



PRACTICE OF EPIDEMIOLOGY

Comparison of Telephone Sampling and Area Sampling: Response Rates and Within-Household Coverage

Donna J. Brogan,¹ Maxine M. Denniston,² Jonathan M. Liff,³ Elaine W. Flagg,⁴ Ralph J. Coates,⁵ and Louise A. Brinton⁶

Random digit dialing is used frequently in epidemiologic case-control studies to select population-based controls, even when both cases and controls are interviewed face-to-face. However, concerns persist about the potential biases of random digit dialing, particularly given its generally lower response rates. In an Atlanta, Georgia, case-control study of breast cancer among women aged 20–54 years, all of whom were interviewed face-to-face, two statistically independent control groups were compared: those obtained through random digit dialing ($n = 652$) and those obtained through area probability sampling ($n = 640$). The household screening rate was significantly higher for the area sample, by 5.5%. Interview response rates were comparable. The telephone sample estimated a significantly larger percentage (by approximately 7%) of households to have no age-eligible women. Both control groups, appropriately weighted, had characteristics similar to US Census demographic characteristics for Atlanta women, except that respondents in both control groups were more educated and more likely to be married. The authors conclude that households contacted through random digit dialing are somewhat less likely to participate in the household screening process, and if they are cooperative, some households may not disclose that age-eligible women reside therein. Investigators need to develop improved methods for screening and enumerating household members in random digit dialing surveys that target a specific subpopulation, such as women. *Am J Epidemiol* 2001;153:1119–27.

case-control studies; data collection; epidemiologic methods; interviews; random digit dialing; sampling studies; selection bias

Case-control studies carried out in a designated geographic area often use population-based controls (1). Random digit dialing (RDD) (2, 3) has been preferred over area probability sampling (APS) (4, 5) for obtaining controls, because of its assumed lower cost and improved interviewer safety (6). However, potential biases of RDD include noncoverage of households without a telephone (7), lower

survey response rates (8), and noncoverage of some household members (9, 10).

Noncoverage of households without telephones is not a major concern in case-control studies that use RDD, since 95 percent of US households have at least one telephone (11) and because case subjects without a telephone can be excluded from analyses (2). The typically lower response rate in RDD is of greater concern, because respondents and nonrespondents may differ more than with other control sampling techniques (12, 13). In addition, there may be differences between household members who are and are not identified in the RDD household screening process.

Most RDD surveys use the telephone both for sampling and for conducting interviews. However, case-control studies may use face-to-face interviewing, no matter how the control subjects are sampled (14). Furthermore, case-control studies, unlike general population surveys, generally have stringent eligibility criteria (e.g., age, gender), involving extensive screening of households in order to locate control subjects (2).

We are aware of only two studies that have compared RDD and APS using equivalent data collection modes (15, 16), and only one of those studies, a relatively small one (15), used face-to-face interviewing. Lele et al. (15) found that RDD and APS controls, stratified by age and sex, did not differ in terms of demographic characteristics, height,

Received for publication December 6, 1999, and accepted for publication September 29, 2000.

Abbreviations: APS, area probability sampling; PUMS, Public Use Microdata Sample; RDD, random digit dialing.

¹Department of Biostatistics, Rollins School of Public Health, Emory University, Atlanta, GA.

²Behavioral Research Center, American Cancer Society, Atlanta, GA.

³Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA.

⁴Department of Medicine, School of Medicine, Emory University, Atlanta, GA.

⁵Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA.

⁶Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.

Correspondence to Dr. Donna Brogan, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322 (e-mail: dbrogan@sph.emory.edu).

weight, or cigarette smoking. In the second study, Aquilino and Wright (16) found that APS and RDD sample controls, all interviewed via telephone, did not differ substantially in terms of demographic characteristics or self-reported drug use.

Regarding coverage of household residents, Maklan and Waksberg (9) found fewer households with only one adult (a female) in their national RDD survey in comparison with the APS-designed Current Population Survey. Massey (10) found that RDD estimated 4.1 percent of households to have children aged 19–35 months, compared with estimates of 5 percent based on several APS surveys.

We examined these sampling methods using data from a case-control study of breast cancer in younger women. Our research provided us with a unique opportunity to compare two statistically independent control groups with each other and with data from the decennial US Census. We assessed whether two sampling techniques, RDD and APS, were equivalent in terms of 1) response rates to household screening and to face-to-face interview and 2) within-household coverage of younger women. Furthermore, we assessed whether RDD and APS gave results comparable to sample US Census data on the demographic and personal characteristics of these women.

MATERIALS AND METHODS

Case-control methods

The Women's Interview Study of Health (WISH) was a multicenter case-control study of breast cancer in younger women (aged 20–54 years) conducted during 1990–1992 in three counties of metropolitan Atlanta (Georgia), in Seattle (Washington), and in several New Jersey counties (17). RDD was used to sample controls at all three sites. Since APS for controls was conducted only in Atlanta, all comparisons of RDD and APS were restricted to this site.

Procedures for selecting the RDD and APS control groups were as comparable as possible. Both control groups were frequency-matched to the anticipated age distribution of cases. The household screener question was the same for both procedures. The same personnel conducted face-to-face interviews for all subjects, using identical survey instruments. Interviewees included 777 cases, 652 RDD controls, and 640 APS controls. Because there were few women of races other than White or Black, White women and women of other races were combined into the category "Nonblack." Furthermore, because few women were aged 20–29 years, analyses comparing RDD and APS interviewed women with US Census data were restricted to women aged 30–54 years ($n = 640$ for RDD and $n = 626$ for APS).

