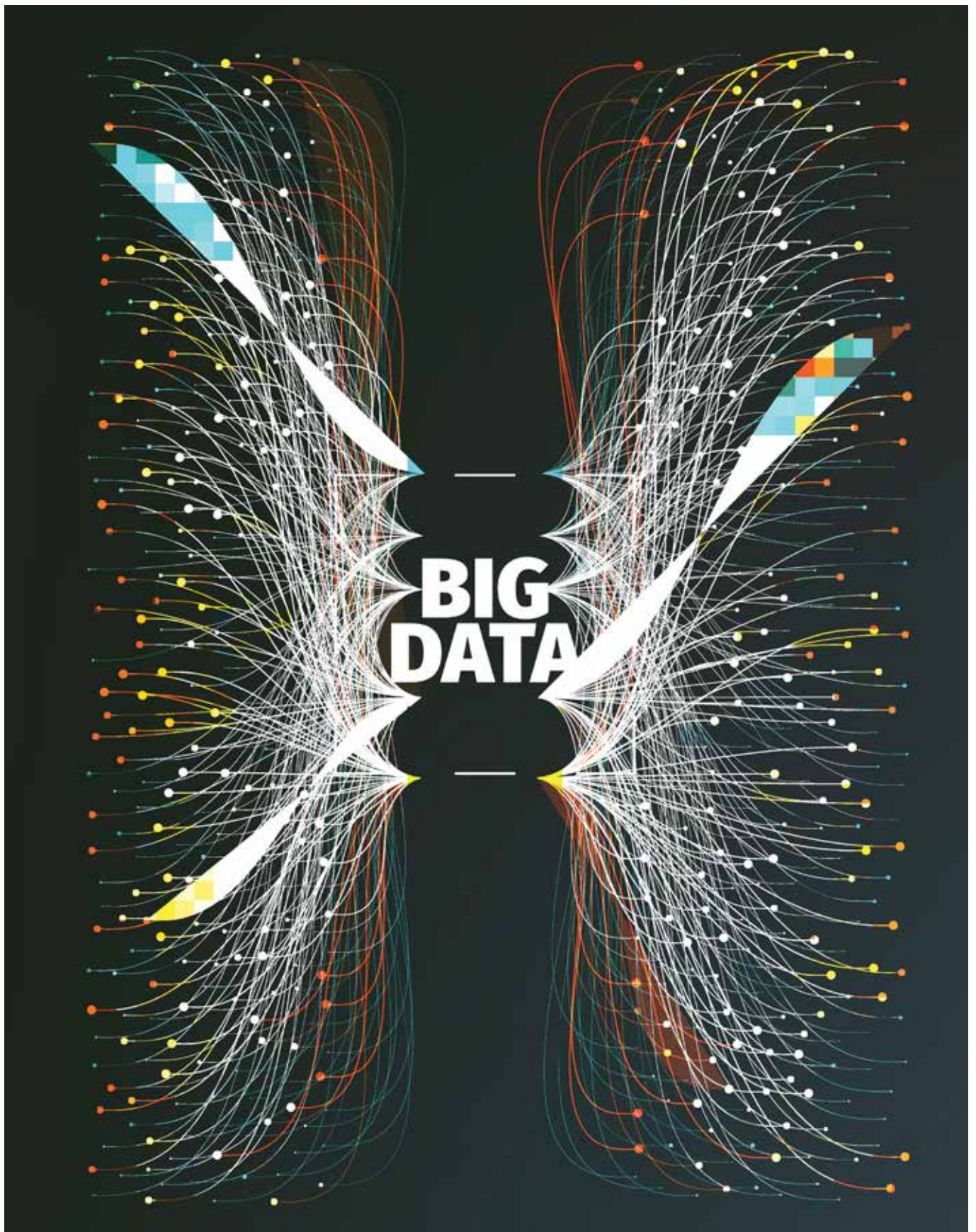
ההחלטה על אופן גיוס המשתתפים שישויו לקבוצת הביקורת, אף היא דורשת חשיבה. גיוס של ביקורות "היסטוריות" ממחקרים קודמים או מהספרות עשוי ליצור הטיה, מכיוון ששיטות הטיפול התומך, שיש להן השפעה סגולית על הפרוגנוזה, משתנות מעת לעת. לכן, מקובל להשתמש בביקורות עדכניות. על מנת לא ליצור הטיית בחירה, מקובל לשייך את משתתפי הניסוי לקבוצה שתקבל את ההתערבות הנבדקת, או לקבוצה שתקבל את התערבות הבקרה (פלסבו או את הטיפול המקובל) באופן אקראי (רנדומלי), מבלי שהחוקר או המשתתף יוכלו להשפיע על בחירה זו. במידה וזה אפשרי, יש חשיבות שהמחקר יהיה כפול-סמיות (double blinded), קרי: המשתתף והחוקר לא ידעו האם המשתתף מקבל את תרופת המחקר או את תרופת-הבקרה (פלסבו או הטיפול המקובל). זאת, על מנת למנוע הטיה באופן המעקב, ההערכה והדיווח. הטיה זו נובעת גם מציפיות המטופל, אך גם מציפיות החוקר. נמצא שמחקרים שנעשו ללא סמיות החוקר ו/או המשתתף נטו לייחס תועלת רבה יותר לטיפול הנחקר בהשוואה למחקרים שבחנו אותו טיפול באופן כפול-סמיות (9).

חיסרון נוסף של ניסויים קליניים היא משך המעקב שאינו ארוך דיו, וגודל המדגם המתוכנן לרוב, על מנת לזהות תועלת. שתי עובדות אלה מקשות על זיהוי תופעות לוואי נדירות, או השפעות מאוחרות המיוחסות להתערבות הנבדקת. זו הסיבה שגם לאחר שתרופה ניסיונית נכנסת לרשימת התרופות בשגר, ממשיכים לעקוב אחר דיווחים על תופעות חריגות בקרב המטופלים בה.

סקירה שיטתית וניתוח-על (מטה-אנליזה)

סקירה שיטתית (systematic review) היא איסוף שיטתי וממצה של תוצאות כל המחקרים שנעשו (כאלה שפורסמו וכאלה שלא) על מנת לתת תשובה על שאלת מחקר מגדרת. החיפוש אחר המידע נעשה על פי מילות מפתח רלבנטיות במאגרי המידע האלקטרוניים הקיימים (כגון: PubMed), ע"י סקירת תקצירים מכנסים, או פניה לחברות תרופות ולחוקרים המובילים בתחום בבקשה לקבל נתונים ממחקרים שלא פורסמו. יש להימנע ככל האפשר מהגבלת שפת הפרסום בעת החיפוש. החיפוש השיטתי אחר כל המידע הקיים בנושא נועד לצמצם את האפשרות להטיה, הנובעת מפרסום סלקטיבי של מחקרים המדווחים על תועלת מההתערבות, וגניזתם של מחקרים בהם לא נמצאה יעילות של ההתערבות (publication bias). לאחר שאותרו מחקרים פוטנציאליים כאלה, הם עוברים הערכה שיטתית על פי פרוטוקול שהוגדר מראש, תוך שימוש בכללים לגבי איזה סוג של מחקרים ייכללו (לדוגמה: רק ניסויים קליניים או גם מחקרים תצפיתיים), ואיך לדרג את המחקרים על פי החוזק המתודולוגי שלהם (למשל הסיכון להטיה). אם ניתן, מבוצע ניתוח-על (meta-analysis), על מנת לקבל מדד מסכם משוקלל אחד של האפקטיביות של ההתערבות. היתרון בסוג זה של מחקר הוא יכולתו לצמצם publication bias, ולאפשר מספיק עוצמה סטטיסטית על מנת לבחון את הקשר בין הטיפול ותוצאים נדירים (למשל תופעות לוואי נדירות או מוות).





Big Data - המהפכה שתשנה את הדרך בה אנו חיים, עובדים וחושבים

ד"ר אייל בנימין

ראש מסלול היזמות במכללה האקדמית תל אביב יפו,
ומנהל אקדמי של תכנית החדשנות בלהב הכשרת מנהלים באוניברסיטת תל-אביב

Big Data הוא מושג מופשט, אשר פרץ לשיחה בסביבות 2011, והפך לאחד המושגים ה'חמים' בתחומי השיווק, המחקר, הניהול והחיווי, תוך שהוא נוגע כמעט בכל תחום העולה על הדעת. למערכות Big Data מגוון שימושים, החל בהנגשת מידע מרשויות ציבוריות, עבור דרך ניהול רצפות ייצור, יישומים חברתיים, פרסום מקוון, וכלה ביישומים למחקר ורפואה. לצד הפוטנציאל העצום של השימוש ב-Big Data בהווה ובעיקר בעתיד, גם מגבלות ואתגרים רבים.

של מידע מסוגים שונים, ולנתח את המידע, בין אם לצורך קבלת החלטות, או לשם הפקת תובנות, בזמן קצר יחסית. עולם זה מתקיים בשלושה מימדים: (א) התהליכים לבניה ושימוש שוטף במערך Big Data, (ב) הטכנולוגיה הנדרשת לצורך כך, (ג) היישום או השימוש אשר מפק תועלת ממערך ה-Big Data שנבנה (4). לדוגמא, יישום Big Data לעולם הקמעונאות יתבסס על איגום מידע צרכני והתנהגותי של קהל מטרה, כגון רשומות כרטיסי האשראי, רשימות הקניות שנערכו, ורשימת אתרי האינטרנט בהם ביקרו הצרכנים; וישמש לצורך חיזוי ביקושים של סגמנטים בקהל, זיהוי מגמות ומתאמים בין התנהגויות וכד'. לצורך כך יהיה צורך בבנייה של מערכת אחסון ושרתי ניתוח מתאימים, ויידרוש מקורות שונים של מידע שזיורם באופן תקין למערכת הניתוח וקבלת ההחלטות.

שלבים בשימוש ב-Big Data

התהליכי ה-Big Data מורכבים מ-4 שלבים: יצירת המידע, איסוף המידע, אחסון המידע וניתוח המידע. שלב יצירת המידע עוסק בהמרת המידע לשטף של נתונים דיגיטליים בפורמט בר אחסון. הנתונים יכולים להגיע מחיישנים, מצלמות, מיקרופונים או כל אלמנט אחר. אמנם חלק גדול מהמידע נוצר דיגיטלי (למשל היסטורית הגלישה באינטרנט), וניתן לאחסון ושימוש כמעט מיידי, אולם סוגי מידע רבים אחרים נדרשים לעבור דיגיטליזציה בטרם ניתן יהיה לעשות בהם שימוש (למשל נתוני מזג אוויר, נתונים ביולוגיים, וכד').

