

UC Berkeley

UC Berkeley Previously Published Works

Title

Conducting density-sampled case-control studies using survey data with complex sampling designs: A simulation study

Permalink

<https://escholarship.org/uc/item/1bx616tc>

Authors

Li, Catherine X
Matthay, Ellicott C
Rowe, Christopher
[et al.](#)

Publication Date

2022

DOI

10.1016/j.annepidem.2021.06.019

Peer reviewed



Published in final edited form as:

Ann Epidemiol. 2022 January ; 65: 109–115. doi:10.1016/j.annepidem.2021.06.019.

Conducting Density-Sampled Case-Control Studies using Survey Data with Complex Sampling Designs: A Simulation Study

Catherine X. Li^{a,b}, Ellicott C. Matthay^c, Christopher Rowe^a, Patrick T. Bradshaw^a, Jennifer Ahern^a

^aDivision of Epidemiology & Biostatistics, School of Public Health, University of California, Berkeley, Berkeley, CA

^bDepartment of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC

^cCenter for Health and Community, University of California, San Francisco, San Francisco, CA

Abstract

Purpose: Population-based surveys are possible sources from which to draw representative control data for case-control studies. However, these surveys involve complex sampling that could lead to biased estimates of measures of association if not properly accounted for in analyses. Approaches to incorporating complex-sampled controls in density-sampled case-control designs have not been examined.

Methods: We used a simulation study to evaluate the performance of different approaches to estimating incidence density ratios (IDR) from case-control studies with controls drawn from complex survey data using risk-set sampling. In simulated population data, we applied four survey sampling approaches, with varying survey sizes, and assessed the performance of four analysis methods for incorporating survey-based controls.

Results: Estimates of the IDR were unbiased for methods that conducted risk-set sampling with probability of selection proportional to survey weights. Estimates of the IDR were biased when sampling weights were not incorporated, or only included in regression modeling. The unbiased analysis methods performed comparably and produced estimates with variance comparable to biased methods. Variance increased and confidence interval coverage decreased as survey size decreased.

Conclusions: Unbiased estimates are obtainable in risk-set sampled case-control studies using controls drawn from complex survey data when weights are properly incorporated.

Keywords

Case-Control Studies; Density Sampling; Risk-Set Sampling; Complex Survey Sampling; Simulation

Corresponding Author: Jennifer Ahern, 2121 Berkeley Way #5201-2, Berkeley, CA 94720, Phone: (510) 643-4350, jahern@berkeley.edu.

Conflicts of Interest: Authors report no conflicts of interest.

INTRODUCTION

Growing availability of “big data,” including electronic medical records and disease surveillance data presents opportunities to analyze comprehensive data on health-related events and outcomes [1]. However, event and outcome data sources may not have a corresponding database from which to gather information about the source population from which cases arose. This poses analytic challenges in using available event and outcome data.

These challenges are particularly relevant to case-control studies [2–6], which require controls that represent the exposure distribution in the source population from which the cases arose [3,7–9]. Historically, methods such as random-digit dialing have been used to randomly sample from the identified source population [2]. However, use of these methods has declined due to low response rates and time and cost constraints [10–12]. The use of hospital- or clinic-drawn controls from increasingly available electronic medical record data may result in biased findings because they are unlikely to represent the exposure experience in the source population [13].

Established population-based surveys are promising sources from which to draw controls. These surveys are conducted frequently, are readily accessible, and collect data on a variety of potential exposures, outcomes, and covariates of interest. However, most of these surveys have complex sampling structures that deviate from simple random samples by using stratification, clustering, multi-stage or multi-phase designs, unequal probability sampling, or multi-frame sampling [14]. Examples of such surveys conducted in the United States include the American Community Survey (ACS) [15] and National Health and Nutrition Examination Survey (NHANES) [16]. Analyses using complex surveys that do not take sampling structure into account could produce biased estimates of measures of association and incorrect estimates of standard errors [17,18]. Several applied studies have combined case data with population-based survey data in both cumulative [19,20] and density [21] case-control analyses. These studies handled the survey weights in different ways, either by ignoring the weights completely [19], running weighted regression using survey weights [20], or expanding the survey dataset by the weights before sampling controls [21]. However, it is not known which of these approaches are valid or optimal.

Previous work on the inclusion of weights in case-control analyses to account for complex sampling of controls has focused primarily on the cumulative case-control design, where controls are sampled at the end of the study period and the case-control odds ratio approximates the cohort odds ratio [6]. In this realm, studies have examined trade-offs between methods that incorporate sampling weights and those that do not for a variety of complex sampling designs (stratified, cluster, multistage, etc.) [22–27]. Other work has addressed methods to increase statistical efficiency when using sampling weights [2,18,28,29]. Notably, work thus far indicates that in the presence of complex sampling of controls for cumulative case-control studies, unweighted methods cannot produce accurate measures of association, and ignoring the weights can lead to bias, underestimated variance, and low confidence interval (CI) coverage [18,26,27,30]. However, it is unclear how these issues translate to case-control studies apart from the cumulative design.

Research on valid methods for incorporating complex sampling weights into density sampled case-control studies is lacking. Density sampling is most commonly operationalized using risk-set sampling in epidemiology, where controls are sampled from those currently at risk at the time incident cases occur. This method of sampling controls facilitates approximation of the incidence density ratio (IDR) from the case-control odds ratio [6,9,31]. Design differences between cumulative and risk-set sampled case-control studies may limit the relevance of the existing literature. The purpose of this study was to use simulations to assess the bias and precision of several approaches to incorporating controls from complex survey data in risk-set sampled case-control studies.

