Applicability of logistic regression with sum-to-zero constraint parameterization in risk assessment for parotid malignancy

CURRENT STATUS: POSTED

Research Square

Shari Messinger Cayetano
University of Miami School of Medicine

smessinger@biostat.med.miami.edu*Corresponding Author*

Kaming Lo
PPD

ORCiD: https://orcid.org/0000-0002-8897-1818

Christopher Fundakowski
Temple University

Zoukaa Sargi
University of Miami School of Medicine

DOI:
10.21203/rs.2.16732/v1

SUBJECT AREAS
Health Economics & Outcomes Research

KEYWORDS
logistic regression, risk assessment, parotid malignancy

Abstract

Background Investigative interest is often to determine how results from a diagnostic tool change the patient’s risk of disease with respect to the overall(naïve) risk at clinical presentation. Logistic regression is popular for data analysis for this type of investigation. However, standard approach, which uses reference cell coding, may not be informative in this setting. This is because this approach compares the risk between two distinct groups.

Methods We considered weighted and unweighted approaches to model parameterization using deviation from means coding for assessing the risk of parotid malignancy, comparing patients with indeterminate fine-needle aspiration biopsy(FNAB) results with the general(naïve) risk among all presenting patients. Results from deviation from means coding and standard reference cell coding were compared.

Results Unweighted coding estimates a two-fold increase in the odds of malignancy with an indeterminate FNAB result compared to the naïve odds at clinical presentation (Odds ratio(OR): 1.97 [95% Confidence Interval(CI): 1.34–2.90], P=0.0006). The weighted approach estimates increased risk (OR: 2.38 [95% CI: 1.45 – 3.89], P=0.0006), more accurately representing the naïve risk at presentation based on the direction of sample imbalance in the study. Using standard reference cell coding, an indeterminate result has a higher risk compared to a negative result, but this does not inform us about the risk with respect to that inherent at clinical presentation.

Conclusions Depending on the investigative interest, it is important to adopt the appropriate coding methodology when logistic regression is applied. In addition, a weighted approach should be considered to account for sample imbalance.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.