

# **פיתוח מודל מבוסס NLP לחיזוי אשפוזים חוזרים בקרב מטופלי נפש**

**מגישים:**

**רן אלמגור - 318955440**

**חן דוד - 315237255**

**רעות תמאם – 206820383**

**שם המנחה: ד"ר זהר ברנט-יצחקי**

1	פיתוח מודל מבוסס NLP לחיזוי אשפוזים חוזרים בקרב מטופלי נפש
3	תקציר:
4	רקע וסקירה ספרותית:
5	שאלת המחקר:
5	שיטות:
5	מקורות הנתונים:
6	תהליך ניקוי ועיבוד הנתונים:
7	שיטות סטטיסטיות תיאוריות והסקה סטטיסטית
8	תוצאות:
16	דיון ומסקנות:
17	בילביוגרפיה:

אשפוזים חוזרים בקרב מטופלים פסיכיאטריים מהווים אתגר קריטי במערכת הבריאות, עם השלכות משמעותיות על איכות החיים של המטופלים, העומס על מוסדות רפואיים, והעלות הכלכלית הגבוהה למערכת. **תופעת "הדלת המסתובבת"** מתארת את המציאות שבה מטופלים משתחררים מאשפוז רק כדי לחזור אליו זמן קצר לאחר מכן, פעמים רבות בשל חוסר מענה טיפולי הולם, היעדר תמיכה משפחתית, או מחסור בשירותי המשך בקהילה. תופעה זו יוצרת מעגל אינסופי שמכביד על המטופלים ועל הצוותים הרפואיים כאחד, ומחייב מציאת פתרונות מבוססי נתונים שיסייעו במניעתה.

במסגרת מחקר זה, בחנו את **הגורמים המנבאים חזרות לאשפוז**, במטרה לזהות דפוסים נסתרים ולפתח כלים שיסייעו בזיהוי מוקדם של מטופלים בסיכון גבוה. לצורך כך, נאסף **מאגר נתונים רפואי רחב-היקף**, הכולל מאפיינים דמוגרפיים, קליניים והיסטוריים אשפוזים. אך לפני שניתן היה להפיק תובנות משמעותיות, התמודדו החוקרים עם **תהליך ניקוי ועיבוד נתונים מאתגר במיוחד**—מאגר הנתונים היה מבולגן, מלא בחוסרים, שגיאות רישום, כפילויות וחוסר עקביות בין משתנים קריטיים. ללא טיפול יסודי, המחקר היה עלול להיכשל. לכן, הושקעו **מאמצים אדירים** בזיהוי ותיקון טעויות, השלמת ערכים חסרים והמרת משתנים לניתוח מתקדם, עד שהתקבל **מאגר מהימן ומדויק**, שיאפשר תובנות משמעותיות.

לאחר שהנתונים עברו **תהליך הכנה קפדני**, בוצעו **ניתוחים סטטיסטיים מתקדמים** שנועדו לחשוף קשרים עמוקים בין מאפייני המטופלים לבין הסיכון לחזרה לאשפוז. במחקר זה יושמו **מבחני חי בריבוע, Mann-Whitney, ספירמן, וכן שיטות סטטיסטיות רבות נוספות**, לצד שיטות ויזואליזציה מתקדמות שהפכו את הנתונים למובנים ונגישים יותר. כל תובנה גובתה בגרפים אינטראקטיביים, שאפשרו לזהות דפוסים חבויים שלא ניתן היה להבחין בהם בניתוחים מסורתיים.

המחקר מספק הצצה לעולם המורכב של אשפוזים חוזרים, ומציע כיוונים חדשים שיכולים לשפר את ההתמודדות עם התופעה – מהבנת הדפוסים ועד יצירת כלים שיאפשרו קבלת החלטות טובה יותר בעתיד.

מערכות בריאות הנפש ברחבי העולם ניצבות בפני אתגר משמעותי, אשר נדמה כי פתרונו עדיין חומק מהידיים: **אשפוזים חוזרים בקרב מטופלים פסיכיאטריים**. למרות התקדמות הרפואה הפסיכיאטרית והשיפור בשירותי השיקום, רבים מהמטופלים המשתחררים מאשפוז חוזרים אליו תוך פרק זמן קצר, במה שמכונה **"תופעת הדלת המסתובבת"**. "חזרות אלו אינן רק עניין סטטיסטי, אלא בעיה המשפיעה על רציפות הטיפול, זמינות המשאבים הרפואיים, איכות החיים של המטופלים והעלות הכלכלית למערכת הבריאות. השאלה הגדולה היא: האם ניתן לחזות מראש מי נמצא בסיכון גבוה לחזרה לאשפוז? ואם כן, כיצד ניתן להתערב מבעוד מועד כדי למנוע זאת?

חוקרים רבים ניסו לפצח את החידה הזו. **מחקר שנערך בגרמניה בשנת 2013** בחן יותר מ-17,000 מטופלים לאורך תקופה של 12 שנים בניסיון להבין אילו גורמים משפיעים על חזרות לאשפוז חוזר. נמצא כי גיל, מגדר, מצב משפחתי ותעסוקה הם בין הגורמים המשמעותיים ביותר, כאשר מטופלים רווקים, מובטלים או חסרי תמיכה חברתית נמצאים בסיכון מוגבר לחזרה. ממצא מעניין נוסף היה כי מגורים בעיר הגדילו את הסיכון לאשפוזים חוזרים, ככל הנראה בשל גורמי לחץ סביבתיים והיעדר קהילה תומכת.

באותה תקופה, **מחקר אחר שנערך בשווייץ בשנת 2019** לקח את החקירה צעד נוסף קדימה, ובחן את הקשר בין סוג האבחנה הפסיכיאטרית לבין שיעורי החזרה לאשפוז. ממצאי המחקר היו ברורים: **מטופלים עם הפרעות פסיכוטיות והפרעות אישיות נמצאים בסיכון הגבוה ביותר לחזרה לאשפוז תוך שנתיים בלבד**, כאשר חלקם אף מתאשפדים מחדש בתוך מספר חודשים. בנוסף, נמצא כי מטופלים שאושפדו בכפייה בפעם הראשונה, סביר להניח שיחוו אשפוזים חוזרים יותר ממטופלים שאושפדו מרצונם.

אך בעוד שמחקרים אלו הצליחו להצביע על מגמות, הם לא הצליחו להציע מודל חיזוי מדויק. מרבית המחקרים שהתבססו על ניתוחים סטטיסטיים מסורתיים לא היו מסוגלים לזהות דפוסים מורכבים בין משתנים שונים, מה שהשאיר חלל מחקרי משמעות. יתרה מכך, **מאמר ישראלי שהתפרסם בשנת 2023** ציין כי **כמות עצומה של מידע טקסטואלי הנמצא בתיקים הרפואיים של המטופלים כלל אינה מנוצלת**, מאחר שמרבית המחקרים התמקדו בנתונים כמותיים בלבד. מידע זה, הכולל דיווחים רפואיים מפורטים, רשומות שיחות טיפוליות והיסטוריה תרופתית, עשוי לספק הבנה עמוקה יותר של הגורמים לחזרה לאשפוז, אך עד כה לא נמצא לו פתרון מחקרי מספק.

כאן נכנס לתמונה המחקר הנוכחי, שמטרתו **לגשר על הפערים הקיימים ולפתח שיטות מתקדמות לניתוח נתוני מטופלים**. בניגוד למחקרים קודמים שהתמקדו בעיקר במשתנים דמוגרפיים וקליניים, מחקר זה מציע גישה המבוססת על ניתוח מעמיק של מידע רפואי תוך שימוש בשיטות סטטיסטיות מתקדמות. באמצעות מאגר נתונים רפואי רחב-היקף, הכולל מידע דמוגרפי, קליני וטקסטואלי, הנתונים עוברים עיבוד, ניקוי וניתוח מקיף במטרה לזהות תבניות חדשות ולספק תובנות רלוונטיות על גורמי הסיכון לאשפוזים חוזרים.

## שאלת המחקר:

האם ניתן לחזות אשפוז חוזר של מטופלי נפש באמצעות מאפיינים קליניים ודמוגרפיים, בשילוב ניתוח טקסט של תיקים רפואיים?

