Published in final edited form as:

Am Stat. 2008 November 1; 62(4): 307–313. doi:10.1198/000313008X364525.

Flexible Frames and Control Sampling in Case-Control Studies: Weighters (Survey Statisticians) Versus Anti-Weighters (Epidemiologists)

Richard F. Potthoff [Senior Research Scientist],

Cancer Statistical Center, Duke University Medical Center, Hock Plaza, 2424 Erwin Road, Suite 802, Durham, NC 27705 (E-mail: potth001@mc.duke.edu)

Susan Halabi [Associate Professor],

Department of Biostatistics and Bioinformatics, Duke University Medical Center, Hock Plaza, 2424 Erwin Road, Suite 802, Durham, NC 27705 (E-mail: susan.halabi@duke.edu)

Joellen M. Schildkraut [Professor], and

Department of Community and Family Medicine and the Duke Comprehensive Cancer Center, Duke University Medical Center, Hock Plaza, 2424 Erwin Road, Room 6036, Durham, NC 27705 (E-mail: schil001@mc.duke.edu)

Beth Newman [Professor]

School of Public Health, Institute of Health and Biomedical Innovation, Queensland University of Technology, Victoria Park Road, Kelvin Grove, Queensland 4059, Australia (E-mail: b.newman@qut.edu.au)

Abstract

We propose two innovations in statistical sampling for controls to enable better design of population-based case-control studies. The main innovation leads to novel solutions, without using weights, of the difficult and long-standing problem of selecting a control from persons in a household. Another advance concerns the drawing (at the outset) of the households themselves and involves random-digit dialing with atypical use of list-assisted sampling. A common element throughout is that one capitalizes on flexibility (not broadly available in usual survey settings) in choosing the *frame*, which specifies the population of persons from which *both* cases and controls come.

Keywords

Bias; List-assisted sampling; Population-based case-control studies; Random-digit dialing; Respondent selection within households

1. INTRODUCTION

Although case-control studies (e.g., Breslow 1996; Kelsey, Whittemore, Evans, and Thompson 1996) often provide the best feasible means for investigating associations of exposure and disease, they are subject to many possible pitfalls. A case-control study is an observational means of investigation that compares cases, who have a disease, with noncases (i.e., controls), who are free of the disease. The most easily defended form of such a study is a *population-based* case-control study, where the cases are all (or a suitable sample) of those that arise in a given population of persons in a given time period and the controls are a suitable sample of noncases from that same population. Population-based case-control studies, as opposed to hospital-based or other types of case-control studies that may be less costly, can eliminate some

potential biases and other pitfalls, but not all of them. Thus any improvements to the design or analysis of population-based case-control studies that can strengthen their advantages and reduce potential objections as to their validity, while also holding down cost, are always desirable.

This article focuses on innovations and improvements in statistical sampling for controls in the design of population-based case-control studies. Its most important contribution is to provide new and better approaches to the thorny problem of choosing a control within a household, when households are drawn through random-digit dialing or perhaps through another means. A second contribution concerns special use of list-assisted sampling for random-digit dialing in obtaining the households from which the controls are selected.

Flexibility in choosing the *frame*—the population of persons from which *both* cases and controls are to be obtained—is a key theme in what follows. Dodge (2003, p. 155) defined a *frame* to be “A list, map or other specification of the units which define a population to be sampled.” A typical study based on a sample survey usually aims to cover a prespecified population, and thus has little latitude in the choice of its frame.

But a population-based case-control study can be seen as quite different and as almost unique in the sense that it can abide extra leeway in the selection of its frame. Coverage of an indicated population may be desirable but not paramount for this type of study. Rather, the main criterion for the frame, not always observed in practice, should be for controls and cases to have the *same* frame insofar as possible. In the choice of the frame itself, though, certain flexibility may be feasible. That flexibility plays a major role in the innovations proposed in this article, and provides an opportunity that has apparently not been fully appreciated, explored, or exploited.

A key feature of these innovations is that they preclude any need for sampling weights. There seems to be a culture collision between survey statisticians and epidemiologists over the use and desirability of weights in analyzing population-based case-control studies. Survey statisticians tend to favor the use of weights; epidemiologists do not. There are pros and cons.