METHODS

Simulated data

We created a “Total Population” dataset of $n = 1,000,000$ individuals based on the 2010 ACS of Californians aged 18 years and over to represent the source population of interest. For this Total Population, we simulated the exposure A and outcome Y from the set of ACS covariates W , comprised of categorical race/ethnicity, age group, sex, and education level. The exposure A was drawn from a Bernoulli distribution with the logit of the probability of exposure given by a linear function that included the covariates W . Event times for the terminal outcome Y were simulated from an exponential distribution with the logarithm of the rate given by a linear function of a subject’s exposure A , and covariate vector W .

We then simulated a ten-year study within the Total Population where all subjects with an event time of ten years or less were considered cases ($Y = 1$), and their simulated event time was recorded as their follow-up time. All subjects with an event time occurring more than ten years after baseline were considered non-cases ($Y = 0$) and their follow-up time was recorded as ten years. We assumed the exposure was defined at the beginning of follow-up and did not vary over time, and that induction time was 0, all subjects had perfect follow-up (no censoring), and that there were no competing risks for the outcome (i.e., the outcome was terminal). We also verified the relationship between the exposure and outcome was substantially confounded in unadjusted models. See Web Appendix 1 for the data-generating mechanisms, and Web Appendix 5 for statistical code.

Survey sampling

To simulate complex surveys, we sampled from non-cases using four different sampling designs (*Simple Random*, *Random Probability*, *Biased Probability*, and *Biased Stratified*) and four different survey sizes ($n = 1,000; 5,000; 15,000; 20,000$). In real-life applications, it may not be possible to confirm only non-cases were sampled into the survey. However, this is unlikely to be a major issue for analyses of rare outcomes. In sensitivity analyses, we relaxed this assumption and allowed complex surveys to sample from the Total Population (cases and non-cases).

Simple Random.—Each non-case from the Total Population had an equal probability of selection into the survey. Weights were the inverse of the probability of selection: n/s , where n is the number of non-cases in the Total Population, and s is the survey size.

Random Probability.—A known probability of selection (p) was randomly generated from a standard uniform distribution ($p \sim \text{Uniform}(0,1)$) for each non-case in the Total Population. Individuals were selected into the survey based on their probability of selection and assigned the survey weight $1/p$.

We expected analyses using controls drawn from the *Simple Random* and *Random Probability* sampling designs would produce unbiased estimates without accounting for survey weights because both result in controls that are randomly sampled without regard for the exposure or outcome. These sampling designs were used in comparison to sampling designs that we expected to generate bias.

Biased Probability.—Non-cases with the exposure were sampled with probability 0.75, and non-cases without the exposure were sampled with probability 0.25. Sampling weights were $P\{A = 1\}/0.75$ for non-cases with the exposure and $P\{A = 0\}/0.25$ for non-cases without the exposure, with $P\{A = a\}$ calculated from prevalence of exposure in the Total Population of non-cases.

Biased Stratified.—Non-cases were sampled from 21 randomly prespecified strata of sizes ranging between 20,000 and 80,000, where smaller strata had higher proportion of exposed individuals. An equal number of individuals were randomly selected from each of the strata and weights were $d/[s/21]$, where d is the stratum size and s is the survey size. Additional detail about how these strata were created and the code used to generate these strata are included in the Appendix (Web Appendices 2 and 5).

Additional detail on survey sampling can be found in Web Appendix 2. We expected analyses using controls drawn from the *Biased Probability* and *Biased Stratified* sampling designs would produce biased estimates if weights were not properly incorporated, because the exposure fraction in the sample of unweighted controls is unlikely to match that of the source population from which the cases arose. This phenomenon is likely to arise in complex surveys such as NHANES that sample based on characteristics associated with exposure—for example, oversampling of specific racial/ethnic minority groups – if membership in the over-sampled group is associated with the exposure.

All sampling designs scaled survey weights for those selected in the survey to sum up to the total number of non-cases in the Total Population, or in sensitivity analyses, to the size of the Total Population.

Analysis methods

We tested four different analysis methods (termed *Unweighted*, *Replicate*, *Sample*, and *Model*) for incorporating complex survey data into case-control studies using risk-set sampling with a ratio of one control per case. Since only one control was selected per case, all descriptions below refer to how controls were sampled across risk-sets. In situations where more than one control is selected per case, we recommend that controls are sampled with replacement across risk sets but without replacement within risk sets [32]. For additional information on how controls could be selected within risk-sets when the ratio

of controls to cases is greater than 1:1, see Web Appendix 2. For all methods, we used conditional logistic regression to account for the pair-matching of cases to controls on time.

Unweighted.—For each case, we risk-set sampled by taking a simple random sample with replacement from only the survey-based controls at the time the case occurred, and then ran an unweighted conditional logistic regression. Survey weights did not impact probability of control selection.

Replicate.—Each survey record was replicated by its corresponding survey weight, rounded to the nearest whole number. For each case, we drew a survey-based control by risk-set sampling with replacement from the expanded survey data and cases still at risk at the time the case occurred, and ran an unweighted conditional logistic regression.

Sample.—We risk-set sampled from controls and cases still at risk with replacement with probability of selection proportional to the individual's survey weight and then ran an unweighted conditional logistic regression.

Model.—We risk-set sampled by taking a simple random sample with replacement from only the survey-based controls without consideration of the survey weights and then ran a weighted conditional logistic regression with the survey weights. For all methods, cases were assigned a weight of 1.