## שיטות:

## מקורות הנתונים:

המאגר שעליו מתבסס מחקר זה התקבל מבית החולים **לב השרון - פרדסיה**, מוסד רפואי פסיכיאטרי מוביל המתמחה בטיפול ושיקום מטופלים עם הפרעות נפשיות. המאגר כלל נתונים מקיפים על מטופלים שאושפזו בין השנים 2020 ועד היום, תוך רישום מפורט של מאפייניהם הדמוגרפיים, הקליניים והחברתיים. עם זאת, המידע התקבל במצב **מאד מבולגן ולא תקין**, וכלל רשומות חלקיות, ערכים סותרים, שגיאות ברישום התאריכים, כפילויות משמעותיות ועמודות טקסט שאינן אחידות. כבר במבט ראשוני היה ברור כי מדובר במאגר הדורש **תהליך יסודי של עיבוד, סינון, תיקון והשלמה**, על מנת להבטיח שהוא יהיה מהימן ומוכן לניתוח סטטיסטי מתקדם. הנתונים כללו משתנים דמוגרפיים כגון: מגדר, ארץ לידה, מצב משפחתי ומספר ילדים, לצד משתנים קליניים קריטיים כמו: מספר האשפוזים הקודמים, משך האשפוז, סיבת הקבלה וסטטוס חוקי של האשפוז (למשל: מרצון או בכפייה). בנוסף, נכללו פרטים על אופן הקבלה והשהות בבית החולים, לרבות תאריכי כניסה ושחרור מהמחלקה, היחידה הרפואית שבה שהה המטופל והפניה להמשך טיפול. רמת אי-הסדר שהתקבלה דרשה **בדיקה קפדנית של כל רשומה ורשומה**, שכן נמצאו אי-התאמות חמורות בין שדות שונים, כמו תאריכי שחרור מוקדמים מתאריכי קבלה, היעדר נתונים מרכזיים על מטופלים, שגיאות הקלדה בסיבת האשפוז ורשומות כפולות שנבעו מהזנת מידע חוזרת במערכות שונות. היה הכרח לבצע **ניקוי מעמיק, סינון מחמיר והשלמת נתונים חסרים בגישות סטטיסטיות מתקדמות**, תוך שימוש במודלים להשלמת נתונים, תיקון סתירות והמרת ערכים טקסטואליים לפורמט תקני לניתוח. מאגר הנתונים עבר תהליך אחידות ושדרוג משמעותי, אשר הפך אותו מכלי לא שמיש וכמעט לא רלוונטי, **לבסיס נתונים תקין, מסודר ומהימן**, שמאפשר ניתוחים סטטיסטיים מדויקים וביסוס מודלים לחקר תופעת האשפוזים החוזרים.

## תהליך ניקוי ועיבוד הנתונים:

תהליך ניקוי ועיבוד הנתונים היווה אתגר משמעותי ודרש שבועות ארוכים של עבודה אינטנסיבית, שבמהלכם הושקעו מאמצים רבים בהפיכת מאגר הנתונים המבולגן, הלא אחיד ורווי בשגיאות למערכת מידע אמינה, מסודרת ומדויקת, המסוגלת לספק ניתוחים סטטיסטיים תקפים ולהוות בסיס מהימן למחקר. כבר בשלבים הראשונים של העבודה התגלו חוסר משמעותיים, כפילויות, חוסר עקביות בתאריכים, טעויות בהזנת מידע ורשומות שאינן עומדות באמות מידה מחקריות, מה שדרש תהליך מעמיק של עיבוד, ניקוי וארגון מחדש.

בשלב הראשון בוצעה המרת עמודות טקסט לערכים מספריים (Dummies) באמצעות One Hot Encoding, צעד קריטי לשיפור איכות הנתונים ולהגברת היעילות של האלגוריתמים החישוביים שישמשו אותנו בהמשך הפרויקט. משתנים קטגוריים כגון אזור מוצא, סטטוס האשפוז, גורם מפנה, סיבת שחרור, למי הופנה המטופל, היחידה הרפואית ועוד רבים הומרו לעמודות בינאריות (0 ו-1), מה שאיפשר ליצור מערך נתונים מובנה ומדויק, המסייע בזיהוי קשרים סטטיסטיים ולישיפור ביצועי המודלים שנבחר בהמשך.

במקביל, בוצעה בדיקת תקינות ואמינות הנתונים, שהתמקדה בשדות קריטיים כגון: תאריכי אשפוז ושחרור, מספרי מקרה רפואיים ושיוך למחלקות שונות. בשלב זה זוהו רשומות עם תאריכים לא עקביים, בהן תאריכי שחרור מוקדמים מתאריכי קבלה, הפרשי זמן לא הגיוניים בין אשפוזים רצופים וטעויות הקלדה שהובילו לחוסר אמינות בתיעוד משך האשפוזים. כדי להתמודד עם בעיה זו, פותח פיצ'ר חדש של גיל המטופל בעת האשפוז, שהתבסס על חישוב הגיל מתוך תאריך הלידה ביחס לתאריך הקבלה לאשפוז. תכונה זו הפכה את הנתונים לברורים, עקביים ומוכנים לניתוח מעמיק.

נוסף על כך, אותרו רשומות עם מספרי מקרה רפואיים לא עקביים, אשר נבעו מתיעוד חוזר ממחלקות שונות או טעויות הקלדה במערכת. חלק מהמטופלים הופיעו עם מספרי מקרה שונים עבור אותו אשפוז, תוצאה של מעבר בין מחלקות שונות או חוסר אחידות בין מערכות הרישום בבית החולים. כדי להבטיח שכל רשומה תייצג באופן מדויק ואחיד את מהלך האשפוז, בוצע תהליך של איחוד נתונים למטופלים, כך שלכל מטופל יוצגו כל האשפוזים החוזרים שלו ברשומות תקינות עם מספר מקרה רפואי עקבי, מה שהקל על הניתוחים הסטטיסטיים ושיפר את איכות המידע המחקרי.

לאחר תיקון נתוני התאריכים ומספרי המקרה, בוצע סיווג מחדש של משתנה סיבת השחרור, כדי להבחין בין מטופלים ששוחררו לביתם, הועברו למחלקות אחרות או למוסדות רפואיים שונים, בהתאם להנחיות הנהלת המחקר. נוסף על כך, נמחקו רשומות של מטופלים שנפטרו או הועברו למוסד רפואי אחר, מאחר שהם אינם רלוונטיים לחקר חזרות לאשפוזים.

בשלב זה, נוספו פיצ'רים חדשים המאפשרים הבנה טובה יותר של תבניות האשפוזים החוזרים. בהתאם להגדרות בית החולים, אשר מגדירות אשפוז חוזר כמצב שבו מטופל חזר תוך חודש או שלושה חודשים, נוצרו משתנים המאפשרים למדוד את מספר האשפוזים החוזרים של כל מטופל, ולבחון את הסיכוי לחזרה לאשפוז תוך פרקי זמן מוגדרים. משתנים אלו היו הכרחיים על

מנת להגדיר ולאפיין אשפוזים חוזרים בצורה תקנית, תוך התאמה לצרכים המחקריים ולמדיניות הרפואית של בית החולים.

נוסף על כך, בוצע **מיפוי מחדש של משתנים רבים**, כולל **חלוקת אזור מוצא למדינות בהתאם לאזור גיאוגרפי, קיבוץ גורמי מפנה ועוד**, למשל: **בתי חולים, קופות חולים ומוסדות רפואיים אחרים** או **מדינות בביטחון מסוימת מופו תחת קטגוריה רחבה יותר**, דבר שיאפשר ניתוח סטטיסטי מדויק יותר תוך שמירה על רמת פירוט נדרשת.

לבסוף, **בוצע סינון יסודי של הרשומות הלא תקינות או מחיקת שורות כמעט ריקות**, אשר לא הכילו מספיק מידע מהותי למחקר. לאחר **כל שלבי העבודה**, מאגר הנתונים **צומצם ליותר ממחצית מכמות השורות המקורית**, תוך שמירה על הנתונים החשובים ביותר לצורך ניתוחים סטטיסטיים אמיתיים. תהליך זה, שנמשך **שבועות ארוכים של עבודה מדויקת וקפדנית**, הפך מאגר ראשוני **מבולגן ומורכב למערכת מידע תקנית, אמינה ומסודרת**, אשר מהווה **תשתית מחקרית מוצקה** להסקת מסקנות מבוססות ולפיתוח תובנות בעלות ערך קליני משמעותי.