In fact, a possible objection to our new methods concerns the frame reduction that they entail as the price that one pays to avoid weights. Though usually the reduction in population coverage would appear to be of limited import, one can still ask about potential distortion.

After a review in Section 2 of older methods and their drawbacks, Section 3 presents the new methods to sample for a control from persons within a household once it has been drawn. Specialized uses of list-assisted sampling, when the drawing of the households themselves uses random-digit dialing, are covered in Section 4. For all the methods in Sections 3 and 4, cases and controls have the same frame and, moreover, no weights are needed. Section 5 discusses frame choice and flexible frames in general terms. Sections 6 and 7 consider the conflict over weights between survey statisticians and epidemiologists, the pros and cons of weighting, and possible bias and distortion from frame reduction. Section 8 gives a summary.

2. EXISTENT METHODS FOR SAMPLING CONTROLS

Useful background information related to the present article can be found in the excellent encyclopedia article by Waksberg (1998), which deals with issues of statistical sampling for controls in population-based case-control studies when random-digit dialing is used. For drawing the controls, we assume two sampling stages, where households are drawn in the first stage and then individuals within those households are chosen in the second stage. The first-stage drawing of households may be by random-digit dialing (covered in Section 4) or by other suitable means. We begin here by dealing with the second stage—specifically with the problem of how to select the controls within the households.

Although there are different existent approaches to address this problem, all of them have drawbacks. One technique is to randomly draw one eligible person (if there *is* at least one) from each household and then weight by the number of eligibles in the household. Although there is no bias, there are disadvantages to weighting (see Section 6).

There are several methods, all with their own limitations, that avoid weights. One can select all the eligibles in each household and thereby avoid both weighting and biased estimates of means, but then nonindependence of two or more controls from the same household entails a violation of common statistical assumptions (Hartge et al. 1984, p. 829; Greenberg 1990, p. 3).

Another possibility is to randomly select only one eligible from each household that has at least one but then not apply any weights. Then there is bias against persons from households with more eligibles. For studies of childhood cancer, Greenberg (1990) contended that the method underrepresents children in households of lower socioeconomic status. He argued that their births tend to be more closely spaced so that they have greater concentration of eligibles in the age range of the study, thus leading to lower probability of their selection as controls.

A different method, suitable only for studies that involve adults of both sexes, is to randomly predesignate each household as “male” or “female” and then choose as control(s) every eligible (if any) of the designated sex but no one of the other sex (Hartge et al. 1984). But some households will contribute multiple controls. There will also be extra recruitment cost, because some “male” households will have no males, and likewise for females.

A method of Potthoff (1994, p. 975) has certain advantages and is not confined to studies with adults of both sexes. But it entails even more extra recruitment expense and also requires an extra contact with each selected control.

3. NEW WAYS TO DRAW CONTROLS WITHIN HOUSEHOLDS

The new methods described in this section do not use weights, never draw more than one control from a household, require no extra contact, and enable unbiased estimation with respect to the chosen frame. As noted earlier, they do involve limited reduction in the frame. For the selection of individuals within households, we show ways to exploit flexible frames for three types of population-based case-control studies: those with (i) adults of one sex, (ii) adults of both sexes, and (iii) children of both sexes.

3.1 Studies With Adults of One Sex

For illustration we consider a study of a disease that afflicts only females, mainly older ones. Our innovation is to define the frame so that, as a result of specified exclusion rules, *no more than one person* (one female) *from any household is in the frame*. This solves the problem of finding a way to sample for controls within households that satisfies all the desired properties simultaneously. Cases are excluded using the same exclusion rules as for controls, thus ensuring (albeit at the cost of losing a few cases) that there is no bias with respect to the chosen frame. With no more than one person per household in the frame, independence of controls automatically holds, there is no inequality of selection probabilities for controls and thus no need to weight, and the one person can be identified in a single contact.

