ehp Environmental Health PERSPECTIVES
ehponline.org
Published by the National Institute of Environmental Health Sciences

Epidemiologic Evaluation of Measurement Data in the Presence of Detection Limits

Jay H. Lubin, Joanne S. Colt, David Camann, Scott Davis, James R. Cerhan, Richard K. Severson, Leslie Bernstein, and Patricia Hartge

doi:10.1289/ehp.7199 (available at http://dx.doi.org/)
Online 13 September 2004

The National Institute of Environmental Health Sciences
National Institutes of Health
U.S. Department of Health and Human Services

Epidemiologic Evaluation of Measurement Data in the Presence of Detection Limits

Jay H. Lubin [1], Joanne S. Colt [1], David Camann [2], Scott Davis [3], James R. Cerhan [4]
Richard K. Severson [5], Leslie Bernstein [6], Patricia Hartge [1]

[1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD
[2] Southwest Research Institute, San Antonio, TX
[3] Fred Hutchinson Cancer Research Center and the University of Washington, Seattle, WA
[4] Mayo Clinic, College of Medicine, Rochester, MN
[5] Karmanos Cancer Institute and Department of Family Medicine, Wayne State University, Detroit, MI
[6] Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School at Medicine at the University of Southern California, Los Angeles, CA

Corresponding author:

Jay Lubin, National Cancer Institute, Biostatistics Branch, 6120 Executive Boulevard, Room 8042, Rockville, MD 20852

Phone number: (301) 496-3357
FAX number: (301) 402-0081
E-mail: lubinj@mail.nih.gov.

Running title: Measurement data with detection limits

Keywords: pesticides, dust, environmental exposure, imputation, missing data, non-Hodgkin lymphoma

Acknowledgment and support:
The authors declare they have no competing financial interest. Support for this study included contracts with the National Cancer Institute: N01-PC-67010, N01-PC-67008, N02-PC-71105, N01-PC-67009, N01-PC-65064.

Abbreviations:

AM	Arithmetic mean
CI	Confidence interval
DL	Detection limit
GC/MS	Gas chromatography/mas spectrometry
GM	Geometric mean
GSD	Geometric standard deviation
LB	Lower bound
MLE	Maximum likelihood estimate
ng/g	Nano-grams per gram
NHL	Non-Hodgkin lymphoma
SE	Standard error
UB	Upper bound

Outline of section headings:

Abstract

Introduction

Example data from a case-control study of NHL and pesticides

Data source
Measurement of carpet dust

Methods and analysis

Regression analysis in control data
Imputation of missing concentrations with application to the fill-in method
Simulation study

Results

Regression analysis using control subjects
Simulation study

Discussion

References

Tables

Figure

Abstract

Quantitative measurements of environmental factors greatly improve the quality of epidemiologic studies, but can pose challenges due to the presence of upper or lower detection limits or interfering compounds, which do not allow for precise measured values. We consider the regression of an environmental measurement (dependent variable) on several covariates (independent variables). Various strategies are commonly employed to impute values for interval-measured data, including assignment of one-half the detection limit to non-detected values, or of "fill-in" values randomly selected from an appropriate distribution. Based on a limited simulation study, we found that the former approach can be biased, unless the percentage of measurements below detection limits is small (5-10 percent). The fill-in approach generally results in unbiased parameter estimates, but may produce biased variance estimates and thereby distort inference when 30 percent or more of the data are below detection limits. Truncated data methods (e.g., Tobit regression) and multiple imputation offer two unbiased approaches for analyzing measurement data with detection limits. If interest resides solely on regression parameters, then Tobit regression can be used. If individualized values for measurements below detection limits are needed for additional analysis, such as relative risk regression or graphical display, then multiple imputation produces unbiased estimates and nominal confidence intervals unless the proportion of missing data is extreme. We illustrate various approaches using measurements of pesticide residues in carpet dust in control subjects from a case-control study of non-Hodgkin lymphoma.