

Impact of Record-linkage Methodology on Performance Indicators and Multivariate Relationships

Kevin M. Campbell, DrPH

Abstract

Program evaluation often requires the linkage of records from independently maintained data systems (e.g., substance abuse treatment and criminal justice). Data entry errors (e.g., misspelled names, transposed digits) complicate the linkage task. In this investigation, three record-linkage algorithms (match–merge, common patient identifier, and probabilistic) are used to link recipients of publicly funded outpatient substance abuse treatment to statewide arrest and death data. The impact of record-linkage algorithm on performance indicators, prevalence indicators (i.e., arrest rates, and death rates), and hazard ratios derived from a multivariate survival analysis predicting risk of arrest following admission to outpatient substance abuse treatment is evaluated. Choice of algorithm substantially impacted estimates of arrest rates (range: year prior to admission, 39.8%–53.4%; year following admission, 24.7%–33.1%). The hazard ratio associated with “prior arrest” as a predictor of arrest following admission to outpatient substance abuse treatment (hazard ratio range = 0.20–0.37, $p < .05$) was also influenced by algorithm choice.

To purchase full text articles, *click here* to visit the Journal of Substance Abuse Treatment website.

Note to Kyle: Document should have WC logo and be in PDF. Link is to [http://www.journalofsubstanceabusetreatment.com/article/S0740-5472\(08\)00089-5/abstract](http://www.journalofsubstanceabusetreatment.com/article/S0740-5472(08)00089-5/abstract)