RDD methods

The Mitofsky-Waksberg technique (18), using banks (clusters) of 100 telephone numbers, provided an equal probability sample of households; any household with $k > 1$ residential telephone numbers was subsampled with probability $1/k$. No stratification of banks was used, and at the

time one area code covered all three Atlanta counties. The screener question for county-eligible residences was: "How many women living in this household (including yourself) are 20 to 54 years old?" All age-eligible women were enumerated, and information on age, name, and address was requested. Subsequently, a sample was selected from the enumerated women; older women were assigned higher selection probabilities so that age frequency-matching to cases could be achieved.

APS methods

Standard multistage APS techniques were used to obtain an equal probability sample of housing units. Primary sampling units were block groups; 180 were selected. Second-stage units were segments (one or more blocks); we selected only one segment per primary sampling unit in order to maximize the geographic dispersion of the sample. After each segment was mapped and housing units were listed, a systematic random sample of approximately 24 housing units was selected.

After an explanatory letter was mailed to each sample housing unit, a female interviewer visited the household, asked the same screening question as that used in RDD, enumerated all eligible women by age and by name or initials, and requested the household's telephone number. A predetermined random selection scheme immediately indicated whether any enumerated women were selected for interview. Details on the APS and RDD methods are provided elsewhere by Brogan (19) and Denniston (20).

Public Use Microdata Sample

The 5 percent Public Use Microdata Sample (PUMS) (21) from the 1990 US Census provided information on household composition and on female residents aged 30–54 years (age, race, marital status, education, number of livebirths, and household income). PUMS is a stratified random sample of 5 percent of US housing units and all persons residing therein; it contains questions beyond those on the standard Census form. PUMS is weighted so that household- or person-based estimates are consistent with the 100 percent Census data. Our PUMS sample size for the three counties was 27,124 households and 14,086 women aged 30–54 years.

Weighting methods for RDD and APS samples

Ordinarily, a population-based control sample is not weighted in case-control analyses of risk factors, although analysts may recognize clustering of subjects (22). However, we weighted our RDD and APS samples so we could make inferences to the three-county area and compare the data with the PUMS data. The final survey weight for each interviewed woman is the number of women in the population whom she represents.

Standard sample survey weighting procedures were based on each sampling method's probabilities of selecting households and enumerated women. Internal (to the sample) weighting adjustments were made for nonresponse at the

household screening and interview stages. We did not post-stratify the RDD and APS samples to Census data, nor did we perform noncoverage adjustment for the RDD sample (7), since we wished to compare RDD and APS with PUMS.

Definition and analysis of response rates

The *screening response rate* is the percentage of households (occupied housing units or residential telephone numbers) with complete screening. "Complete" was defined as a response to the screener question and enumeration of all age-eligible women. "Partial screening," in which the names and/or addresses of age-eligible women were not given, occurred only in RDD. Partial screening was defined as not complete if a woman in the household was selected for interview but could not be located (23, 24); otherwise, it was considered complete. The *interview response rate* is the percentage of selected eligible women who were interviewed; some selected women were determined later to be ineligible because of age or county of residence. The *overall survey response rate* is the product of the screening and interview response rates. Chi-squared tests were used to compare the screening, interview, and overall survey response rates for the APS and RDD samples, using unweighted analyses and assuming all households and women to be statistically independent.

Definition and analysis of within-household coverage

A linear contrast was used to compare the APS, RDD, and PUMS samples according to their estimates of the percentage of households that had no women aged 20–54 years. The unit of analysis was the household for all three samples, with households being weighted and clustered in APS and PUMS (25). Since the RDD contractor did not provide data with which to classify all RDD screened households by telephone bank (primary sampling unit), we accounted for the clustering of RDD households by using the design effect approach (4). We estimated the variance of our RDD point estimate by assuming simple random sampling of households and then multiplied the simple random sampling variance by an assumed design effect of 2.0. The design effect for the RDD sample is unlikely to have been larger than 2.0, since the Mitofsky-Waksberg design tends to have small intracluster correlation coefficients (18, 23).

Characteristics of women

For a given characteristic, we used weighted and clustered chi-squared analyses to test the null hypothesis that the RDD, APS, and PUMS samples made inference to the same population of women aged 30–54 years. If this null hypothesis was rejected, linear contrasts were used to determine which of the three samples differed from each other. For residential telephone coverage, only the APS and PUMS samples were compared. Analyses accounted for weighting as well as clustering of women in telephone bank, segment, and housing units for the RDD, APS, and PUMS samples, respectively. SUDAAN (26) was used for all weighted and clustered analyses of households or interviewed women.

RESULTS

Response rates

The APS screening response rate was 94.9 percent (3,150/3,318), with contact problems and refusal being the most common reasons for screening failure (table 1). Of the 802 women selected for interview in APS, 794 were eligible. The interview response rate was 80.6 percent (640/794), with refusal ($n = 105$) being the most common reason for noninterview. The overall APS survey response rate was 76.5 percent (0.949×0.806).

