Estimation of Attributable Number of Deaths and Standard Errors from Simple and Complex Sampled Cohorts

Barry I. Graubard, Mitchell H. Gail, Katherine M. Flegal, David F. Williamson

Abstract

Estimates of the attributable number of deaths (AD) from all-causes can be obtained by first estimating population attributable risk (AR) adjusted for confounding covariates, and then multiplying the AR by the number of deaths determined from vital mortality statistics that occurred for a specific time period. Proportional hazard regression estimates of adjusted relative hazards obtained from mortality follow-up data from a cohort or a survey is combined with a joint distribution of risk factor and confounding covariates to compute an adjusted AR. Two estimators of adjusted AR are examined, which differ according to the reference population that the joint distribution of risk factor and confounders is obtained. The two types of reference populations considered: (i) the population that is represented by the baseline cohort and (ii) a population that is external to the cohort. Methods based on influence function theory are applied to obtain expressions for estimating the variance of the AD estimator. These variance estimators can be applied to data that range from simple random samples to (sample) weighted multi-stage stratified cluster samples from national household surveys. The variance estimation of AD is illustrated in an analysis of excess deaths due to having a non-ideal body mass index using data from the second National Health and Examination Survey (NHANES) Mortality Study and the 1999-2002 NHANES. These methods can also be used to estimate the attributable number of cause-specific deaths or incident cases of a disease and their standard errors when the time period for the accrual of is short.

1 National Cancer Institute, Bethesda, MD
2 National Center for Health Statistics, Hyattsville, MD
3 Centers for Disease Control and Prevention, Atlanta, GA, USA