The exclusion rules that lead to at most one female per household could be formulated in different ways. The best way may vary by country or ethnic group, because of differing demographics. The following rules illustrate what may be suitable for a study of females aged 20 to 74 in the United States. First, let the frame exclude any female 20 to 74 if there is an *older* female in the same household who is 65 to 74. The aim is to hold down costs through

easier recruitment of controls in the hardest-to-get (oldest) age groups, but with only limited loss of cases. Second, let the frame exclude any female 20 to 64 if there is a *younger* female in the same household who is at least 20. The aim is to reduce loss of cases among the lowest-incidence (youngest) age groups, but with only limited higher cost of recruiting controls. Given that only a small percentage of households have more than one female 20 to 74, the percentage of females 20 to 74 whom the frame excludes will also be low.

Each potential case must be checked to see if she is in the frame and removed from the study if not. A case is dropped if her household at her time of diagnosis had an *older* female who was then 65 to 74, or a *younger* female who was then at least 20 if the case was then under 65.

For controls, the recruitment questionnaire for a household finds the oldest female 65 to 74 (at time of recruitment) if there is one. Otherwise, it selects the youngest female 20 to 64.

3.2 Studies With Adults of Both Sexes

If a study covers adults of both sexes, consider a hypothetical (but reasonably realistic) population where every 100 households are distributed as follows:

F 20, M 12, FF 4, MM 2, FM 56, FFM 3, FMM 3

Here “FF 4,” for example, means that 4 of the 100 households each contain two adult females and no adult males. The 100 households contain 174 adults ($32 \times 1 + 62 \times 2 + 6 \times 3$). Suppose first that one defines the frame so that it includes just one person from each household. Although no household will be without a potential control, 43% of adults (74 out of 174) are excluded from the frame, with attendant loss of cases that will probably be so heavy as to be unacceptable.

An alternative, though, is to define a male frame so that there is at most one male per household and a female frame with at most one female per household, and then, when recruiting controls, randomly predesignate each household as “male” or “female,” with probability 1/2 for each sex. The one male (if there is one) would then be drawn from each “male” household and the one female (if any) from each “female” household. This random-predesignation method is like that of Hartge et al. (1984) mentioned earlier, except that the problem of multiple controls in a household is eliminated by defining extra members of either sex out of the frame. Although 19% of households will be without a potential control [half of $(20 + 12 + 4 + 2)$ out of 100], only 7% of persons [$(4 + 2 + 3 + 3)$ out of 174] are excluded from the frame—far better than 43%.

Further improvement is possible, though. Define the frame so that there are at most *two* adults (of any sex breakdown) per household, rather than one. (For a household with more than two adults, one could specify the two for the frame based somehow on age or on both age and sex, e.g.) In control recruitment from a household with two adults in the frame, one of them is chosen randomly (with probability 1/2 for each). But to avoid control overrepresentation of adults from one-adult households, a special step is required: Discard such persons with probability 1/2 to offset the fact that otherwise their likelihood of selection would be twice that of other adults in the frame. Thus, 16% of households will not yield a potential control [half of $(20 + 12)$ out of 100], and 3% of persons [$(3+3)$ out of 174] are excluded from the frame. Both figures are a bit better than for the random-predesignation method (16% versus 19% and 3% versus 7%).

3.3 Studies With Children of Both Sexes

For studies of children of both sexes within a given age range, consider a hypothetical (but again plausibly realistic) population where every 100 households with at least one child in the age range are distributed as follows:

B 30, G 30, BB 8, BG 16, GG 8, BBB 1, BBG 3, BGG 3, GGG 1

The meaning is as before, except that B stands for boy and G for girl. If the frame is defined so as to include just one child per household, then none of the 100 households are without a potential control, but many children, 32% (48 out of 148), are excluded from the frame.

The modified random-predesignation method—which in its original form (Hartge et al. 1984) seems not to have been proposed anyway for studies of children, apparently with good reason (Greenberg 1990)—likewise fails to yield very favorable results. The modification is as in Section 3.2, so that there is a male frame with at most one boy per household and a female frame with at most one girl per household. Not only are 39% of the households without a potential control [half of $(30 + 30 + 8 + 8 + 1 + 1)$ out of 100], but also 18% of children $[(8 + 8 + 1 \times 2 + 3 + 3 + 1 \times 2)$ out of 148] are excluded from the frame.