The RDD screening rate was 89.4 percent ($((4,572 + 355 - 42)/5,464)$, based on 4,572 complete screens and 355 partial screens (table 1). The most common reason for RDD screening failure was refusal (6.3 percent; 304 refusals to interview plus 42 incomplete partial screens). Of the 898 women selected for interview, 818 were eligible, yielding an RDD interview response rate of 79.7 percent (652/818); refusal ($n = 113$) was the primary reason for noninterview. The overall RDD response rate was 71.2 percent (0.894×0.797).

In the RDD sample, 7.2 percent (355/4,927) of all household screenings were "partial," and 12.3 percent (355/2,894) of eligible household screenings were "partial." The partial screenings did not substantially lower the RDD screening response rate, because no women were selected for interview in most of those households.

The APS screening response rate was significantly higher ($p = 0.001$) than that of RDD by 5.5 percent (table 1). Screening failure due to noncontact was somewhat higher in RDD (3.8 percent) than in APS (3.2 percent), and the screening refusal rate was significantly higher ($p = 0.001$) for RDD (6.3 percent) than for APS (1.7 percent). The interview response rates for the two samples were remarkably similar. Because of differential screening rates, the overall survey response rate was significantly higher ($p = 0.001$) for APS (76.5 percent) than for RDD (71.2 percent).

Within-household coverage for RDD, APS, and PUMS

The estimated percentages of households with no women aged 20–54 years were 41.3 (95 percent confidence interval: 39.3, 43.2) for RDD, 34.8 (95 percent confidence interval: 31.5, 38.3) for APS, and 33.5 (95 percent confidence interval: 32.9, 34.0) for PUMS. The APS and PUMS samples yielded consistent confidence intervals, but the RDD point estimate was higher and its confidence interval had no overlap with the APS and PUMS confidence intervals. In the linear contrast comparing the average of the APS and PUMS point estimates with the RDD point estimate, the result differed significantly from zero ($z = 5.33, p < 0.0001$).

Compared with the APS and PUMS samples, the RDD screening process appears to have missed the identification of approximately 11 percent ($((0.66 - 0.59)/0.66)$) of eligible households. This drives up the cost of RDD by increasing the sample size of screened households, in effect "substituting" eligible households who identify themselves as such for those eligible households who identify themselves as ineligible.

TABLE 1. Outcome of household screening using two different sampling methods, by sample type, metropolitan Atlanta, Georgia, 1990–1992*

Outcome	No. (or %) of housing units/telephones with outcome	
	Area probability sampling	Random digit dialing
Total no. of housing units/telephones	3,804	12,033
Total no. noneligible	486	6,542
Not a housing unit (APS†)	25	
Vacant housing unit (APS)	461	
Nonresidential number (RDD‡)		2,176
Outside of three-county study area (RDD)		408
Nonworking number (RDD)		3,563
No answer, nonresidential (RDD)		395
Total no. nonallocated (RDD)		49
Total no. of households not screened	168	515
No one ever at home or unavailable (APS)	106	
No answer, residential number (RDD)		112
Language problem	6	27
Refusal	56	304
Maximum contact (RDD)		72
Total no. of households screened	3,150	4,927
Complete screen	3,150	4,572
No eligible woman in household	1,097	2,033
Eligible woman/women in household	2,053	2,539
Partial‡ screen	0	355
Total no. of households eligible	3,318	5,464§
Percentage of eligible households		
Refusal	1.7	6.3
Language problem	0.2	0.5
Contact problem	3.2	3.8
Successful screen	94.9	89.4

* Data were obtained from a case-control study of breast cancer among women aged 20–54 years (17).

† APS, area probability sampling; RDD, random digit dialing.

‡ Numbers and ages of women were obtained but their names and/or addresses were not obtained. Forty-two of these women were selected for interview but could not be identified for interview.

§ The total number of households eligible for screening was 5,464.25, with 22.25 of the “nonallocated” households counted as eligible and not contacted.

Women's characteristics in RDD, APS, and PUMS

The estimated percentages of Black women given by the three samples did not differ significantly ($p = 0.17$): PUMS = 34.9 percent, APS = 33.5 percent, and RDD = 28.8 percent (table 2). The samples estimated equivalent distributions for age ($p = 0.60$; table 2), number of livebirths ($p = 0.71$; table 3), and household income ($p = 0.75$; table 3). When analyses were stratified by race, the samples still esti-

mated equivalent distributions for age (table 2), number of livebirths (table 3), and household income (table 3).

The estimated percentages of women who had graduated from high school (the graduation rate) differed significantly among the three samples ($p < 0.0005$): RDD = 93.2 percent, APS = 91.0 percent, and PUMS = 88.1 percent (table 4). Linear contrasts indicated that the RDD and APS rates did not differ significantly ($z = 1.05$, $p = 0.29$), but the PUMS graduation rate was significantly lower than the mean of the RDD and APS rates ($z = 3.71$, $p < 0.0002$). Stratified analyses by race showed the same findings (table 4), as did analyses of educational attainment as a multilevel ordinal variable (data not shown).

The estimated distributions of women by marital status differed significantly ($p < 0.0005$) among the three samples (table 5). Linear contrasts indicated that RDD and APS did not differ significantly in terms of the estimated percentage of married women ($z = 0.12$, $p = 0.90$), whereas the PUMS estimate of 59.7 percent was significantly lower than the average (67.6 percent) of the RDD and APS estimates ($z = 3.03$, $p < 0.003$). In analyses stratified by race (table 5), the samples differed significantly with regard to marital status for Nonblacks ($p < 0.0005$) but not for Blacks ($p = 0.30$). Follow-up linear contrasts for Nonblacks showed the same results as those for all races.