The *Unweighted* and *Model* methods risk-set sampled from only survey-based controls while the *Replicate* and *Sample* methods risk-set sampled from survey-based controls and cases still at risk. Since *Unweighted* and *Model* did not incorporate survey weights into risk-set sampling (i.e., cases and unique survey based controls would have equal probability of selection), cases could not be incorporated because there would be an overwhelmingly large number of cases selected by random sampling during risk-set sampling (50–80% of controls selected would be cases), and thus controls would not represent the exposure distribution in the Total Population.

Of the four methods tested, we expected *Unweighted* and *Model* would be biased and perform poorly, but these are approaches used in previous applied research [19,20]. While *Unweighted* was expected to be biased because the survey weights were not taken into account at all, *Model* was expected to be biased because survey weights were not taken into account during control sampling, and only implemented in regression (where the weighted pseudo-population of survey-based controls would be unlikely to capture the exposure fraction of the Total Population). We compared how these approaches performed against *Replicate* and *Sample*, which we expected to be unbiased.

Performance metrics

Over 1000 simulations, we assessed performance by estimating: (1) relative bias, $(\overline{IDR_{Est}} - IDR_{True}) / IDR_{True} \times 100\%$, where $\overline{IDR_{Est}}$ was the mean estimated IDR over 1000 simulations and IDR_{True} was the true IDR in the Total Population, (2) variance, (3) mean-squared error, $MSE = bias^2 + variance$, where $bias = \overline{IDR_{Est}} - IDR_{True}$ and lower MSE

indicates a better estimate, and (4) 95% CI coverage, the proportion of simulations that had 95% CIs that captured the true IDR.

All analyses were conducted in R version 3.6.0 [33]. Risk-set sampling was implemented using a modified version of the “ccwc” function in the “Epi” package [34] to incorporate sampling with probability proportional to survey weights (see Web Appendix 5).

RESULTS

Simulated data

Characteristics of the Total Population are presented in Web Table 1. Briefly, 47.8% of the population was exposed, 2.0% had the outcome, and the true IDR was 1.99.

Unweighted

Results for the Unweighted analysis are presented in Table 1. As expected, this method was unbiased only in the Simple Random and Random Probability sampling designs. We observed high relative bias (−69.7% to −19.1%) under the Biased Probability and Biased Stratified sampling designs (Figure 1(a)). Variance tended to increase as survey size decreased for all sampling designs (Web Figure 1(a)). The MSEs for biased sampling designs (Biased Probability, Biased Stratified) were at least one order of magnitude higher compared to unbiased sampling designs (Simple Random, Random Probability) using the Unweighted method. The 95% CI coverage for the Simple Random sampling design was only close to 95% at the largest survey size ($n = 20,000$) and decreased as survey size decreased. At the largest survey size, we achieved 95% CI coverage for the Random Probability sampling design, but CI coverage similarly decreased as survey size decreased. We observed poor coverage (0% to 0.5%) for the Biased Probability and Biased Stratified sampling designs across all survey sizes using this method (Figure 2(a)).

Replicate

The Replicate method was unbiased for all sampling designs (Table 2, Figure 1(b)). Variance was similar across sampling designs for a given survey size though tended to be the highest for Random Probability sampling. Variance and MSE both increased as survey size decreased (Web Figure 1(b) and 2(b)). CI coverage of 95% was achieved under Simple Random and Biased Probability sampling at the largest survey size ($n = 20,000$). Coverage became increasingly poor as survey size decreased, with coverage as low as 35% for the smallest survey size ($n = 1,000$).

Sample

The Sample method was unbiased for all four sampling designs (Table 3). Variance and MSE tended to increase as survey size decreased for given sampling design and was the lowest overall for the Biased Probability sampling design. Biased Probability sampling attained over 95% CI coverage for the larger two survey sizes ($n = 20,000$ and $n = 15,000$), and close to 95% CI coverage for the $n = 5,000$ survey size (Figure 2(b)). Biased Stratified sampling also attained 95% CI coverage at the largest survey size. Coverage was still increasingly poor at smaller survey sizes.

Model

We observed the most bias using the Model method, with magnitudes upwards of 100% (Table 4, Figure 1(d)). Variance was lowest for Biased Probability, and several orders of magnitude larger for Simple Random, Random Probability, and Biased Stratified sampling (Web Figure 1(d)). MSE was several orders of magnitude higher for all four sampling designs we considered compared to the other analysis methods (Web Figure 2(d)). There was 0% 95% CI coverage for this method (Figure 2(d)). Estimates for the smallest survey size ($n = 1,000$) are not reported for several sampling designs because the regression models did not converge due to extreme sampling weights.

Sensitivity analyses

In sensitivity analyses where cases were included in survey sampling, performance was slightly worse (more biased, less precise), but not substantively different (Web Appendix 4).

DISCUSSION

This simulation study shows that unbiased estimates of the IDR can be obtained from risk-set sampled case-control studies using complex survey data if survey weights are accounted for preceding or during risk-set sampling of controls (Replicate and Sample methods). Results from the Replicate and Sample methods were comparable to results obtained using the Unweighted method under the Simple Random and Random Probability sampling designs, which both resulted in controls that were unbiased random samples of the non-cases in the Total Population. However, CI coverage tended to decrease with decreasing survey size regardless of analysis method and sampling design. When survey sampling was not independent of exposure status (Biased Probability and Biased Stratified sampling; as would be expected in many complex surveys), analytic methods that did not take weights into account during risk-set sampling were biased (Model and Unweighted) and had poor CI coverage.