יג דאטה (Big Data) הוא מושג מופשט, אשר פרץ לשיחה בסביבות 2011, והפך לאחד המושגים ה'חמים' בתחומי השיווק, המחקר, הניהול והחיווי, תוך שהוא נוגע כמעט בכל תחום העולה על הדעת. בספרם "Big Data - המהפכה שתשנה את הדרך בה אנו חיים, עובדים וחושבים" (1), מתארים ויקטור מאייר-שוברגר וקנט קוקייר את Big Data בתור "היכולת לעבד כמויות עצומות של מידע, לנתח אותו בעצמות גבוהה, ולהפיק מסקנות חדשות, חלקן מפתיעות". Big Data הצטרף למושגים קודמים כגון מסדי נתונים (Databases) או מחסני נתונים (Data Warehouses), וסולל את הדרך לנושאים נוספים כגון אנליזות מידע (Data Analytics), אנליזות חיזוי (Predictive Analytics), מדעי מידע (Data Scientists) ועוד. תחום ה-Big Data אינו מחליף או מייתר את השימוש במאגרי המידע הקלאסיים והמוכרים, אלא מייצג רובד נוסף בהתייחסות למידע הנצבר ולאופן השימוש בו (2). מסדי הנתונים משמשים לאחסון סדור של נתונים, לרוב במבנה טבלאי יחסי (Relational Database), ומאפשרים אחסון ושליפה של הנתונים ליישום ספציפי (למשל ניהול מלאי או פציינטים בקליניקה). לעומתם, מחסני הנתונים מאגדים נתונים ממספר מסדי נתונים, ומאפשרים שליפה, הצלבה, וניתוח מידע ממספר מקורות. אולם בסביבה הדיגיטלית של ימינו נוצרו הזדמנויות ודרישות נוספות מעולם המידע, שרובן נוגעות לשלושה מאפייני ליה: נפח המידע המאוחסן הצומח למימדי ענק (Volume), התפתחות מגוון סוגי המידע, הכוללים כיום למשל גם תמונות וקבצי קול (Variety), ומהירות הנגישות למידע או לנגזרותיו, שהיא נגזרת מכמות ומגוון המידע (Velocity). מאפיינים אלו הם ליבת המעבר לעולם ה-Big Data לצורך יצירת ערך (Value) חדש (3). מכאן, שמערכות Big Data מסוגלות לאחסן כמויות עצומות





ההחלטה על קצב העדכון תלויה במידה רבה בסוג הערך הנדרש ממערכת ה-Big Data. יישומי זמן אמת, כגון החלטה על הפרסומת שתוצג בעמוד אינטרנט או החלטת ניהוג לרכב אוטונומי, יידרשו מערכת איסוף מהירה ועדכנית; מערכות המחפשות חוקיות או סיביותיות בתוך המידע, יוכלו להסתפק במערך איסוף איטי יותר

תת בעיות הנשלחות למחשבים שונים במקבץ, וצמצום הפתרונות המתקבלים מכל מחשב לתוך פתרון אחד של הבעיה המרכזית. ניתוח מידע Big Data מתקשר גם לשני מושגים נוספים שהם 'בינה מלאכותית' (Artificial Intelligence – AI) המתייחסת לאלגוריתמיקה המדמה בינה טבעית (לרוב אנושית) בתחומי קבלת החלטות, זיהוי ואיתור תבניות; וכן ל'למידה מכונה' (Machine Learning – ML) המתייחסת להתנהגות מכונה, אשר אינה מבוססת תכנות יישיר אלא תוצאה של הישגות, חיזור (Iterations) ותוצאות קודמות. שני התחומים קיימים ומפותחים כבר משנות התשעים, אולם עד לאחרונה יושמו בסביבות קלאסיות, והמעבר ליישום מבוצע בנפחים גדולים האו אתגר חדש. מצד שני, ליישומי AI ו-ML במערכות Big Data יש פוטנציאל להפקת תובנות חדשות כמעט בכל תחום (10). אחת הדוגמאות המוכרות לשימוש ב-AI ו-ML הוא כמובן סירי של אפל, והמקבילות לה (אלכסה באמאזון וקורטנה במיקרוסופט).

מדוע זה טוב?

למערכות Big Data ישנן כמה חזקות משמעותיות אשר חשוב להתייחס אליהן. הראשונה שבהן היא גמישות ביחס למבנה הנתונים הזורם אל תוך המערכת. בעוד מערכות יחסיות (Relational Database) מצפות לקבל את המידע בפורמט ידוע מראש, מערכות Big Data מסוגלות לאכסן את המידע בפורמט המגיע אליהן בפועל. דבר המוריד את הצורך לצפות מראש את כל סוגי המידע העתידיים, או לחילופין, לעדכן את כל מבנה מסד הנתונים בהיתן מידע מסוג חדש. לצד הגמישות באופן בו נשמרים הנתונים, ולמבנה העלויות הנגזר מכך, מצטרפת גמישות יחסית בפיתוח האפליקציות הנדרשות. עם התפתחות התחום, ירדה כמות הקוד הנדרשת כדי לפתח אפליקציה. פעילות קהילת המפתחים סביב הקוד הפתוח הובילה לשכלול ועלייה ביעילות וגמישות כתיבה הקוד, כמו גם ביכולת לפיתוח משותף של פרויקטים (Collaboration) (11). חוזקה נוספת היא סולמיות או מדרגיות (Scalability/סקליליות) המערכת. היכולת האינהרנטית של מערכות Big Data לפעול באחסון/ניתוח מבוצע מאפשרות שמירה על אותה רמת ביצועים עם הגדלת יחידות העיבוד, זאת בניגוד למערכות קלאסיות אשר

מערך האיסוף מזרים את המידע שנוצר אל מערכות האכסון. במקרים בהם המידע נוצר בסמוך למערכות האכסון, למשל בקווי הייצור של מפעל, תהליך האיסוף הוא מובן מאליה. אולם במקרים בהם המקורות הם מרוחקים (אנשים הנמצאים בביתם) או ניידים (מכוניות חכמות) יש צורך בבניית מנגנון המזרים בשטף רציף, או במנות, את המידע העדכני אל אמצעי האחסנה. במהלך האיסוף יש לשמור גם על הלימה ואינטגרציה של המידע שנאסף עם מבנה ומנגנוני האחסנה הקיימים והמידע שכבר נמצא בהם. ההחלטה על קצב העדכון תלויה במידה רבה בסוג הערך הנדרש ממערכת ה-Big Data. יישומי זמן אמת, כגון החלטה על הפרסומת שתוצג בעמוד אינטרנט או החלטת ניהוג לרכב אוטונומי, יידרשו מערכת איסוף מהירה ועדכנית; מערכות המחפשות חוקיות או סיביותיות בתוך המידע, יוכלו להסתפק במערך איסוף איטי יותר.

אחסון המידע הוא שלב ביניים מהותי בין האיסוף לעיבוד. באופן טבעי אכסון מידע לשימושי Big Data דורש נפחים עצומים, בסדרי גודל של פטהביטים (10^{15} bit) ואף יותר של מידע. מכאן שסוגיית האכסון דורשת התייחסות חדשה לאמינות, זמינות, ומהירות הגישה למידע. התפתחות ה-Big Data הובילה לפיתוח של גישות חדשות לאכסון המידע כגון מסדי אחסנה מבוזרת (Distributed Storage), (5) המאפשרים פיצול האחסנה לשרתים/אזורים שונים, שיטות עיבוד חדשות (6), ומעבר למסדים מכווני מסמכים (Document Oriented Database), המאפשרים אחסון מידע מובנה חלקית (Semi Structured Database).

ניתוח מידע Big Data מיועד לשתי תפוקות מרכזיות: חיזוי מדויק ככל האפשר של תופעה, ו/או הפקת תובנות לגבי קשרים בין אלמנטים שונים. שתי תפוקות נוספות, המתאפשרות מגודל המאגר, הינן היכולת לאתר תבניות משמעותיות בתת-קבוצות, וכן היכולת לאתר מכנים משותפים באופן החוצה תת קבוצות במאגר, גם כאשר יש שונות גבוהה בין הפרטים (8). היות ומדובר בכמויות מידע ומהירות עיבוד אשר מעבר ליכולת מעבד בודד בטכנולוגיה הקיימת היום ובציפייה לביצועים סבירים, הפך הניתוח לאתגר כשלעצמו. תשתית העיבוד המקובלת היא פתרון בעיה באמצעות עיבוד מבוזר על גבי מקבץ (Cluster) של מחשבים. הגישה הנפוצה כיום היא שימוש בטכנולוגית MapReduce אשר פותחה על ידי צוות מגוול (9), ומשמשת את החברה במגוון אפליקציות. כמשתמע משמה, MapReduce מורכבת משני שלבים: מיפוי הבעיה לסדרה של



(סקירת קבצי MRI) לאורך זמן, ובהצלבה עם גורמי מידע אחרים לתת תובנות המותאמות אישית. מיזמים כגון DayTwo (19) העוסק בתזונה מותאמת אישית או Patients like me (20) העוסק בהתמודדות עם מחלות יתום (Orphan disease) (מחלות עם מספר חולים קטן אשר אינו מצדיק את ההשקעה במחקר לתרופות), ניגשים לבעיות רפואיות מוכרות באמצעות איסוף פרטני, והשוואה לקבוצה גדולה ומגוונת.

מגבלות של Big Data

לפלטפורמות Big Data, על כל מעלותיהן, ישנן גם מספר מגבלות. למשל מערכת Google Flu Trends (21) החווה את מגמות התפשטות השפעת על בסיס ניתוח מילות חיפוש מסוימות בגוגל, התבררה עם הטיה חזקה לאורך השנים אי דיוקים בחיזוי, כנראה בגלל הטיית הנובעות מהשפעות צולבות של מידע עונתי אחר והדינמיות של האלגוריתמים השונים המעורבים בחיזוי (22).

בויד וקראפורד (23) מתארים מספר אתגרים הנובעים מ-Big Data: ראשית, אימוץ רחב של יישומי Big Data ככלי עבודה עשוי להשפיע על מהות המחקר, תפיסת המציאות, ואף להשפיע על המציאות עצמה, בדיוק כפי שהשימוש בחשבונאות בעולם העסקי הפך מכלי המייצג את הפעילות לכלי המשפיע עליה, ועל החלטות המנהלים מעצם קיומו. שנית, השימוש בכמות עצומה של מידע דיגיטלי עשוי לייצור אשליית דיוק ואובייקטיביות, בעוד שלב ליבו של העניין שהוא הפרשנות ועומק הניתוח, אשר אמורים להגן בפני זיהוי תבניות והקשרים חסרי משמעות, הנובעים רק מעצם גודלו של בסיס המידע. איכות מקורות המידע המזינים את האלגוריתמים עשויה גם היא להיות אבן גוף בפעילות Big Data, קרי השימוש בטכנולוגיה חדשה ועוצמתית אינה מסירה את האחריות הנדרשת מהחוקרים בכל הקשור לתוקף מחקר, כגון גודל ורלוונטיות המדגם ובחינת מידת הרלוונטיות של Big Data לשאלת המחקר. למשל האם ניתוח מגמות הנובע מהמידע הזורם באפליקציות סטטיסטיות אכן מייצג את האוכלוסייה?

