

קידום מחקר רפואי בינלאומי מבוסס AI

מודל OMOP- CDM

ניר מקובר
גיא ליבנה
הדס אשל גבע

תוכן עניינים

3.....	קהילת OHDSI – כדאי להכיר!	1.
6.....	התועלת למחקר וה-AI הרפואי	2.
8.....	היכרות עם ישויות המידע במודל ה-OMOP	3.
10.....	יתרונות בשילוב OMOP ו-Unstructured Data	4.
11.....	פתרונות טכנולוגיים לייעול תהליכי המחקר	5.
15.....	ההבדל בין סטנדרט המידע FHIR ל-OMOP	6.
16.....	דוגמאות למחקרי OHDSI ברחבי העולם ובכנרת	7.
18.....	הייחודיות של מיזם כנרת (Kineret)	8.

1. קהילת OHDSI – כדאי להכיר!

מטרה וחזון

קהילת OHDSI (Open Health Data Science Initiative) היא קהילה בינלאומית של יזמים, חוקרי דאטה רפואי מארגוני בריאות או אקדמיה, גופי מחקר וחברות בתעשייה, אשר נוסדה במטרה לקדם את הרפואה באמצעות מחקרים עתירי דאטה.

הקהילה הקימה רשת עולמית למחקר רפואי המבוססת על 3 עקרונות: אחדות מבנה נתונים, אחדות טרמינולוגיה רפואית ושיתוף מחקרים על בסיס פלטפורמה למחקר רב מרכזי. מטרת OHDSI היא להעמיק את השימוש במידע רפואי, לאפשר ביצוע מחקרים מקומיים ובינלאומיים ע"י חוקרים מנוסים ממגוון ארצות ומוסדות וכן להפיק תובנות בעולמות הבריאות באמצעות רשת בינלאומית של מחקר רפואי בר השוואה, כאמצעי להשגת רפואה מצוינת.

קהילת OHDSI הוקמה בארה"ב בשנת 2014 ומאז התרחבה באופן מרשים למעל ל-80 מדינות ברחבי העולם.

הקהילה כוללת מעל ל-3,500 גופים וחברות המשתפים מידע למחקרים רפואיים מבוססי מידע. מדינת ישראל, באמצעות מיזם כנרת, הוסיפה 6 מרכזים רפואיים לרשת, נכון לסוף שנת 2023.



פעילות קהילת OHDSI

הקהילה מבצעת מגוון רחב של פעולות ליצירת רשת עולמית לשיתופי פעולה מחקריים והטמעת סטנדרט רפואי בינלאומי, תוך פיתוח, תחזוקה והרחבת פתרונות טכנולוגיים מותאמים, ייזום מפגשים פיזיים ווירטואליים להדרכה וייעוץ ועוד.

מחקר ושיתוף ידע



פורטל מחקרים

פורטל ידע הכולל מחקרים והצעות למחקר. הפורטל פתוח לחוקרים מכל העולם, המבקשים ליזום מחקרים ו/או להצטרף למחקרים קיימים בכל תחום רפואי.

פרסום מדריכים וסיוע בביצוע מחקרים רב-מרכזיים תוך דגש על סטנדרטיזציה ותאימות כלל עולמית.



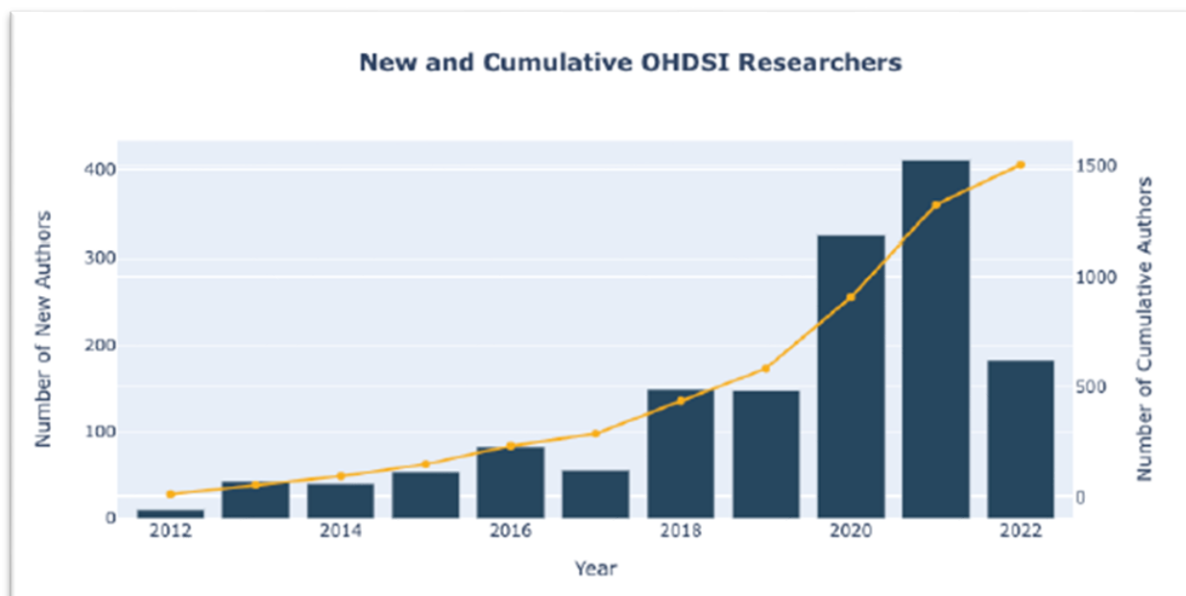
פורומים קהילה

פורומים פעילים למחקר, בהם מתקיימים דיונים במגוון נושאים, כגון: מחקרים רפואיים, טרמינולוגיה רפואית, ניהול נתונים ושיפורים בכלי ה Open source של הקהילה.

קיום סדנאות וכנסים לקידום שיתוף של נתונים רפואיים באמצעות כלים מתקדמים.

המערכת הפתוחה של OHDSI מאפשרת לחוקרים לפרסם ולשתף מחקרים, קוד ותובנות עם כל חברי הקהילה. השיתוף מקדם שיפור ידע ויכולות ניתוח ברמה הבינלאומית.

התרשים מציג את קצב גידול החוקרים השותפים לקהילה לאורך השנים:



פתרונות טכנולוגיים

מגוון כלי Open source פותחו ע"י הקהילה כדי לייעל ולסייע במחקר ושיתופי פעולה. מבין שלל הכלים נציין את ATLAS המאפשר לחוקר לבצע מחקר סטטיסטי בסיסי ללא צורך בידע תכנותי או בתמיכת IT ואת HADES, הכולל חבילות מוכנות של R לביצוע מחקרי פרדיקציה ו-AI מתקדמים. פירוט על כלי OHDSI מופיע בהמשך.