The estimated percentages of women living in households with a telephone were 97.4 for PUMS and 97.5 for APS ($p = 0.92$); no statistically significant differences were found in analyses stratified by race. Thus, approximately 97 percent of women aged 30–54 years lived in households with a telephone (99 percent of Nonblacks and 94 percent of Blacks).

DISCUSSION

The findings of our study are consistent with other observations of a higher household screening rate in APS compared with RDD (27). Our APS and RDD screening rates were close to or exceeded other rates reported (9, 15, 16, 23, 24). Our face-to-face interview response rates were comparable for APS and RDD, and at 80 percent they compared favorably with those of similar studies (15, 23, 24). Aquilino and Wright (16) obtained differential telephone interview rates for their RDD and APS samples (74 percent and 80 percent, respectively), which suggests a more favorable response from households that are first contacted in person rather than by telephone.

Several factors may have contributed to the higher screening rate with APS. First, household noncontact was somewhat larger in RDD than in APS, probably because of the increasing use of answering machines and caller identification and call blocking services (27). Second, the screening refusal rate was substantially higher with RDD, similar to the findings of Aquilino (28); it is easier to refuse survey participation over the telephone than to an interviewer at your doorstep (27, 29). Third, the contact with households by letter prior to the interviewer's visit may have made APS households more likely to be screened. As Groves and Couper (27) also reported, our interviewers believed that the

TABLE 2. Estimated distribution of women aged 30–54 years according to age and race, by sample type (weighted/clustered analysis), metropolitan Atlanta, Georgia, 1990–1992

Race, sample type, and age group (years)	Sample size	Percentage	Standard error
All races	$\chi^2 = 4.56\ddagger$ (6 df); $p = 0.60$		
PUMS*			
30–39	6,985	50.30	0.44
40–44	3,207	22.45	0.37
45–49	2,255	15.79	0.32
50–54	1,639	11.46	0.28
RDD*			
30–39	118	48.50	2.69
40–44	153	23.66	1.76
45–49	172	14.76	1.30
50–54	197	13.08	1.08
APS*			
30–39	129	50.55	4.61
40–44	166	23.07	2.41
45–49	193	15.40	1.79
50–54	138	10.97	1.35
Blacks	$\chi^2 = 10.66$ (6 df); $p = 0.10$		
PUMS			
30–39	2,646	56.04	0.75
40–44	996	20.85	0.62
45–49	645	13.29	0.52
50–54	482	9.82	0.44
RDD			
30–39	38	51.47	4.97
40–44	49	26.69	3.49
45–49	39	11.87	2.16
50–54	44	9.97	1.72
APS			
30–39	55	49.04	4.98
40–44	73	28.99	3.32
45–49	56	13.41	1.90
50–54	36	8.56	1.52
Nonblacks	$\chi^2 = 4.73$ (6 df); $p = 0.58$		
PUMS			
30–39	4,339	47.22	0.54
40–44	2,211	23.31	0.45
45–49	1,610	17.13	0.41
50–54	1,157	12.34	0.35
RDD			
30–39	80	47.29	3.25
40–44	104	22.43	2.03
45–49	133	15.93	1.66
50–54	153	14.35	1.37
APS			
30–39	74	51.31	6.40
40–44	93	20.09	2.99
45–49	137	16.41	2.55
50–54	102	12.19	2.02
All races	$\chi^2 = 3.58$ (2 df); $p = 0.17$		
PUMS—% Black	9,317	34.92	0.44
RDD—% Black	470	28.83	3.19
APS—% Black	406	33.52	5.31

* PUMS, Public Use Microdata Sample; RDD, random digit dialing; APS, area probability sampling.

† Chi-squared test for independence of sample type and age group (percentages) or race (percentages).

advance letter made their first household visit less difficult. Reverse directories could be used to send advance letters to RDD-selected telephone numbers before calling (11, 29), but this procedure adds considerable expense; in addition, Brick and Collins (30) were able to match only 45 percent of their sampled telephone numbers to addresses. Fourth, a request to enumerate all household residents or all age-eligible household residents generally leads to increased screening response with APS (31) as compared with RDD (11, 30, 32), probably because the interviewer visits the household. Harlow and Hartge (33) reported that the RDD screening rate was 16 percent lower when the full names of women aged 30–69 years were requested, compared with not asking for full names. This may have been a factor in our study, since the APS procedure accepted the initials of eligible women if the screening respondent was reluctant to give names. If a woman with initials only was selected for interview, the APS enumeration interviewer immediately requested her name. Probability sampling of women for interview at the time of RDD screening, rather than 1–3 months later as in our study, could reduce the need for identifying information such as name and/or address at RDD screening and increase the RDD screening response rate.

Because we minimized the number of screening questions, we had limited data on nonscreened households and on households with no age-eligible women. Hence, it was not possible to determine whether the RDD/APS differential in screening rate was related to household characteristics such as race, household composition, or socioeconomic status. With increasing concern about privacy and personal security, the lower RDD screening rate, as well as the non-negligible “partial” screening rate, may be partly attributed to reluctance to reveal household information over the telephone, particularly when questioned about younger (aged 20–54 years) female residents.