בעיות אלו מצטרפות לסוגיות נוספות של הקשרים (Context) הנובעים מהתאמת המידע העשיר לתוך מודלים מתמטיים; סוגיות הנובעות מאתגרים אתיים של שימוש במידע הנאגר למטרה אחת ליעוד אחר לחלוטין באופן העשוי לפגוע באדם או קהילה; וסוגיות הנובעות מזמינות מידע מסוים וייחודי בידי גופים, כגון חברות. גופים רבים, חלקם מסחריים גרידא, שולטים בנגישות אל המידע ו/או בחתכים ומידע אותו הם מעוניינים לחלוק עם אחרים ובכך עשויים להשפיע על הממצאים הנחשפים לחברה בכללותה.

העתיד של Big Data

למרות ההתפתחות העצומה של תחום ה-Big Data, זוהי כנראה רק ההתחלה. בעוד ישנם העושים שימוש ביכולות חדשות אלו בכדי להתמודד עם האתגר עתיק היומין של חיזוי מזג אוויר (24), אחרים חוששים מעתיד בו יכולות ה-Big Data מסכנות את עצם קיומן של דמוקרטיות כפי שאנו מכירים אותן כיום (25). בכל מקרה, כמו בתחומים אחרים, אנו נמצאים בעינה של מהפכה. ה-Big Data מניע שינויים בחברה, במדע ובכלכלה, אך בו זמנית גם ניוון מהם. תחומים חדשים כמו גרפיקה מרובת נתונים (Big Graph Data) (26) יאפשרו הצגה בהירה וברורה יותר של מגמות ותבניות, אוטונומיים כגון רכבים, אלגוריתמים ורובוטים יעשו שימוש מקוון בהשוואת מצב קיים למספר עצום של השוואות לקבלת החלטות (27). אך סקירה של התחום והמגמות (28) מעלה שעדיין ישנם אתגרים רבים לפנינו.

המאמר, כולל רשימת המקורות, מופיע באתר מכון תנובה למחקר

לרוב מגיעות אל נקודה בה תוספת של יחידות עיבוד מעלות את הביצועים במידה שולית בלבד. יכולת זו מאפשרת גמישות תכנונית, כמו גם השקעה מדורגת בהתאם לצרכי המידע המתפתחים.

מהירות הינה עוד סיבה טובה להשתמש במערכות Big Data. אולם לא תמיד. מחקרים מראים שלמערכות Big Data, ובמיוחד למסדים מכווני מסמכים (לרב ממשפחת NoSQL), יש יתרון משמעותי כאשר מדובר במידע בנפחים גבוהים או מספר גדול מאוד של פעולות, ביחס למסדי נתונים רלציוניים, אולם יש לציין כי היתרון אינו אחיד על פני מגוון הפעולות ומסדי הנתונים השונים.

עלות היא מרכיב משמעותי במערכות נתונים. עלותה הכוללת של מערכת Big Data מכילה מרכיבים רבים של חומרה, תוכנה, כ"א וכד', אולם מרכיב מסד הנתונים עצמו ניתן למידור באמצעות שימוש בתוכנות קוד פתוח. ישנו מגוון פתרונות רחב הזמין כיום לארגונים המעוניינים להיכנס לתחום בעלות סבירה כגון: Couch, MongoDB, Cassandra, DB ואחרים (13).

שימושים ל-Big Data

כאמור, למערכות Big Data מגוון שימושים. החל בהנגשת מידע מרשויות ציבוריות, עבור דרך ניהול רצפות ייצור, יישומים חברתיים, פרסום מקוון, וכלה ביישומים למחקר ורפואה. קצרה היריעה מלתאר את כולם, אולם ישנם כמה תחומים מרכזיים בהם ההשפעה כבר ניתנת: בעולם הניהול העסקי נוצרת הסטת ערך הנוצר ממידע עמוק אל ליבת העסקים (14). אם בעבר מערכות המידע היוו בעיקר מערכת תיעוד למטרות תפעוליות וחשבונאיות כגון ניהול מלאי, חיזוי מלאי, תמחור ודיווח, הרי שבארגונים בהם נוצרת כמויות גדולות של מידע, כגון מערכות קמעונאיות, כרטיסי אשראי וכד', מאפשרות מערכות Big Data לייצר ערך מוסף משמעותי וחדש, אשר לא היה קיים בעבר בתצורה של ניתוח התנהגויות, חיזוי והשפעה על התנהגות צרכנים ושווקים. שחקני סחר הגדולים כגון אמזון ואחרים משתמשים ביכולות אלו כדי ליצור יתרונות תחרותיים ולביצור מקומם בשוק. בעולם הפיננסים והכלכלה משמשים נתוני סחר, נתוני בורסה ונתונים כלכליים אחרים כבסיס לניתוח, מדידה וקבלת החלטות כגון החלטות השקעה, סחר, ייעוץ וניהול סיכונים (15). ניתוח בתחום כלכלת המקור, בהם מספר הפרמטרים צומח באופן מעריכי עם הגדלת המודל, הופכים להיות מעשיים יותר עם השימוש בכלי Big Data (16).

יישומים בעולם הרפואה

עולם הרפואה היישומי איטי יותר באימוץ ה-Big Data לחיי היום יום ביחס למערכות הפיננסיות. הסיבות לכך נובעות כנראה משמרנות טבעית של התחום בשינוי פרוטוקולים, לצד היעדר היסטוריה של תשתיות מידע שניתן לבנות עליה מערכי Big Data, וכמובן נושאים רגולטוריים כגון סודיות רפואית, זכות שימוש ובעלות על המידע, חשיפה לבעלי עניין מסחריים וכד'. יחד עם זאת, בעולם המחקר הרפואי נעשה כבר שימוש ביכולות Big Data בתחומים כגון גנטיקה ונירולוגיה. מגבלות המשאבים של מעבדות המעוניינות לבנות מודלים ולנתח מידע גנומי, מובילות אותן להתחבר למאגרים מרכזיים, כגון אלו של השוק האירופאי המשותף ואחרים, כדי לבצע מחקר. חשוב לציין שלמרות השימוש הרב של חוקרים בתחום הביולוגיה והגנטיקה במערכות Big Data, הם עדיין נאלצים להתמודד עם אתגרים רבים (17).

המעורבות הרגולטורית הנמוכה ברמת הפרט, לצד המודעות לבעיות הנובעות מחשיפה של נתונים אישיים, חברו ליכולות הטכנולוגיות של חיישני גוף זעירים ותקשורת מקוונת, אפשרו את התפתחות ה"אני הכמותי" (Quantified Self) (18). גישה זו מאפשרת לאסוף מידע מן האדם הבודד, בין אם באמצעות חיישנים (מד צעדים הינו הדוגמה הפשוטה ביותר), דגימות (הפרשות, טמפ', משקל ועוד) או הזנה



עולם הרפואה

היישומי איטי יותר

באימוץ ה-Big

Data לחיי היום יום

ביחס למערכות

הפיננסיות. הסיבות

לכך נובעות כנראה

משמרנות טבעית

של התחום בשינוי

פרוטוקולים, לצד

היעדר היסטוריה של

תשתיות מידע שניתן

לבנות עליה מערכי

Big Data, וכמובן

נושאים רגולטוריים

כגון סודיות רפואית,

זכות שימוש

ובעלות על המידע,

חשיפה לבעלי עניין

מסחריים וכד'

השימוש בביג דאטה בשירות הרפואה

פרופ' גדעון קורן, מאיה עדולמי ופרופ' ורדה שלו
מכון קאהן - מכבי למחקר וחדשנות

השנים האחרונות מתאפיינות בפיתוחן של שיטות מחשוב, שמסוגלות לנתח כמויות נתונים ענקיות (Big Data), ולהבחין בגילויים רפואיים ומדעיים שלא התגלו בניתוחי נתונים קלאסיים עד היום. ה"ביג דאטה" ולמידת המכונה עתידים לשנות את עולם הרפואה מן היסוד, ולהביא לפריצות מדעיות בכל הנוגע לפיתוח רפואה מותאמת אישית, גילוי מוקדם, דיוק מוגבר באבחנות ומניעת מחלות. אלו יתאפשרו על בסיס אלגוריתמים סטטיסטיים מתוחכמים המסוגלים "ללמוד" כמויות אדירות וסוגים שונים של נתונים (כגון שדות מובנים, טקסט חופשי, תמונה וקול) ולהסיק מסקנות מדויקות, במהירות ולעתים ללא הקשר ידוע מראש. העושר העצום של מאגרי המידע הרפואיים הדיגיטליים בישראל, בשילוב עם חוקרים וחברות העוסקים בפיתוח אלגוריתמים מורכבים, מעמידים את ישראל בחזית המחקר העולמית של ביג דאטה.

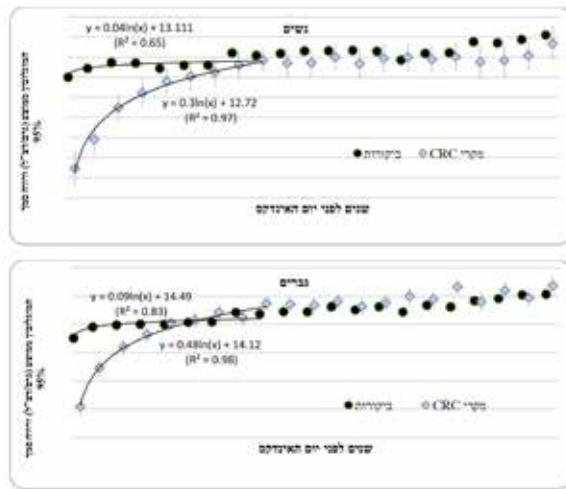
Machine Learning, Big Data ומה שביניהם

בעולם נתוני העתק, מקובל לתאר את המידע לפי 5 קטגוריות: **Volume**: כשמו כן הוא, מדובר בכמויות המידע. ההערכה היא שמדי חודש נוצרים כ-2.5 Exabytes של נתונים חדשים (Exabyte אחד מכיל מיליארד Gigabytes). **Variety**: סוגים שונים של מידע. דאטה מסורתי הוא דאטה מובנה (structured data) אשר מאורגן בטבלאות מוגדרות מראש. מידע לא מובנה כולל למעשה את כל היתר - נתוני טקסט חופשי, תמונות, קבצי PDF, הקלטות שמע, נתונים מתוך מכשירים לבישים (מדי סוכר, מדי דופק כדוגמת שעוני כושר) ועוד. **Veracity**: איכות ואמינות המידע - עד כמה ניתן להסתמך על המידע בקבלת החלטות מקצועיות. נכון להיום, אחד מכל שלושה מנהלים לא סומכים על המידע וניתוחו בקבלת החלטות, והמשמעות היא שרוב ההחלטות מתקבלות שלא על בסיס מידע, אלא על סמך אינטואיציה, היכרות עם הארגון וכיו"ב. **Velocity**: מהירות ההגעה של המידע החדש המצטבר מדי יום במערכות שיש לעבד. מקובל לומר כי 90% מהמידע בעולם הצטבר בשנתיים האחרונות (1). מדי יום מצטברות כמויות אדירות של מידע במערכות

