

# GT data mining

## לתעשיית הביומד

אדית אורי

חברת Procedureware

טל': +972-3-5232164

דואל: edit@actcom.co.il

© All Rights Reserved, Edith Ohri  
2010

מצגת מכנס מידע  
ברפואה 20 דצמ' 2010,  
מערכות מידע רפואיות  
מושב 9

# המטרה: להפוך נתונים לידע

להפיק ערך ממאגרי נתונים, להשתמש בהם באסטרטגיה  
ובתפעול, לצורכי מחקר, ייעול, חיזוי, בקרה, תמיכה  
בהחלטות, שיפור שירות, ומינהלה.  
(דוגמאות בהמשך)



ידע

מידע



# פיתוח פתרון GT data mining

- ☆ אדית אורי, מהנדסת תעשייה וניהול MSc, בעלי חב' Procedureware, פיתחה את הפתרון, במהלך ייעול של ריצפת הייצור.
- ☆ GT, המוצר המרכזי של החברה עוסק באבחון נתוני תפעול, מכאן התפיסה המיוחדת שלו, הרואה בנתונים משאב ייצור, ובהפקת ידע – תהליך עבודה סדיר שיש לנהל באפקטיביות.
- ☆ יישומים של GT בישראל, ארה"ב וסינגפור, כוללים את הנושאים: ביוטק, BI, Fraud Detection, מחקר, סקרים, ייצור, איכות, ואספקה.
- ☆ לחברה שת"פ בנושאי טכנולוגיות עזר לקשישים ומידע בתי חולים.
- ☆ אדית אורי משתתפת בפורומים של כריית נתונים, ומנהלת שניים מהם באתר DWH ובלשכת המהנדסים.



# מה הקושי בניתוח נתונים?



אילוסטרציה

- שטפון נתונים וסיבוך \*
- ריבוי גורמים \*
- דינאמיקה \*
- אי תאימות בין מקורות \*
- איכות לא מספקת \*
- חוסר בפרטים חשובים... \*

ת': הפתרון חייב להתגבר על קשיים מעשיים כגון אלה, בכך עוסק GT ←

# GT data mining

GT הוא פתרון כריית נתונים שיודע לאבחן בעצמו דפוסי התנהגות בנתונים, ולהגדיר קשרים חדשים סמויים או חריגים, מגמות, גורמים ואינדיקטורים.

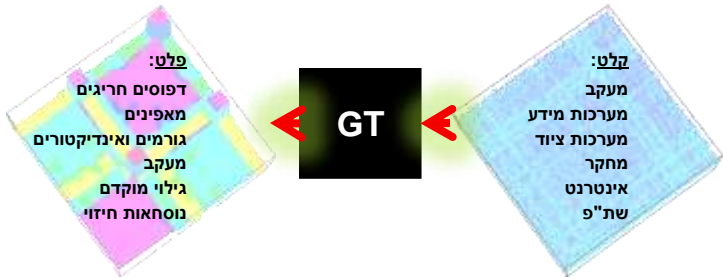
## תחומי ביומד עם ערך מוסף גבוה ל-GT:

- בקרה על שירותי רפואה
- מעקב מטופלים
- מחקר ופיתוח תרופות וטיפולים, מחקר קליני
- ניטור מגפות
- שיפור שירות, כספים ומנהלה
- ניתוח סקרים ו-analytics.



# מודל שירות GT

אנליטיקה ובידול מידע



# דוגמאות ערך מוסף ביישום GT ◀

דוגמא ממחקר על Osteoarthritis - שקפים 8-16

והמחשות קצרות מפרויקטים שונים - שקפים 17-26



# גילויים במחקר מחלת Osteoarthritis

מטרות:

- (1) למצוא את קצב התקדמות המחלה,
- (2) להגדיר ביומרקר'ס.