תחום הרגולציה ומחקר יישומי

מחקר יישומי נדרש לעבור תהליך אישור רגולטורי כדי לייצר שינוי במערכת הבריאות. OHDSI קיבלה הכרה של ה-FDA כפורמט הגשה מאושר. בזכות זאת ניתן להגיש תוצאות מחקר שבוצע בתשתית OHDSI (מבנה OMOP וטרמינולוגיה אחידה) כחלק מדיווח תוצאות ניסוי קליני ל-FDA. הכרה זו כבר מיושמת הלכה למעשה ע"י חברות תרופות וחברות ציוד רפואי המשלבות מחקרים על דאטה רפואי אמיתי (Real World Data) בהגשת מוצריהן לאישור רגולטורי.

"קיבלנו החזר משמעותי על ההשקעה בזמן ובמשאבים הנדרשים להקמת פלטפורמה עקבית באמצעות OMOP-CDM.

בזכות הפלטפורמה זכינו בהבנה משופרת שהשגנו של מסדי הנתונים התצפיתיים שלנו; באיכות המשופרת של הנתונים; וביעילות המוגברת שהושגה בביצוע הניתוחים שבהם תמכנו".

Erica A Voss et al , Feasibility and utility of applications of the common data model to multiple, disparate observational health databases, *Journal of the American Medical Informatics Association*, Volume 22, Issue 3, March 2015, Pages 553–564



2. התועלת למחקר וה-AI הרפואי

קיימת תמימות דעים כי ביצוע מחקרים ופיתוח פתרונות AI הם כח מניע לרפואה טובה, חכמה ויעילה יותר. אם קיימת הסכמה כל כך רחבה על הפוטנציאל, עולה השאלה - מדוע עדיין מתקשים יזמים וחוקרים רבים להוציא לפועל מחקרים מוצלחים ולתקף את תוצאותיהם?

התשובה נעוצה בשלושה חסמים מרכזיים:

1 איכות המידע – אנחנו מדברים על סטטיסטיקה, מודלי חיזוי ולמידת מכונה (Machine Learning). כל אלה מתבססים על למידה מהעבר. אם העבר אינו מיוצג בצורה נכונה, כך גם תוצאות החיזוי לעתיד. Garbage in – Garbage Out. היעדר תהליכי בקרה ואיכות מידע במאגרים רבים משפיעים על איכות ומהימנות המסקנות.

2 גודל אוכלוסיית המחקר – אנו עדים למחקרים רבים המבוססים על מדגמים קטנים. מחקרים נעשים בדרך כלל על בסיס הנתונים של ארגון בריאות בודד. כאשר מדובר על מחלות נדירות או על פרוצדורות רפואיות לא שכיחות, האתגר הוא כפול.

3 גיוון אוכלוסיית המחקר – יש מעט מחקרים ויזמות המתוקפים בין תרבויות, אזורי זמן, אזורים אקלימיים ועוד. ככל שהמחקרים והיזמות ממוקדים גיאוגרפית באזור מסוים, היכולת לאמץ אותם בצורה רחבה יותר (ולקבל אישורים רגולטוריים מקומיים) קטנה. יש לכך גם משמעות רפואית, אך לא פחות מכך גם משמעות עסקית עבור חברות ויזמים.

על מנת לענות על שלושת החסמים המרכזיים הללו, OHDSI ביצעה אינטגרציה מלאה בין: ייזום מודל סטנדרטי לניהול מידע רפואי למחקר (OMOP-CDM), החלת טרמינולוגיה רפואית אחידה ופיתוח כלים לביצוע מחקר רפואי שיתופי.

האינטגרציה של שלושת הרכיבים, מספקת יתרונות משמעותיים לחוקר:

סטנדרטיזציה של אחסון המידע (מבנה הנתונים והטרמינולוגיה)

1 מבנה טבלאות ושדות סטנדרטי וגלובלי: מבנה OMOP מציע מבנה נתונים אחיד וגלובלי של נתונים קליניים, אדמיניסטרטיביים וכלכליים. המבנה האחיד מייצר קוהרנטיות כלל עולמית בנתונים, מאפשר להריץ שאילתות ומחקרים באתרים שונים ללא צורך לבצע התאמות לכל מחקר ומאפשר להשוות תוצאות ברחבי העולם ללא חסמי הכניסה הגבוהים.

שפה קלינית אחידה: שפות שונות וקטלוגים רפואיים שונים בין במדינות ולעיתים בארגונים שונים בתוך אותה מדינה ועלויות גבוהות הכרוכות בביצוע, מהווים את המחסום העיקרי למחקר רב מדינות. ב OHDSI קיימת שפה קלינית גלובלית אחידה ונבנו כלים להמרה של כל מונח רפואי לשפה זו, למשל: אבחנות ופרוצדורות נרשמות בקודי SNOMED, תרופות ב RxNorm, בדיקות מעבדה ב LOINC וכו'. כיום ישנם מעל ל-3,500 מקורות מידע הממופים לקטלוגים הנבחרים, מה שמאפשר שילוב נתונים מישראל, אנגליה, קוריאה, סינגפור וכו' ללא מאמץ ובזמן קצר.

2

ביצוע פשוט ויעיל של מחקרים רב-מרכזיים

גודל האוכלוסייה הנחקרת: השימוש בכלי OHDSI, המבוססים על מבנה וטרמינולוגיה רפואית אחידים, מספק לארגוני הבריאות ולחוקרים ברחבי העולם פוטנציאל לגישה לנתונים רפואיים אחידים ומובנים מראש לניתוח רב מרכזי. שילוב נתונים ממקורות שונים באמצעות אפשרויות אינטגרציה מתקדמות, מאפשר להגדיל ולגוון את האוכלוסייה הנחקרת, וכפועל יוצא לייצר מחקרים איכותיים ומהימנים יותר.

תיקוף מודלים: השימוש בכלי OHDSI מאפשר לתקף מודל שאומן בארגון בריאות אחד, בארגונים שונים ברחבי העולם. הדבר מקנה פרספקטיבה רחבה ומאפשר לשפר את איכות ומהימנות המחקרים.