בנים האחרונות מתקיים שיח משמעותי בקהילה הרפואית אודות השינויים וההבטחות שנושא עמו תחום נתוני עתק (Big Data). תחום עיסוק חדשני זה הינו בעל פוטנציאל להשפעה משמעותית ולמעשה - שינוי מן היסוד של האופן שבו מתנהלים המחקר והטיפול הקליני בתחומי הרפואה בשנים הקרובות. עולם הרפואה מתבסס כיום על ניסיון קליני שנצבר לאורך מאות שנים, תצפיות על חולים וממצאי מחקר מדעי בסיסי. עם השנים נוצרו פריצות דרך רפואיות רבות, והתפתחו טכנולוגיות רפואיות חדשניות. המורכבות הגדלה והולכת של גוף הידע הקליני בעולם הרפואה, במקביל לכניסת המחשבים לעולם הרפואה, הביאה לייצור כמויות אדירות של מידע דיגיטלי בפורמטים שונים - נתוני עתק, טכנולוגיות מבוססות נתוני עתק יכולות לסייע בגילוי מוקדם ומניעה של מחלות, לשנות את הדרך בה אנו מבצעים סקר לאיתור מוקדם של מחלות, להפוך את הטיפול הקליני למותאם אישית ומדויק יותר, לצמצם טעויות אנוש, לשפר את איכות החיים של המטופלים ורמת השירות, ולהעלות את טיב המחקר הרפואי המתקיים. מטרת מאמר זה היא להדגים כיצד, ובאילו עולמות תוכן, יכולים כלים מבוססים נתוני עתק לתרום לרופא המטפל ולחוקר. במאמר נסקור מושגי יסוד בתחום נתוני העתק ולמידת מכונה, נציג את השימושים בתחומים אלו בעולם העסקי ובעולם הרפואה, תוך מתן דוגמאות ממחקרים וטיפולים חדשניים, אשר מתקיימים כיום בישראל ובעולם.



גרף 1: רמות המוגלובין ממוצעות במשך 10 שנים לפני אבחנת סרטן מעי-גס, בהשוואה בין חולים ובריאים



נתוני עתק בעולם העסקי

בעידן הדיגיטלי הנתונים עוברים לחזית העסקית, בין אם מדובר בתעשיות מסורתיות או בחברות חדשניות, מתחומי הבנקאות, הטכנולוגיה, קמעונאות ותעשיית התרופות. הנתונים זוכים לכבוד הראוי להם, והופכים למשאב משמעותי ביכולת הארגון לייצר יתרון תחרותי, לחשוף תובנות עסקיות, ולייצר מוצרים חדשניים ומקורות הכנסה חלופיים (1). השימוש המסחרי המוכר ביותר כיום בתחום נתוני העתק הינו בעולמות השיווק והמכירות - הצעת מוצרים לרכישה פוטנציאלית על פי רכישות קודמות, והשוואתן לרוכשים דומים במשתנים דמוגרפיים, חברתיים וכיו"ב מהווה תשתית משמעותית בעולמות ה-e-commerce. לאחרונה, צוטט חבר הנהלה בכיר ב-Citibank באומרו "האינפורמציה חשובה בדיוק כמו הכסף" (7).

השימוש במידע מאפשר גם יצירת מודלים עסקיים חדשים: חברת UBER למשל, ערכה פיילוט בבריטניה, בו ניטרה את רמות זיהום האוויר בעיר באמצעות התקנת מד זיהום אוויר במכוניות אשר בשירותה. בכך הפכה UBER לתחנת ניטור הזיהום הגדולה והמדויקת ביותר בעולם, עם יכולות שימוש חדשות למידע שנאסף. באנלוגיה לעולם הרפואי, מידע הנאסף במערכות הבריאות לצרכי טיפול משמש באופן משני גם כמידע לצרכי מחקר. במקומות רבים בעולם, גם דגימות ביולוגיות הנאספות בכל מקרה, מאוחסנות ומשמשות לצרכים דומים, במה שמכונה ביו-בנק. דגימות אלו הן בעלות ערך מחקרי וכלכלי שלא יסולא בפז.

Patients like me הוא דוגמא לאתר, בו חולים מספקים מרצונם מידע אודות מצבם הרפואי בכדי לסייע לחולים אחרים, למחקר ולטיפול במצבם. למעלה מ-500,000 איש חולקים באתר מידע אודות מחלתם, הטיפול בהם, התסמינים שהם חווים ועוד. האתר הופך את הנתונים לתובנות עבור חולים אחרים במצב דומה, ומאפשר לחולים המעוניינים בכך ליצור קשר ישיר עם חולים אחרים במצבם. הרעיון הוא להעצים את החולה, לסייע לו בקבלת מידע על מצבו הרפואי, ועל החלטות אודות הטיפול בו. הרישום והשימוש באתר הוא ללא תשלום, כאשר מודל ההכנסות מבוסס על שיתופי פעולה מחקריים ועסקיים עם חברות תרופות, ציוד רפואי, משל וכו', על בסיס המידע. המידע אודות תופעות לוואי של תרופות, למשל, משותף עם חוקרי מנהל המזון והתרופות של ארה"ב ועם מחקרים נוספים (8).

(במהלך דקה נשלחים כ-150 מיליון מיילים, 2.4 מיליון שאלות בגוגל, 21 מיליון הודעות ב-whatsapp וכו'. בעולם האקדמי בלבד מתפרסמים כ-8,000 מאמרים ביום). עקב הצורך לקבל תשובות מהירות, בזמן אמת, נדרשת יכולת ניתוח גבוהות בכדי לעבד את המידע. Value: יצירת ערך עסקי ומטרות ברורות באמצעות תובנות כתוצאה מהמידע.

המידע לבדו אינו מספיק. על מנת להפוך את המידע לידע (Knowledge) בר שימוש עבור הארגון, ולהפיק ממנו תובנות, יש צורך בניתוח סטטיסטי חדשני, המסוגל למצוא במידע הקשרים ודפוסים, שבדרך כלל קשה לאתר באופן אחר.

פיטר סונדגרד, סמנכ"ל בכיר בחברת המחקר גרטנר, כותב במאמר שהתפרסם ב-2015 ב-Forbes: "נתוני עתק הם הנפט של המאה ה-21. למרות כל הערך שלהם, הנתונים הם טיפשים. הם אינם עושים כלום עד שאתה לומד כיצד להשתמש בהם. גם הנפט הגולמי הוא חסר ערך עד שמזקקים אותו והפכים אותו לדלק. גרסת הביג דאטה של נפט מזוקק - אלגוריתמים ייחודיים שנועדו לפתור בעיות מסוימות וניתן לתרגם אותם להחלטות ופעולות. זהו המתכון הסודי של הארגונים המצליחים בעתיד. הבהלה לזהב של העידן הדיגיטלי תהיה ממוקדת כיצד ניתן לעשות דברים בעלי ערך עם הנתונים" (2).

Machine Learning

Machine Learning (למידת מכונה) היא תחום המחקר המאפשר למחשבים את היכולת ללמוד חוקיות מנתונים, מבלי להיות מתוכנתים באופן ספציפי (3). מודלים מבוססים למידת מכונה מסוגלים ללמוד עצמם למצוא הקשרים, לזייק ולטייב מגמות ללא הקשר ידוע מראש. זאת בניגוד למודלים מבוססי חוקיות ידועה, אליהם מוזנים הכללים מבעוד מועד ומיושמים לאחר מכן. באופן זה, נחשפות בפנינו אפשרויות חדשות למציאת קשרים בלתי צפויים בין משתנים וזיהוי תבניות שימושיות (4). למידת המכונה כוללת בתוכה אלגוריתמים מודרניים שונים - שיטות סטטיסטיות מגוונות, אשר מסוגלות להסתמך על כושר החישוב הגבוה של מחשבים כיום, על מנת למצוא את אותם קשרים ודפוסים חדשים בנתונים (5). שיטות אלו פורצות דרך באפשרות ליצור מודלים לקשרים בעלי התפלגות שאינה ליניארית, או אינה דומה להתפלגות ידועה, שימוש במידע הכולל ערכים חסרים, וניצול מספר רב מאוד של משתנים לצורך חיזוי או זיהוי דפוסים. בניגוד למודלים המיועדים לשימוש ידני אנושי, ומחייבים מספר מועט של פרמטרים ושיטות חישוב קלות, אין צורך להגביל מודלים מחשבוניים, וניתן להעשיר במספר רב של משתנים ופירוט רב אודותיהם, ובכך לאפשר להם רבדים רבים יותר ורב מימדיות.

למידה עמוקה (Deep Learning)

למידה עמוקה (Deep Learning) היא אחת מהשיטות המובילות בתחום למידת המכונה, אשר נשענת על רעיון חיקוי פעולת המוח האנושי באופן ממוחשב, באמצעות רשתות נוירונים מלאכותיות. נתונים הנוספים כל העת למערכות אלו מאפשרים לשפר ולטייב את האלגוריתם הנבנה במהלך הזמן, בדומה למוח האנושי המשפר ומעדכן את הבנתו ככל שלומד מידע חדש. שיטות הלמידה העמוקה מיושמות בהצלחה בעולמות תוכן רבים כולל ניתוח תמונה, זיהוי קול וניתוח שפה טבעית (natural language processing) ועוד (6).

מרבית האלגוריתמים המודרניים ללמידת מכונה סובלים ממגבלה משמעותית ביכולת הפרשנות. תוצר מודל הניבוי מכונה לעתים רבות "קופסא שחורה" (Black Box), על שום הקושי להתחקות אחר גורמי הסיכון או משקלניהם, אשר הביאו בסופו של דבר למסקנות או ההקשרים שנמצאו במודל. מודל למידת מכונה מספק לרוב הקשרים תוצאתיים, אך לא סיבתיים - מדוע הקשר מסוים מתקיים ולא אחר. על כן, במובנים רבים הוא יכול לשמש כלי עזר לקבלת החלטות אך האלמנט האנושי - מומחה תוכן, תמיד יידרש בכדי לאתר טעויות אפשריות, להבין ולהכווין את התוצאות ולקבל החלטות מושכלות על בסיסן.





טבלה 1: יחס הסיכויים לאפשרות ממצאים מסוכנים במעי הגס באמצעות ספירת דם בודדת לעומת מספר ספירות דם

חלון זמן (חודשים)	טווח גילאים	גישה לספירות דם קודמות	יחס הסיכויים (CI 95%)
180-30	75-50	כן	29.34 [26.83, 32.09]
180-30	75-50	לא. ספירת דם בודדת.	21.2 [19.04, 23.32]

נתוני עתק בעולם הרפואה

עולם הרפואה עתיד להשתנות ולהרוויח מן ההתקדמות בתחום נתוני עתק יותר מכל. כבר היום, הדוגמאות המעניינות והמעשירות ביותר לשימושים יישומיים בנתוני עתק הינן בתחומי הרפואה הקלינית והמחקר הרפואי. מערכות אלו ישנו פרדיגמות ודרכי חשיבה שמלווים אותנו שנים רבות. עבור רופאים, הן יוכלו לשמש ככלי עזר מתקדם לקבלת החלטות המבוססות על מודלים סטטיסטיים מורכבים. עבור החוקר, הן יאפשרו ביצוע מחקרים פורצי דרך, הן בתוכנם, והן באופן ביצועם.