Substantial improvement of both percentages comes about through a definition of a frame that includes up to *two* children (of any sex breakdown) per household. In the same vein as before (Section 3.2), the procedure for drawing controls randomly picks one of these two (from any household where there *are* two) but randomly discards half the children from one-child households so as to bring about equal probability of selection of all children in the frame. Potential controls fail to exist in just 30% of the households [half of $(30 + 30)$ out of 100], and only 5% of children $[(1 + 3 + 3 + 1)$ out of 148] are excluded from the frame. For households with more than two children, the two to be included in the frame could be defined using any objective criterion (based, e.g., on age, sex, birthday, or alphabetical ordering of first name).

For the studies either of adults of both sexes (Section 3.2) or of children of both sexes, the superiority of our proposed method over the (modified) random-predesignation method is not dependent on the particular numbers used in the two illustrations above. Rather, regardless of the numbers, our method yields both smaller losses of controls and fewer frame exclusions. This is because, on both of these criteria, it is better than the random-predesignation method for multiperson single-sex households (FF and MM in the first illustration; BB, GG, BBB, and GGG in the second) and no worse for the remaining households.

Finally, the basic rules for cases are, of course, the same for any frames with up to two persons per household as for those with one person per household. That is, any case not in the frame, based on household composition at the time of diagnosis, is to be left out of the study.

4. SPECIAL TELEPHONE SAMPLING OF HOUSEHOLDS

Our second innovation applies flexible frame definition again and deals with random-digit dialing in sampling households for the first stage of selecting the controls. It involves the use of *list-assisted sampling* (Brick, Waksberg, Kulp, and Starer 1995). Although this type of sampling finds extensive use in different sorts of standard telephone surveys, its use for population-based case-control studies is a bit more complex and entails some unique features.

In the United States, any telephone number has 10 digits, of which the first three make up the area code, the next three compose the prefix (sometimes also called the “exchange”), and the last four constitute the suffix. Each group of 100 numbers whose first eight digits are the same may be called a *bank*. A *working bank* is a bank for which a database or list (periodically

obtained from the latest telephone directories by commercial organizations) indicates that there is at least one listed residential (land-line) number. Nationwide, such banks have been estimated to cover about 96% of all residential numbers and miss 4% (Brick et al. 1995). Most of the missed 4% are probably in newly activated banks populated mainly by recent movers.

Working banks form the basis for the frame in list-assisted sampling. From each (randomly) selected working bank one draws a single telephone number, regardless of whether it turns out to be residential. Thus the chosen numbers do not need to be weighted differentially.

A working bank (and the associated frame) could be defined based on a minimum of (say) two or three listed residential numbers rather than one. Then, however, there would be a drop not only in dialing costs for controls (because fewer nonresidential numbers would be dialed), but also in percentage of coverage of residential numbers, with attendant loss of cases (because more numbers would be outside the frame).

For population-based case-control studies, the initial steps in the use of list-assisted sampling are like those in its other applications. One first defines the frame geographically by choosing a list of combinations of area code and prefix (first six digits of the telephone number), and then, within each chosen combination, one excludes any of the 100 banks (of 100 numbers each) that are not classified as working banks. Choosing the list of six-digit combinations is straightforward if a study covers entire state(s): One just picks every combination in a state. But if a study covers only part of a state, some specialized choices and precautions may be needed, because geographical boundaries for a combination can cross county lines or other borders.

One does encounter some singular features, though. First, instead of buying just enough telephone numbers to recruit the needed number of controls, one has to buy at least one (and probably only one) randomly selected telephone number from the 100 numbers in each working bank within each chosen six-digit combination. One thereby gets not only the telephone numbers (probably more than needed) to dial to recruit controls, but also the full frame itself (obtained by disregarding the last two digits of each telephone number). We were able to purchase this type of sample from Marketing Systems Group (Genesys Sampling Systems), Fort Washington, PA.

Second, with the full data thus available to ensure that cases and controls come from the same frame, one checks all cases and excludes any who are not in the frame. *At the time of diagnosis*, the case must have had a residential telephone number whose first eight digits appear in the purchased frame. Otherwise, the case is dropped.

