The Utility of Prediction Models to Oversample the Long Term Uninsured

Steven B. Cohen and William Yu

Agency for Healthcare Research and Quality, Rockville MD, U.S.A.

Estimates of the health insurance status of the U.S. civilian population are critical to policymakers and others concerned with access to medical care and the cost and quality of that care. Health insurance helps people get timely access to medical care and protects them against the risk of expensive and unanticipated medical events. When estimating the size of the uninsured population, it is important to consider the distinction between those uninsured for short periods of time and those who are uninsured for several years. Given the risk of exposure to high out of pocket medical expenditures faced by the long term uninsured and associated economic and health related consequences, this population subgroup is of particular relevance to health policy considerations. Consequently, a prediction model that can accurately identify the long term uninsured is an important analytical tool. These models have particular relevance as statistical tools to facilitate efficient sampling strategies that permit the selection of an over-sample of individuals likely to be uninsured for long periods of duration in the future. This study provides a summary of the development of prediction models to identify the long term uninsured under age 65 and includes an evaluation of its potential utility as an oversampling strategy for use in a national longitudinal medical care expenditure survey. In addition, this type of modeling effort enhances the ability to discern the causes of high health care expenses and the characteristics of the individuals who incur them. This feature also applies to prediction models that can accurately identify those individuals with persistently low or average levels of expenditures.