התחום הראשון בו ניכרים ניצני השינוי הוא טיפול רפואי מותאם אישית, ודיוק מוגבר באבחון ובטיפול: בשנות ה-70 של המאה הקודמת החלו להתפתח בעולם הרפואה מערכות תומכות החלטה קליניות (decision support systems). מטרתן היא לסייע לרופא לשפר את איכות הטיפול, למנוע טעויות רפואיות, לנהל באופן מושכל את משאבי הבריאות, ולהעלות את רמת השירות הניתן למטופלים. כיום, מערכות תומכות החלטה אלו מתבססות על כללים מוסכמים (rule based) בכדי לספק מידע ומסקנות לגבי כלל החולים במצבים שהוגדרו והוזנו מראש למערכות המידע. לצד התועלת הרבה של מערכות אלו, הכללים לפיהן הן עובדות אינם ספציפיים לחולה מסוים, אלא לקבוצות חולים כפי שהוגדרו מראש. כמו כן, מערכות אלו אינן מערכות לומדות, ואינן מתאימות את עצמן באופן אוטומטי לשינויים בגוף הידע הקיים, או למאפייני האוכלוסייה המטופלת, ודורשות תחזוקה אינטנסיבית ומתמדת. המהפכה המתהווה עם הכנסת נתוני עתק משנה את הסטטוס קוו. על בסיס מערכות למידת מכונה מבוססות נתוני עתק ניתן להתמודד עם כמויות אדירות של מידע, ולהסיק מסקנות מתוך הדאטה הקיים באופן בלתי תלוי, לעתים אפילו ללא ידע קליני או מחקר קודם; מערכות אלו מאפשרות "עידון" והתאמה של הכללים המקובלים לאוכלוסייה ספציפית או למצבים ספציפיים, הוספת כללים חדשים המבוססים על המידע הנחקר ומציאת חוקיות שונה ולעתים בלתי צפויה בין פריטי המידע מזו שהייתה ידועה עד כה.

במחקר שפורסם לאחרונה, בהתבסס על מאגר מידע בריטי שמכיל כ-380,000 מטופלים, ניסו החוקרים לנבא אירועים קרדי-וסקולריים. הבסיס להשוואה היה ההנחיות הקליניות המקובלות של איגוד בריאות הלב האמריקאי, הכוללות גורמי סיכון מסורתיים כגון גיל, כולסטרול, לחץ דם וסוכרת. החוקרים הרחיבו את המודל כך שיכלול גורמי סיכון נוספים מתוך התיקים הרפואיים, והפעילו מספר שיטות למידת מכונה (random forest, logistic regression, gradient boosting, neural networks) על מנת לנבא אירוע לבבי תוך עשור. כל המודלים חזו מעט יותר אירועים לבביים ביחס להנחיות

נתוני עתק הם
הנפט של המאה
ה-21. למרות
כל הערך שלהם,
הנתונים הם
טיפשיים. הם אינם
עושים כלום עד
שאתה לומד כיצד
להשתמש בהם.
גם הנפט הגולמי
הוא חסר ערך עד
שמזקקים אותו
והפכים אותו לדלק.
גרסת הביג דאטה
של נפט מזוקק-
אלגוריתמים
ייחודיים שנועדו
לפתור בעיות
מסוימות וניתן
לתרגם אותם
להחלטות ופעולות

הקליניות המקובלות, כאשר מודל הלמידה העמוקה (neural networks) חזה באופן המיטבי מביניהם - 7.6% יותר אירועים. מעניין לציין שהמודל החדש דירג מחלות נפש חמורות כגורם סיכון משמעותי, על אף שאינן נכללות בהנחיות המקובלות של איגוד הלב האמריקאי. לעומת זאת, סוכרת לא היוותה גורם משמעותי במודל החדש, אף שבהנחיות היא נחשבת לגורם סיכון משמעותי (9). עובדה זו עשויה להעלות תהיות בקרב קלינאים לגבי המודל, ומדגימה את הבעייתיות המכונה Black Box בהקשרים של שימוש בכלי נתוני עתק. על כן, נכון להיום, על מנת להסיק מסקנות ולחפש הקשרים בתוך המידע, נדרש צוות רב מקצועי, הכולל מומחי תוכן ומומחי אלגוריתמיקה ולמידת מכונה, בכדי לנטרל שימוש בלתי מדויק במערכות, והסקת מסקנות שגויה.

תחום נוסף בעולם הרפואה בו ניתוח נתוני עתק יתפוס מקום משמעותי כבר בשנים הקרובות הוא בפינוח תמונה המבוסס על קשר עין - מוח, כגון רדיולוגיה ופתולוגיה. התועלות בתחומים אלו נובעות מיכולת הדיוק הגבוהה של מערכות מבוססות נתוני עתק, ומיכולתן לזהות ולהצביע אלמנטים מסוימים בתמונה שעין אנושית אינה יכולה להבחין בהם. הדבר דומה מאוד לתחום הנהיגה האוטונומית, הסיבה שמערכות כגון Mobileye ודומותיה מצליחות היא בזכות שתי יכולות שמשתלבות: היכולת "לראות" ולנתח באופן בלתי אמצעי את הסביבה והגורמים האפשריים לתאונה, בשילוב עם האפשרות לפעול על סמך ניתוח סטטיסטי בלתי מוטא, ולנטרל את החסרונות האנושיים של הנהג כגון עייפות או הסחות דעת בזמן הנהיגה. פוטנציאל השימושים הקליניים והמחקריים בתחום זה הינו גבוה בכל הנוגע לרמת הדיוק ויעול האבחון אך נכון להיום, אינו יכול לבטל את הצורך בגורם האנושי בכדי לקבל החלטות.

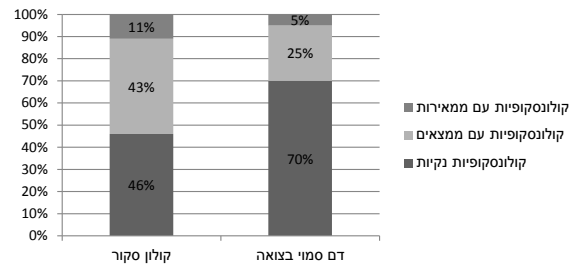
לאחרונה פורסמו תוצאות ניסוי של חטיבת המחקר של חברת Google בתחום הרטינופטיה - פגיעה ברשתית העין אצל חולי סכרת, ומהגורמים המובילים בעולם לעיוורון, שניתנת למניעה באבחון מוקדם. האלגוריתם שפותח בגוגל לניתוח תמונות רשתית העין הצליח להגיע לאותה רמת דיוק באבחון כמו זו של הרופאים המומחים, והמשמעות היא שבעתיד ניתן יהיה להשתמש בו ככלי עזר בכדי להפחית עומס על הרופאים, לאתר ולטפל מוקדם וביעילות במחלה (10).

מחקר עדכני בתחום זה נוגע לזיהוי ואבחנה של נגעים עוריים באופן אוטומטי. סרטן העור הינו הסרטן הנפוץ ביותר, ואובחן עד היום באופן ויזואלי ובאמצעות ביופסיה. המחקר, שנערך באוניברסיטת Stanford, אימן בשיטות למידת מכונה את האלגוריתם שפותח בכדי להבחין באופן בינארי בין נגעים ממאירים לכאלו שאינם ממאירים (קרצינומה של תאי הבסיס - הנפוצה ביותר מבין סוגי סרטן העור לעומת קרטוזיס סבוראי שפיר, וכן מלנומה ממאירה - הקטלנית ביותר מבין סוגי סרטן העור לעומת נבוס שפיר). האימון כלל 129,450 תמונות ובהן 2,032 מחלות שונות ותוצאותיו נבחנו אל מול 21 רופאים מומחי עור. בשני סוגי האבחנות האלגוריתם הגיע לתוצאות זהות לאלו של רופאי העור. תוצאות אלו מדגימות שבעתיד, ניתן יהיה לאבחן מרחוק נגעים עוריים (אולי אף באמצעות טלפון סלולרי בלבד), להוריד עלויות האבחון, להביא לאבחון מהיר יותר ולהנגישו לאוכלוסייה רחבה יותר, שאולי לא הייתה מגיעה אליו בדרכים אחרות (11).

אחת הדוגמאות הידועות ביותר לשימושי נתוני עתק ברפואה היא של IBM Watson. Watson ותוכנות אחרות, המשתמשות בשיטות למידה עמוקה, הוכיחו יכולת עיבוד וסיכום טובה של כמויות טקסט גדולות, בפרט לסקירת ספרות מדעית. Watson מסוגל לקרוא שפה טבעית במהירות אדירה (כ-25 מיליון מאמרים רפואיים בשבוע) מבלי להחסיר אף פרט, וכן לסרוק את כלל אתרי האינטרנט הרלוונטיים. כתוצאה מכך, נפתחות בפני הרופא המטפל אפשרויות טיפול רחבות יותר בהתבסס על המידע הרב המצטבר כל העת ברחבי העולם. במחקרים שנערכים כעת ב-Lineberger Comprehensive Cancer Center נדרש ווטסון להציע חלופות טיפול במצב הדורש קבלת החלטה



גרף 2: תוצאות בדיקות קולונוסקופיה שהתקיימו על סמך בדיקת דם סמוי בצואה או על סמך קולון סקור פרוספקטיבי



קלינית (ללא סטנדרטי טיפולי ברור). מתוצאות ראשוניות עולה, כי המערכת מציעה אופציות טיפוליות דומות לאלו שהומלצו על ידי ועדת המומחים של המרכז הרפואי, אך בכ-30% מהמקרים המערכת הציעה אופציות נוספות לטיפול שלא הועלו על ידי ועדת המומחים. בדרך כלל מדובר בטיפולים חדישים, שאושרו רק לאחרונה או שמצויים במסגרת מחקרים קליניים (12).

ביג דאטה במכבי - אלגוריתם לאבחון מוקדם של סרטן המעי הגס

החל מתחילת שנות ה-90, החלה מכבי שירותי בריאות לעבור תהליך מחשוב מקיף, כולל הקמת מאגרי מידע דיגיטליים. הייחודיות במאגר המידע של מכבי הוא בקימו של בסיס נתונים מרכזי המכיל מיליוני רשומות אודות כ-25% מאוכלוסיית המדינה באופן מייצג, לאורך שנים רבות, ותוך איסוף של כל סוגי המידע (אבחנות, בדיקות מעבדה, מדדים, נתונים דמוגרפיים, הדמיות ועוד). שילוב של גורמים אלו מעמיד את ישראל ואת מכבי שירותי בריאות בחזית המחקר העולמית ומהווה כר פורה לשיתופי פעולה, למחקרים חדשניים ופריצות דרך מדעיות ומחקריות.

להדגמת הפוטנציאל העצום של מחקר ביג דאטה ברפואה, נביא בפניכם מחקר פורץ דרך שהתנהל במכון המחקר של מכבי, ואשר הביא לשינוי משמעותי ביכולת לזיהוי מוקדם של סרטן המעי הגס.