Third, we note that in fields where list-assisted sampling is already widely used, the lack of 100% residential telephone coverage could be seen as a minor drawback (Brick et al. 1995), but as more of one than in the application depicted here. Moreover, the drawback for standard telephone surveys has become greater in recent years. That is because the frames in effect for such surveys may now exclude not only households with land-line numbers outside of working banks (or with no telephone at all) but also households with cellular telephones only. The increasing frequency of cell-telephone-only households poses growing problems for these surveys (Nielsen Media Research 2005). But for population-based case-control studies, the effect is less. For them, excluded households can duly be defined out of the frame for both cases and controls, given that having the same frame for both is of more concern than fuller coverage.

One important issue relates indirectly to the telephone frame. Let $F(\geq 1)$ denote the number of (landline) residential telephone numbers at which a household can be reached. In recruitment of controls, there is greater chance of dialing a household the higher its F is. There are two

ways to avoid selection bias that would favor households with higher F ; both require asking the value of F . One way is to retain every household and weight each control by $1/F$, but then weights are encountered. The other way, which requires no weighting and thus appears better for the present type of application, is to discard with probability $(F - 1)/F$ any household reachable at more than one number (e.g., Hartge et al. 1984; Greenberg 1990; Waksberg 1998).

5. FRAME CHOICE AND FLEXIBLE FRAMES

For standard sampling studies in general, the *frame* is (conceptually, at least) simply the list of all units from which the sample is to be drawn and, ideally, also coincides with the units that the study is supposed to represent. Thus, there is little flexibility in defining the frame, and difficulties may arise. For example, if a telephone survey is supposed to represent the whole population of a given state, one could define the frame to either exclude or include persons without (landline) telephones and then run the survey accordingly. With exclusion, one suffers selection bias, whereas with inclusion, one incurs extra cost in using special means to identify and contact people with no telephones or with cellular telephones only.

Population-based case-control studies also have as their goal the coverage of a predefined population. But if the main criterion for the frame is to try to have it the same for cases and controls, then there is more leeway in the choice of the frame. More specifically, in defining the frame for a population-based case-control study, one can (as in Sections 3 and 4) duly exclude parts of the population that one might otherwise include, provided that the exclusions are for both cases and controls. For studies that recruit controls through random-digit dialing, this principle was earlier suggested (e.g., Greenberg 1990; Kelsey et al. 1996, p. 202; Waksberg 1998), and applied by other investigators, to remove cases who had no telephones (though such removal of cases has perhaps not been done nearly often enough, judging from lack of mention in published reports).

The same basic principle (exclusion of cases who would not have been eligible to be controls, and vice versa) also governs if controls are drawn randomly from other frames or rosters. Such frames include driver's-license lists, Medicare files for persons 65 and over, state resident lists, birth-certificate listings, telephone directories, commercial city directories, voter lists, and commercial marketing lists. Not everyone will be on such lists. Cases and controls are drawn from the same frame, though, if all cases not on the list are ultimately excluded (and if any potential control who would not have been eligible to be a case is also excluded). Of course, data must be available to find out whether cases are on the list. Many of these types of lists have been used for case-control studies.

6. TO WEIGHT OR NOT TO WEIGHT

Our proposals in this article have been crafted so as to avoid the need for any weighting. As such, they may find more favor with epidemiologists than with survey statisticians.

Methods for weighted analysis of population-based case-control studies with complex sample designs are available and are used in practice (Korn and Graubard 1999, chap. 9; Scott and Wild 2001). Statistical software packages that handle complex sample surveys may be applied. Weighted logistic-regression analyses can be run. Also possible are analyses that take into account any clustering that may be present in the design (Graubard, Fears, and Gail 1989).

But it appears that, whether for better or worse, epidemiologists are rarely inclined to apply weights or use related complex tools of analysis in population-based case-control studies, as others have observed as well. Scott and Wild (2001, p. 400) acknowledged (but not approvingly) that “complex sampling is often ignored in the analysis” of such studies. Brogan

et al. (2001, p. 1120) stated that “Ordinarily, a population-based control sample is not weighted in case-control analyses of risk factors.” In a listing of “Typical requirements for selecting a sample of population-based controls,” DiGaetano and Waksberg (2002, p. 772) included a provision to “Permit analysis without special software” that is “required for handling data from complex sample surveys.” They stated that analyses “are frequently done without using sample weights or ‘special’ software.”