RDD estimated a significantly higher percentage of households to have no women aged 20–54 years (41 percent), compared with either APS or PUMS (35 percent and 34 percent, respectively). We conclude that RDD screening respondents underreport households that contain at least one woman aged 20–54 years. Our finding is consistent with that of Massey (10) for households containing young children in the National Immunization Survey and with that of Maklan and Waksberg (9) for adult females in a national survey. Massey (10) concluded that the RDD underestimation was due not to exclusion of households without a telephone but rather to failure of households to identify age-eligible children as a way of declining survey participation.

A possible alternate explanation for the higher percentage of ineligible households with RDD is that eligible households were overrepresented among the nonscreened households. If we make the extreme assumption that all RDD households refusing screening ($n = 304$) were eligible, then the RDD screening response rate would have been 95.0 percent, and the estimated percentage of ineligible households would have been 38.9 percent (95 percent confidence interval: 36.9, 40.9); this confidence interval still lies above the PUMS confidence interval. Thus, this alternate explanation cannot totally explain the higher percentage of ineligible households found by RDD.

TABLE 3. Estimated distribution of women aged 30–54 years according to number of livebirths and household income, by sample type (weighted/clustered analysis), metropolitan Atlanta, Georgia, 1990–1992

Livebirths				Income			
Race, sample type, and no. of livebirths (ever)	Sample size	Percentage	Standard error	Race, sample type, and annual household income	Sample size	Percentage	Standard error
All races	$\chi^2 = 5.40\ddagger$ (8 df); $p = 0.71$			All races	$\chi^2 = 5.10\ddagger$ (8 df); $p = 0.75$		
PUMS*				PUMS			
0	3,347	24.34	0.39	<\$20,000	2,272	16.32	0.33
1	2,697	19.11	0.35	\$20,000–\$34,999	2,803	20.21	0.36
2	4,473	31.19	0.41	\$35,000–\$49,999	2,628	19.02	0.36
3	2,147	15.25	0.32	\$50,000–\$69,999	2,823	20.00	0.36
≥4	1,422	10.11	0.27	≥\$70,000	3,560	24.44	0.38
RDD*				RDD			
0	100	21.48	2.61	<\$20,000	85	14.75	2.34
1	113	20.31	2.35	\$20,000–\$34,999	115	19.76	2.38
2	204	31.64	2.58	\$35,000–\$49,999	124	19.47	2.28
3	133	16.10	1.87	\$50,000–\$69,999	105	17.37	2.14
≥4	90	10.47	1.63	≥\$70,000	185	28.66	2.74
APS*				APS			
0	112	18.94	2.77	<\$20,000	109	18.13	3.76
1	106	16.07	2.54	\$20,000–\$34,999	102	19.63	4.91
2	210	41.52	5.51	\$35,000–\$49,999	112	23.64	5.12
3	119	14.82	1.99	\$50,000–\$69,999	116	17.26	2.52
≥4	79	8.65	1.58	≥\$70,000	169	21.34	3.31
Blacks	$\chi^2 = 2.97$ (8 df); $p = 0.94$			Blacks	$\chi^2 = 5.95$ (8 df); $p = 0.65$		
PUMS				PUMS			
0	837	18.08	0.60	<\$20,000	1,480	30.40	0.71
1	1,020	21.37	0.62	\$20,000–\$34,999	1,283	27.02	0.70
2	1,381	28.82	0.69	\$35,000–\$49,999	896	19.14	0.63
3	777	16.18	0.56	\$50,000–\$69,999	704	14.81	0.56
≥4	754	15.55	0.55	≥\$70,000	406	8.63	0.46
RDD				RDD			
0	25	18.95	4.57	<\$20,000	50	31.22	5.17
1	37	22.45	4.02	\$20,000–\$34,999	37	26.44	4.94
2	34	22.35	4.19	\$35,000–\$49,999	30	16.72	3.87
3	33	16.14	3.05	\$50,000–\$69,999	23	13.94	3.11
≥4	41	20.12	4.57	≥\$70,000	21	11.69	3.53
APS				APS			
0	32	19.13	3.93	<\$20,000	80	40.97	7.45
1	37	19.31	4.06	\$20,000–\$34,999	47	19.21	3.57
2	64	28.84	3.65	\$35,000–\$49,999	38	20.69	4.72
3	42	17.13	3.20	\$50,000–\$69,999	27	11.09	2.56
≥4	45	15.58	3.49	≥\$70,000	23	8.03	2.81

Table continues

Note that failure to answer honestly about household composition is not the same thing as refusing to provide screening information. In fact, failure to disclose age-eligible female residents of the household is likely to increase the apparent screening response rate, since the screening respondent is able to successfully complete the screen by providing false information, as opposed to refusing to participate in the screening process.