סיפור פיתוח ויישום הקולון סקור

בסוף שנת 2006, אובחן חולה בן 70 בסרטן המעי הגס. בדומה לחולים אחרים, גם חולה זה לא התייצב לבדיקות השגרתיות של דם סמוי בצואה, למרות הזמנות חוזרות ונשנות לעשותן. בדומה לכ-60% מהחולים, גם חולה זה אובחן בשלב מתקדם של המחלה אחרי שסבל מכאבי בטן ונפיחות. בזמן האבחנה היתה מחלתו מפושטת עם גרורות לריאות ולעצמות, והוא נפטר תוך שנתיים מאבחנתו. לאחר האבחנה בדקה רופאת המשפחה שלו את כל ממצאי התיק הרפואי כדי לוודא ששום סימן או סמן מעבדתי לא הוחמץ. בשנים טרם אבחנתו נעשו לחולה 5 ספירות דם וכולן היו בתחום הנורמה, ועל כן לא הביאו לחשד כלשהו. אולם בהסתכלות יותר מדוקדקת הסתבר שספירות הדם השגרתיות של המטופל הדגימו ירידה קטנה ומתמשכת ברמות ההמוגלובין לאורך 3 שנים טרם האבחנה, למרות שנותרו במסגרת הטווח הנורמלי (ללא אנמיה) ולא עוררו חשד.

כתוצאה מהתצפית הקלינית של חולה אחד, הותנע במכון המחקר של מכבי מחקר אפידמיולוגי רטרוספקטיבי, שנועד לבדוק אפשרות לאיתור מוקדם של חולי סרטן המעי הגס, באמצעות תוצאות ספירות הדם השגרתיות ובפרט רמות ההמוגלובין. במסגרת המחקר נבדקו תוצאות בדיקות המעבדה של מטופלים שאובחנו

בגילאי 45-75 בין השנים 2004-2009, ואכן נמצא כי כ-3 שנים טרם האבחנה בסרטן המעי הגס, החלו רמות ההמוגלובין לרדת בתוך טווח הנורמה (גרף 1) (13).

בשלב זה התפתח הרעיון להשתמש בשיטות של למידת מכונה, שתדגמנה בהמשך, ולפתח מודל פרוספקטיבי לאיתור והתרעה מבעוד מועד. לצורך כך נוצר שיתוף הפעולה עם חברת מדיאל - Medial CS, אשר פיתחה אלגוריתם ("קולון סקור" Colon Score) הבוחן את כלל המדדים המצויים בספירות הדם, בשילוב גיל ומין. אימון המודל ותיקופו נערך בישראל על בסיס המידע של מכבי שירותי בריאות (180,000 איש), ובהמשך בוצע גם תיקוף חיצוני על בסיס מאגר מידע בריטי (אוכלוסייה של כ-35,000 איש).

אופן פיתוח ה"קולון סקור" מדגים היטב את האפקטיביות שבשימוש בנתוני עתק ולמידת מכונה: בעוד שהמחקר הראשון התקיים בשיטות המקובלות, תוך התבססות על תצפית בודדת והיפותזה הנוגעת לירידה בהמוגלובין לאורך זמן, האלגוריתם שפותח לקח בחשבון פרמטרים רבים נוספים, ומצא הקשרים שלא נצפו מראש. במהלך הפיתוח נמצא כי האלגוריתם תומך גם בזיהוי על פי ספירת דם בודדת ברמה סבירה, אם כי פחותה מאשר במספר ספירות עוקבות (ראה טבלה 1).

לאור תוצאות המחקר החיוביות, הוטמע האלגוריתם בתחילת שנת 2016 במכבי שירותי בריאות. הוא נחשף לנתוני ספירות הדם של מבטחי הקופה בגילאים 50-75 שלא עברו בדיקות סקר, ומתריע בפני הרופא כאשר מתקבל ציון קולון סקור גבוה (0.5% העליונים עפ"י האלגוריתם) כלומר סיכון מוגבר לממצא ממאיר. במהלך החודשים הראשונים לאחר הטמעת הקולון סקור, אותרו כ-50 חולים רלוונטיים בחודש וההפניות לבדיקות קולונוסקופיה עלו ב-1% (14). התוצאות מראות כי ברמת סגוליות של 99.6%, האלגוריתם מביא לתוצאות טובות יותר מאלו של בדיקות דם סמוי בצואה (גרף 2): בסיון שנעשה באמצעות קולון סקור התגלו ממצאים ממאירים ב-11% מבדיקות הקולונוסקופיה, לעומת 5% בלבד בסיון באמצעות בדיקות דם סמוי בצואה. נכון לחודש יוני 2017, אותרו באמצעות ה"קולון סקור" 104 חולים עם ממצאים (פוליפים) ו-25 חולים עם ממצאים. כאמור, מדובר בחולים שאותרו למרות שלא ביצעו בדיקות סקר.

מכבי שירותי בריאות וחברת מדיאל רשמו פטנט על המודל החדשני, אשר ממשיך לעורר עניין ברחבי העולם.

לראשונה בישראל, המכון מפתח פלטפורמה טכנולוגית חדשנית אשר תנגיש את המידע הקיים במסד הנתונים של מכבי שירותי בריאות עבור חוקרים מכל מקום - באמצעות האינטרנט. המידע שיונגש באמצעות הפלטפורמה יעבור תהליכי אנונימיזציה ו"ערבול נתונים" - מה שיבטיח למאגר נתונים סינטיטי אך מייצג סטטיסטית המאפשר שמירה על חסיון הפרט.

סיכום

עולם הרפואה הקלינית והמחקרית דרש מאז מעולם התמודדות עם כמיוות עצומות של מידע מסוגים שונים. ככל שמצבם הרפואי של החולים נעשה מורכב יותר, כמויות המידע גדלות והטכנולוגיה מתקדמת, עולמות נתוני העתק ולמידת המכונה הופכים לרלוונטיים ואף הכרחיים לרופאים קליניים ולחוקרים, לטובת שיפור הרפואה (5). השינוי לא יהיה מהיר וחד כפי שקרה בתעשיות אחרות. נוכחות הרופא, שיקול דעתו ויכולותיו המקצועיות יישארו החלק המרכזי והמהותי בטיפול בחולה ואף יתעצמו ויתחדדו עם התפתחות מגמות אלו. עבור המטופלים והמחקר הרפואי, מדובר בבשורה של ממש שתשפר את הגישה והאיכות לשירותים הרפואיים.

המאמר, כולל רשימת המקורות, מופיע באתר מכון תנובה למחקר

מעניין לציין

שהמודל החדש

דירג מחלות נפש

חמורות כגורם

סיכון משמעותי על

אף שאין נכללות

בהנחיות המקובלות

של איגוד הלב

האמריקאי.

לעומת זאת,

סוכרת לא היותה

גורם משמעותי

במודל החדש, אף

שבהנחיות היא

נחשבת לגורם סיכון

משמעותי



מיקרוביום - ביג דאטה, אתגר ענק, פוטנציאל עצום

מיה אלחל

מייסדת אש - גוף מדיה המתמחה בחדשנות בבריאות ורפואה,
מנהלת אקדמית ומרצה בקורס עתיד הבריאות ביחידה להכשרת מנהלים במרכז
הבינתחומי הרצליה ועורכת אתר החדשות והטכנולוגיה Longevity Science

**חוקרים, רופאים, תזונאים, מהנדסים, ממצאים, מפתחי קוד פתוח,
ואזרחים מעורבים סביב העולם מחוללים יחד מהפכה במחקר המיקרוביום
וביישומיו בבריאות האדם, בחקלאות, במדע המזון ובתחומים חיוניים
אחרים.**

**כל אחד יכול לקחת חלק קטן במאמץ, וההשקעה עשויה להשתלם לנו
בעתיד הלא רחוק, באיכות ותוחלת בריאותנו, ובקיימות סביבת חיינו.**

ככל שחוקרים משתמשים בקבוצות נתונים גדולות יותר, כך הם יכולים לתכנן מחקרים מעמיקים ואיכותיים יותר, לשאול ולענות על שאלות מורכבות יותר ברמת דיוק גבוהה. לרשותם טכנולוגיות ריצוף, שהולכות ומשתכללות, ומגבירות את תפוקת הנתונים בקצב מדהים, אך התוצאה היא גם גידול אדיר בדרישה ליכולות ניתוח נתונים (4). זהו אתגר שמחייב שילוב כוחות, וגם בתחום הזה אנו רואים התקדמות משמעותית בשנים האחרונות. פרוטוקולים ששימשו את ה-Human Microbiome Project (5) ופריקטים דומים אומצו על ידי חוקרים ברחבי העולם, ביניהם חישוב בענן (6), והם מסייעים בעריכת מטא-אנליזה ושילוב ממצאים.

מאמצים ראשוניים לאפיין ולהבין את המיקרוביום האנושי הבריאות תוך שימוש בטכניקות ריצוף מתקדמות עוררו יותר שאלות מתשובות (7). הם הובילו לגידול מואץ במחקר, אשר חשף קשרים בין המיקרוביום לבין מחלות שונות - החל מהשמנת יתר (8) ומחלות מעיים דלקתיות ועד מחלות לב וכלי דם, עור ודיכאון (9) בין היתר. לגבי מחלות אלו ואחרות השאלה האם מדובר בקשר סיבתי היא בלב מאמצי המחקר ובשלב הבא שילוב שיטות מחקר, למשל בין נתוני הריצוף הגנטי לבין טכניקות ריצוף חלבונים, ריצוף של ה-RNA ואפיון מטבוליסטים עשוי לשמש מפתח להתקדמות נוספת בהבנת השפעות המיקרוביום על מניעת והתפתחות מחלות.