### שיטות סטטיסטיות תיאוריות והסקה סטטיסטית:

כחלק מהשלב האנליטי במחקר, בוצע **ניתוח סטטיסטי תיאורי והסקה סטטיסטית**, שנועד להעריך קשרים מובהקים בין מאפייני המטופלים לבין הסיכוי שלהם להתאשפז **מלכתחילה** וכן **לחווה אשפוז חוזר**. ניתוח זה התבצע תוך השוואה לנתוני האוכלוסייה הכללית בישראל, בהתאם לנתוני הלשכה המרכזית לסטטיסטיקה (הלמ"ס), במטרה להבין האם קיימת **חריגה בשיעורי האשפוז בקרב קבוצות מסוימות**, בהשוואה להתפלגות הדמוגרפית הכללית.

השלב הראשון של הניתוח כלל **בדיקה של משתנים דמוגרפיים מרכזיים**, בהם **מגדר, מצב משפחתי, גיל המטופל בעת האשפוז ומספר הילדים**, והערכת תרומתם האפשרית לסיכון של אדם להתאשפז מלכתחילה ולחזור לאשפוז לאחר שחרורו. הסטטיסטיקה התיאורית אפשרה לבחון האם **קבוצות מסוימות באוכלוסייה מצויות בסיכון גבוה יותר להיכנס למסגרת אשפוזית**, בהשוואה להתפלגות באוכלוסייה הכללית בישראל, ובכך להבין האם קיימת **שונויות דמוגרפית מובהקת** בין האוכלוסייה המתאשפזת לבין כלל האוכלוסייה.

לאחר מכן, **בוצע פילוח של הנתונים לפי מצב משפחתי ומגדר**, והוערכו הקשרים בין משתנים אלו לבין סוגי האשפוזים החוזרים. הסטטיסטיקה בוצעה בשלושה מישורים מרכזיים:

- **חזרה לאשפוז באופן כללי** – נבדק האם קיים קשר בין מגדר או מצב משפחתי לבין הנטייה לחוות חזרה לאשפוז בכל נקודת זמן.
- **חזרה לאשפוז תוך חודש** – בהתאם להגדרת בית החולים, נבחן הקשר בין מאפייני המטופלים לבין חזרה מהירה לאשפוז תוך 30 ימים משחרור.
- **חזרה לאשפוז תוך שלושה חודשים** – נבדק האם יש קשר מובהק בין משתנים דמוגרפיים לבין חזרה לאשפוז בטווח של 90 ימים לאחר השחרור, בהתאם להגדרת בית החולים.

כדי לבחון קשרים סטטיסטיים בין המשתנים הדמוגרפיים לבין חזרה לאשפוז, בוצע **מבחן Chi-Square Test** (בריווע), אשר אפשר לבדוק האם **מגדר או מצב משפחתי מהווים גורם מובהק סטטיסטי בסבירות של מטופל לחזור לאשפוז**. מבחן זה איפשר **לזהות קבוצות אוכלוסייה הנמצאות בסיכון גבוה יותר** ולבחון האם קיימת תלות סטטיסטית בין מאפייני המטופל לבין סיכוייו לחזור לאשפוז.

לצד זאת, בוצע מבחן Mann-Whitney אשר השווה גיל מטופלים בין אלו שחזרו לאשפוז לבין אלו שלא, במטרה לזהות הבדלים מובהקים סטטיסטית בין קבוצות שונות ולבחון האם גיל בעת האשפוז מהווה משתנה קריטי בזיהוי אוכלוסיות בסיכון גבוה. מאחר שהנתונים אינם מתפלגים בצורה נורמלית, מבחן זה היה הבחירה המתאימה ביותר להערכת הבדלים סטטיסטיים, תוך שמירה על תקפות מחקרית גבוהה.

בנוסף, נבדק מבחן ספירמן (Spearman's Rank Correlation Coefficient) לצורך הערכת הקשר בין מספר הילדים של המטופל לבין הסיכוי לחזרה לאשפוז. מבחן זה נבחר מאחר שהוא מודד קשר בין שני משתנים רציפים או סדרתיים מבלי להניח התפלגות נורמלית, מה שמתאים במיוחד לנתונים הדמוגרפיים שנבחנו במחקר זה.

כדי לוודא האם הנתונים מתפלגים נורמלית, בוצע מבחן שפירו-וילק (Shapiro-Wilk Test), אשר מאפשר לבדוק האם משתנים רציפים כמו גיל המטופל או משך האשפוז מתפלגים בצורה נורמלית. המבחן הראה כי ההתפלגות אינה נורמלית, ולכן הוחלט להשתמש במבחנים לא פרמטריים כמו Mann-Whitney וספירמן, אשר מתאימים יותר לתנאים אלו.

בנוסף להשוואת הנתונים בתוך מאגר המטופלים, בוצעה השוואה לנתוני הלמ"ס, במטרה לבדוק האם שיעורי האשפוזים החוזרים חורגים מההתפלגות הדמוגרפית הכללית של האוכלוסייה הישראלית. ההשוואה נועדה לזהות האם קבוצות מסוימות מתאפיינות בשיעור אשפוזים גבוה מהמצופה בהתבסס על שיעורן באוכלוסייה, דבר שעשוי להעיד על השפעות מערכתיות, חברתיות או רפואיות שיש להביא בחשבון.

תהליך הניתוח הסטטיסטי בוצע בזהירות רבה, תוך בקורות מחמירות ואימות חוזר של תוצאות המבחנים, על מנת להבטיח דיוק ומהימנות בניתוחי הקשרים בין המשתנים השונים. שלב זה מהווה צעד קריטי במחקר, שכן הוא מאפשר לזהות מאפיינים דמוגרפיים מרכזיים שעשויים לשמש כאינדיקטורים לזיהוי מוקדם של מטופלים בסיכון גבוה לחזרה לאשפוז, ומעניק בסיס מדעי מוצק להמשך פיתוח הכלים האנליטיים בפרויקט.

## תוצאות:

להלן ממצאי המחקר, המוצגים באופן שיטתי ומעמיק, תוך שימוש בסטטיסטיקות תיאוריות והסקת מסקנות מבוססות מבחנים סטטיסטיים. תוצאות אלו מספקות הבנה רחבה ומעמיקה של הקשרים הדמוגרפיים והקליניים הקשורים לאשפוזים פסיכיאטריים חוזרים, ומאפשרות לזהות דפוסים חשובים אשר עשויים לסייע בעתיד בגיבוש אסטרטגיות למניעת חזרות לאשפוז.

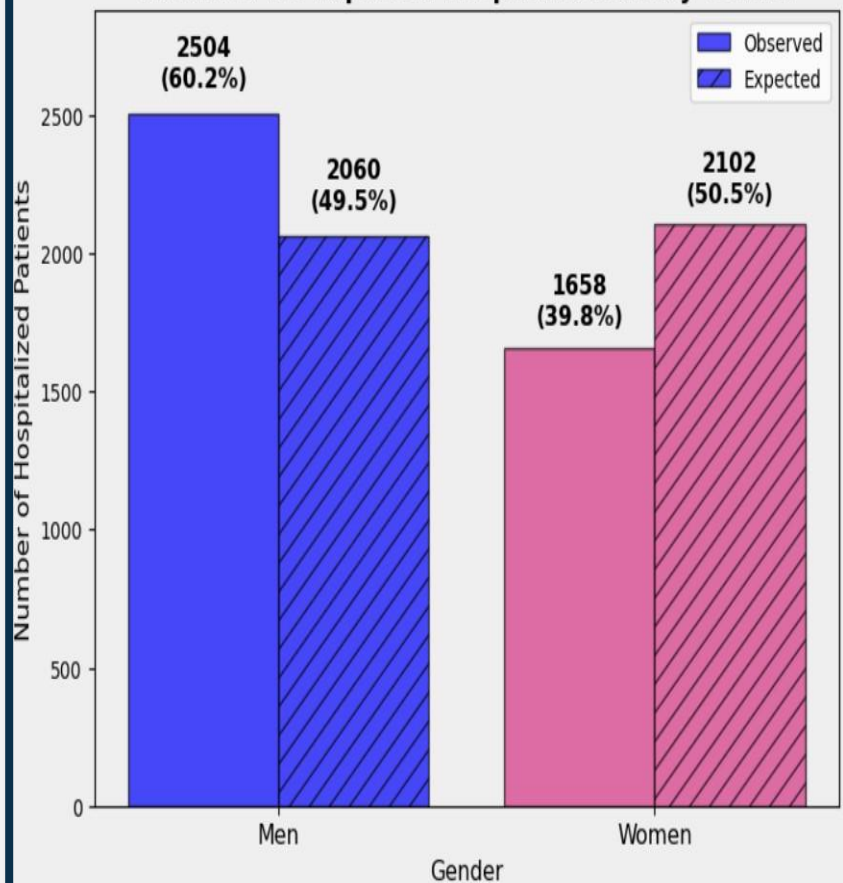
תחילה, יוצגו התפלגויות האשפוזים לפי מגדר, אשר יאפשרו לבחון האם קיימים הבדלים משמעותיים בין גברים לנשים בשיעורי האשפוז הכללי. לאחר מכן, ינותחו שיעורי האשפוזים החוזרים לפי מגדר, ולאחר מכן יוצגו תוצאות הבוחנות את הקשר בין מגדר לבין חזרה לאשפוז בפרקי זמן קצרים – תוך חודש ותוך שלושה חודשים מהשחרור הקודם.