Although RDD estimates were always lower for percentage of women who were Black, most of the comparisons of RDD, APS, and PUMS with regard to racial distribution did

not reach statistical significance, probably because of the large and moderate design effects in APS and RDD, respectively. The three samples differed consistently regarding education, overall and stratified by race, with RDD and APS being similar and yielding consistently higher high school graduation rates than PUMS. Perhaps less educated women were less likely to be identified in the household screening process or to agree to be interviewed, whether in RDD or APS. Madigan et al. (34), in their study of RDD respondents and nonrespondents in the Women's Interview Study of Health, found respondents to be more highly edu-

TABLE 3. Continued

Livebirths				Income			
Race, sample type, and no. of livebirths (ever)	Sample size	Percentage	Standard error	Race, sample type, and annual household income	Sample size	Percentage	Standard error
$\chi^2 = 9.10$ (8 df); $p = 0.33$				$\chi^2 = 3.91$ (8 df); $p = 0.87$			
Nonblacks				Nonblacks			
PUMS				PUMS			
0	2,510	27.70	0.49	<\$20,000	792	8.76	0.31
1	1,677	17.90	0.41	\$20,000–\$34,999	1,520	16.56	0.41
2	3,092	32.46	0.50	\$35,000–\$49,999	1,732	18.95	0.44
3	1,307	14.75	0.39	\$50,000–\$69,999	2,119	22.79	0.47
≥4	668	7.19	0.28	≥\$70,000	3,154	32.93	0.51
RDD				RDD			
0	75	22.50	3.16	<\$20,000	35	8.08	2.18
1	76	19.44	2.93	\$20,000–\$34,999	78	17.06	2.62
2	170	35.41	3.23	\$35,000–\$49,999	94	20.58	2.75
3	100	16.09	2.33	\$50,000–\$69,999	82	18.76	2.71
≥4	49	6.56	1.11	≥\$70,000	164	35.53	3.41
APS				APS			
0	80	18.84	3.58	<\$20,000	29	6.58	2.07
1	69	14.44	3.08	\$20,000–\$34,999	55	19.84	7.13
2	146	47.91	7.23	\$35,000–\$49,999	74	25.12	7.27
3	77	13.65	2.37	\$50,000–\$69,999	89	20.38	3.61
≥4	34	5.16	1.38	≥\$70,000	146	28.07	4.99

* PUMS, Public Use Microdata Sample; RDD, random digit dialing; APS, area probability sampling.

† Chi-squared test for independence of sample type and livebirths (percentages) or household income (percentages).

TABLE 4. Estimated percentage of high school graduates among women aged 30–54 years, by race and sample type (weighted/clustered analysis), metropolitan Atlanta, Georgia, 1990–1992

Race and sample type	Sample size	% of high school graduates	Standard error
$\chi^2 = 18.90$ † (2 df); $p < 0.0005$			
All races			
PUMS*	12,413	88.11	0.29
RDD*	572	93.25	1.14
APS*	556	91.04	1.76
$\chi^2 = 8.06$ (2 df); $p = 0.02$			
Blacks			
PUMS	3,773	79.90	0.62
RDD	135	88.20	2.82
APS	174	83.90	3.08
$\chi^2 = 7.54$ (2 df); $p = 0.02$			
Nonblacks			
PUMS	8,640	92.52	0.30
RDD	437	95.30	0.97
APS	382	94.65	1.96

* PUMS, Public Use Microdata Sample; RDD, random digit dialing; APS, area probability sampling.

† Chi-squared test for independence of sample type and high school graduation (percentages).

cated, supporting the possibility of an effect due to interview response.

Among Blacks, there was no difference between marital status distributions estimated by the three samples. Among

Nonblacks, however, RDD and APS were similar but estimated a significantly higher percentage of currently married women and a significantly lower percentage of single (never married) women than did the PUMS sample. Perhaps a greater reluctance of single Nonblack women to be screened and/or interviewed, whether sampled by RDD or by APS, could explain part of this difference.

Our study used Mitofsky-Waksberg RDD sampling, whereas list-assisted RDD sampling (35) is generally used today, because it is assumed to be less expensive and less difficult logistically. Our findings regarding response rates, within-household coverage, and characteristics of interviewed women should apply equally well to list-assisted RDD, for several reasons. First, list-assisted RDD, without truncation, has the same sampling frame coverage of residential telephone numbers as does the Mitofsky-Waksberg method. Second, list-assisted RDD, with truncation of zero banks, has minimum sampling frame undercoverage of residential telephone numbers (35). Third, contacted households and individuals selected for interview are not aware of which RDD technique was used to reach them.

Although we were not able to make precise cost comparisons between RDD and APS, our impression is that the cost of APS sampling would compare favorably to that of RDD sampling in our situation. The relatively small geographic area of three adjacent metropolitan counties reduced the typical APS field cost of counting and listing housing units. In contrast, the small geographic area increased typical RDD sampling costs, because it was necessary to screen for only three counties in an area code that covered approximately

TABLE 5. Estimated distribution of women aged 30–54 years according to marital status, by race and sample type (weighted/clustered analysis), metropolitan Atlanta, Georgia, 1990–1992

