Utilizing the time-interval dependence of odds and risk ratios to analyze the correlation of hydrochlorothiazide and statins usage with the risk of skin cancer

Ethnan Weinstock, Rudi Chen, Tingting Xu, and Ioannis Ch. Psachalidis

Bridgewater-Raritan High School, 600 Garretson Road, Bridgewater, New Jersey 08807

Boston University Division of Systems Engineering, 8 St. Mary’s Street, Boston, Massachusetts 02215

Objective

Primary Objective: Quantify the relationship between hydrochlorothiazide and statins usage and the risk of a diagnosis of nonmelanoma skin cancer among Americans.

Secondary Objective: Use the time-interval dependency of odds ratios and risk ratios to further analyze the relationship.

Key Terms

- Nonmelanoma skin cancer (NMSC): the most common form of cancer in the world.
- Basal cell carcinoma (BCC): most common subtype.
- Squamous cell carcinoma (SCC): second most common subtype.
- Hydrochlorothiazide (HCTZ): drug that treats high blood pressure, heart failure, and edema. Used by 10+ million Americans annually.
- Statins: cholesterol-lowering drug class reducing risk of heart attack and stroke. Used by 25+ million Americans annually.
- Odds ratio (OR): a quantitative measure of correlation obtained from a case-control study.
- Risk ratio (RR): a quantitative measure of causation obtained from a cohort study.
- Case-control/cohort study: two methods of determining the relationship between a risk factor and an outcome.

Introduction

Correlation of Hydrochlorothiazide and Statins Usage with the Risk of Skin Cancer

What do the results mean in or prior to 2009 and a NMSC diagnosis? What is it 2+ years before the diagnosis? What are the odds/risk that a HCTZ or statin user receives an NMSC diagnosis 2-5 years later? What do the results mean 5 years later.

In the “population-level attributable proportion” (AP), most of the increased risk of NMSC in the general population can be attributed to statin usage. • 7.5% of BCC diagnoses from 2011-2014 (1.3 million out of 17.28 million) can be attributed to statin usage.

To the American Heart Association recommending that 70+ million Americans use statins (up from 36 million), the results of this project imply a potential negative risk for an increased number of NMSC diagnoses among Americans.

The “population-level attributable proportion” (AP) measures how much of increased risk of NMSC has been increased by the use of statins. It is used to estimate the impact of a risk factor on the occurrence of a disease in a population.

We conclude that ORs and RRs have a time dependency. The formula for conditional logistic regression is:

\[
\text{OR} = \frac{\text{probability of NMSC in case} \times \text{probability of statin use in control}}{\text{probability of NMSC in control} \times \text{probability of statin use in case}}
\]

We create an increased risk for BCC with Statin Usage Over Time. Increased Risk for BCC with Statin Usage Over Time

We conclude what is the risk ratio for the 2 models are largely supported based on statistically-significant Odds and RRs. In some situations, case-control and cohort studies are similar, and in others they are significantly different.

We would like to thank Prof. Psachalidis for introducing me to the exciting world of university research and his work with healthcare and machine learning; Rudi Chen and Tingting Xu for teaching me the skills I needed for this project and helping me along the way; Dr. G. Bastani from BUC for providing insight into how cancer diagnoses are recorded; and BU and Dr. Martin for organizing the RISE program.

Methods and Hypotheses

Time-Interval Dependency

ORs and RRs have a time interval dependency that comes from the method of calculation, conditional logistic regression. The formula for conditional logistic regression is:

\[
P(Y|X) = \frac{1}{1 + e^{-\beta X}}
\]

The model is based on the conditional probability formula which is defined by Brookes et al. [12]

\[P(Y|X) = \frac{e^{\beta X}}{1 + e^{\beta X}}\]

What is it (OR)? measures the relationship between the incidence of NMSC and HCTZ/statins usage in the general population.

What is the hypothesis? The resulting ORs/Rrs should be relatively constant over the different intervals (within the 95% confidence interval [CI] of the largest time interval).

What do the results mean? the ratio of the odds(OR)/risk(RR) that someone with an NMSC diagnosis had taken HCTZ or statins 2 years before the diagnosis to the odds/risk that someone with an NMSC diagnosis had not taken HCTZ or statins.

What is it (RR)? measures the relationship between usage of statins in or prior to 2009 and a NMSC diagnosis 2-5 years later.

What is the hypothesis? the resulting ORs/Rrs should increase over time because the cumulative dosage would increase.

What do the results mean? the ratio of the odds(OR)/risk(RR) that a HCTZ or statin user receives an NMSC diagnosis 2-5 years later to the odds/risk that a HCTZ or statin user does not receive an NMSC diagnosis 2-5 years later.

Conclusions

The main limitation of this study was the small dataset, consisting of only 900 NMSC diagnoses. The small dataset is a possible explanation for: • lack of correlation of HCTZ with NMSC. • the number of statistically-significant Odds (14/40) and RRs (11/40).

Future work would include: • Analyzing the correlation of HCTZ and statins with NMSC using a bigger dataset. • Testing the hypothesis on other datasets to determine validity.

Acknowledgements

I would like to thank Prof. Psachalidis for introducing me to the exciting world of university research and his work with healthcare and machine learning; Rudi Chen and Tingting Xu for teaching me the skills I needed for this project and helping me along the way; Dr. G. Bastani from BUC for providing insight into how cancer diagnoses are recorded; and BU and Dr. Martin for organizing the RISE program.

References