מילת אזהרה, על חלוקת משתנה רציף, לקטגוריות

חלוקת נתונים רציפים, כמו, משקל, לקטגוריות, נפוצה במחקרים רפואיים.

הסיבה המובילה לחלוקה זו, היא: הצורך בתצוגה פשוטה של תוצאות המחקר.

חלוקה לא זהירה של משתנה רציף לקטגוריות, עלולה לגרום להסקה מוטעית מהנתונים.

בטור הרצ"ב, של הוורד ויינר, מודגם כיצד חלוקות שונות של אותו משתנה רציף, מובילות למסקנות סותרות, על הקשר בין שני משתנים רציפים: עבור חלוקה מסוימת נמצא קשר חיובי בין המשתנים ועבור חלוקה אחרת נמצא קשר שלילי.

אם יש צורך לחלק משתנה רציף לקטגוריות, חובה לקבוע מהן הקטגוריות, לפני שלב איסוף הנתונים, ועל פי שיקולים קליניים; כמו למשל, חלוקת BMI לקבוצות של: רזון יתר, תקין, השמנה קלה והשמנה חולנית.

קביעת הקטגוריות על פי נתוני המחקר, כמו למשל חלוקה רבעונית, עלולה להוביל לטעויות בניתוח הסטטיסטי.

ניתן למצוא את כל הפינות הסטטיסטיות ב:

<http://homenew.clalit.org.il/SITES/HOSPITALS/RMC/BEILINSON/STAT/Pages/Desktop800.aspx>

בברכה