Race, sample type, and marital status	Sample size	Percentage	Standard error
All races	$\chi^2 = 25.08\ddagger$ (6 df); $p < 0.0005$		
PUMS*			
Married	8,509	59.71	0.44
Widowed/separated	1,012	7.24	0.23
Divorced	2,572	18.51	0.35
Single	1,993	14.54	0.32
RDD*			
Married	428	67.88	2.75
Widowed/separated	52	7.65	1.58
Divorced	115	16.89	2.07
Single	45	7.59	1.55
APS*			
Married	405	67.26	4.31
Widowed/separated	44	4.96	1.17
Divorced	127	17.17	2.41
Single	50	10.60	2.81
Blacks	$\chi^2 = 7.21$ (6 df); $p = 0.30$		
PUMS			
Married	2,039	43.14	0.76
Widowed/separated	705	14.40	0.53
Divorced	1,008	21.24	0.62
Single	1,017	21.21	0.63
RDD			
Married	81	41.07	4.48
Widowed/separated	27	16.41	4.22
Divorced	36	27.81	4.75
Single	26	14.71	3.56
APS			
Married	99	43.75	5.48
Widowed/separated	33	10.61	2.75
Divorced	58	25.83	3.84
Single	30	19.82	6.15
Nonblacks	$\chi^2 = 26.29$ (6 df); $p < 0.0005$		
PUMS			
Married	6,470	68.60	0.52
Widowed/separated	307	3.39	0.20
Divorced	1,564	17.05	0.41
Single	976	10.96	0.36
RDD			
Married	347	78.75	2.71
Widowed/separated	25	4.09	1.31
Divorced	79	12.46	1.90
Single	19	4.70	1.61
APS			
Married	306	79.12	4.14
Widowed/separated	11	2.12	0.84
Divorced	69	12.81	2.63
Single	20	5.95	2.26

* PUMS, Public Use Microdata Sample; RDD, random digit dialing; APS, area probability sampling.

† Chi-squared test for independence of sample type and marital status (percentages).

half of the state. Furthermore, the travel costs for face-to-face home interviews were probably greater with the RDD sample, because the addresses tended to be more geographically dispersed in comparison with APS.

RDD and APS are not the only methods that can yield a population-based sample for a case-control study. Other sampling frames include Medicare enrollees (for older persons), licensed drivers within a state, addresses with utility (electric, gas, or water) service, and address records at county property tax offices. The strengths, limitations, and costs of each potential sampling frame depend on many factors, including the size of the geographic area under study and the characteristics of the study population therein.

Our study had several strengths. It was designed so that the RDD and APS sampling and screening procedures were as identical as possible. Fieldwork was conducted in 1990–1992, allowing comparison with the 1990 US Census sample as an external standard. Our sample size of approximately 650 in each control group was much larger than that in the Washington State study (15) and, in addition, included a sizable minority population. However, our study was limited by being 1) restricted to only one metropolitan, southern geographic area, 2) restricted to women aged 20–54 years, and 3) unable to provide precise estimates of the cost differential between RDD and APS.

Even though list-assisted methods promise greater efficiency (11, 35, 36), RDD sampling is becoming more difficult and expensive. Several secular trends in the United States have contributed to declining survey response rates (27), particularly in the initial contact phase for RDD. These trends include: 1) more households containing single adults, 2) more households with all adults in the labor force, 3) the increased mobility of the population, 4) household residents' spending more time away from home, 5) increased time commitments of household members, and 6) increased hostility to telemarketers. Massey et al. (8) noted that only a small percentage of RDD surveys currently obtain a true overall response rate of more than 70 percent, and surveys that involve household screening have even lower overall response rates.

Furthermore, rapid changes in communication technology are having a dramatic impact on the cost of constructing and using RDD sampling frames. The proliferation of fax machines, beepers, home computers, and cellular/mobile phones means that the sampling frame contains a much larger percentage of nonresidential telephone numbers. A growing number of people possess only a cellular/mobile phone, so exclusion of cell phone telephone exchanges from the sampling frame results in noncoverage.

It is clear from this study and from several others (8, 13) that research is needed on improving the household screening and enumeration process in RDD, especially for surveys that require detailed information about household composition prior to the sampling of persons within the household (8). Such research should separate the household contact phase (an adult answers the phone) from the household cooperation phase (an adult provides information on household composition), since different factors and approaches seem to be relevant. Groves and Couper (27) make a compelling argument for research on "tailoring" approaches to

households in both the contact and cooperation stages, using any known information about the household to personalize repeated attempts to obtain contact and cooperation. It may be necessary for research on both RDD and APS to evolve in this direction so that scientifically acceptable response rates can be maintained.

ACKNOWLEDGMENTS

This research was partially supported by grant DAMD17-94-J-4461 (Principal Investigator: Dr. Donna J. Brogan) from the US Army's Breast Cancer Research Program, by an Interagency Personnel Agreement (Principal Investigator: Dr. Donna J. Brogan) from the National Cancer Institute, and by contract N01-CP-95604 (Principal Investigator: Dr. Jonathan M. Liff) from the National Cancer Institute.

The authors are grateful for the contributions of the study interviewers, managers, and other staff.