הקושי: למחלה עדיין אין תרופות, חסרים מחקרים קליניים שיספקו חומר לכריית נתונים, וגם - משך המחלה עשרות שנים ולאור זה, מעקב מטופלים יחיד מכיל מידע מועט יחסית למידע המצוי ברשת.

להלן מוצגת אפשרות חלופית לניצול מידע חופשי  
מהאינטרנט





## המשך – שימוש בנתוני אינטרנט למחקר

מקור הנתונים – הרשות האמריקאית לדיווח על תופעות לוואי בחיסונים –  
VAERS

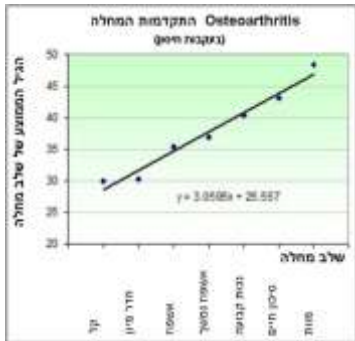
בחרנו מהאתר 1160 דיווחים של תופעות לוואי של המחלה עקב חיסון.  
ברשומה של דיווח יש 8 שדות בלב, המתארים את החיסון, תסמינים עקריים,  
דרגת חומרה, טווח גילים של המטופלים, וכמות הדיווחים שהתקבלו בטווח זה.

<u>היתרון של נתונים מהרשת</u>	<u>החסרון</u>
נתונים חופשיים, אמינים, מאוכלוסייה רחבה ומגוונת.	מעט שדות לרשומה, ללא תיאור פרטני של המטופלים ומחלתם.





# המשך – קצב התקדמות מחלה, Osteoarthritis



את הדרישה הראשונה למציאת "קצב התקדמות המחלה", ניתן לסכם פשוט מתוך הנתונים וההגדרות של VAERS: המחלה מתקדמת בשלב אחד כל 3 שנים.

← הנה תרשים התוצאות

**מסקנה: נתונים עקיפים  
מהרשת הינם שימושיים  
ביותר**

## המשך – ביומרקר'ס, Osteoarthritis

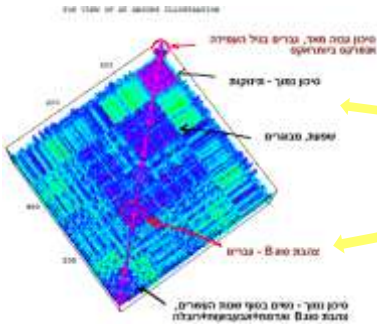
השלב הבא הוא כריית נתונים עם GT, להגדרת דפוסי ההתנהגות והגורמים המשפיעים.

(הביומרקר'ס יהיו חלק מהגדרות אלה)

השתמשנו בחלק מהנתונים, כאשר יתרם נשמר לצורך מבחן אימות.



# המשך - GT לגילוי ביומרקרס



**נמצאו 2 דפוסים יציבים שמאפיניהם:**

← חיסונים נגד שפעת ואנתראקס

← חיסונים נגד צהבת מסוג B, ואבעבועות+אדמת+חורבלה

המחשה של פלט GT - דפוסי ההתנהגות שאותרו



# המשך – GT ביומרקר'ס Osteoarthritis

ניתוח הדפוסים מציף אינדיקטורים להידרדרות המחלה:

סיכון גבוה: בעיות בחוט השדרה אצל גברים מבוגרים, בחיסון אנתראקס ביותראקס. ☠

נשים צעירות בחיסון נגד צהבת מסוג B או החיסון המשולש נגד אבעבועות. \*

כל תסמין של Osteoarthritis אחרי החיסונים הבאים: \*

- צהבת מסוג B (Hep-B)
- אבעבועות+אדמת+רובלה (MEASLES + MUMPS + RUBELLA)
- אנתראקס ביותראקס (Anthrax Biothrax)



# המשך – Osteoarthritis GT, קבוצות מסוכנות

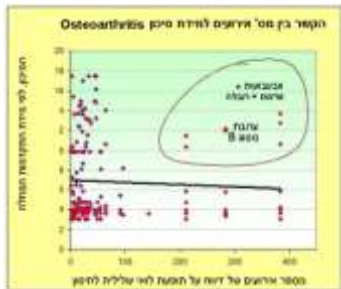
אובחנו תת-קבוצות חריגות, שבהן יש שילוב של סיכון גבוה וריבוי מקרים מדווחים.