Widespread reluctance to use weighting is found even in settings where failure to do it is difficult to defend. For example, upon examining state political polls, Mitofsky (1993) noted disapprovingly that the vast majority of them did not apply weights to compensate for unequal selection probabilities caused either by unequal numbers of eligibles in a household or by multiple telephone numbers at which the household was reachable. In a close election, of course, accuracy could be critical for such polls.

Korn and Graubard (1995) also addressed the estimation bias that can result from not applying sample weights that are needed to offset unequal selection probabilities. They presented examples (from a health survey) in which such bias is large.

The idea of using weights in population-based case-control studies would be hard to sell to epidemiologists, who have customarily eschewed weighting. Yet inertia alone hardly provides a solid argument for not weighting. There are, however, better reasons to shy away from weighting *if* one can legitimately avoid it. They all play a role and need to be considered.

First is the obvious point that designs that avoid the need for weighting entail a narrower need for skilled statistical assistance and a statistical analysis that is simpler. Second, without weights there will be greater simplicity in presenting and interpreting results both in professional reports and (where applicable) for the popular media: Descriptions can correctly refer to whole numbers of persons. Third, weighting entails a loss of statistical efficiency (e.g., Korn and Graubard 1999, sec. 4.4), although the loss may be partly offset or more than offset by cost savings that stem from use of a sampling design that has weights.

Finally, and perhaps of greatest importance in some situations, a wider variety of statistical tests is available if there are no sample weights. In fact, when planning a study, one may not even yet know whether a lengthier menu of tests will be of benefit.

Usual logistic-regression analysis, based on asymptotic approximation, can be carried out in case-control studies when there is weighting (e.g., Korn and Graubard 1999, sec. 3.6). Asymptotic logistic regression, however, can be problematic even when there are no weights (Mehta and Patel 1995; King and Ryan 2002) and could become more so with the added complication of weights. But if (and only if) there are no sample weights to deal with, a remedy—exact conditional logistic regression (Mehta and Patel 1995)—does exist if needed. Moreover, even if a logistic-regression problem (without any weights) is so large that computer time for the exact calculations is prohibitive, results that are almost exact can be obtained with Monte Carlo methods (Mehta, Patel, and Senchaudhuri 2000).

For analyzing case-control studies with 1:1 matching, where each stratum has one case and one control, (asymptotic) logistic regression can be applied in a certain way if there is no weighting (e.g., Hosmer and Lemeshow 2000, chap. 7) but, unlike the situation of the previous paragraph, is not available at all if there are sample weights to contend with. Exact conditional logistic regression can likewise be run for 1:1 matching if weights are absent but not if they are present. For matching other than 1:1, asymptotic but not exact analysis is available without weights, but neither can be run if there are weights. Thus studies with matching may provide fertile ground for use of our weight-free control sampling methods.

In their standard form based on asymptotic approximation, other statistical tests are also unavailable [or at least virtually so—see Korn and Graubard (1999, sec. 3.10)] if there are sample weights. These tests include many nonparametric tests, among which are well-known rank tests.

Suppose that one wishes to use the Wilcoxon (Wilcoxon–Mann–Whitney) two-sample rank-sum test statistic to assess whether cases and controls differ with respect to a variable that is heavily skewed or is otherwise highly nonnormal. A sample of controls with unequal sample weights will not fit the framework and will thus prevent use of the test. If a study uses strata, either based on frequency matching or otherwise, then a linear combination of such Wilcoxon statistics across strata faces the same barrier.

Sample weights likewise interfere if one wants to run a Spearman test or a Kruskal–Wallis test. The former might (e.g.) test for association between a continuous exposure variable and an ordered categorical variable whose classes are the control category plus subgroups of cases distinguished by the severity of their disease. The latter might (e.g.) test for association between a continuous exposure variable and a categorical variable that has three or more unordered classes each composed of specified cases or controls.

For the Wilcoxon, Kruskal–Wallis, and Spearman tests as well as other rank tests, the option of calculating exact conditional results has been developed (Agresti 1992; Mehta and Patel 1998). Monte Carlo methods can be used for large problems. But if there are sample weights to grapple with, then the exact tests and the corresponding Monte Carlo procedures, like the standard asymptotic approximations, cannot be applied.