3

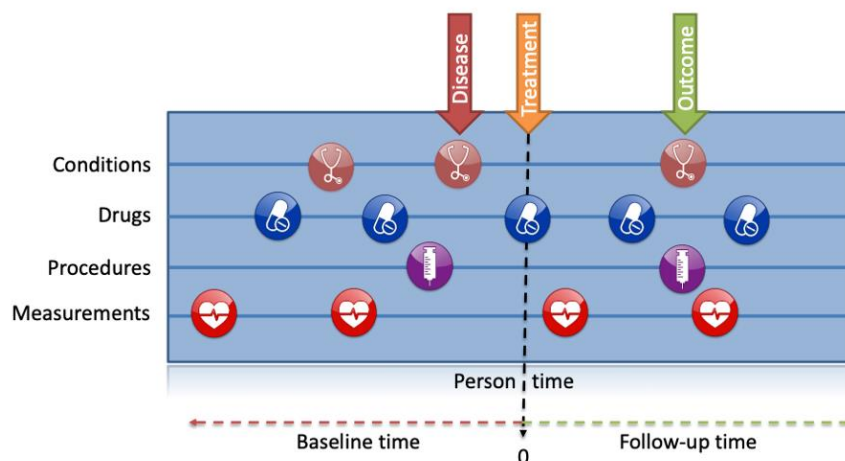
גמישות והתאמה אישית

שינוי ולמידה מתמיד: OHDSI מקיפה את כל תחום בריאות הציבור ומנסה לתת מענה לכלל הצרכים - טיפולים בבתי חולים ובקהילה, תביעות רפואיות וסקרים רפואיים. הקהילה אימצה תקנים שמתאימים למגוון רחב של מחקרים בתחומי רפורמה ושינוי מדיניות רפואית, מה שמאפשר לחוקרים להתאים את המודל לצרכיהם המשתנים. לדוגמא: אישור חיסוני קורונה ע"י ה FDA בוצע לאחר ניתוח תוצאות מחקר OHDSI במגוון מדינות כדי לייצר מבט על כלל עולמי.

4

פונקציונליות רחבה לדרישות מחקר מורכבות

אופן ארגון המידע לפי מודל הנתונים של OMOP וכלי המחקר של הקהילה מאפשרים לחוקרים להגדיר, בצורה פשוטה, נקודת בסיס, טווחי זמן ותנאי הכללה והחרגה המבוססים על אוסף של אירועים התלויים אחד בשני (לדוגמה: חילוף מידע אודות מטופלים אשר קיבלו תרופה מוגדרת X חודשים אחרי ניתוח ונרשמה להם אבחנה Y בטווח של שנה מקבלת התרופה, ובאותה תקופה זוהו ערכים חיוביים של סמן מסוים בבדיקת מעבדה).



5

קהילה פעילה וערוצי תמיכה וסיוע

שיתוף הידע בקהילה ומתן תמיכה, סיוע ולייווי, הינן מטרות מרכזיות המאפשרות למידה משותפת ושיפור מיומנויות ותוצרים.

3. היכרות עם ישויות המידע במודל ה-OMOP

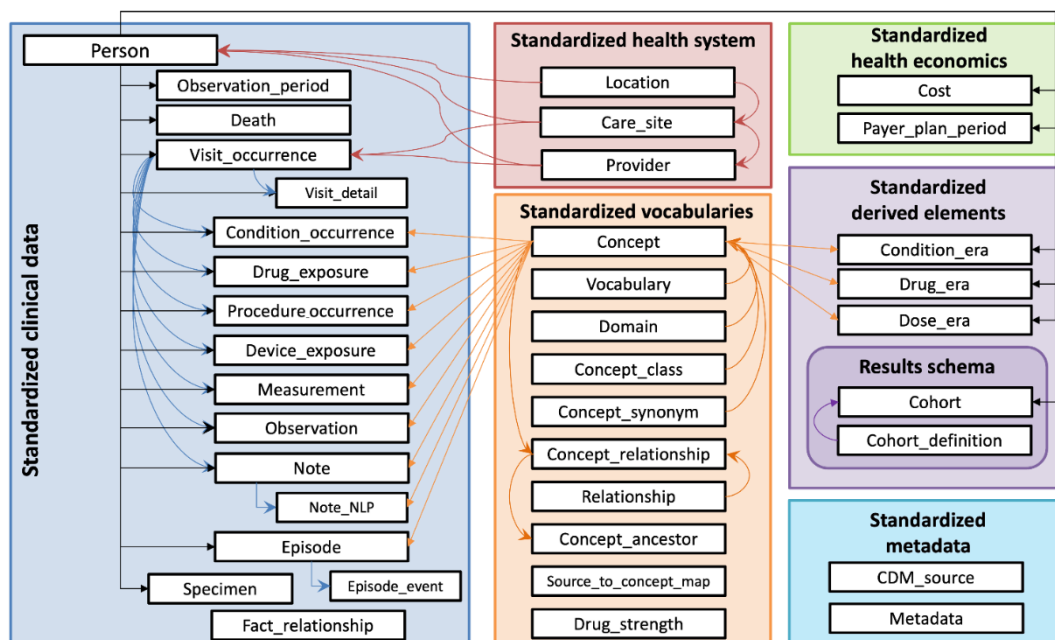
מודל המידע של ה OMOP מכיל את ישויות המידע העיקריות והמרכזיות בניהול רשומה ומידע רפואי. זוהי אבן היסוד ונקודת המוצא של כל מחקר ומודל AI, אליה מצטרפים מידעים נוספים (מובנים ולא מובנים) בהתאם לשאלת המחקר.

מרכיבי המודל:

Standardized Clinical Data - המידע הקליני

הישויות המרכזיות:

- מידע דמוגרפי
- ביקורים
- ניתוחים ופרוצדורות רפואיות
- אבחנות
- מדדים חיוניים
- גורמי סיכון
- תרופות
- דימות
- בדיקות מעבדה
- טקסטים (פענוחים, סיכומי ביקור ועוד)



Standardized Health System – מידע ארגוני-ניהולי

הישויות המרכזיות:

- מיקום הטיפול (המחלקה, המרפאה וכד')
- מוסד בו בוצע הטיפול (ארגון הבריאות)
- הצוות המטפל (התמחות, תחום עיסוק וכד')

Standardized Health Economics – מידע כספי

הישויות המרכזיות:

- עלויות הפעולות הרפואיות
- הסכם התקשרות ותוכנית ביטוח רפואי

Standardized Vocabularies (וקבוצת Concept Sets)

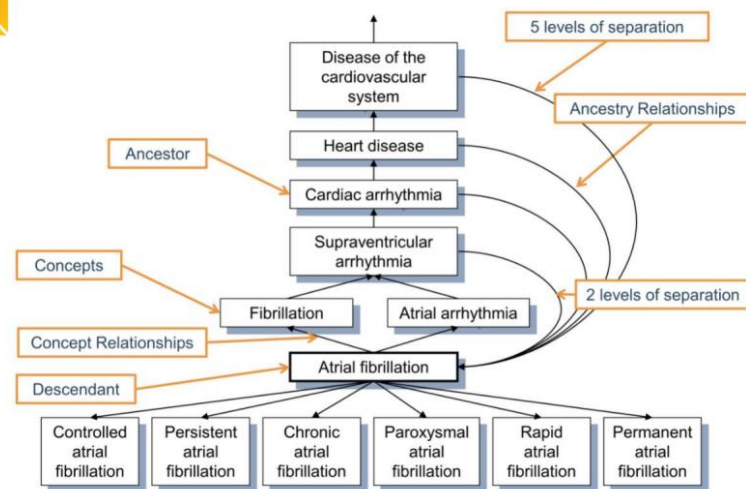
המודל משתמש במילון מוסכם וסטנדרטי של מונחים, הידועים גם בשם Standardized Vocabularies. שימוש זה מבטיח סטנדרטיזציה והתאמה ישראלית של מונחים קליניים, כמו תרופות, אבחנות ופעולות.

