445

BIAS FROM MATCHING ON AGE AT DEATH OR CENSOR IN NESTED CASE-CONTROL STUDIES.

*M J Hein, J A Deddens, M K Schubauer-Berigan (National Institute for Occupational Safety and Health, Cincinnati, OH 45226)

Background: Incidence density sampling in nested case-control studies frequently matches controls to cases based on attained age because age is one of the most important risk factors for mortality. Additional matching criteria such as sex, race or year of birth have been used in order to reduce confounding. Recently it was suggested that an additional matching criterion be age at death or censor with eligible controls having an age at death or censor within a specified number of years of the case's age at death. Methods: We used simulated occupational cohorts to evaluate the potential for this additional matching criterion to bias the estimated hazard ratio (HR) from Cox proportional hazards regression. An additional objective was to investigate issues related to exposure lagging; therefore, we compared estimated HRs from analyses including and excluding lagged out workers. Results: The use of the additional matching criterion resulted in downwardly biased HR estimates. In these simulations, when risk was related to a lagged cumulative exposure, the estimated HRs from analyses including lagged out workers (assigned zero lagged cumulative exposure) were generally similarly or less biased than estimated HRs from analyses excluding lagged out workers. Conclusion: Incidence density sampling with matching based on attained age plus age at death or censor introduces bias and is not recommended for nested case-control studies. The observed bias was not unexpected since information used to select controls must be known at the time of selection. Since there can be no way of predicting the age of death or censor for the controls, this information cannot be used for selecting controls in incidence density sampling.

446-S

ANALYZING DATA SUBJECT TO A LIMIT OF DETECTION: COMMON PRACTICES IN CURRENT RESEARCH.

*L A Rosenthal, N J Perkins, B W Whitcomb, E F Schisterman (DESPR, NICHD, NIH, 6100 Executive Blvd., Bethesda, MD 20852)

Data in epidemiological studies are often subject to a limit of detection (LOD) due to instrument sensitivity. Affected datasets complicate data analysis and, as a result, often lead to biased results and conclusions. In this paper, the authors review common practices in epidemiological research to evaluate such censored data, including omission of and substitution for missing values as well as maximum likelihood estimations where appropriate. The authors investigated the effect of the different methods with respect to estimation of mean and variance; coefficients in both linear and logistic regressions were considered as well. Data for estrogen levels from the BioCycle study and interleukin-6 levels from a study of cytokine levels in women with and without spontaneous abortion were used for analysis. Through implementation of these methods with the two data sets possessing low (15.2% missing) and moderate (31.2% missing) degrees of censoring, the authors compared estimates of mean and variance of biomarkers and regression coefficients under each approach to handling censored data. Estimations of the mean and variance were subject to the method used. Similarly, beta coefficients in both linear and logistic regressions were substantially affected; with moderate censoring there was as much as a three-fold difference between estimates of the relation between IL-6 and birth weight. The examples demonstrate the degree to which the approach impacts the results of such analysis. The approach selected among these commonly used methods for handling data with an LOD strongly affects results, which could explain discrepant findings in the literature, and has implications for proper inference.

447-S

GENERALIZED ROC CURVE INFERENCE FOR A BIOMARKER SUBJECT TO A LIMIT OF DETECTION AND MEASUREMENT ERROR.

*N J Perkins, E F Schisterman, A Vexler (DESPR, NICHD, NIH, 6100 Executive Blvd., Bethesda, MD 20852)

The receiver operating characteristic (ROC) curve is a tool commonly used to evaluate biomarker utility in clinical diagnosis of disease, especially during biomarker development research. Assessment of emerging biomarkers is often hindered by random measurement error and limits of detection, masking discriminatory ability by negatively biasing the estimates of ROC curves and subsequent area under the curve. Methods have been developed to correct the ROC curve for each of these sources of bias but here we develop a method by which the ROC curve is corrected for both simultaneously through replicate measures and maximum likelihood. Our assertions of asymptotic unbiasness and nominal coverage probability of our proposed point and confidence interval estimates, respectively, for the generalized ROC criterion are confirmed via simulation study. As motivation, these methods are applied to two potential discriminators of women with and without preeclampsia, demonstrating the positive effects of properly planning and correcting for a limit of detection and measurement error during biomarker development.

448

FLEXIBLE, MULTIVARIABLE MODELS TO ESTIMATE COMPLETE LIFE TABLES.

*B Rachet, C Maringe, L Woods, L Ellis (London School of Hygiene & Tropical Medicine, London, UK, WC1E 7HT)

Complete life tables, specific to geographic area, socioeconomic status or ethnicity, are useful tools in demography and public health, and essential in relative survival analysis, which uses the background mortality provided by those life tables. The CONCORD study recently demonstrated the noticeable impact on cancer survival estimates in the USA of using State- and ethnicity-specific life tables instead of national life tables. Such life tables are often based on small populations, leading to unstable background mortality estimates. Several approaches have been proposed to smooth these estimates, including the widely used four-parameter relational model and the Elandt-Johnson method. We developed two new multivariable approaches which exploit the flexibility of spline functions to model the age-specific mortality rates by various factors. We compared our models with the conventional approaches, using both real and simulated data. The goodness-of-fit of our models was comparable to the best conventional approach with large populations. With populations less than half a million, overall and age-specific goodness-of-fit both improved, and the cancer survival estimates were more accurate. The increasing interest in cancer survival for more tightly defined populations broadens the utility of such models.

ISSN 0002-9262
PRINTED IN THE U.S.A.

Supplement to:

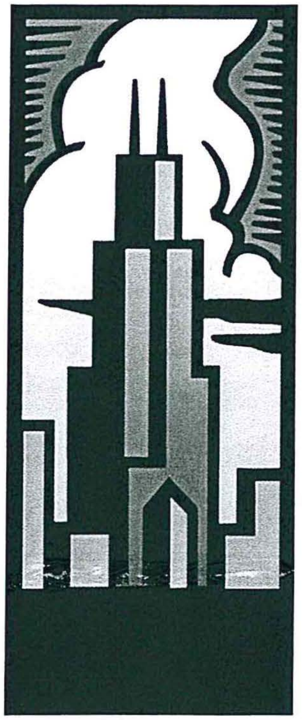
American Journal of Epidemiology

Volume 167 Number 11 June 1, 2008
www.aje.oxfordjournals.org

Abstracts of the 41st Annual Meeting

Society for Epidemiologic Research

Chicago, Illinois, June 24–27, 2008



Published for the Johns Hopkins Bloomberg School of Public Health by Oxford University Press Sponsored by the Society for Epidemiologic Research

Founded in 1920 by W. H. Welch and W. H. Howell as the *American Journal of Hygiene* at the Johns Hopkins School of Hygiene and Public Health

OXFORD OPEN

OXFORD JOURNALS
OXFORD UNIVERSITY PRESS