In case-control studies that have strata, each stratum may have a 2×2 table that shows the numbers of exposed cases and controls and unexposed cases and controls (as in the setup for the Mantel–Haenszel test). One may wish to test the homogeneity of the odds ratios across strata; one may want to obtain confidence limits for the common odds ratio (and test that it equals 1) if homogeneity can be assumed. For such situations, exact conditional results and their Monte Carlo counterparts are available (e.g., Volsett, Hirji, and Elashoff 1991; Mehta and Patel 1998). But they cannot be applied if the controls are sampled through a design that involves weights.

Software for performing the various types of exact conditional inference described above is available in (e.g.) SAS, SPSS or, especially, the products of Cytel Software Corporation. The software also provides Monte Carlo options that can be used for problems that are too large.

As a final example, an unusual type of analysis in case-control studies may use certain between-person distance measurements to try to assess whether cases are geographically clustered (Armitage, Berry, and Matthews 2002, p. 713). If the controls have sample weights, though, such an analysis cannot be done.

To weight or not to weight? If the sampling design for the controls draws them with unequal but known relative probabilities, then any analyses that ignore the associated weights will be faulty (and only those types of analyses that allow weights will be permissible to run). But if in the first place one chooses a design (like those proposed here) that avoids the need for weights, then the benefits noted above can be realized.

On the other hand, the designs that avoid weights do so only because they have frame exclusions. Are these exclusions acceptable? That issue is the topic of the next section.

7. BIAS?

Survey statisticians typically take great care to try to prevent or minimize bias in sample surveys. For cost and other reasons, the sampling designs often use unequal probabilities of selection. Then weights need to be applied or else parameter estimation will be subject to bias.

For our proposed schemes for sampling controls, frame reduction is the price that one has to pay to avoid weights. For a usual sample survey, selection bias stemming from frame exclusion would be a cause for concern. But does that also hold for case-control studies?

In overall remarks about exclusions, Kelsey et al. (1996, p. 208) stated that “Although excluding potential cases and controls limits generalizability, the validity of the comparison between cases and controls ... must take high priority.” They then went on to note the importance of the “general principle that the same exclusion criteria should be applied to cases and controls.”

Frame exclusion in case-control studies will not cause difficulty if, with respect to a potential explanatory variable, there is no *interaction* between case-control status and the variable(s) (e.g., lack of a telephone, or within-household status as set forth in Section 3) that distinguish inclusion from exclusion. That is, one supposes that any *differences* between cases and controls are the same in the excluded part of the population as in the included part. For example, interaction does (does not) exist if the case-versus-control difference in blood pressure, a possible explanatory variable, is (is not) different in the telephone and nontelephone populations, where the latter is excluded.

One might ask why it is more important to have the same frame for cases and controls than to have full representation of the population from which the cases arise. The answer is that a different frame for cases and controls can lead to a biased estimate of the case-versus-control main effect, or of the odds ratio, even if no interaction exists, whereas incomplete population representation will cause bias only if (and to the extent that) the interaction effect does exist.

One may want to be wary about frame exclusion if appreciable interaction is suspected. Furthermore, one may also want to avoid it where a study is aimed not only at assessing case-control differences but also, perhaps tangentially, at providing unbiased estimation of values that refer to the whole population (included and excluded combined). Otherwise, though, frame exclusion may be both appropriate and beneficial for the schemes covered in Sections 3 and 4 that sample for controls without a need for weights. When the schemes are used, however, one should describe the frame exclusions and report the estimated exclusion percentages.

8. SUMMARY

In the context of sampling for controls in population-based case-control studies, survey statisticians tend not to see the downside of weights, whereas epidemiologists tend to improperly disregard them when they arise. Provided that small frame reductions are tolerable, the methods of this article provide statistically valid ways, without the use of weights, to select a control within a household (by limiting the frame so that it includes at most one or two persons from a household) and to use list-assisted sampling to select the households themselves.

Acknowledgments

This work was partially supported by grants CA76016, CA33601, and 5P30-CA14236 from the National Cancer Institute. The authors thank the referee and the associate editor for their helpful comments.

