

Case-control studies are widely used to detect gene-environment interactions in the etiology of complex diseases. Many variables that are of interest to biomedical researchers are difficult to measure at the individual level, e.g. nutrient intake, cigarette smoking exposure, long-term toxic exposure. Measurement error may cause bias in parameter estimates, thus masking key features of data. We develop a Bayesian methodology for analysis of case-control studies for the case when measurement error is present in an environmental covariate and the genetic variable has missing data. This approach offers several advantages. It allows prior information to enter the model to make estimation and inference more precise. The environmental covariates measured exactly are modeled completely nonparametrically. Further, information about the probability of disease can be incorporated in the estimation procedure to improve quality of parameter estimates. A unique feature of the procedure under investigation is that the analysis is based on a pseudo-likelihood function therefore conventional Bayesian techniques may not be technically correct. We propose an approach based on Markov Chain Monte Carlo sampling as well as a computationally simple method based on an asymptotic posterior distribution. Simulation experiments demonstrated that our method produced parameter estimates that are nearly unbiased even for small sample sizes. An application of our method is illustrated using a population-based case-control study of the association between calcium intake with the risk of colorectal adenoma development.