

Outcome-Dependent Sampling and Its Application to Environmental Health Studies

Haibo Zhou, PhD^{1,2} and David B. Peden, MD^{2,3,4}

The Department of Biostatistics¹, the School of Public Health, the Center for Environmental Medicine, Asthma and Lung Biology², the Departments of Pediatrics³ and Medicine⁴, the School of Medicine, the University of North Carolina at Chapel Hill, Chapel Hill, NC, U.S.A.

Outcome-dependent sampling (ODS) is a cost effective way to enhance study efficiency. The case-control design for binary outcomes is a mainstay of epidemiology research. As the field of epidemiology expanding and evolving, an increasing number of studies are conducted using the ODS design with a “continuous” outcome. In an ODS design, observations made on a judiciously chosen subset of the base population can provide nearly the same statistical efficiency as observing the entire base population. Different statistical inference procedures are needed in order to reap the benefits of such sampling. We discuss recently developed inference procedures that account for the ODS design and illustrate the method with a data set from environmental health study on evaluating the influence of the C-159T single nucleotide polymorphism of the CD14 gene promoter on lung function in smokers with chronic bronchitis.