הישויות המרכזיות:

- רשימת מונחים רפואיים
- קשרים בין מונחים רפואיים
- היררכיות בין מונחים רפואיים
- המרה של מונחים לא סטנדרטים לסטנדרט בינלאומי



Concepts Hierarchy



למידע נוסף על OMPO – נא סרקו:



4. יתרונות בשילוב OMOP ו-Unstructured Data

השימוש בבינה מלאכותית (AI) בעיבוד מידע רפואי לא מובנה הינו נדבך חשוב במחקרים בתחום הבריאות והרפואה.

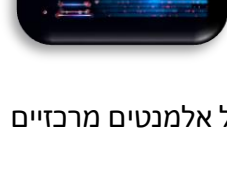
ניתן לסווג מחקרים אלו למספר תחומים עיקריים:



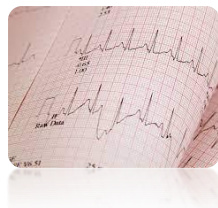
Medical imaging analysis – שימוש בטכניקות כמו למידה עמוקה לניתוח נתונים לא מובנים מתמונות רפואיות כגון רנטגן, MRI, סקירות CT, במטרה לסייע באבחון ואף בחיזוי התקדמות מחלה.



Genomic data analysis - כולל שימוש ב-AI להבנת המבנה הגנומי, שינויים גנטיים והשלכותיהם על מחלות. מחקרים אלו מחפשים סמנים גנטיים שיכולים לסייע ברפואה מותאמת אישית.



Natural Language Processing (NLP) - מחקרים אלו עוסקים בחילוץ מידע משמעותי מנתונים טקסטואליים לא מובנים כמו סיכומי ביקור ורשומות קליניות אחרות, תמלילים רפואיים ורישומי מטופלים. טכניקות NLP משמשות לזיהוי וקטלוג של אלמנטים מרכזיים כמו תסמינים, אבחנות, טיפולים ותוצאות למטופלים.



Monitor and Wearable Data Analysis - מחקרי AI על כמות רבה של נתונים ממכשירים נישאים ומערכות ניטור מטופלים. נתונים אלו כוללים קצב לב, רמות פעילות, דפוסי שינה ועוד, ומשמשים לניתוח ולניטור בריאות, תהליכי שיקום והחלמה, וחיזוי בעיות בריאות פוטנציאליות.

סוגי מחקר אלו טומנים בחובם מגוון אתגרים שונים, אך שני אתגרים המשותפים לכל התחומים הם הפיכת המידע הלא מובנה למידע מובנה מותאם ללמידת מכונה, והצורך בהשלמה של מידע נוסף על מנת לשפר את תקפות תוצאות המודלים והמחקרים.

מחקר משולב של מידע לא מובנה ומידע במבנה OMOP מספק מענה לשני אתגרים אלו:

שילוב של המידע הלא מובנה עם מידע מובנה במבנה OMOP סטנדרטי מייצר אחידות עם כלל התיק הרפואי המובנה של המטופל, בפורמט נגיש וברור למודלים אנליטיים, ומעשיר את הרשומות הרפואיות הטאבולריות.

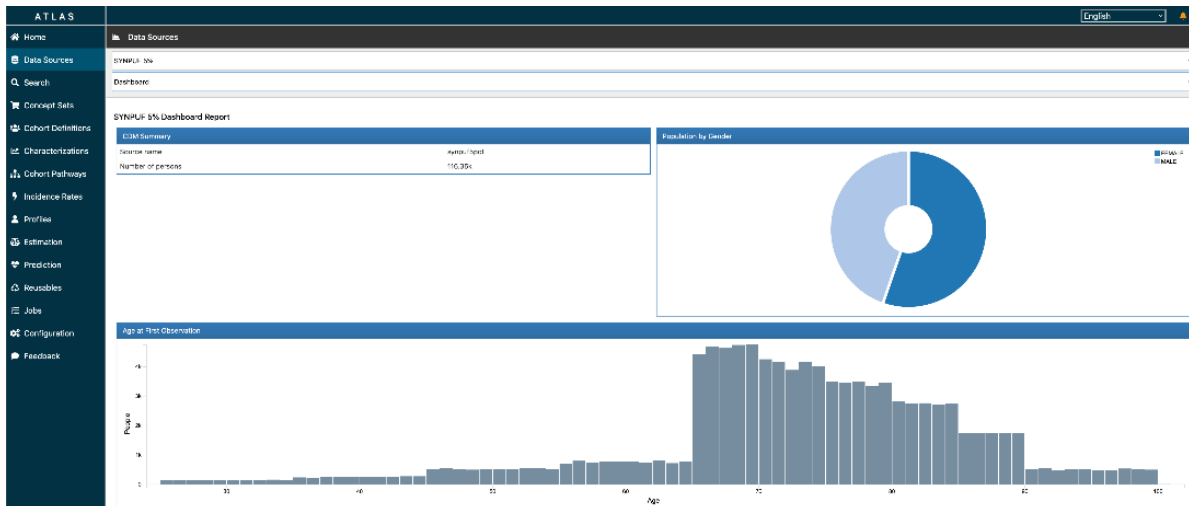
שיפור תקפות תוצאות מחקר דורש, לרוב, הרחבה של אוכלוסיית המחקר למקורות נוספים או הצלבה עם מידע נוסף ממערכות אחרות. חיבור מידע לא מובנה למידע מובנה בתצורה סטנדרטית ומוכרת בכל העולם מאפשר גם להרחיב את אוכלוסיית המחקר לארצות אחרות ואוכלוסיות מגוונות וגם להוסיף משתנים מהרשומה הרפואית להגברת דיוק המודלים וטיוב התוצאות.