בהמשך, ינותחו הקשרים שבין מצב משפחתי לבין האשפוזים, כאשר תחילה תוצג התפלגות האשפוזים הכללית לפי סטטוס משפחתי, ולאחר מכן יבחנו הקשרים בין מצב משפחתי לבין חזרה לאשפוז בכלל, ובהמשך גם חזרה לאשפוז תוך חודש ותוך שלושה חודשים.

לבסוף, יוצגו התוצאות הבוחנות את השפעת מספר הילדים של המטופל על שיעור האשפוזים החוזרים, וכן הקשר בין גיל המטופל לבין חזרה לאשפוז, במטרה לבדוק האם קיימת מגמה דמוגרפית המנבאת סיכון מוגבר לאשפוזים חוזרים.

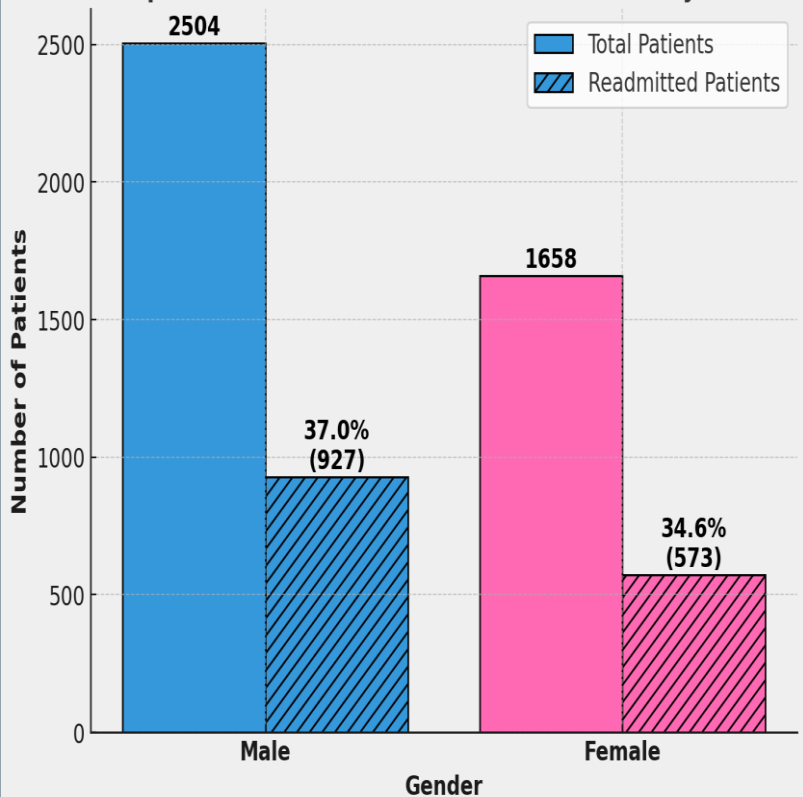


Observed vs. Expected Hospitalizations by Gender



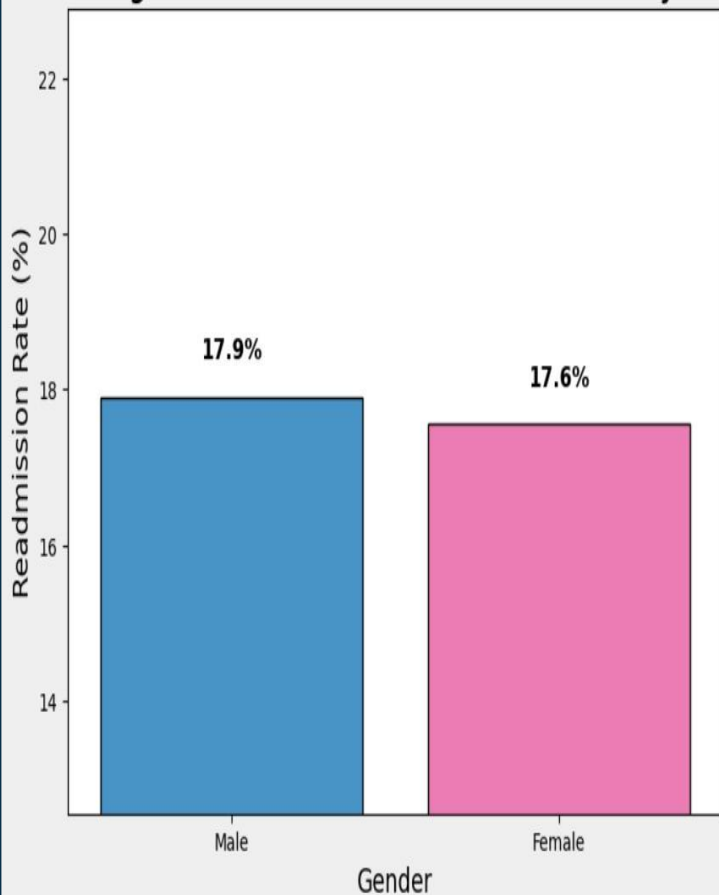
פרשנות	תוצאה מספרית	מדד סטטיסטי
המדגם כולל 4,162 מטופלים שאושפזו במהלך התקופה.	4,162	מספר המטופלים שאושפזו
נמצא הבדל מובהק בין גברים לנשים במספר האשפוזים.	95.2	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך קטן מ-0.05, ולכן קיימת מובהקות סטטיסטית.	$0.01 < \text{קטן מ}$	ערך ק רמת מובהקות
לפי ההתפלגות הצפויה של הנתונים (ה"מס).	2,060 (49.5%)	מספר אשפוזים צפוי (גברים)
נמצא כי מספר האשפוזים בפועל גבוה מהצפוי.	2,504 (60.2%)	מספר אשפוזים בפועל (גברים)
לפי ההתפלגות הצפויה של הנתונים (ה"מס).	2,102 (50.5%)	מספר אשפוזים צפוי (נשים)
נמצא כי מספר האשפוזים בפועל נמוך מהצפוי.	1,658 (39.8%)	מספר אשפוזים בפועל (נשים)