הערה: VAERS מוציא משימוש חיסונים מסוכנים כאלה ←

סיבה אפשרית שהם לא הוצאו:

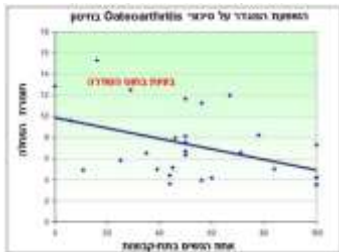
(א) מס' מקרים נמוך מתוך כלל מקבלי החיסון, או

(ב) VAERS לא הבחינו בתופעה כי היא נבלעת בשונות הסטטיסטית הכללית.



**GT מאתר גם דפוסים חריגים וסמויים**

# המשך – Osteoarthritis GT אפקט מגדרי



המשך drildown חושף קשרים חדשים נוספים.

לדוגמא:

**נשים** רגישות פחות מגברים, כללית, לסיכוי Osteoarthritis עקב חיסון.

**GT מוצא תופעות מענינות גם אם לא נשאל עליהן במישרין**

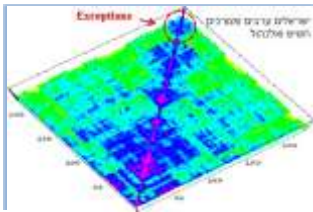


# דוגמא – סיכוני מחלה, מעקב מטופלים

נתון: כ-250 מטופלים במעקב, ב-2 מרכזים רפואיים.

כ-50 שדות, על הסטוריה בריאותית, פרטים אישיים, והרגלי שימוש בסמים ואלכוהול.

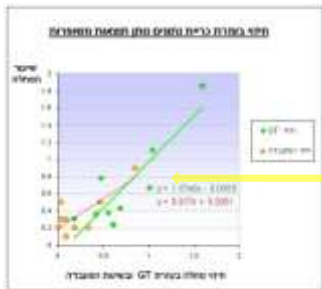
הרופא-חוקר מבקש לחדד את הידע על גורמי גורמי המחלה.



## מוקדי סיכון חדשים שהתגלו ב-GT

- 1) **דפוס חדש**: ישראלים ערבים צרכני חשיש ואלכוהול.
- 2) **גורם חדש כללי**: "יהודים מבוגרים אלמנים" שלא נמצאים בסיכון ידוע.

## המשך דוגמא – ניטור סיכוני מחלה



– החיזוי של GT (בירוק) תואם לתוצאות האמת –  
הגרף הוא כמעט  $45^\circ$

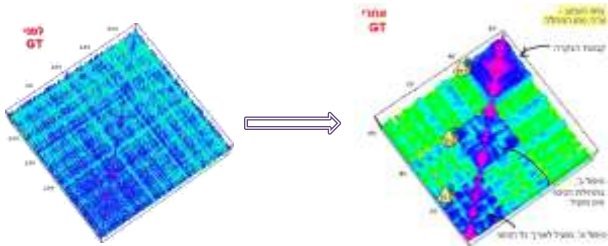
השוואה בין חיזוי עם GT לחיזוי  
בשיטה המעבדתית:

GT מתייחס גם לנתוני הסביבה,  
להרגלי חייהם של מטופלים,  
ולמצבים שבהם אין בדיקות מעבדה.  
כתוצאה, גדל מספר הפרמטרים  
הנלקחים בחשבון, והאבחון משתפר.