5. פתרונות טכנולוגיים לייעול תהליכי המחקר

קהילת OHDSI פועלת באמצעות מספר כלים ותוכנות קוד פתוח שפותחו ע"י משתמשים מכל העולם. להלן רשימה חלקית:

ATLAS

כלי ויזואלי המאפשר ניהול וניתוח של נתוני בריאות בסטנדרט OMOP. הוא מספק יכולת ייצוג גרפי של נתונים ומאפשר למשתמשים לבצע שאילתות מתקדמות ולהתמקד באזורים שונים של הנתונים. השימוש בכלי הוא פשוט ואינו מצריך ידע פיתוחי. הערך המשמעותי של הפתרון, בשילוב עם מודל ה-OMOP, הוא היכולת להגדיר את ה-Cohort, על כל מאפייניו, בצורה פשוטה בארגון בריאות אחד, לייצא את ההגדרות ולהריץ אותן אצל כל אחד מהשותפים ברחבי העולם, מבלי להידרש לתהליך תיאום ארוך ומורכב לפני שיוצאים לדרך.



מומלץ לסרוק ולצפות בסרטון קצר (03:52 דק') כדי להכיר את הכלי ואת יתרונותיו:



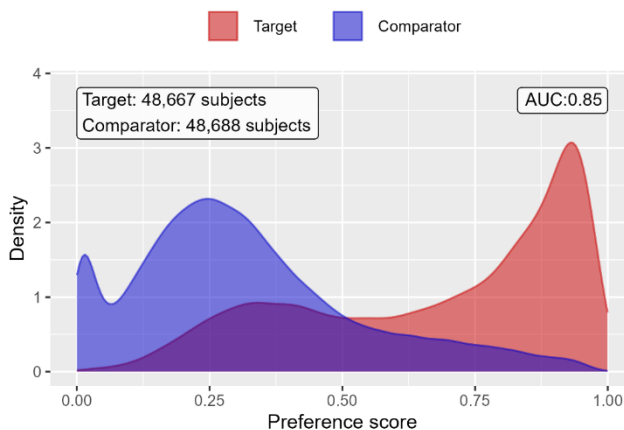


אוסף חבילות R (קוד פתוח) המספקות פונקציות שניתן לשלב ביניהן כדי לבצע מחקר תצפיתי מלא. מדובר בחבילות R הניזונות ממודל נתונים במבנה OMOP ליצירת הערכות, סטטיסטיקות תומכות, תרשימים וטבלאות. החבילות מתממשקות ישירות עם הנתונים במאגר וניתן להשתמש בהן לביצוע ניתוחים מתקדמים ומתוקננים לאיפיון אוכלוסייה, הערכת השפעה ברמת אוכלוסייה וניבוי ברמת המטופל.

שתי חבילות R השימושיות במחקרים רבים הן:

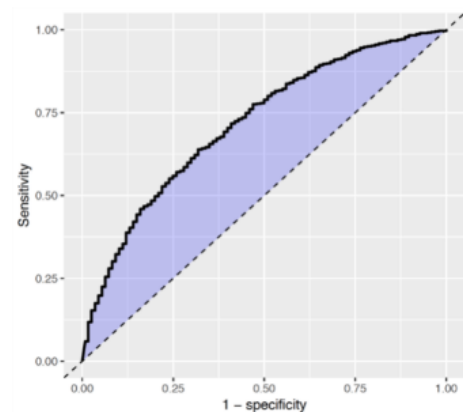
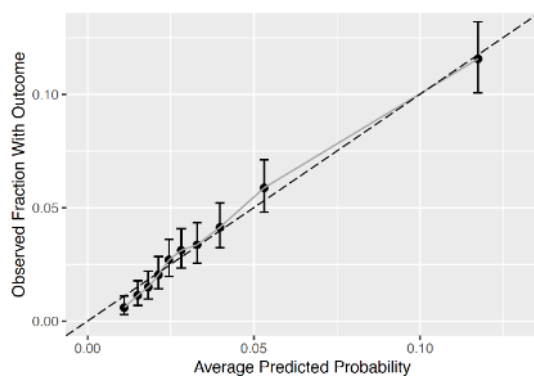
CohortMethod

סט חבילות של קוד R ספציפי לתיקוף תקינות קוהורטים. זהו כלי ייעודי המשתמש במודלים סטטיסטיים, כגון: רגרסיה סדורה (Regularized Regression), רגרסיית לוגיסטית, רגרסיית פואסון ורגרסיית קוקס על מגוון רחב של נתונים מסבירים ובוחן תקינות מודל וקוהורט בהיבט של סיבה ותוצאה.



Population-Level Estimation (PLE)

חבילת R לביצוע מודל חיזוי ברמת מטופל. הייחודיות של חבילה זו באה לידי ביטוי ביכולת לפתח ולאמת מודלי חיזוי עבור מגוון רחב של קבוצות יעד ותוצאה, בשימוש במגוון רחב של משתנים מסבירים ואלגוריתמים, ובאפשרות להתאמה אישית ואימות המודלים.



Data Quality Dashboard

אחד האתגרים בהפקת תובנות ופיתוח פתרונות AI בעולם הרפואי הוא איכות המידע. מאגר מידע מחקרי המובסס על OMOP מיוצר ומוגש למחקר לאחר שימוש במנוע אנליטי להערכת איכות הנתונים. הפתרון מנתח את כלל הטבלאות והשדות במאגר המידע המחקרי ומבצע למעלה מ-1,500 בדיקות איכות.



DATA QUALITY ASSESSMENT

IBM MARKETSCAN MULTI-STATE MEDICAID DATABASE

DataQualityDashboard Version: 1.0.0
Results generated at 2020-08-24 15:44:34 in 3 hours

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1849	6	1855	100%	281	6	287	98%	2130	12	2142	99%
Conformance	550	13	563	98%	80	0	80	100%	630	13	643	98%
Completeness	322	5	327	98%	12	0	12	100%	334	5	339	99%
Total	2721	24	2745	99%	373	6	379	98%	3094	30	3124	99%

Athena

Athena פותח בכדי לתמוך בחיפוש מונחים רפואיים (Concepts) ולבחון כיצד מוגדרים מונחים רפואיים בקטלוג הסטנדרטי ומה הקשרים בין המונחים השונים. הכלי מוטמע כחלק מהאטלס ומסייע בהגדרת אוכלוסיות מחקר וטיוב נתונים.