אחת הדוגמאות לעוצמת השילוב בין מספר טכניקות היא פרויקט המיקרוביום האנושי - Human Microbiome Project (או HMP) של ה-National Institute of Health האמריקני (NIH) שמטרתו לקדם את האפיון של המיקרוביום האנושי, כדי לשפר את ההבנה שלנו לגבי האופן שבו המיקרוביום משפיע על בריאות וחולי בבני אדם, ועבור נשים, הוקם פרויקט מיוחד (10) להבנת השפעות המיקרוביום הוגנלי על בריאותנו. בעת פרסום השלב הראשון שלו,

קר המיקרוביום נמצא על סף פריצת דרך מבחינת היקף הנתונים המופקים ורמת הספסיפיקציה. חסידי חקר המיקרוביום מאמינים שבעתיד הקרוב נגלה שבעוד ששנים נהגנו לומר "אנחנו מה שאנחנו אוכלים", בעצם "אנחנו מה שאנו סופגים ומפיקים ממה שאנו אוכלים". במילים אחרות, בזכות הקשר הסמביוטי בין האדם לקהילות המיקרוביום שהוא מארח, למעשה "אנחנו מה שאנחנו מאכילים את המיקרוביום שלנו". לשיטתם, חלקם של עשרות טריליוני האורגניזמים החיים על עורנו ובגופנו הוא כה משמעותי, לא רק בתהליכים מטאבוליים, אלא גם בהשפעתם על מערכת החיסון והתנהגויותינו (1), שהם כמפעיל השולט באדם כבובוה על חוטים. התשובות שתעלינה מהמחקר הענף בשנים הקרובות עשויות לשנות את האופן שבו חשבנו על בריאות עד כה, אך כגודל הפוטנציאל הגלום בנתונים, כך גדול גם האתגר בניתוחם. כדי לבחון נתונים בהיקף כה גדול באופן מדויק, ולהפיק מהם תובנות והמלצות מעשיות בתזונה וברפואה, אנו זקוקים לחדשנות והתייעלות בכלי עיבוד נתונים, לאימוץ סטנדרטים, ולשיתוף פעולה חוצה מכוני מחקר, אוניברסיטאות ומדינות וגם להתגייסות הרגולטור והציבור. עשינו דרך ארוכה בזמן קצר להפליא מזיהוי מבודד וצר, למאמץ אנליטי עולמי מרוכז ורחב היקף לריצוף מערך המיקרוביום בגוף האדם, וזוהי רק ההתחלה. מאז טבע ג'ושוע לדרברג את המונח 'microbiome' ב-2001 (2) הלך תחום המחקר וגדל במהירות עצומה מבחינת הטרונגויות הנתונים והיקפם. בתחילת הדרך, הגישות לאפיון המיקרוביום התבססו על שיטות זיהוי ממוקדות-מטרה, ובעיקר עסקו באיתור פתוגנים. אך בשנים האחרונות, שילובן של טכנולוגיות ריצוף מחולל מהפיכה בתחום, ומאפשר לחוקרים לדגום ולאפיין קהילות מיקרוביאליות עצומות (3) באופן ישיר, ולהעמיק את הכרותנו עם אלה התומכים בתהליכים מטאבוליים חיוניים.





אחד האתגרים
בפרויקטים מסוג
זה הוא ריבוי
הדגימות: בעוד
שרוב מחקרי
המיקרוביום
הקיימים כוללים
מאות או אלפי
דגימות, המיזמים
האזרחיים כוללים
מספר הולך וגדל
של דגימות,
ובמקרים מסוימים
מספרן עולה על
עשרת אלפים
דגימות

מיזמים לשיתוף הציבור במחקר בתחום המיקרוביום - כמו ה-American Gut Project תרמו תרומה משמעותית לתחום על ידי דמוקרטיזציה של מחקר המיקרוביום, וסיפקו מערכי נתונים בקנה מידה גדול, שיכולים לשמש כמסגרות השוואתיות למחקרים אחרים. אזרחים מהשורה מסייעים למדע על ידי שליחת דגימות מגופם, מחיות המחמד שלהם, או מהסביבה שבה הם חיים, וגם תורמים את הכספים הדרושים למימון העיבוד של הדגימות. אחד האתגרים בפרויקטים מסוג זה הוא ריבוי הדגימות: בעוד שרוב מחקרי המיקרוביום הקיימים כוללים מאות או אלפי דגימות, המיזמים האזרחיים כוללים מספר הולך וגדל של דגימות, ובמקרים מסוימים מספרן עולה על עשרת אלפים דגימות. כמו כן, היות והנתונים נמסרים על ידי האנשים עצמם, יש צורך בכמויות משמעותיות של נתונים נוספים כדי לאפיין נכונה את מקור המדגם. ה-American Gut Project אוסף מאות משתנים על משתתפי המחקר, לרבות מידע תזונתי מפורט, באמצעות שאלון מתוקן על תדירות הצריכה של מזונות מסוימים. ניתוח כל המשתנים הללו אינו פשוט, ואחד הפתרונות הוא הרחבת המשמעות של citizen science בדרך של מיקור חוץ להמונים של ניתוח הנתונים עצמו.

מיקרוביום, ביג DATA ואפידמיולוגיה

החוקרים גתרי וגרדי (15) מתארים את התועלת שבשימוש בטכניקות מתקדמות לריצוף מלא כדי להבין את ההתפרצויות של מחלת השחפת מסוג Mycobacterium. ריצוף של הגנום המלא של פתוגן ספציפי יכול לחשוף את נתיב הזיהום, לרבות של חולה 0 בהתפרצות, על ידי כך שהוא מאפשר לחוקרים לעקוב אחר מוטציות ממספר זנים המבודדים מחולים נגועים. לא רק התחקות אחר המקור ושרשרת הפצה (16), אלא גם טיפול. העמידות האנטיביוטית של שחפת מסוג Mycobacterium היא בעיה קלינית

בשנת 2012 היה זה פרויקט ריצוף המיקרוביום האנושי הגדול ביותר. אלפי דגימות שנאספו מ-242 מבוגרים בריאים (11) ממספר אתרים שונים בגוף: מעברי האף, חלל הפה, העור, דרכי העיכול והמערכת האורוגניטלית, רוצפו תוך שימוש במגוון שיטות, והתוצאה היתה נתונים בהיקף של 2.9 ג'יגה-בייט לדגימה וקרב ל-2 טרה-בייט של מידע לקבוצת הנתונים כולה. בסך הכל, נוצרו במסגרת הפרויקט עד כה-14.23 טרה-בייט של נתונים, והם זמינים, ברוח הגישה הפתוחה, לציבור הרחב באתר של ה-HMP (12).

שפע ונגישות הנתונים הזניק שלל חברות פרטיות בתחום ואת ההשקעות בהן (13). בהן כאלה המפתחות מוצרים למיקרוביום כפרוביוטיקה ופרה-ביוטיקה, חברות מחקר המייעדות את תוצאות מחקריהן לשיתופי פעולה פוטנציאליים עם חברות תרופות, וחברות הפונות ישירות אל הצרכן, במודל הדומה לשירותי ריצוף גינום אנושי כ-23andMe, ומציעות שירותי ריצוף מיקרוביום והמלצות תזונה - ביניהן הישראלית Viome ו-DayTwo. ידע פרקטי רב נצבר במחקרי תעשיית המיקרוביום שהולכת ומתפתחת ומושכת השקעות במאות מיליונים.

פרויקט המיקרוביום האנושי, ופרויקטים אחרים העוסקים במיקרוביום ברחבי העולם, יצרו בסיס חשוב להבנת המיקרוביום החיים בגופו של כל אחד מאיתנו. אולם עד כה, יכולתו של הציבור לקחת חלק פעיל במחקרים כאלה היתה מוגבלת. כעת, ה-American Gut (14) מאפשר לנו (בינתיים לאמריקאים אך יש גם גרסה בריטית - British Gut) להצטרף לפרויקט המחקר, ולהשוות את המיקרוביום שבמעיים שלנו לאלה המצויים במעיים של אלפי אנשים אחרים בארה"ב וברחבי העולם. ה-American Gut הוא פרויקט המבוסס על עקרונות הקוד הפתוח. הנתונים על המיקרוביום שלנו נאספים לצורך הבנה, והם ישותפו הן עם המשתתפים האחרים במחקר, והן עם מדענים אחרים. בתמונה 1 דוגמה למיקרוביום של היסטוריון המזון הנודע ומחבר רבי המכר מייקל פולן, המוכר בעיקר בזכות הספר "דילמת השפע".





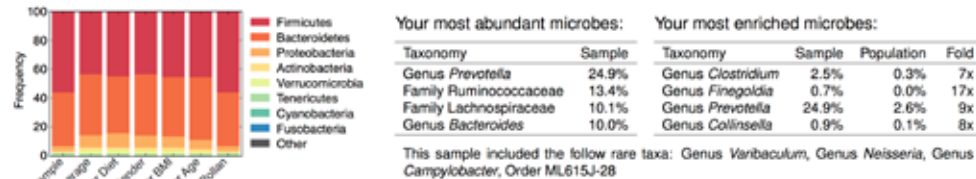
תמונה 1: המיקרוביום של היסטוריון המזון הנודע, ומחבר רבי המכר, מייקל פולן



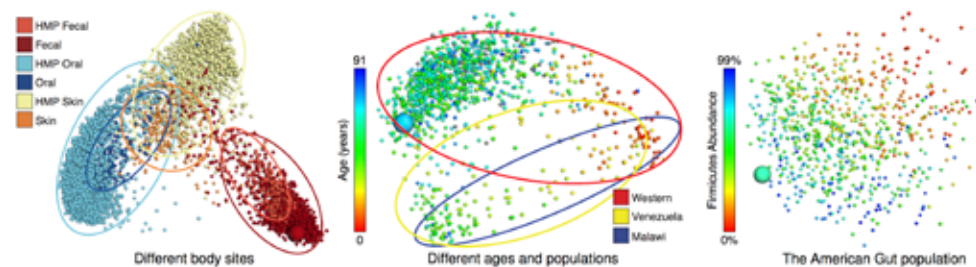
YOUR AMERICAN GUT SAMPLE

MICHAEL POLLAN

What's in your American Gut sample?



How do your gut microbes compare to others?



עשינו כבר דרך

בהבנת חשיבות

המיקרוביום,

אך דרושה עוד

התקדמות עד

שנוכל לתרגם

את התוצאות

במעבדה לטיפולים

ולהמלצות ממשיות

לשינוי הרגלים

בחיי היומיום

אתגרים ומבט קדימה

גוף האדם נושא מגוון רחב של מיקרואורגניזמים, והם מהווים את החלק הארי מהחומר הגנטי המצוי בגופנו. הכלים והטכנולוגיות שאיפשרו את התפתחות המחקר במיקרוביום עד כה ממשיכים להשתפר בקצב מסחרר. השימוש במרצפים ניידים בעתיד יביא לכך שכמות הנתונים שתיאסף על ידי חוקרי מיקרוביום תגיע להיקפים עצומים עוד יותר. לא ניתן יהיה להתעלם מהנתונים האלה כאשר בוחנים את בריאותו של אדם, ועשינו כבר דרך בהבנת חשיבות המיקרוביום. אך דרושה עוד התקדמות עד שנוכל לתרגם את התוצאות במעבדה לטיפולים ולהמלצות ממשיות לשינוי הרגלים בחיי היומיום.

בכדי להגיע ליישומים פרקטיים נדרשים נהלים סטנדרטיים אוניברסליים - לרבות פרוטוקולים לאיסוף, טיפול, אחסון ועיבוד של דגימות, כמו גם חדשנות רגולטורית. על מנת לרתום אל המאמץ בעלי ענין מתחומים שונים ואת הציבור חשוב ליצור דוחות קלים להבנה של ניתוח הנתונים, וכמובן יהיה צורך להתאים את ההכשרה של קלינאים בתחום. בנוסף, הצורך להפחית את העלויות של עיבוד וניתוח הדגימות. שיפורים בכל אלו יחברו יחד לקידום הבנת קשרים משמעותיים בין המיקרוביום לבין הגוף המארח, וניתן יהיה למנף את הקשרים הללו לגבהים חדשים בבריאות.

