

## שיטות שונות לניתוח מערכי נתונים גיאוגרפיים במחקר אפידמיולוגי

עופר שטיינברג

בהנחיית: פרופ' בוריס א. פורטנוב, הפקולטה לניהול, אוניברסיטת חיפה

פרופ' משנה חאלד כרכבי, שירותי בריאות כללית והפקולטה לרפואה, טכניון

### תקציר

נתוני תחלואה במחקר אפידמיולוגי מייצגים לעיתים קרובות אזורים ללא תכונות נוספות. השיטות סטטיסטיות אשר זמינות לניתוח מערכי נתונים גיאוגרפיים הינן מוגבלות. במחקר זה אנו משווים מספר שיטות אנליטיות לניתוח נתוני תחלואת אסטמה בקרב ילדים המתגוררים באזור חיפה רבתי במדינת ישראל. השיטות בהן נשתמש להשוואה זו הינן: רגרסיה לינארית של אומדן אזורי, רגרסיה לוגיסטית, מודל ממוצע בייסיני, רגרסיה פואסונית, וכן כלי אנליטי מתקדם יותר, "קרנל כפול". מטרת מחקר זה הינה לבחון ולקבוע אילו מבין הגישות הנבחנות יעזרו לנו לקבל מידע נוסף אודות גורמים סביבתיים המשפיעים על דפוסי תחלואה מקומיים.

הנתונים התקבלו מקופת חולים "כללית" ומכילים כ-24,115 רשומות אודות ילדים בגילאי 6-14 (%2.62 הינם אסטמתיים). כלל הנתונים מופו בעזרת תוכנת ArcGIS, כאשר כל מיקום גיאוגרפי קושר למשתנים סוציו-אקונומיים על פי אזור המגורים, נתוני זיהום האוויר, קרבה לאזורי תעשייה, כבישים ועוד. לאחר מכן, נבחן הקשר בין היארעות אסטמה לבין משתנים אלה בכל אחת מהשיטות.

תוצאות הניתוחים הסטטיסטיים מציגים כי בעוד שמודל הרגרסיה הלוגיסטית ומודל הרגרסיה הלינארית של אומדן אזורי אינם מובהקים ( $p > 0.1$ ), שיטת ה-"קרנל כפול" הינה מובהקת ועוזרת לזהות מספר גורמים סביבתיים להסבר דפוסי תחלואת האסטמה. כמו כן, מודלי הרגרסיה הפואסונית והבייסינית נמצאו מובהקים גם כן, אם כי פחות ממודל ה-"קרנל הכפול".

על אף מובהקות מודל ה-"קרנל הכפול", שיטה זו הינה רגישה ביחס ליתר השיטות הסטטיסטיות שנבחנו במחקר זה, על ידי העובדה כי בשיטה זו אין תלות של תיחום מוגדר מראש של האזורים הגיאוגרפיים המסייע ליצור שכבה מתמשכת של סיכון לתחלואה המכסה את אזור המחקר כולו. אנו ממליצים לעשות שימוש בשיטה זו לניתוח בסיסי נתונים מסוג זה, בהם ישנם מספר מועט של אזורים סטטיסטיים לניתוח ונתוני התחלואה בהם נדירים למדי.