REFERENCES

1. Wacholder S, Silverman DT, McLaughlin JK, et al. Selection of controls in case-control studies. II. Types of controls. *Am J Epidemiol* 1992;135:1029-41.
2. Waksberg J. Random digit dialing sampling for case-control studies. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*. New York, NY: John Wiley and Sons, Inc, 1998:3678-82.
3. Potthoff RF. Telephone sampling in epidemiologic research: to reap the benefits, avoid the pitfalls. *Am J Epidemiol* 1994;139:967-78.
4. Kish L. *Survey sampling*. New York, NY: John Wiley and Sons, Inc, 1965.
5. Hansen MH, Hurwitz WN, Madow WG. *Sample survey methods and theory*. Vol 1. Methods and applications. New York, NY: John Wiley and Sons, Inc, 1953.
6. Groves RM, Kahn RL. Surveys by telephone: a national comparison with personal interviews. New York, NY: Academic Press, Inc, 1979.
7. Massey JT, Botman SL. Weighting adjustments for random digit dialed surveys. In: Groves R, Biemer P, Lyberg L, et al, eds. *Telephone survey methodology*. New York, NY: John Wiley and Sons, Inc, 1988:143-60.
8. Massey JT, O'Connor D, Krotki K. Response rates in random digit dialing (RDD) telephone surveys. In: 1997 proceedings of the Survey Research Methods Section, American Statistical Association. Arlington, VA: American Statistical Association, 1998:707-12.
9. Maklan D, Waksberg J. Within-household coverage in RDD surveys. In: Groves R, Biemer P, Lyberg L, et al, eds. *Telephone survey methodology*. New York, NY: John Wiley and Sons, Inc, 1988:51-69.
10. Massey JT. Estimating the response rate in a telephone survey with screening. In: 1995 proceedings of the Survey Research Methods Section, American Statistical Association. Arlington VA: American Statistical Association, 1996:673-7.
11. Casady RJ, Lepkowski JM. Telephone sampling. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*. New York, NY: John Wiley and Sons, Inc, 1998:4498-511.
12. Aquilino WS, Losciuto LA. Effects of interview mode on self-reported drug use. *Public Opin Q* 1990;54:362-95.
13. Groves RM, Lyberg LE. An overview of nonresponse issues in telephone surveys. In: Groves R, Biemer P, Lyberg L, et al, eds. *Telephone survey methodology*. New York, NY: John Wiley and Sons, Inc, 1988:191-211.
14. Olson SH, Kelsey JL, Pearson TA, et al. Evaluation of random digit dialing as a method of control selection in case-control studies. *Am J Epidemiol* 1992;135:210-22.
15. Lele C, Holly EA, Roseman DS, et al. Comparison of control subjects recruited by random digit dialing and area survey. *Am J Epidemiol* 1994;140:643-8.
16. Aquilino WS, Wright DL. Substance use estimates from RDD and area probability samples: impact of differential screening methods and unit nonresponse. *Public Opin Q* 1996;60:563-73.
17. Brinton LA, Daling JR, Liff JM, et al. Oral contraceptives and breast cancer risk among younger women. *J Natl Cancer Inst* 1995;87:827-35.
18. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc* 1978;73:40-6.
19. Brogan D. *Methodology for case-control studies of breast cancer*. (Final report, US Army Breast Cancer Research Program). Atlanta, GA: Biostatistics Department, Rollins School of Public Health, Emory University, 1997. (Grant DAMD-17-94-J-4461).
20. Denniston MM. *A comparison of area and random digit dialing sampling*. (Master's thesis). Atlanta, GA: Emory University, 1997.
21. Bureau of the Census, US Department of Commerce. *Census of Population and Housing, 1990: Public Use Microdata Sample U.S. technical documentation*. Washington, DC: Bureau of the Census, 1992.
22. Graubard B, Fears T, Gail M. Effects of cluster sampling on epidemiologic analysis in population-based case-control studies. *Biometrics* 1989;45:1053-71.
23. Hartge P, Brinton LA, Rosenthal JF, et al. Random digit dialing in selecting a population-based control group. *Am J Epidemiol* 1984;120:825-33.
24. Hartge P, Cahill JI, West D, et al. Design and methods in a multi-center case-control interview study. *Am J Public Health* 1984;74:52-6.
25. Brogan D. Software for sample survey data: misuse of standard packages. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*. New York, NY: John Wiley and Sons, Inc, 1998:4167-74.
26. Shah BV, Barnwell BG, Bieler GS. *SUDAAN user's manual, version 6.4*. 2nd ed. Research Triangle Park, NC: Research Triangle Institute, 1996.
27. Groves RM, Couper MP. *Nonresponse in household interview surveys*. New York, NY: John Wiley and Sons, Inc, 1998.
28. Aquilino WS. Telephone versus face-to-face interviewing for household drug use surveys. *Int J Addict* 1992;27:71-91.
29. Dillman DA. Call-backs and mail-backs in sample surveys. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*. New York, NY: John Wiley and Sons, Inc, 1998:466-8.
30. Brick JM, Collins MA. A response rate experiment for RDD surveys. In: 1997 proceedings of the Survey Research Methods Section, American Statistical Association. Arlington, VA: American Statistical Association, 1998:1052-7.
31. Warnecke RB. Sampling frames. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*. New York, NY: John Wiley and Sons, Inc, 1998:3935-9.
32. Oldendick RW, Bishop GF, Sorenson SB, et al. A comparison of the Kish and last birthday methods of respondent selection in telephone surveys. *J Offic Stat* 1988;4:307-18.
33. Harlow BL, Hartge P. Telephone household screening and interviewing. *Am J Epidemiol* 1983;117:632-3.
34. Madigan P, Troisi R, Potischman N, et al. Characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women. *Int J Epidemiol* 2000;29:793-8.
35. Casady RJ, Lepkowski JM. Stratified telephone survey designs. *Surv Methodol* 1993;19:103-13.
36. Brick JM, Waksberg J, Kulp D, et al. Bias in list-assisted telephone samples. *Public Opin Q* 1995;59:218-35.