**GT מדויק יותר**

# דוגמא – GT בפיתוח תרופה

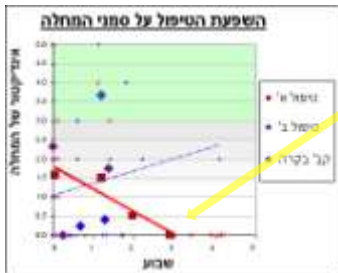
התרופה נבדקה בתנאים שטח קשים, הנתונים באיכות לא מספיקה למסקנות. המשקיעים רוצים לדעת, האם הטיפול החדש (טיפול א') הינו יעיל או לא.



GT מגלה בהירות, שהטיפול החדש מועיל

# המשך דוגמא – GT במסקנות לגבי תרופה חדשה

## Drilldown

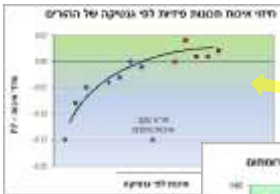


### מסקנות

- (א) הטיפול הנבדק הינו אפקטיבי (מסלק את המחלה תוך 3 שב')  
 (ב) נחשפה תופעה חדשה, בעלת מחזור של שבועיים, שמסבירה את הפיזור הרב בתוצאות שקיבל החוקר...

**GT מסוגל להפיק ידע חדש מנתונים שנחשבים לקשים**

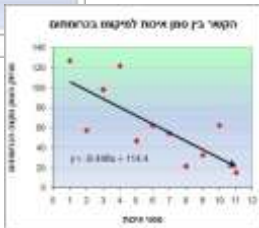
# דוגמא – מחקר בגנטיקה



נתונים: 900 רשומות, 93 מישתנים

החוקר מחפש דרך לחזות הורשה גנטית

GT מוצא קשר בין גנטיקה של האבות לאיכויות של הבנות

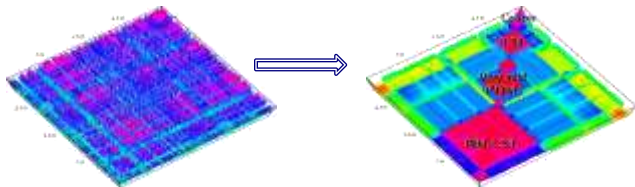


בנוסף: נמצא קשר בין איכויות למיקום הסמן בכורמזום

GT מסוגל להסביר מספר מישתנים גדול מאד

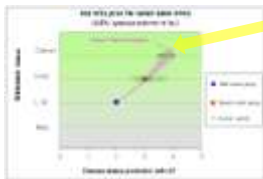
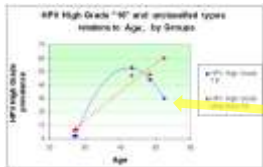
# דוגמת מחקר HPV

**נתונים:** פחות מ-200 רשומות ו-10 שדות (לא מלאים), מבדיקות מעבדה של וירוסים גורמי סרטן צוואר הרחם – HPV, אצל נשים בקבוצת סיכון. מפתחי ערכת הטיפול בודקים את עדיפותו באיתור סיכוני סרטן הרחם



GT מגדיר את דפוסי התפתחות הסרטן

# המשך דוגמת מחקר HPV



## ממצאים חדשים

1. **יורוס-16 מסוכן פחות;**
2. **ערכת HPV דורשת השלמה**

(בתרשים רואים שהוריאנס המסומן בורוד, מתרחב בשלב הסופי, כלומר יש התפרצות לא נחזית של סרטן, בגלל יורוסים מסוכנים שלא כלולים בקוקטייל של HPV).

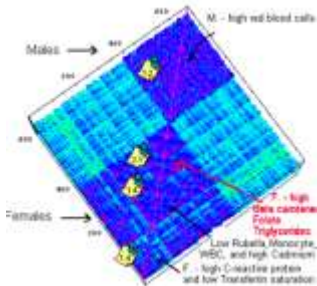
**GT עוזר לפענח את התופעות בשלב מוקדם, ומקצר את ה"זמן לשוק"**



# דוגמא – אבחון דיכאון לפי סמני-דם

נתון: מידע חופשי באינטרנט, כולל 5000 רשומות וכ-50 מישתנים.