Comparison of Total and Readmitted Patients by Gender



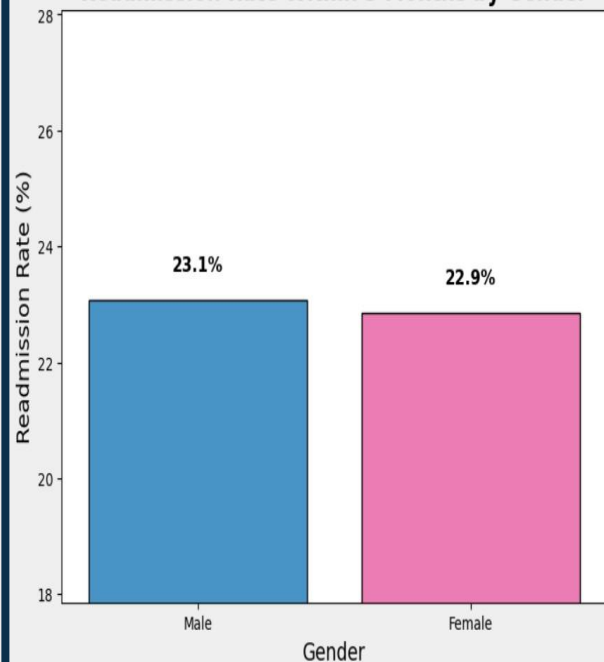
פרשנות	תוצאה מספרית	מדד סטטיסטי
המדגם כולל 4,162 מטופלים שאושפזו במהלך התקופה.	4,162	מספר המטופלים הכולל
1500 מטופלים חזרו לאשפוז חוזר מהמדגם.	1,500	מספר המטופלים שחזרו
נמצא ערך נמוך יחסית של חי-בריבוע, המעיד על קשר חלש בין מגדר לאשפוז חוזר.	2.52	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך ק גדול מ-0.05, ולכן אין מובהקות סטטיסטית.	0.11	ערך ק רמת מובהקות
מספר הגברים שאושפזו במדגם.	2,504	מספר הגברים שאושפזו
מספר הנשים שאושפזו במדגם.	1,658	מספר הנשים שאושפזו
כמות הגברים שחזרו לאשפוז לפחות פעם אחת.	927 (37%)	מספר הגברים שחזרו לאשפוז
כמות הנשים שחזרו לאשפוז לפחות פעם אחת.	573 (34.6%)	מספר הנשים שחזרו לאשפוז

Percentage of Readmissions Within One Month by Gender



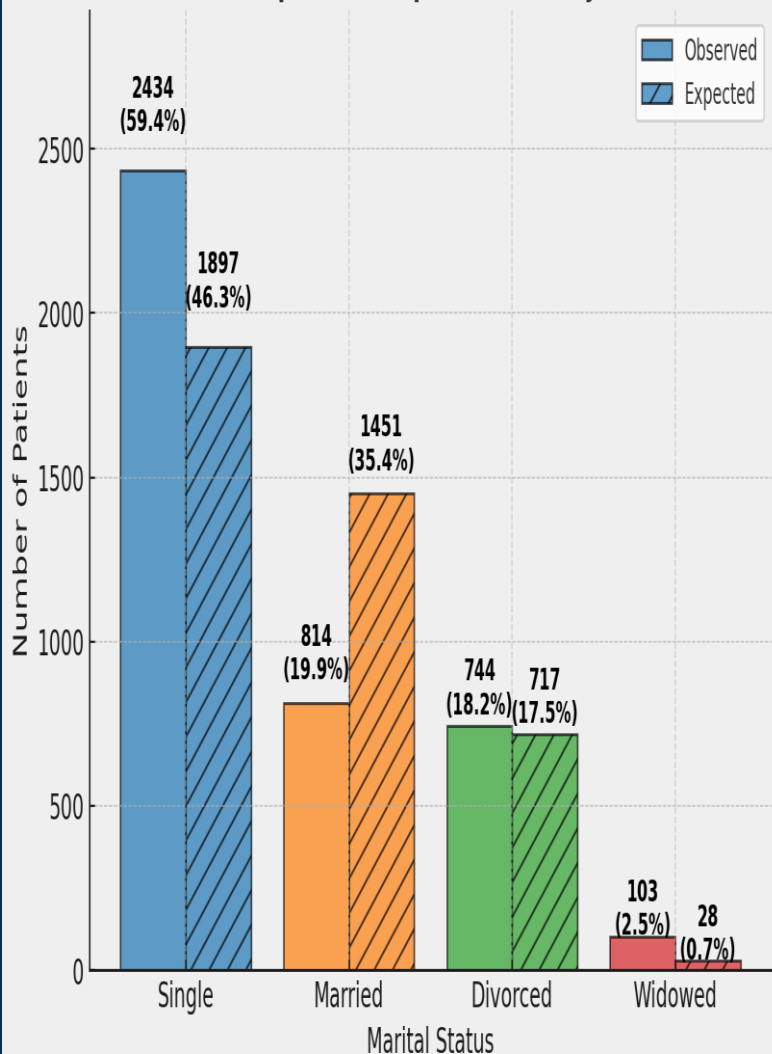
פרשנות	תוצאה מספרית	מדד סטטיסטי
המדגם כולל 4,162 מטופלים שאושפזו במהלך התקופה.	4,162	מספר המטופלים הכולל
739 מטופלים חזרו לאשפוז אחרי חודש.	739	מספר המטופלים שחזרו
נמצא ערך נמוך יחסית של חי-בריבוע, המעיד על קשר חלש בין מגדר לאשפוז חוזר אחרי כחודש.	0.05	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך ק גדול מ-0.05, ולכן אין מובהקות סטטיסטית.	0.81	ערך ק רמת מובהקות
מספר הגברים שאושפזו במדגם.	2,504	מספר הגברים שאושפזו
מספר הנשים שאושפזו במדגם.	1,658	מספר הנשים שאושפזו
כמות הגברים שחזרו לאשפוז לפחות פעם אחת אחרי חודש.	448 (17.9%)	מספר הגברים שחזרו לאשפוז
כמות הנשים שחזרו לאשפוז לפחות פעם אחת אחרי חודש.	291 (17.6%)	מספר הנשים שחזרו לאשפוז