ATHENA									
SEARCH BY KEYWORD									
atrial fibrillation									
Download Results									
Show by 15 items Total 3,769 items									
DOMAIN	ID	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB	
CONCEPT	313217	49436004	Atrial fibrillation	Clinical Finding	Standard	Valid	Condition	SNOMED	
CLASS	45611600	D001281	Atrial Fibrillation	Main Heading	Non-standard	Valid	Condition	MeSH	
VOCAB	4064452	164889003	ECG: atrial fibrillation	Clinical Finding	Standard	Valid	Condition	SNOMED	
VALIDITY	42539346	762247006	Preexcited atrial fibrillation	Clinical Finding	Standard	Valid	Condition	SNOMED	
	4154290	282625002	Paroxysmal atrial fibrillation	Clinical Finding	Standard	Valid	Condition	SNOMED	
	44807374	81640100000105	Atrial fibrillation excluded	Context-dependent	Standard	Valid	Observation	SNOMED	
	4119602	233911009	Non-rheumatic atrial fibrillation	Clinical Finding	Standard	Valid	Condition	SNOMED	
	4047554	134377004	Atrial fibrillation monitoring	Procedure	Standard	Valid	Observation	SNOMED	
	42689664	1066831000000104	Atrial fibrillation detected	Clinical Finding	Standard	Valid	Condition	SNOMED	
	44788803	196371000000102	Atrial fibrillation resolved	Clinical Finding	Standard	Valid	Condition	SNOMED	

Usagi נועד לעזור במיפוי אוטומטי של מונחים בין מערכות קודים שונות, ובפרט למיפוי קודי אבחנות. הכלי משמש בעיקר אנשי IT ורפואה בתהליך המרת קודים מקומיים לסטנדרט הבינלאומי.

Usagi - testIcpCodesMapping.csv

Status	Source code	Source term	Frequency	CodeText	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard con.	Parents	Children	Comment
Approved	K87.00	Hypertension...	694195	Hypertensie ...	0.81	316856	Hypertensive ...	Condition	Clinical Findi...	SNOMED	38341003	S	1	27	
Approved	L99.00	Other diseases...	680422	Andere ziekte...	0.47	0	Unmapped						0	0	Too generic
Approved	D01.00	Abdominal p...	678588	Gegeneralis...	0.61	197988	Generalized ...	Condition	Clinical Findi...	SNOMED	102614006	S	1	0	
Unchecked	S99.00	Skin disease...	675817	Andere ziekte...	0.75	4317258	Disorder of s...	Condition	Clinical Findi...	SNOMED	95320005	S	2	193	
Unchecked	T85.00	Hypothyroid...	667283	Hypotheorid...	1.00	4113642	Hypothyroid...	Condition	Clinical Findi...	SNOMED	286910004	S	1	0	

Source code	Source term	Frequency	CodeText
S99.00	Skin disease, other	675817	Andere ziekte(n) huid/subcutis

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
4317258	Disorder of skin	Condition	Clinical Finding	SNOMED	95320005	S	2	193

Remove concept

Search

Query

☒ Use source term as query

☐ Query:

Filters

☐ Filter by user selected concepts

☐ Filter by concept class: 2-dig nonbill code

☒ Filter standard concepts

☐ Filter by vocabulary: ABMS

☒ Include source terms

☒ Filter by domain: Condition

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.75	Skin disease	4317258	Disorder of skin	Condition	Clinical Finding	SNOMED	95320005	S	2	193
0.65	Skin Disease, Fungal	137213	Dermal mycosis	Condition	Clinical Finding	SNOMED	14560005	S	3	12
0.57	AIDS with skin dise...	4224566	Skin disorder associated with AIDS	Condition	Clinical Finding	SNOMED	421394009	S	2	2
0.56	Chronic skin disease	4134132	Chronic disease of skin	Condition	Clinical Finding	SNOMED	128236002	S	2	26
0.55	Disease, Otolgic	378161	Disorder of ear	Condition	Clinical Finding	SNOMED	25906001	S	4	43
0.55	Disease, Hers	4163346	Glycogen storage disease, type VI	Condition	Clinical Finding	SNOMED	29291001	S	2	0
0.55	Other peripheral va...	321052	Peripheral vascular disease	Condition	Clinical Finding	SNOMED	400047006	S	1	44
0.55	Other peripheral va...	4119612	Lower limb ischemia	Condition	Clinical Finding	SNOMED	233961000	S	2	3
0.55	Disease, Ormond	4176725	Retroperitoneal fibrosis	Condition	Clinical Finding	SNOMED	49120005	S	1	3
0.54	Pathological fractur...	73571	Pathological fracture	Condition	Clinical Finding	SNOMED	268029009	S	1	21
0.52	Disease, Tooth	4122115	Tooth disorder	Condition	Clinical Finding	SNOMED	234947003	S	3	58
0.52	Disease, Lip	135858	Disorder of lip	Condition	Clinical Finding	SNOMED	90678009	S	3	35
0.51	Disease, Olfier	4113600	Multiple congenital avostosis	Condition	Clinical Finding	SNOMED	254044004	S	6	0

Replace concept Add concept

Comment:

Approved / total: 5 / 1037 8.8% of total frequency

Approve

Analytical Tools

לאורך השנים פותחו ומפותחים בימים אלה פתרונות מתקדמים נוספים המאפשרים לחוקרים להשתמש במודל לניתוח נתונים רפואיים.

כל הפתרונות הינם פתרונות קוד פתוח (Open-Source) ופתוחים לשימוש הכלל ללא עלות!



לכלים נוספים – ניתן לסרוק את הקוד:

6. ההבדל בין סטנדרט המידע FHIR ל-OMOP



FHIR (Fast Healthcare Interoperability Resources) הוא מודל נתונים בינלאומי המפותח על ידי חברי קהילת FHIR ואשר נתמך על ידי ארגון התקנים HL7. המטרה העיקרית לשמה נבנה ה-FHIR היא תפעולית, מתוך רצון לאפשר שיתוף והנגשת נתונים בצורה שקופה, פשוטה ובטוחה, בתוך מערכות המידע בארגון הבריאות, בין ארגוני הבריאות ובין שירותים ואפליקציות עבור המטופלים והצוות המטפל. הייחודיות של התקן נעוצה בעובדה שהיא מגדירה מראש מבנה מוסכם של מסרים בין הצדדים הקרויים APIs (Application Programming Interface) ובכך מקלה על האינטגרציה ביניהם.

רבים שואלים "מה ההבדל בין סטנדרט המידע FHIR ל-OMOP" ו"אם הארגון בחר לקדם ולהשקיע ב-FHIR האם נכון גם להשקיע ב-OMOP"? התשובה היא שלא רק ששני התקנים אינם סותרים אחד את השני, אלא הם מחזקים ואף משלימים זה את זה. בעוד שימוש ב-FHIR וה-Resource המובנים שלו משמשים בעיקר כאמצעי להעברת מידע ואינטגרציה בין מערכות ברמה תפעולית של תיק המטופל, ה-OMOP, קהילת OHDSI וכלי המחקר שלה משמשים כאקוסיסטם שלם לצרכי מחקר ופתרונות מידע רפואי אנליטי. ניתן להניח שבעתיד הלא רחוק ארגונים יבנו את מאגרי ה-OMOP שלהם מעל תשתיות ה-FHIR ובכך יתייבצו ויועמקו השימושים של שני התקנים זה לצד זה.