Although the Replicate and Sample methods performed similarly in our simulations (low bias, similar degrees and patterns in variance and coverage), the Sample analysis method seemed to have lower computational burden. The Replicate method was previously implemented in applied work [21] and therefore tested in our simulation, but the Sample method may be preferable over Replicate because it does not require rounding of weights preceding replication of survey subjects, which may introduce some bias. While we did not see meaningful differences in bias due to rounding in our simulations, the Sample method attained slightly higher 95% CI coverage at the larger survey sizes compared to the Replicate method. Prior to our modification of the “ccwc” R function used for risk-set sampling, selection probabilities could not be incorporated into the risk-set sampling process using this function. This may still be the case for other statistical software, in which case the investigator might opt to use the Replicate method.

To our knowledge, this is the first study to compare different methods of incorporating control data from complex surveys into risk-set sampled case-control studies. We compared several different types of survey sampling designs in combination with four realistic

methods of analysis, including three (Replicate, Model, Unweighted) that were previously employed in applied studies aiming to use complex survey data as a source for controls in case-control analyses in both cumulative [19,20] and density [21] case-control designs. Our findings for risk-set sampled case-control studies are also consistent with previous work that addressed complex-sampled controls in analyses of cumulative case-control studies that found that improper handling of survey weights leads to bias, underestimated variance, and low CI coverage [18,26,27,30].

There were several limitations to this study. Our results only apply to the simulations and analytic strategies we tested. There may be further variation in performance for different data-generating mechanisms. Future work might consider data-generating mechanisms and study designs beyond what was implemented here, including incorporating time-varying exposures, non-zero induction times [35], and longitudinal studies that might be sources for controls [36,37]. Other analytic methods of incorporating risk-set sampled controls from complex surveys may exist that we did not consider, including some methods that helped improve efficiency in cumulative designs [2,18,28,29]. The survey sampling designs used in this study were not as complex as those in typical population-based surveys such as ACS or NHANES. These surveys use multiple stages of sampling, including cluster sampling, which was not implemented here, as well as weighting adjustments to account for non-response and under/over-coverage. While we expect the Model and Unweighted analysis methods will continue to be problematic in these scenarios, future research is needed to confirm that Replicate and Sample remain unbiased, since similarities within clusters may not be captured by a sampling weight alone. Additional variance may also be induced by estimation of non-response and coverage adjustments that may not be accounted for in estimation.

While our study considered a cross-sectional survey conducted at a single time point as the source for controls, multiple cross-sectional surveys including those with multi-year windows could be used to reconstruct exposures over time in the study base. For example, for cases that occurred between 2015–2018, we could use continuous NHANES 2015–2016, 2017–2018, and 2019–2020 to sample controls from. For a case occurring in 2016, we would want to use the 2017–2018 survey to sample controls from, and for a case occurring in 2018, we would want to use the 2019–2020 survey. Implicit in this approach is the assumption that the survey-based controls are not (and have not been) cases at the time they are surveyed.

Further work is needed to address several issues that came to light in the current study. While CI coverage was close to 95% in the larger survey sizes, CI coverage decreased substantially below 95% in small survey sizes regardless of analysis method or sampling design. This decrease could be driven by the decreasing number of unique survey-based controls available per case at smaller survey sizes in accordance with the equation for relative efficiency of nested case-control studies compared to full cohort analyses, $R/(R + 1)$, where R is the ratio of unique controls per case [38]. While the ratio of unique survey-based controls to cases was approximately 1:1 in the largest survey size, it was only 1:4 in the smallest survey size. The decrease in coverage could also be due to smaller proportions of the Total Population surveyed at small survey sizes. The population-based surveys currently conducted in the United States typically survey proportions of the population close to

or smaller than the smallest survey size we considered (approximately 0.1% of the Total Population). It is unclear which of these factors is more important. For rare outcomes, it may be feasible to find large enough complex-sampled surveys to ensure appropriate CI coverage. Otherwise, additional research is needed to address potential statistical corrections that can improve coverage in situations with small survey sizes (e.g. bootstrapping [39]). There are also additional case-control design features (matching, ratio of cases to controls, etc.) that could further improve statistical efficiency [40]. It is also important to note that several steps (survey sampling, Random Probability generation, risk-set sampling) in our simulation could have produced uncertainty that may not be appropriately accounted for in inference and variance estimates.

We show that risk-set sampled case-control studies combining case data with complex survey data holds promise as an analytic method in epidemiology and other “big data” applications in studying human health and disease for a variety of exposures and outcomes. Additional work in this area will improve the methodological rigor of future case-control studies to produce more accurate and precise findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Source of Funding:

This work was supported by grant DP2HD080350 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development and the National Institutes of Health Office of the Director. CXL received funding support from the National Heart, Lung, and Blood Institute (NHLBI) through the Genetic Epidemiology of Heart, Lung, and Blood Traits (GenHLB) Training Program at the University of North Carolina at Chapel Hill (T32HL129982).

Data and Code Availability:

Data for this study were derived from sources that are publicly available online. Statistical software code used for the simulated data and analysis is provided in Web Appendix 5.