Readmission Rate Within 3 Months by Gender

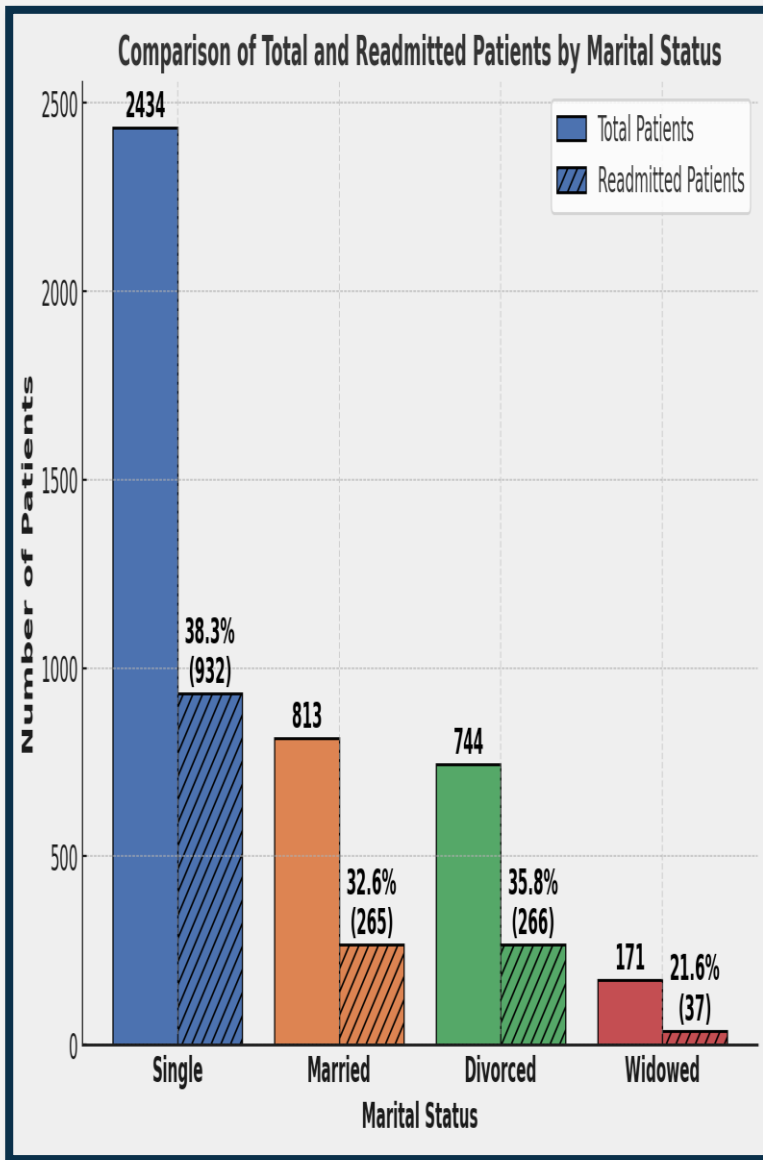


פרשנות	תוצאה מספרית	מדד סטטיסטי
המדגם כולל 4,162 מטופלים שאושפזו במהלך התקופה.	4,162	מספר המטופלים הכולל
957 מטופלים חזרו לאשפוז אחרי חודש.	957	מספר המטופלים שחזרו
נמצא ערך נמוך יחסית של חי-בריבוע, המעיד על קשר חלש בין מגדר לאשפוז חוזר אחרי כשלושה חודשים.	0.01	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך ק גדול מ-0.05, ולכן אין מובהקות סטטיסטית.	0.81	ערך ק רמת מובהקות
מספר הגברים שאושפזו במדגם.	2,504	מספר הגברים שאושפזו
מספר הנשים שאושפזו במדגם.	1,658	מספר הנשים שאושפזו
כמות הגברים שחזרו לאשפוז לפחות 3 חודשים.	578 (23.1%)	מספר הגברים שחזרו לאשפוז
כמות הנשים שחזרו לאשפוז לפחות 3 חודשים.	379 (22.9%)	מספר הנשים שחזרו לאשפוז

Observed vs. Expected Hospitalizations by Marital Status

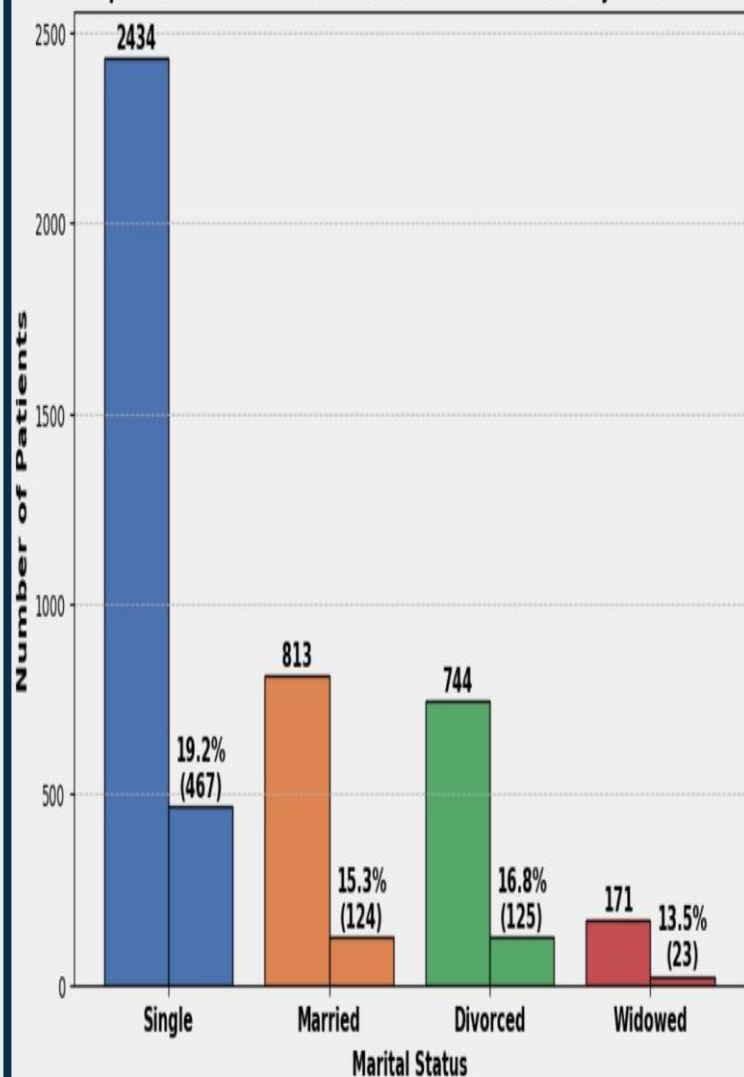


פרשנות	תוצאה מספרית	מדד סטטיסטי
נמצא הבדל מובהק בין מצבים משפחתיים במספר האשפוזים.	624.56	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך קטן מ-0.05, ולכן קיימת מובהקות סטטיסטית.	$< 0.01$ קטן מ	ערך קרמט מובהקות
לפי ההתפלגות הצפויה של הנתונים (הל"מס).	2434 (59.4%)	מספר אשפוזים צפוי (רווקים)
לפי ההתפלגות הצפויה של הנתונים (הל"מס).	145 (19.9%)	מספר אשפוזים צפוי (נשואים)
לפי ההתפלגות הצפויה של הנתונים (הל"מס).	744 (18.2%)	מספר אשפוזים צפוי (גרושים)
לפי ההתפלגות הצפויה של הנתונים (הל"מס).	103 (2.5%)	מספר אשפוזים צפוי (אלמנים)
נמצא כי מספר האשפוזים בפועל גבוה מהצפוי.	1897 (46.3%)	מספר אשפוזים בפועל (רווקים)
נמצא כי מספר האשפוזים בפועל נמוך מהצפוי.	1451 (35.4%)	מספר אשפוזים בפועל (נשואים)
נמצא כי מספר האשפוזים בפועל גבוה מהצפוי.	717 (17.5%)	מספר אשפוזים בפועל (גרושים)
נמצא כי מספר האשפוזים בפועל גבוה מהצפוי.	28 (0.7%)	מספר אשפוזים בפועל (אלמנים)



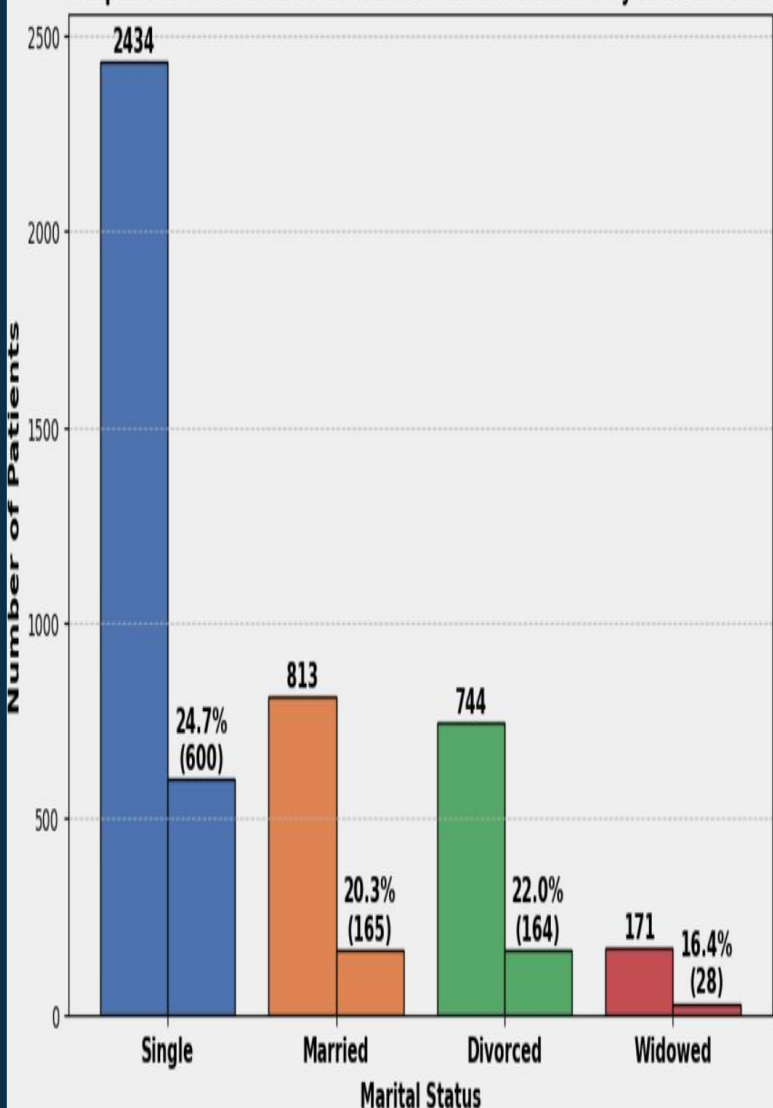
פרשנות	תוצאה מספרית	מדד סטטיסטי
נמצא הבדל מובהק בין מצבים משפחתיים לחזרה לאשפוז חוזר.	24.95	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך קטן מ-0.05, ולכן קיימת מובהקות סטטיסטית.	$0.01 < \text{קטן מ}$	ערך קרמט מובהקות
מספר הרווקים הכולל שאושפזו במדגם.	2,434	מספר הרווקים שאושפזו
מספר הנשואים הכולל שאושפזו במדגם.	813	מספר הנשואים שאושפזו
מספר הגרושים הכולל שאושפזו במדגם.	744	מספר הגרושים שאושפזו
מספר האלמנים הכולל שאושפזו במדגם.	171	מספר האלמנים שאושפזו
כמות הרווקים שחזרו לאשפוז לפחות פעם אחת.	932 (38.3%)	מספר הרווקים שחזרו לאשפוז
כמות הנשואים שחזרו לאשפוז לפחות פעם אחת.	265 (32.6%)	מספר הנשואים שחזרו לאשפוז
כמות הגרושים שחזרו לאשפוז לפחות פעם אחת.	266 (35.8%)	מספר הגרושים שחזרו לאשפוז
כמות האלמנים שחזרו לאשפוז לפחות פעם אחת.	37 (21.6%)	מספר האלמנים שחזרו לאשפוז