ניתן להמשיל את היחס בין שני תקני המידע ליחס שבין מערכות תפעוליות למחסני נתונים/ פתרונות BI. גם בעולמות ה-BI עלו בעבר שאלות סביב הנחיצות בהקמת מחסני נתונים ייעודיים לצרכים של אנליזה ודשבורדים ביחס להפעלת שאילתות ודוחות ישירות מעל המערכות התפעוליות. נכון להיום ההסכמה החד-משמעית היא כי צורת ארגון הנתונים, קלות השימוש, הביצועים, הגמישות והכלים הייחודיים ה"רובים" מעל מחסני נתונים מהווים יתרון עצום לאנליסטים ומפתח להצלחה עסקית. דומה הדבר גם סביב השימוש ב-OMOP.



במסמך שפורסם בנובמבר 2021 ע"י HDR UK (Health Data Research), ארגון לאומי לקידום מחקר רפואי באנגליה ניתן למצוא חיזוק לכך. המסמך בוחן את הסטנדרטים השונים הקיימים בעולם המידע הרפואי ואף כולל סקר שבוצע בין מגוון ארגוני בריאות, מוסדות אקדמאים, גופי מחקר וחברות טכנולוגיה בתחום ה-AI. המלצת הארגון היא לקדם בארגונים ובחברות את השימוש בשני הסטנדרטים במקביל.

7. דוגמאות למחקרי OHDSI ברחבי העולם ובכנרת

במרצת השנים בוצעו אלפי מחקרים עצמאיים ומשותפים ברחבי העולם במסגרת קהילת OHDSI על גבי מאגרי מידע מבוססי תקן ה-OMOP.

דוגמאות נבחרות למחקרים:



פיתוח ותיקוף של מודל פרוגנוסטי החוזה טרנספורמציה דימומית סימפטומטית בשבץ איסכמי חריף בקנה מידה ברשת OHDSI



מטרת המחקר הייתה לפתח ולאמת מודל חיזוי לסיכון של שינוי דימומי (Hemorrhagic Transformation) בתוך 30 יום לאחר אירוע שבץ מוחי איסכמי.

שינוי דימומי הינה תופעה מורכבת ורב-גורמית בשלב האקוטי של שבץ מוחי, ולעיתים קרובות גורמת לאבחנה שגויה. זיהוי גורמי סיכון וחיזוי מוקדם של שינוי דימומי בשבץ מוחי איסכמי חריף תורמים לא רק לבחירת משטר טיפולי אלא, חשוב מכך, לשיפור האבחנה של אירוע מוחי איסכמי חריף.

המחקר הינו מחקר רטרוספקטיבי רב-מרכזי שבו פותח מודל רגרסיה לוגיסטית Lasso עם קבוצת נתונים גדולה מרשומות רפואיות אלקטרוניות אמריקאיות שאורגנו לפי מבנה נתונים משותף (OMOP-CDM). תיקוף המודל לפני מעבר לשימוש קליני בוצע על 10 קבוצות נתונים (Real World Data) הכוללות רשומות EHR של מטופלים מאמריקה, אירופה ואסיה.

מסקנות: המודל שפותח יכול לזהות מטופלים בעלי סיכון גבוה יותר לשינוי דימומי ממוצע האוכלוסייה, עם שטח מתחת לעקומה (AUC) של 0.78. השימוש בסטנדרט OMOP-CDM הדגים תקן נתונים אשר מתאים לשימוש משני ב-EMR במחקר קליני רב-מרכזי לפיתוח מודל חיזוי פרוגנוסטי ואימותו. בעתיד, שילוב מודל זה עם מערכות מידע קליניות יסייע לקלינאים לקבל החלטות טיפול הנכונות עבור מטופלים עם שבץ מוחי איסכמי.



קונטקסטואליזציה של אירועי לזואי בעלי עניין מיוחד כדי לאפיין את שיעורי ההיארעות הבסיסיים ב-24 מיליון חולים עם COVID-19 על פני 26 מאגרי מידע: מחקר עוקבה רטרוספקטיבי רב לאומי



מטרת המחקר הינה לאפיין את שיעורי ההיארעות של אירועים חריגים (AEIs) הקשורים לחיסוני COVID-19 לאחר הדבקה בנגיף SARS-CoV-2, ולהשוות אותם לשיעורים היסטוריים באוכלוסייה הכללית.

המחקר התבסס על קוהורט רב-לאומי עם נתונים ממרפאות ראשוניות, תיקי רפואה אלקטרוניים ותביעות ביטוח שהותאמו למודל נתונים משותף. הנתונים עבור מחקר זה נאספו בין ה-1 בינואר 2017 למאי 2022. 16 ה-AEIs הנפוצים שנבחרו מראש היו: אוטם שריר הלב חריף, אנפילקסיס, דלקת התוספתן, שיתוק הפנים על שם בל, טרומבוזיס ורידי עמוק, קרישה תוך-צינית מפושטת,

דלקת המוח והחוט השדרה, תסמונת גיאן-בָּרֶה, שבץ מוחי המורגי, שבץ מוחי לא המורגי, טרומבוציטופניה תלויית חיסון, דלקת שריר הלב/דלקת קרום הלב, נרקולפסיה, תסחיף ריאתי, מ yelitis רוחבית וטרומבוזיס עם טרומבוציטופניה.

יחס שיעורי ההיארעות הסטנדרטיים לפי גיל ומין (SIR) הוערכו כדי להשוות את השיעורים לאחר COVID-19 לשיעורים לפני המגפה בכל אחד מבסיסי הנתונים.

התוצאות הראו שונות משמעותית לפי גיל בשיעורי ה-AESIs, כאשר חלקם גדלו בבירור עם הגיל אך אחרים הראו מגמה הפוכה. הבדלים נצפו גם בין בסיסי הנתונים השונים עבור אותם אירועי בריאות וחתכי גיל-מין. כל ה-AESIs שנחקרו הראו באופן עקבי יותר שכיחות בקבוצות שלאחר COVID-19 לעומת הקבוצות קודמות.

ממצאינו מרמזים על כך שכל ה-AESIs נפוצים יותר לאחר COVID-19 מאשר באוכלוסייה הכללית. אירועי קריש דם היו נפוצים במיוחד, ויותר מפי 10. נדרש מחקר נוסף על מנת לקשר סיבוכים שלאחר COVID-19 בטווח הארוך יותר.



לרשימת פרסומים נוספים והצעות למחקר ומחקרים בתהליך:

בישראל נבנה תקן OHDSI לראשונה על ידי [מיזם כנרת](#) אשר עלה לאוויר במהלך שנת 2021. מאז הפעלתו בוצעו מגוון של מחקרים על ידי חוקרים מבתי החולים השותפים במיזם ובימים אלה משתתפים חוקרים וחוקרות מהמרכזים הממשלתיים בצוותי מחקר בינלאומיים.