Abbreviations:

CI	confidence interval
IDR	incidence density ratio
ACS	American Community Survey
NHANES	National Health and Nutrition Examination Survey

REFERENCES

- [1]. Murdoch TB, Detsky AS. The Inevitable Application of Big Data to Health Care. *JAMA* 2013;309:1351–2. 10.1001/jama.2013.393. [PubMed: 23549579]
- [2]. Kalton G, Piesse A. Survey research methods in evaluation and case-control studies. *Stat Med* 2007;26:1675–87. 10.1002/sim.2796. [PubMed: 17278183]

- [3]. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of Controls in Case-Control Studies: I. Principles. *Am J Epidemiol* 1992;135:1019–28. 10.1093/oxfordjournals.aje.a116396. [PubMed: 1595688]
- [4]. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of Controls in Case-Control Studies: II. Types of Controls. *Am J Epidemiol* 1992;135:1029–41. 10.1093/oxfordjournals.aje.a116397. [PubMed: 1595689]
- [5]. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of Controls in Case-Control Studies: III. Design Options. *Am J Epidemiol* 1992;135:1042–50. 10.1093/oxfordjournals.aje.a116398. [PubMed: 1595690]
- [6]. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Lippincott Williams & Wilkins; 2008.
- [7]. Miettinen O. Estimability and estimation in case-referent studies. *Am J Epidemiol* 1976;103:226–35. 10.1093/oxfordjournals.aje.a112220. [PubMed: 1251836]
- [8]. Kupper LL, McMichael AJ, Spirtas R. A Hybrid Epidemiologic Study Design Useful in Estimating Relative Risk. *J Am Stat Assoc* 1975;70:524–8. 10.2307/2285927.
- [9]. Greenland S, Thomas DC. On the Need for the Rare Disease Assumption in Case-Control Studies. *Am J Epidemiol* 1982;116:547–53. 10.1093/oxfordjournals.aje.a113439. [PubMed: 7124721]
- [10]. Curtin R, Presser S, Singer E. Changes in Telephone Survey Nonresponse over the Past Quarter Century. *Public Opin Q* 2005;69:87–98. 10.1093/poq/nfi002.
- [11]. Blumberg SJ, Luke JV. Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, January-June 2017. *Natl Cent Health Stat* 2017;13.
- [12]. Guterbock T, Benson G, Lavrakas P. The Changing Costs of Random Digital Dial Cell Phone and Landline Interviewing. *Surv Pract* 2018;11:3168. 10.29115/SP-2018-0015.
- [13]. Berkson J Limitations of the Application of Fourfold Table Analysis to Hospital Data. *Biom Bull* 1946;2:47–53. 10.2307/3002000.
- [14]. *Complex Sample Surveys*. *Encycl. Surv. Res. Methods*, 2455 Teller Road, Thousand Oaks California 91320 United States of America: Sage Publications, Inc.; 2008. 10.4135/9781412963947.n78.
- [15]. National Research Council. *Realizing the Potential of the American Community Survey: Challenges, Tradeoffs, and Opportunities*. Washington, DC: The National Academies Press; 2015. 10.17226/21653.
- [16]. National Center for Health Statistics (U.S.). *National health and nutrition examination survey*. 2013.
- [17]. Pfeiffermann D The use of sampling weights for survey data analysis. *Stat Methods Med Res* 1996;5:239–61. 10.1177/096228029600500303. [PubMed: 8931195]
- [18]. Scott A, Wild C. Case-Control Studies with Complex Sampling. *J R Stat Soc Ser C Appl Stat* 2001;50:389–401.
- [19]. Wiebe DJ. Homicide and suicide risks associated with firearms in the home: A national case-control study. *Ann Emerg Med* 2003;41:771–82. 10.1067/mem.2003.187. [PubMed: 12764330]
- [20]. Kleck G, Hogan M. National Case-Control Study of Homicide Offending and Gun Ownership. *Soc Probl* 1999;46:275–93. 10.2307/3097256.
- [21]. Matthey EC, Farkas K, Skeem J, Ahern J. Exposure to Community Violence and Self-harm in California: A Multilevel, Population-based, Case-control Study. *Epidemiology* 2018;29:697–706. 10.1097/EDE.0000000000000872. [PubMed: 29889134]
- [22]. Scott AJ, Wild CJ. Fitting Logistic Models Under Case-Control or Choice Based Sampling. *J R Stat Soc Ser B Methodol* 1986;48:170–82.
- [23]. Scott A, Wild C. On the Robustness of Weighted Methods for Fitting Models to Case-Control Data. *J R Stat Soc Ser B Stat Methodol* 2002;64:207–19.
- [24]. Langholz B, Goldstein L. Conditional logistic analysis of case-control studies with complex sampling. *Biostatistics* 2001;2:63–84. 10.1093/biostatistics/2.1.63. [PubMed: 12933557]
- [25]. Fears TR, Gail MH. Analysis of a Two-Stage Case-Control Study with Cluster Sampling of Controls: Application to Nonmelanoma Skin Cancer. *Biometrics* 2000;56:190–8. 10.1111/j.0006-341X.2000.00190.x. [PubMed: 10783795]
- [26]. Lumley T *Complex Surveys: A Guide to Analysis Using R*. 1st ed. Wiley; 2010.