Comparison of Total and 1-Month Readmitted Patients by Marital Status

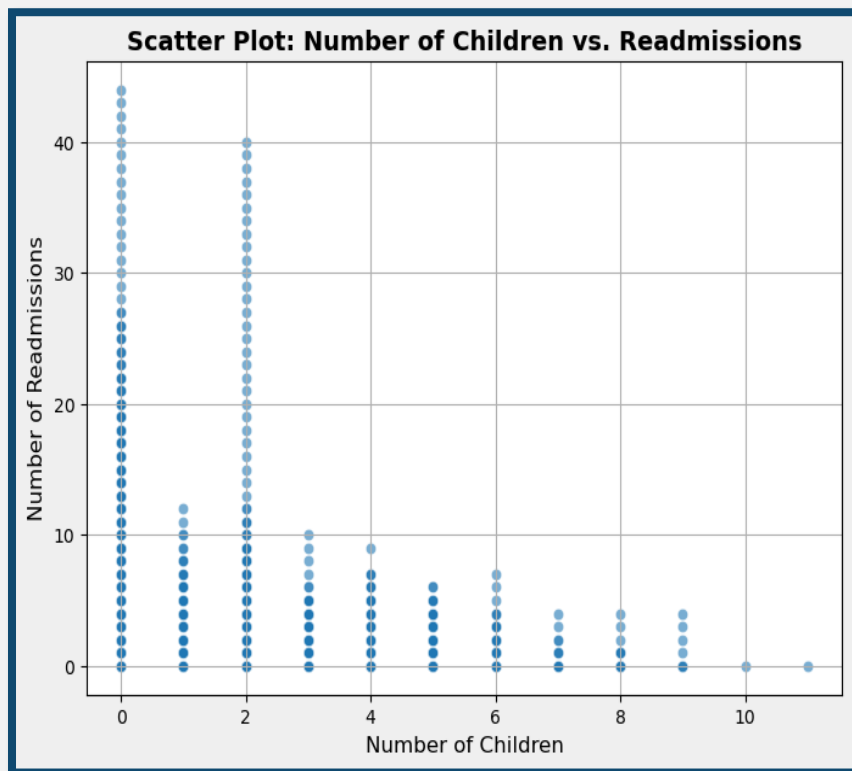


פרשנות	תוצאה מספרית	מדד סטטיסטי
נמצא הבדל בין מצבים משפחתיים לחזרה לאשפוז חוזר אחרי חודש.	9.54	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך קטן מ-0.05, ולכן קיימת מובהקות סטטיסטית.	0.02	ערך קרמט מובהקות
מספר הרווקים הכולל שאושפזו במדגם.	2,434	מספר הרווקים שאושפזו
מספר הנשואים הכולל שאושפזו במדגם.	813	מספר הנשואים שאושפזו
מספר הגרושים הכולל שאושפזו במדגם.	744	מספר הגרושים שאושפזו
מספר האלמנים הכולל שאושפזו במדגם.	171	מספר האלמנים שאושפזו
כמות הרווקים שחזרו לאשפוז לאחר כחודש.	467 (19.2%)	מספר הרווקים שחזרו לאשפוז
כמות הנשואים שחזרו לאשפוז לאחר כחודש.	124 (15.3%)	מספר הנשואים שחזרו לאשפוז
כמות הגרושים שחזרו לאשפוז לאחר כחודש.	125 (16.8%)	מספר הגרושים שחזרו לאשפוז
כמות האלמנים שחזרו לאשפוז לאחר כחודש.	23 (13.5%)	מספר האלמנים שחזרו לאשפוז

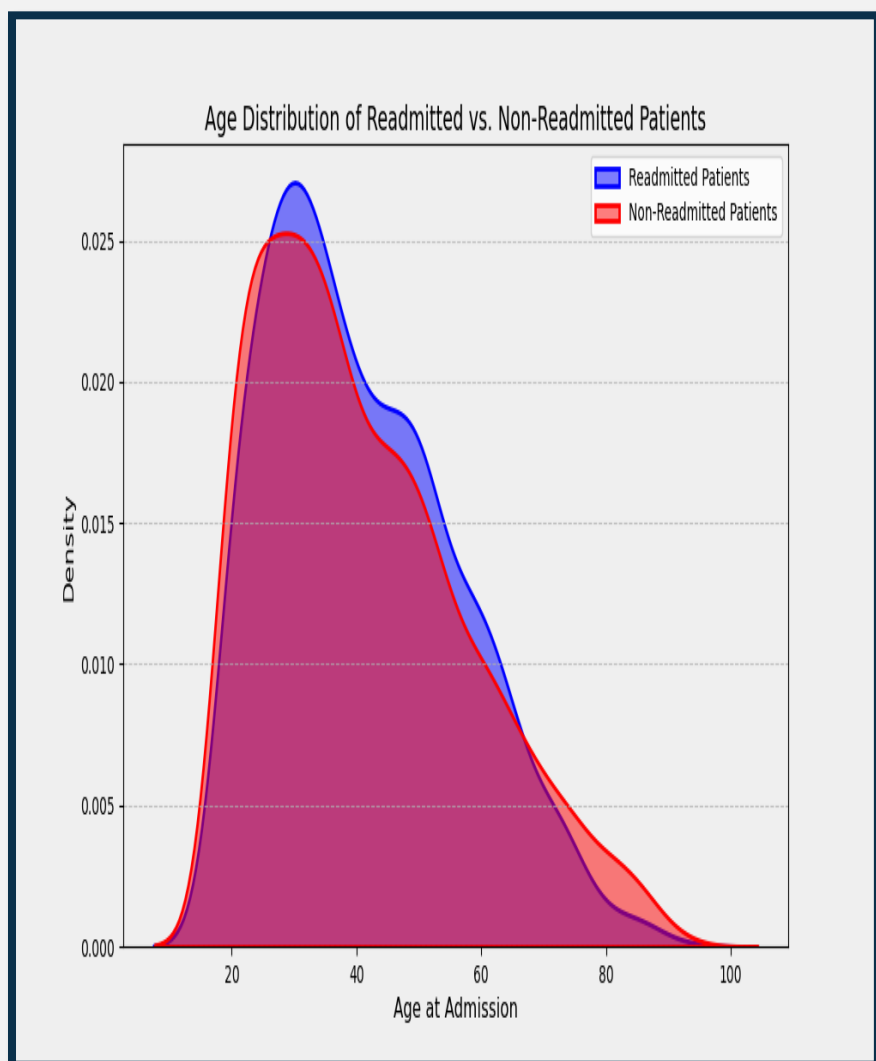
Comparison of Total and 3-Month Readmitted Patients by Marital Status