שפע זה של מידע גם יאפשר לחוקרים להבין כיצד פועלות קהילות של חיידקים ואופן האינטראקציה ביניהן, ויוביל להתפתחויות פורצות דרך לא רק בתחום בריאות בני האדם, אלא גם בחקלאות (19), בדלק ביולוגי (20) וביישומים רבים אחרים.

המאמר, כולל רשימת המקורות, מופיע באתר מכון תנובה למחקר

לא פשוטה; הטכניקות הנוכחיות מבוססות על תרבויות, וקבלת התוצאות עלולה לקחת עד 8 שבועות. במקרה זה, כאשר המוטציות האחריות על העמידות לתרופות מאופיינות היטב, והאיכות של מסדי הנתונים הרלוונטים גבוהה, ריצוף הגנום כולו עשוי להפחית את משך הזמן של קבלת התוצאות לימים בודדים. אחד מתוצרי הלוואי החיוביים של השימוש בריצוף גנום שלם לצורך מעקב ואבחון מהיר של התפרצויות של מחלות הוא היווצרותו של מסד נתונים גנומי, העשוי לסייע לפיתוח תרופות חדשות, ולשיפור האפיון של מחלות.

חקר המיקרוביום ותוחלת הבריאות האנושית

הנתונים על מיקרוביום, והתובנות שנפיק מהם, עשויים להשפיע על תוחלת החיים הבריאים (Healthspan) ממש. לאנשים החיים מעל לגיל 90 יש דפוסים אחרים של הרכב המיקרוביום במעיים מאשר לכלל האוכלוסייה. שאלת מיליון דולר היא האם חיידקים אלה הם חלק מהסיבה שהאנשים הללו מאריכים חיים יותר מרוב האנשים. בינתיים מדענים גילו כי לחיידקים מסוימים יש השפעה על תוחלת החיים של תולעת בשם *C. elegans*. בפרט הם גילו 12 גרסאות שונות של החיידק *E. coli* שקבעו את תוחלת החיים של התולעים הללו, ואפילו עיכבו אצלן את הצמיחה של גידולים הקשורים לגיל (17). מוקדם מדי לדעת איך התוצאות בחסרי חוליות יתורגמו ליונקים אך זהו רמז ראשוני לכך שכל שנבין יותר את חיידקי המעיים שלנו, נדע אולי כיצד נוכל להאריך את גם תוחלת החיים הבריאים שלנו. ראיות אנקדוטיות לכך אולי כבר קיימות עבור הזדקנות ה-Mice (18).



חנה, בת 50. בשנתיים האחרונות, מאז הפסקת המחזור החודשי החלה לסבול מעצירות כרונית. כל הבדיקות שלה - תקינות. היא ניסתה את כל המשלשלים הקיימים בשוק ללא הועיל. לאחר ביקור אצל רופא גסטרו שהעלה אפשרות לתסמונת המעי הרגיז/ש - IBS, היא קראה ואספה מהאינטרנט מידע רב אודות ההפרעה. הסימפטומים שהיא סובלת מהם: יציאה קשה אחת לשבוע לערך, גזים ונפיחות ביום-יומיים לפני ההתרוקנות בלבד. אין שינויים ביציאות, לא נצפה קשר למצבי לחץ או מתח יוצאי דופן, ואין סימפטומים נוספים המצדיקים את האבחון הסופי של IBS. למרות זאת, חנה נאחזת בהצעה זו ומגיעה אליי עם בקשה מפורשת: עזרי לי ליישם את הדיאטה דלת FODMAP.

דיאטה דלת FODMAP היא הדיאטה הנחקרת ביותר בתחום של תסמונת המעי הרגיז. יחד עם זאת, זוהי דיאטה לא קלה ליישום, ולא מומלצת לאורך זמן. מחסור במגוון רחב של מזונות, ביניהם פירות וירקות רבים, דגנים וקטניות, יכול ליצור חסרים תזונתיים, ולפגוע במצבו התזונתי התקין של המטופל. לאחר שסברתי שיש לשים דגש על אלמנטים אחרים בטיפול התזונתי של חנה, ניסיתי להניע אותה מרצונה לדבוק בדיאטה דלת FODMAP, ובמקומה הצעתי לה תפריט עשיר בסיבים תזונתיים, מים, המלצתי על פעילות גופנית, שמפעילה את השרירים של הבטן ופוגה הגוף התחתון, וכן תרגול של דמיון מודרך. חנה התעקשה שיש לה IBS. היא הציגה שיחות מפורומים אינטרנטיים, שלפה דפים מודפסים עם מחקרים אודות יעילותה של דיאטה דלת FODMAP לתסמונת המעי הרגיז, וביקשה ממני להרכיב לה תפריט קבוע לפי גישה זו.

כיצד אוכל לשכנעה לנקוט באסטרטגיה תזונתית וטיפולית המתאימה לה יותר ממה שמשכנע אותה יותר בקריאה מהאינטרנט?

דבר הפסיכותרפיסט: אמיר זמורה MA, מרכז "דרך הבטן" מורה למדיטציה ומיינדפולנס.

סיפור המקרה מתאר מצב בו המטופלת לא קוראת נכון את מצבה, ומבקשת טיפול שלא יתמוך בהבראתה ועשוי אף לפגוע בה. אם כך נשאלת השאלה מדוע היא עושה זאת? מדוע קשה לה לראות את העמדה שמציעה הדיאטנית? התשובה עשויה להימצא בזירה שאינה קשורה למזון, אלא לביטחון, שליטה ואמון.

יש אנשים שסדר וביטחון חשוב להם כדי להתמודד עם הכאוס בחיים. יש כאלו שהאמון קשה להם, והם מחליפים אותו בשליטה. יש רבים שעושים גם וגם. לעיתים הצורך בשליטה הוא ביטוי של סטרס מתמשך, דיסטס, והוא כשלעצמו גורם שעשוי להשפיע על הופעה או החמרת מחלות. כשהדיסטס בא לידי ביטוי כסימפטומים פיזיים או כמחלה, גובר הערעור של הסדר שאדם ניסה לבנות בחייו והסטרס גובר יחד איתו. במקום כזה התפיסות והדעות של האדם ממלאות את הצורך בביטחון, וכל ניסיון לשנות את דעותיו נתפס על ידו כאיום קיומי.

עלינו להבין שזהו ביטוי של מנגנון ההישרדות המובנה בנו. המוח ההשרדתי אוהב שגרה. כשהוא מכיר מצב, הוא מניח שסיכויי ההישרדות גבוהים, הוא יתפוך אותו להרגל ויספק רציונליזציות שישבירו למה עדיף לדבוק בו. גם אם ההיגיון אומר אחרת. כיצד זה קורה?

סטרס גורם לשני תהליכים פיזיולוגיים עיקריים: האחד הוא פיזי משאבים ממרכז הגוף, והפנייתם אל הידיים והרגליים כדי לתמוך במצבי לחימה או בריחה. תהליך דומה קורה גם במוח - הסטרס מחליש את המוח החושב היצירתי שנמצא בקורטקס הקדמי, ומפנה את המשאבים לבלוטת ההשרדות הקדומה - האמיגדלה. בלוטה זו יודעת לחשוב בשחור או לבן - חיים או מוות, כדי שנפעל באופן אוטומטי - בריחה או לחימה. זהו מנגנון יעיל שעוזר לנו לשרוד במקרים של איום ברור ומוחשי. הבעיה מתחילה כשהסטרס הופך למצב מתמשך, וכל דבר נחוה אצלנו כאיום. היצמדות לתפיסה מסוימת כמוה כעצירות.

אז מה עושים? נרגעים. כשהגוף נרגע, המשאבים חוזרים למרכז הגוף, הזרימה הטבעית חוזרת לסדרה, והמחשבה הופכת שוב יצירתית, שמודעת למגוון האפשרויות שקיים בחיים במגע שבין שחור ללבן. נחזור למטופלת שלנו - כיצד אם כן עוזרים לה להירגע. קודם כל מודעים למצב ולא לוחצים עליה לשינוי. זה עשוי להרגיע אותנו כמטפלים. חוסר

הבנה למצבה מצידנו עשוי לגרום לשיפוט או ביקורת כלפיה מה שיגביר את הסטרס אצלה. היא עשויה להגיב בתוקפנות, אנחנו נילחץ ומכאן המצב רק ידרדר. מה עוזר לנו להירגע? נשימה עמוקה, השהיית תגובה והקשבה. קבלה שלנו את מצבה עשויה להרגיע אותה ולייצר אצלה הקשבה.

מכאן אני מציע ארבעה צעדים שעשויים לעזור למטופלת להגיע לשכנוע עצמי. בסיסם בשינוי האסטרטגיה מדחיפה למשיכה. כאמור כל ניסיון לשנות את דעתה של המטופלת יחוה מצידה כאובדן שליטה, ערעור הביטחון ויתקל בהתנגדות.

לכן הצעד הראשון הוא יצירת הסכמה. למשל על המטרה המשותפת. מטרה שבסופה היא תרגיש טוב יותר, ותגיע להבראה בדרך הקצרה ביותר האפשרית.

הצעד השני הוא ריגוש. הוא נועד לעקוף את המנגנון המנטלי שרואה את החיים בשחור ולבן. למשל לשאול אותה כיצד היא תרגיש פיזית כשהעצירות תשתחרר. לא מה היא חושבת שיקרה, מה בא לה שיקרה. כדי לרגש ניתן למשל לשאול כיצד זה ישיפיע על משפחתה, על החלומות שלה. לתת לה להרגיש זאת.

עכשיו הזמן להניח בפניה את השאלה: "מה אם היית יודעת שיש אפשרות נוספת, להשיג את המטרה באופן קל ופשוט יותר?"

וכאן הצעד השלישי והחשוב ביותר הוא להשהות את הרצון שלנו לגרום לה לשנות את דעתה מיד. פשוט לשתוק ולחכות שההחלטה הזו תבשיל בתוכה. רק כשהיא מבקשת לשמוע על מה מדובר אפשר לעבור לצעד הרביעי, והוא כולל את הצגת הדרך החדשה ואת הצידוקים שתומכים בהחלטה. "מחקרים מוכיחים ש...", "מטופלים במצבך בוחרים בדרך הזו" וכד'.

לסיכום, הדרך המוצעת היא מעבר משכנוע - לשכנוע עצמי. היא כוללת הכרה במצב הסטרס בו נמצאת המטופלת ובכך שרק רגיעה תחזיר אותה למצב בו תהיה קשובה.

לסיכום, ארבעת הצעדים הם:

- הסכמה על המטרה המשותפת.
- ריגוש שמחבר אותה למצב בו היא תהיה כשהמטרה תושג.
- הצעה והשהיה שמאפשרת לה להשתכנע בעצמה ולבקש את הפתרון.
- צידוק המאפשר לה להסביר לעצמה מדוע שינתה את דעתה. זה אולי דורש מעט תרגול אבל בהחלט שווה את המאמץ.



Review



מגזין מכון תנובה למחקר

כל גיליונות ה-Review לרשותכם בפורמט דיגיטלי
באתר מכון תנובה למחקר:
www.tnuva-research.co.il