- [27]. Graubard BI, Fears TR, Gail MH. Effects of cluster sampling on epidemiologic analysis in population-based case-control studies. *Biometrics* 1989;45:1053–71. [PubMed: 2611317]
- [28]. Li Y, Graubard BI, DiGaetano R. Weighting methods for population-based case-control studies with complex sampling. *J R Stat Soc Ser C Appl Stat* 2011;60:165–85. 10.1111/j.1467-9876.2010.00731.x.
- [29]. Landsman V, Graubard BI. Efficient analysis of case-control studies with sample weights. *Stat Med* 2013;32:347–60. 10.1002/sim.5530. [PubMed: 22833421]
- [30]. Lawless JF, Kalbfleisch JD, Wild CJ. Semiparametric Methods for Response-Selective and Missing Data Problems in Regression. *J R Stat Soc Ser B Stat Methodol* 1999;61:413–38.
- [31]. Prentice RL, Breslow NE. Retrospective studies and failure time models. *Biometrika* 1978;65:153–8. 10.1093/biomet/65.1.153.
- [32]. Lubin JH, Gail MH. Biased Selection of Controls for Case-Control Analyses of Cohort Studies. *Biometrics* 1984;40:63–75. 10.2307/2530744. [PubMed: 6375751]
- [33]. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
- [34]. Carstensen B, Plummer M, Laara E, Hills M. *Epi: A Package for Statistical Analysis in Epidemiology*. 2019.
- [35]. Strömberg U Does Induction Time Have Any Bearing on Definition of Study Base? *Epidemiology* 1994;5:356–9. [PubMed: 8038254]
- [36]. Schildcrout JS, Haneuse S, Tao R, Zelnick LR, Schisterman EF, Garbett SP, et al. Two-Phase, Generalized Case-Control Designs for the Study of Quantitative Longitudinal Outcomes. *Am J Epidemiol* 2020;189:81–90. 10.1093/aje/kwz127. [PubMed: 31165875]
- [37]. Schildcrout JS, Schisterman EF, Mercaldo ND, Rathouz PJ, Heagerty PJ. Extending the case-control design to longitudinal data: stratified sampling based on repeated binary outcomes. *Epidemiol Camb Mass* 2018;29:67–75. 10.1097/EDE.0000000000000764.
- [38]. Goldstein L, Langholz B. Asymptotic Theory for Nested Case-Control Sampling in the Cox Regression Model. *Ann Stat* 1992;20:1903–28.
- [39]. Efron B, Tibshirani R. *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy*. *Stat Sci* 1986;1:54–75.
- [40]. Thomas DC, Greenland S. The relative efficiencies of matched and independent sample designs for case-control studies. *J Chronic Dis* 1983;36:685–97. [PubMed: 6630405]

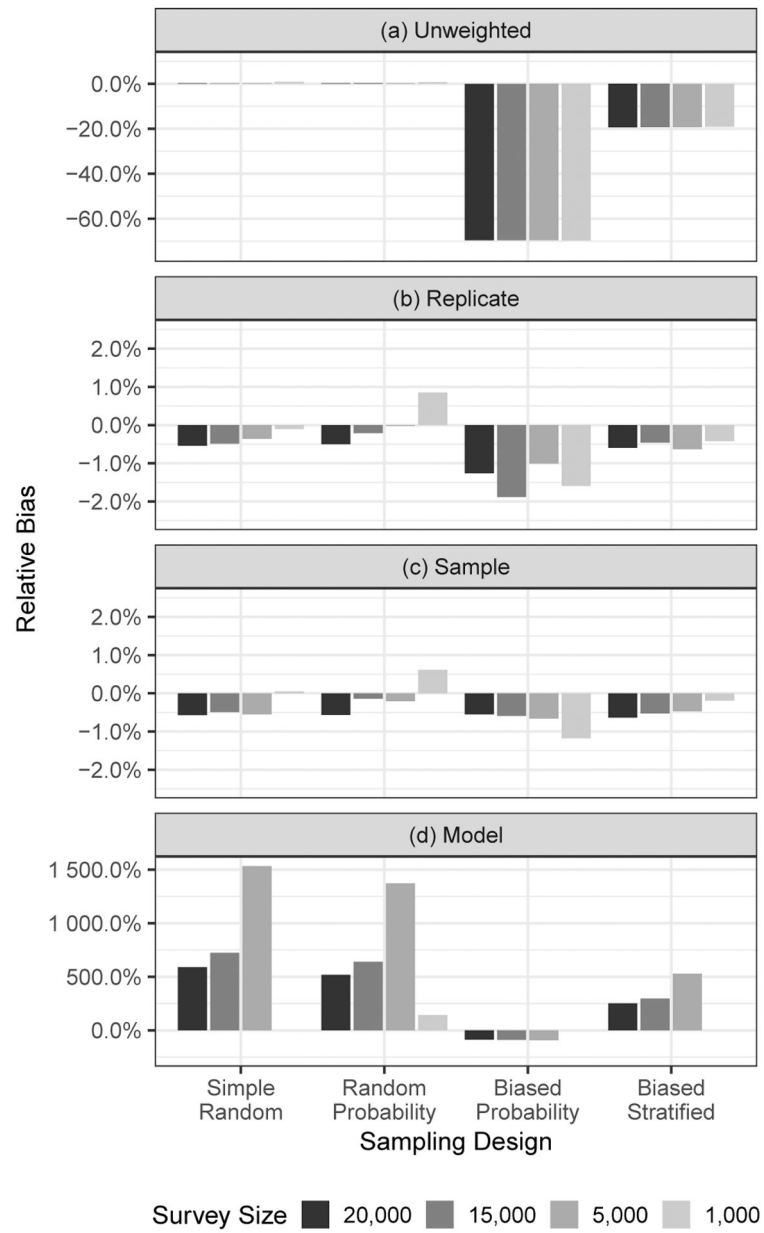


Figure 1. Bar graph of simulation results for relative bias (%) across all sampling designs and analysis methods (1000 iterations).