החוקר מבקש למצוא סמני דם.

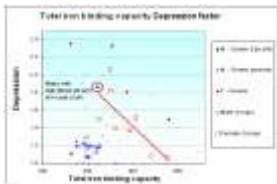


מסקנה ראשונה:

נמצא הבדל ברור בין דפוסי ההתנהגות בגברים ובנשים.

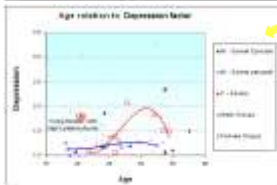


# המשך דוגמא – אבחון דיכאון לפי סמני-דם



מסקנות לגבי גורמי דכאון:

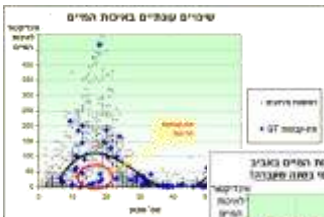
- ערך נמוך של הגורם **Total Iron Binding**
- **נשים** חשופות יותר לדכאון, ובגיל גבוה במספר שנים מגברים.



- גורמים משפיעים ייחודיים לנשים: רמה גבוהה של **Triglycerides** מחסור בויטמין E ...

**GT חושף מידע סמוי בעל ערך רב**

# דוגמא - אקולוגיה של מאגרי מיים



דרוש למצוא במאגר בקרה אקולוגית, מה גורם לתופעה חריגה מדאיגה.

GT מגדיר את התופעה ובנוסף – מאתר **סמנים מוקדמים שלה.**


אחד הסמנים מוצג כאן  
רואים שהחריגים נמצאים  
בקצה הסקאלה של  
ההסבר הכללי.


**GT מפענח חריגים**





# מדגם יישומי כריית נתונים, בעולם


Strand Life Sciences & Clinigene - **ביו-מרקרים** של סיכון לסכרת. הדבר הביא נתח שוק חדש בארה"ב והודו. 


חברת Omron - **אבחון רפואי** של לחץ-דם, לבתי חולים ושירותי רפואה, ולמטופלים בביתם. 

חברת Meridian Medical Technologies עם המכשיר CardioBeeper, **דיאגנוסטיקה מרחוק**, כולל **אבחון מוקדם** של מחלות. 

בארץ: 

מרכז רפואי סורוקה - **אינדיקטורים** לדלקת ריאות 

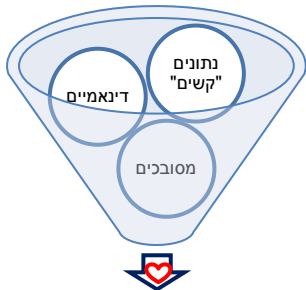
בי"ח הילל יפה - **איתור סיכונים** 

בי"ח איכילוב, מרכז סוראסקי – **שפור שרות** 

חב' **Perlis** – **אבחון מוקדם ומונע לקשישים בביתם או תחת השגחה**, ועוד. 



# שטפון הנתונים הוא הזדמנות עסקית!



**GT DATA MINING**

- < GT מסוגל ליצר ידע מיידי מנתוני תפעול.
- < בלי דרישות מוקדמות לאינטגרציה או תכנה, ובלי להכביד על מערכת הלקוח.
- < היחיד שפותח על בסיס מדעי הניהול לניצול יעיל של משאבי מידע ושילוב תוצריהם בעבודה.

# תודה למשתתפים

אדית אורי

Procedureware חברת

+972-3-5232164 טל:

edit@actcom.co.il דואל:

[http://www.ilsa.org.il/companies\\_life\\_science\\_company.asp?ID=827](http://www.ilsa.org.il/companies_life_science_company.asp?ID=827)

*~ Thanks for your time ~*