References

- Agresti A. A Survey of Exact Inference for Contingency Tables (with comments). *Statistical Science* 1992;7:131–177.
- Armitage, P.; Berry, G.; Matthews, JNS. *Statistical Methods in Medical Research*. Vol. 4. Oxford, UK: Blackwell Science; 2002.
- Breslow NE. Statistics in Epidemiology: The Case-Control Study. *Journal of the American Statistical Association* 1996;91:14–28. [PubMed: 12155399]
- Brick JM, Waksberg J, Kulp D, Starer A. Bias in List-Assisted Telephone Samples. *Public Opinion Quarterly* 1995;59:218–235.
- Brogan DJ, Denniston MM, Liff JM, Flagg EW, Coates RJ, Brinton LA. Comparison of Telephone Sampling and Area Sampling: Response Rates and Within-Household Coverage. *American Journal of Epidemiology* 2001;153:1119–1127. [PubMed: 11390332]
- DiGaetano R, Waksberg J. Commentary: Trade-Offs in the Development of a Sample Design for Case-Control Studies. *American Journal of Epidemiology* 2002;155:771–775. [PubMed: 11943696]
- Dodge, Y., editor. *The Oxford Dictionary of Statistical Terms*. New York: Oxford University Press; 2003.
- Graubard BI, Fears TR, Gail MH. Effects of Cluster Sampling on Epidemiologic Analysis in Population-Based Case-Control Studies. *Biometrics* 1989;45:1053–1071. [PubMed: 2611317]
- Greenberg ER. Random Digit Dialing for Control Selection: A Review and a Caution on Its Use in Studies of Childhood Cancer. *American Journal of Epidemiology* 1990;131:1–5. [PubMed: 2403464]
- Hartge P, Brinton LA, Rosenthal JF, Cahill JJ, Hoover RN, Waksberg J. Random Digit Dialing in Selecting a Population-Based Control Group. *American Journal of Epidemiology* 1984;120:825–833. [PubMed: 6334439]
- Hosmer, DW.; Lemeshow, S. *Applied Logistic Regression*. Vol. 2. New York: Wiley; 2000.
- Kelsey, JL.; Whittemore, AS.; Evans, AS.; Thompson, WD. *Methods in Observational Epidemiology*. Vol. 2. New York: Oxford University Press; 1996.
- King EN, Ryan TP. A Preliminary Investigation of Maximum Likelihood Logistic Regression Versus Exact Logistic Regression. *The American Statistician* 2002;56:163–170.
- Korn EL, Graubard BI. Examples of Differing Weighted and Unweighted Estimates From a Sample Survey. *The American Statistician* 1995;49:291–295.
- Korn, EL.; Graubard, BI. *Analysis of Health Surveys*. New York: Wiley; 1999.
- Mehta CR, Patel NR. Exact Logistic Regression: Theory and Examples. *Statistics in Medicine* 1995;14:2143–2160. [PubMed: 8552893]
- Mehta, CR.; Patel, NR. Exact Inference for Categorical Data. In: Armitage, P.; Colton, T., editors. *Encyclopedia of Biostatistics*. Chichester, UK: Wiley; 1998. p. 1411–1422.
- Mehta CR, Patel NR, Senchaudhuri P. Efficient Monte Carlo Methods for Conditional Logistic Regression. *Journal of the American Statistical Association* 2000;95:99–108.
- Mitofsky WJ. The State of State Election Polls. *Chance* 1993;6(1):9–16.
- Nielsen Media Research. Cell Phone Sampling Summit II, February 3–4, 2005. 2005 [(accessed March 26, 2007)]. <http://www.nielsenmedia.com/cellphonesummit/cellphone.html>
- Potthoff RF. Telephone Sampling in Epidemiologic Research: To Reap the Benefits, Avoid the Pitfalls. *American Journal of Epidemiology* 1994;139:967–978. [PubMed: 8178785]
- Scott AJ, Wild CJ. Case-Control Studies With Complex Sampling. *Applied Statistics* 2001;50:389–401.
- Vollset SE, Hirji KF, Elashoff RM. Fast Computation of Exact Confidence Limits for the Common Odds Ratio in a Series of 2×2 Tables. *Journal of the American Statistical Association* 1991;86:404–409.
- Waksberg, J. Random Digit Dialing Sampling for Case-Control Studies. In: Armitage, P.; Colton, T., editors. *Encyclopedia of Biostatistics*. Chichester, UK: Wiley; 1998. p. 3678–3682.