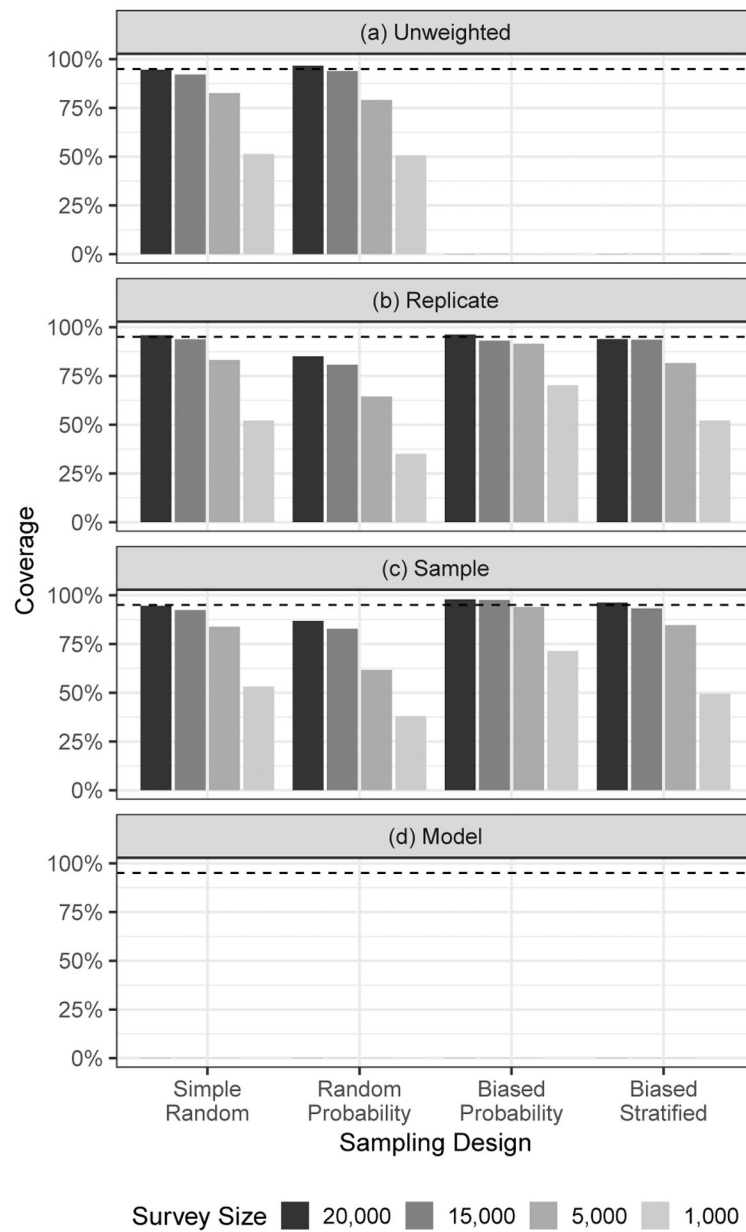


Figure 2. Bar graph of simulation results for 95% confidence interval coverage across all sampling designs and analysis methods (1000 iterations). Dashed line at 95% confidence interval coverage.

Table 1.

Simulation Results for Unweighted Analysis Method (1000 Iterations)

Sampling Design	Survey Size	Relative Bias ^a	Variance	MSE ^b	Coverage
Simple Random	20,000	0.17%	2.4×10^{-3}	2.4×10^{-3}	94.6%
Simple Random	15,000	0.24%	2.9×10^{-3}	2.9×10^{-3}	92.1%
Simple Random	5,000	0.30%	5.4×10^{-3}	5.5×10^{-3}	82.6%
Simple Random	1,000	0.87%	2.2×10^{-2}	2.2×10^{-2}	51.3%
Random Probability	20,000	0.22%	2.4×10^{-3}	2.4×10^{-3}	96.6%
Random Probability	15,000	0.35%	2.6×10^{-3}	2.7×10^{-3}	93.9%
Random Probability	5,000	0.32%	5.9×10^{-3}	5.9×10^{-3}	79.1%
Random Probability	1,000	0.75%	2.1×10^{-2}	2.1×10^{-2}	50.6%
Biased Probability	20,000	-69.6%	1.7×10^{-4}	1.9×10^0	0%
Biased Probability	15,000	-69.6%	1.7×10^{-4}	1.9×10^0	0%
Biased Probability	5,000	-69.6%	2.1×10^{-4}	1.9×10^0	0%
Biased Probability	1,000	-69.7%	5.9×10^{-4}	1.9×10^0	0%
Biased Stratified	20,000	-19.4%	1.3×10^{-3}	1.4×10^{-1}	0%
Biased Stratified	15,000	-19.3%	1.8×10^{-3}	1.5×10^{-1}	0%
Biased Stratified	5,000	-19.3%	3.3×10^{-3}	1.5×10^{-2}	0%
Biased Stratified	1,000	-19.1%	1.3×10^{-2}	1.6×10^{-1}	0.5%

MSE, mean squared error;

$$^a \text{Relative Bias} = \frac{(\overline{IDR}_{Est} - IDR_{True})}{IDR_{True}} \times 100\%$$

$$^b \text{MSE} = \text{Bias}^2 + \text{Variance}$$

Table 2.