דוגמא למחקר דאטה שבוצע בכנרת על בסיס נתוני מודל ה-OMOP



האם ל"פרדוקס ההשמנה" יש תאריך תפוגה? מחקר עוקבה רטרוספקטיבי.

"פרדוקס ההשמנה" מתייחס להשפעה מגנה של אינדקס מסת גוף גבוה יותר (BMI) על תמותה בחולי מחלות זיהומיות חריפות. עם זאת, ההשפעה ארוכת הטווח של הפרדוקס עדיין לא ברורה.

נערך מחקר רטרוספקטיבי של חולים שאובחנו עם מחלות זיהומיות חריפות הנרכשות בקהילה במרכז הרפואי שמיר, ישראל (2010–2020).

החולים קובצו לפי BMI: תת משקל, משקל תקין, עודף משקל והשמנה בדרגות I–III. שיעורי תמותה לטווח קצר וארוך הושוו בין קבוצות אלו.

תוצאות: מתוך 25,226 החולים, שיעורי התמותה לטווח קצר (90 יום) ולטווח ארוך (שנה אחת) היו גבוהים יותר בקבוצות בתת משקל ובקבוצות משקל תקין בהשוואה לאחרים.



מסקנות: במחקר מקיף זה, השמנת יתר הייתה קשורה באופן עצמאי לירידה בתמותה בטווח הקצר והארוך. התוצאות הבלתי צפויות הללו מעוררות חקירה נוספת של תופעה נוגדת אינטואיציה זו.

8. הייחודיות של מיזם כנרת (Kineret)

רקע על המיזם

כנרת הוא מיזם דיגיטלי חדשני, המהווה פריצת דרך אסטרטגית על ידי ועבור רשת המרכזים הרפואיים הממשלתיים, במסגרתו נבנתה תשתית טכנולוגית ואנושית לצד תהליכי עבודה תומכים, למיצוי הפוטנציאל של שימושים משניים במידע ולקידום חדשנות ופיתוח מודלי AI – ML מבוססי מידע.

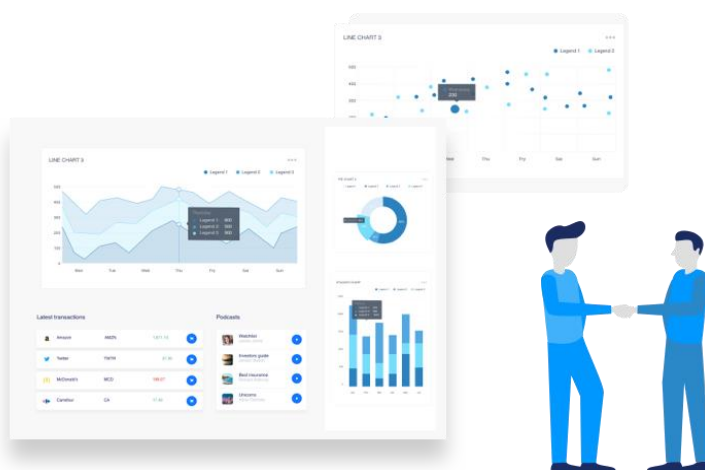
לצורך כך נבנה Data Lake המכיל מידע רפואי של בתי החולים הממשלתיים הבאים: ברזילי, שמיר, הלל יפה, בני ציון, המרכז הרפואי לצפון והמרכז הרפואי לגליל.

השימוש בפלטפורמת כנרת מאפשר ליזם לקבל גישה רחבת היקף לרשומות רפואיות ומידע מגוון מ 15 עד 20 השנים האחרונות מרשת בתי החולים הגדולה בישראל.

המידע המנוהל בכנרת נחלק לשניים:

- מידע מובנה - חלק ניכר מהמידע הנגיש מופה למודל הבינלאומי OMOP. המידע הזמין כחלק מהמודל הינו: פרוצדורות, אבחונים, תרופות, בדיקות מעבדה, סימני חיים, רגישויות, כל מקרי האשפוז של המטופל לאורך השנים או הביקורים במרפאות.
- מידע לא מובנה - הכולל את כל יתר המידע: נתוני טקסט חופשי, נתונים כלכליים, תמונות, קול, קבצי PDF, נתוני ניטור ומכשור רפואי ועוד (שליפת מידע זה תתבצע לאחר בחינת צרכים).

הערך של מיזם כנרת לעולם המחקר והיזמות הרפואית



מחקרים ויוזמות מבוססים מידע הם החזית של עולם החדשנות הרפואית. שימוש חכם במידע מהמרכזים הרפואיים עם יזמים וגורמים מחקריים הוא בעל פוטנציאל לקדם את מערכת הבריאות, להשיג פריצות דרך מדעיות, לקדם את איכות הטיפול ולשפר את השירות למטופלים.

כנרת מגישה לחוקרים וליזמים מידע רפואי נרחב, מסייעת בידע ובניסיון של המומחים, תוך הסרת חסמים והיכרות עמוקה עם צרכי השטח.

פרויקט כנרת הוקם כדי לשפר, למקסם ולקדם כל תהליך רפואי אפשרי ובכך לשפר את איכות חייהם של מיליוני מטופלים בישראל ומחוצה לה.

ב 20 השנים האחרונות נצבר מידע רפואי, תפעולי וכלכלי עצום בבתי החולים הממשלתיים בישראל. הפוטנציאל לא מומש בשל מורכבות השימוש במידע מבזר, ממערכות שונות, שאינו אחיד ואינו נגיש, היעדר תשתיות מתאימות, קשיים רגולטוריים וחוסר בכוח אדם מיומן.

כנרת מנגישה מידע מותמם ומטויב וכן מספקת שימוש בחדרי מחקר מתקדמים בענן.

בנוסף, המיזם מציע שירותים תומכים כגון: גישה למומחי תוכן קליניים, סיוע בתכנון מחקר והגשות רגולטוריות, מנתחי מערכות, תמיכה של Data Scientist, אפשרות להפעיל אתר בטא בבתי החולים ועוד.

קהל היעד כולל חוקרים בבתי החולים ובאוניברסיטאות, יזמים, חברות הזנק, תעשייה ותרופות.

מיזם כנרת מרכז חומר שנאסף במשך שנים ארוכות ומטרתו אחת: להיות מאגר איכותי, אמין ובעל ערך עבור חוקרים ויזמים.



פרטים נוספים באתר כנרת:





ניר מקובר

nir.makover@moh.gov.il



גיא ליבנה

guy.livne@moh.gov.il



הדס אשל גבע

hadas.eshel-geva@moh.gov.il