פרשנות	תוצאה מספרית	מדד סטטיסטי
נמצא הבדל מובהק מצב משפחתי לחזרה לאשפוז חוזר.	11.73	סטטיסטי חי-בריבוע ( $\chi^2$ )
ערך ק קטן מ-0.05, ולכן קיימת מובהקות סטטיסטית.	0.01	ערך ק רמת מובהקות
מספר הרווקים הכולל שאושפזו במדגם.	2,434	מספר הרווקים שאושפזו
מספר הנשואים הכולל שאושפזו במדגם.	813	מספר הנשואים שאושפזו
מספר הגרושים הכולל שאושפזו במדגם.	744	מספר הגרושים שאושפזו
מספר האלמנים הכולל שאושפזו במדגם.	171	מספר האלמנים שאושפזו
כמות הרווקים שחזרו לאשפוז לאחר כ 3 חודשים.	600 (24.7%)	מספר הרווקים שחזרו לאשפוז
כמות הנשואים שחזרו לאשפוז לאחר כ 3 חודשים.	165 (20.3%)	מספר הנשואים שחזרו לאשפוז
כמות הגרושים שחזרו לאשפוז לאחר כ 3 חודשים.	164 (22%)	מספר הגרושים שחזרו לאשפוז
כמות האלמנים שחזרו לאשפוז לאחר כ 3 חודשים.	28 (13.5%)	מספר האלמנים שחזרו לאשפוז



פרשנות	תוצאה מספרית	מדד סטטיסטי
קיים קשר שלילי חלש מאוד בין מספר הילדים לבין הסבירות לחזרה לאשפוז.	-0.08	מתאם ספירמן (r)
ערך קטן מ-0.05, ולכן קיימת מובהקות סטטיסטית.	$< 0.01$ קטן מ	ערך ק רמת מובהקות



פרשנות	תוצאה מספרית	מדד סטטיסטי
Mann-Whitney U מבחן-בוצע להשוואת גיל אשפוז בין מטופלים שחזרו לאשפוז לבין אלו שלא. הערך הגבוה נובע מגודל המדגם הגדול.	2,027,676.00	Mann-Whitney U
ערך ק גדול מ-0.05, ולכן אין הבדל מובהק סטטיסטית	0.51	ערך ק רמת מובהקות

## דיון ומסקנות:

ממצאי המחקר מספקים תובנות חשובות על הגורמים המשפיעים על חזרות לאשפוזים בקרב מטופלים פסיכיאטריים. במחקר זה נבחנו משתנים דמוגרפיים וקליניים מרכזיים והשפעתם על שיעור האשפוזים החוזרים, תוך שימוש בניתוחים סטטיסטיים מתקדמים. התוצאות מדגימות כי ישנם מאפיינים דמוגרפיים מסוימים המשפיעים על הסיכון לחזרה לאשפוז, אך לא בכל המקרים נמצאה מובהקות סטטיסטית, מה שמעורר עניין ומחייב חקירה מעמיקה יותר.

ניתוח ההבדלים המגדריים באשפוזים חוזרים מצביע על כך ששיעור האשפוזים בקרב גברים גבוה מהצפוי ביחס לשיעורם באוכלוסייה, בעוד ששיעור האשפוזים בקרב נשים נמוך מהצפוי. עם זאת, לא נמצאה מובהקות סטטיסטית בין מגדר לבין חזרה לאשפוז תוך חודש או שלושה חודשים. ממצא זה מעיד על כך שאף על פי שגברים מאושפדים בשיעור גבוה יותר, הסיכון שלהם לחזור לאשפוז חוזר אינו שונה באופן מובהק מזה של נשים. עובדה זו יכולה להצביע על כך שהבדלים המגדריים באשפוז הראשוני אינם מתרגמים ישירות להבדלים בסיכויי החזרה לאשפוז, דבר שדורש ניתוח מעמיק יותר של משתנים חיצוניים כגון סוג האבחנה הפסיכיאטרית או גישה לשירותי תמיכה בקהילה.

בהקשר של המצב המשפחתי, נמצא כי מטופלים רווקים, גרושים ואלמנים מצויים בסיכון גבוה יותר לחזרה לאשפוז חוזר בהשוואה לנשואים. ממצא זה עולה בקנה אחד עם מחקרים קודמים שהצביעו על כך שתמיכה משפחתית משחקת תפקיד מרכזי בשיקום נפשי ומקטינה את הסיכון לחזרה לאשפוזים. הבדל זה עשוי להעיד על כך שמערכת תמיכה רחבה מהווה גורם מגן משמעותי, המסייע ליציבות רגשית ויכולת הסתגלות טובה יותר לאחר האשפוז. יחד עם זאת, נדרשת בחינה מעמיקה יותר של האינטראקציה בין משתנים דמוגרפיים שונים, כמו גם השפעתם של גורמים מערכתיים דוגמת גישה לשירותי בריאות נפש, זמינות מטופלים בקהילה והשתייכות למסגרות תמיכה.

ניתוח הקשרים בין מספר הילדים לבין חזרה לאשפוז הצביע על מגמות מעניינות אך ללא מובהקות סטטיסטית גבוהה. אף על פי שההשערה הראשונית הייתה שמספר הילדים ישפיע על חזרה לאשפוז, נראה כי גורמים נוספים כמו תמיכה משפחתית רחבה יותר, מצב כלכלי וגורמים פסיכו-סוציאליים עשויים להיות רלוונטיים יותר להבנת הקשר. ייתכן כי השפעת מספר הילדים היא עקיפה, ונובעת ממשתנים אחרים שאינם נמדדו במחקר הנוכחי, כגון מעורבות בני המשפחה בחיי המטופל, עוצמת הקשרים המשפחתיים או תחושת הבדידות והבידוד החברתי.

בהתבסס על הממצאים, נרצה להמשיך ולחקור היבטים נוספים שיוכלו להעמיק את הבנתנו בנוגע לגורמים המשפיעים על חזרות לאשפוזים פסיכיאטריים. שילוב של מידע טקסטואלי מתיקים רפואיים באמצעות שיטות עיבוד שפה טבעית (NLP) עשוי לחדד את יכולת הניבוי ולאפשר איתור דפוסים חבויים במידע הקליני. כמו כן, בחינת גורמים סביבתיים נוספים, כגון מצב סוציו-אקונומי, מצב



חברתי, סטטוס משפחתי והשפעתם על שיעורי האשפוזים החוזרים, תאפשר הבנה רחבה יותר של המניעים המערכתיים המשפיעים על האשפוזים. לצד זאת, יש צורך בבחינת גורמים קליניים כגון סוג התרופות שניתנו למטופלים, שיטות טיפול, אופן ומשך הטיפול והשפעתם על שיעורי החזרה לאשפוזים, שכן משתנים אלו עשויים להיות מכריעים בקביעת הצלחת ההתערבות הטיפולית.

לבסוף, בשלב הבא של המחקר יבוצע שימוש במודלים של סיווג (Classification Models) כדי לשפר את יכולת החיזוי של חזרות לאשפוז. נרצה להשתמש במודלים מתקדמים של למידת מכונה כגון Random Forest, Logistic Regression, אשר מאפשרים לזהות תבניות מורכבות במידע הקליני והדמוגרפי. שילוב טכניקות כגון בחירת תכונות (Feature Selection) והנדסת תכונות (Feature Engineering) יכולים להעצים את המודלים ולהתאים אותם טוב יותר למציאות הקלינית. פיתוח מערכות חיזוי מתקדמות מסוג זה עשוי לשפר את יכולת הזיהוי המוקדם של מטופלים בסיכון גבוה לאשפוז חוזר, ולאפשר יישום תוכניות מניעה מותאמות אישית שיתרמו להפחתת עומסי האשפוזים ולשיפור איכות חייהם של מטופלים רבים.

## בילביוגרפיה:

1. Frick, U., Frick, H., Langguth, B., Landgrebe, M., Hübner-Liebermann, B., & Hajak, G. (2013). *The revolving door phenomenon revisited: Time to readmission in 17,415 patients with 37,697 hospitalisations at a German psychiatric hospital*. PLoS ONE, 8(10), e75612.
2. Lay, B., Kawohl, W., & Rössler, W. (2019). *Predictors of compulsory re-admission to psychiatric inpatient care*. Frontiers in Psychiatry, 10, 120.
3. Beer Sheva Mental Health Center. (2023). *A follow-up on factors associated with readmission of patients in Beer Sheva Mental Health Center during 2015-2018*. Harefuah, 162(9-10), 513-517.