Simulation Results for Replicate Analysis Method (1000 Iterations)

Sampling Design	Survey Size	Relative Bias ^a	Variance	MSE ^b	Coverage
Simple Random	20,000	-0.54%	2.2×10^{-3}	2.3×10^{-3}	95.8%
Simple Random	15,000	-0.49%	2.5×10^{-3}	2.6×10^{-3}	93.8%
Simple Random	5,000	-0.36%	5.3×10^{-3}	5.3×10^{-3}	83.2%
Simple Random	1,000	-0.10%	2.1×10^{-2}	2.1×10^{-2}	52.1%
Random Probability	20,000	-0.50%	4.6×10^{-3}	4.7×10^{-3}	85.0%
Random Probability	15,000	-0.21%	5.9×10^{-3}	6.0×10^{-3}	80.7%
Random Probability	5,000	-0.02%	1.2×10^{-2}	1.2×10^{-2}	64.5%
Random Probability	1,000	0.85%	5.2×10^{-2}	5.2×10^{-2}	35.0%
Biased Probability	20,000	-1.26%	1.7×10^{-3}	2.3×10^{-3}	96.2%
Biased Probability	15,000	-1.88%	1.6×10^{-3}	3.0×10^{-3}	93.1%
Biased Probability	5,000	-1.01%	2.8×10^{-3}	3.2×10^{-3}	91.5%
Biased Probability	1,000	-1.59%	8.1×10^{-3}	9.1×10^{-3}	70.2%
Biased Stratified	20,000	-0.60%	2.6×10^{-3}	2.7×10^{-3}	93.9%
Biased Stratified	15,000	-0.46%	2.9×10^{-3}	3.0×10^{-3}	93.6%
Biased Stratified	5,000	-0.66%	5.1×10^{-3}	5.3×10^{-3}	81.6%
Biased Stratified	1,000	-0.42%	2.2×10^{-2}	2.2×10^{-2}	52.1%

MSE, mean squared error

$$^a \text{Relative Bias} = \frac{(\overline{IDR}_{Est} - IDR_{True})}{IDR_{True}} \times 100\%$$

$$^b \text{MSE} = \text{Bias}^2 + \text{Variance}$$

Table 3.

Simulation Results for Sample Analysis Method (1000 Iterations)

Sampling Design	Survey Size	Relative Bias ^a	Variance	MSE ^b	Coverage
Simple Random	20,000	-0.57%	2.5×10^{-3}	2.7×10^{-3}	94.5%
Simple Random	15,000	-0.49%	2.9×10^{-3}	2.9×10^{-3}	92.4%
Simple Random	5,000	-0.55%	5.0×10^{-3}	5.2×10^{-3}	83.9%
Simple Random	1,000	0.05%	2.1×10^{-2}	2.1×10^{-2}	53.2%
Random Probability	20,000	-0.57%	4.6×10^{-3}	4.7×10^{-3}	86.8%
Random Probability	15,000	-0.14%	5.8×10^{-3}	5.8×10^{-3}	82.8%
Random Probability	5,000	-0.20%	1.3×10^{-2}	1.3×10^{-2}	61.7%
Random Probability	1,000	0.61%	4.7×10^{-2}	4.7×10^{-2}	38.0%
Biased Probability	20,000	-0.55%	1.7×10^{-3}	1.8×10^{-3}	97.9%
Biased Probability	15,000	-0.59%	1.9×10^{-3}	2.0×10^{-3}	97.6%
Biased Probability	5,000	-0.66%	2.6×10^{-3}	2.7×10^{-3}	94.0%
Biased Probability	1,000	-1.18%	7.8×10^{-3}	8.4×10^{-3}	71.4%
Biased Stratified	20,000	-0.64%	2.2×10^{-3}	2.4×10^{-3}	96.2%
Biased Stratified	15,000	-0.53%	2.7×10^{-3}	2.8×10^{-3}	93.3%
Biased Stratified	5,000	-0.47%	4.9×10^{-3}	4.9×10^{-3}	84.6%
Biased Stratified	1,000	-0.19%	2.1×10^{-2}	2.1×10^{-2}	49.4%

MSE, mean squared error

$$^a \text{Relative Bias} = \frac{(\overline{IDR_{Est}} - IDR_{True})}{IDR_{True}} \times 100\%$$

$$^b \text{MSE} = \text{Bias}^2 + \text{Variance}$$

Table 4.

Simulation Results for Model Analysis Method (1000 Iterations)

Sampling Design	Survey Size	Relative Bias ^a	Variance	MSE ^b	Coverage
Simple Random	20,000	590.%	1.5×10 ⁰	1.4×10 ²	0%
Simple Random	15,000	725%	2.8×10 ⁰	2.1×10 ²	0%
Simple Random	5,000	1530%	3.6×10 ¹	9.6×10 ²	0%
Simple Random	1,000				<i>c</i>
Random Probability	20,000	518%	1.2×10 ⁰	1.1×10 ²	0%
Random Probability	15,000	640.%	2.2×10 ⁰	1.6×10 ²	0%
Random Probability	5,000	1370%	2.6×10 ¹	7.7×10 ²	0%
Random Probability	1,000	144%	4.0×10 ²	4.0×10 ²	0%
Biased Probability	20,000	-87.4%	5.2×10 ⁻⁴	3.0×10 ⁰	0%
Biased Probability	15,000	-89.1%	4.7×10 ⁻⁴	3.1×10 ⁰	0%
Biased Probability	5,000	-93.5%	3.3×10 ⁻⁴	3.5×10 ⁰	0%
Biased Probability	1,000				<i>c</i>
Biased Stratified	20,000	252%	4.4×10 ⁻¹	2.6×10 ¹	0%
Biased Stratified	15,000	298%	7.2×10 ⁻¹	3.6×10 ¹	0%
Biased Stratified	5,000	529%	6.2×10 ⁰	1.2×10 ²	0%
Biased Stratified	1,000				<i>c</i>

MSE, mean squared error

$$^a \text{Relative Bias} = (\overline{IDR_{Est}} - IDR_{True}) / IDR_{True} \times 100\%$$

$$^b \text{MSE} = \text{Bias}^2 + \text{Variance}$$

^cResults not presented because many simulations did not converge due to extreme weights