THE WOMEN'S CARE STUDY: THE RESULTS OF AN EXPERIMENT IN NEW SAMPLING DESIGNS FOR RANDOM DIGIT DIALING SURVEYS

Lou Rizzo, Ralph DiGaetano, Diane Cadell, Westat Inc.
Westat Inc., 1650 Research Boulevard, Rockville, Maryland 20850
Dale W. Kulp
GENESYS Sampling Systems, Inc.

Keywords: 1+banks, tritones, undercoverage

1. Introduction

1.1 The Women's CARE Study

The Women's Contraceptive and Reproductive Experiences (Women's CARE) study is a population-based case-control study of the association between oral contraceptive use and the risk of breast cancer in women. This study is being carried out by the National Institute for Child Health and Human Development (NICHD). The RDD screening portion of the study is being conducted by Westat under contract to the Centers for Disease Control and Prevention (CDC). The RDD telephone component of the survey involves selecting a monthly sample of women to serve as controls. This sample matches in size and relative distribution across specified age-race categories the expected number of breast cancer cases who are interviewed monthly in five sites across the country: Los Angeles, Seattle, Detroit, Philadelphia, and Atlanta. Methodological findings from this RDD component are the focus of this paper.

Several factors had to be considered in the development of the RDD sample design. To permit the planned analytic approach, the sample of controls within age-race categories had to be with equal probability. Any clustering of telephone numbers had to be limited, as effects of such clustering on estimates of variance would not be reflected using the planned analysis procedures. A requirement of the protocol was that there be no more than one month between the identification of a control and its delivery to the field site for interview. This is a relatively quick turnaround period for an RDD sample. Finally, coverage of residential telephone numbers in the five sites had to be as complete as possible. Taking these factors into account, the following sample design was developed and implemented.

1.2 The Sample Design for the Identification of Controls for Women's CARE

Two standard RDD sample designs were considered and rejected. The standard Mitofsky-Waksberg methodology, where residential clusters of phone numbers are identified in a two-stage process, can be time-consuming and relatively costly, and can result in somewhat variable monthly sample sizes. The RDD procedure often characterized as "list-assisted" was also considered. Under this design, the sample frame includes only those banks of 100 numbers (area code, exchange, and next two digits) with at least a specified number of listed residential numbers. (See, for example, Tucker, Casady, and Lepkowski (1993)). The list-assisted approach was the least costly approach considered and nationally provides roughly 96 to 97 percent coverage of the general population, but was considered unsatisfactory in terms of coverage for Women's CARE.

Instead, it was decided to use the following design: an equal probability sample of phone numbers across all banks in vitually all exchanges associated with a site, thus providing as complete coverage as feasible. The banks would be classified into three categories for assessment purposes: (1) the 1+ banks (those with at least one listed residential number); (2) the zero banks (those with no listed residential numbers); and (3) banks in new exchanges (those with no listing information available).

In carrying out the survey, residential numbers are screened for the presence of women eligible to serve as controls. All eligible women found in screened households over the course of a month are placed in their respective age-race categories and sampled at rates expected to yield targeted sample sizes.

The choice of this sample design was based on the use of a technology to screen sampled phone numbers for a tritone signal. Each sampled telephone number that was not listed as business or residential was to be checked twice for a tritone signal. Those with at least one tritone signal would be screened from the sample as nonworking and thus would not have to be worked by an interviewer. It was projected that this would result in survey costs comparable to those for the Mitofsky-Waksberg approach, but the survey process would be quicker. However, there were some uncertainties as to the effectiveness of the tritone screening approach, including the percentage of numbers that would be identified as nonworking by the tritone screener and the percentage of residential numbers that would be inappropriately associated with a tritone signal. Both of these issues are addressed in this paper, as are estimates of coverage and residential number rates by site and bank type.

2. The Women's CARE RDD Samples

The Women's CARE RDD sample used for this portion of the analysis was drawn in October 1994 and was fielded as six monthly subsamples from November 1994 through April 1995. The overall sample sizes for each site are given in Table 2-1 below.

Table 2-1. Site RDD sample sizes and percentages of numbers in each bank type[1]

	Atlanta	Detroit	Phila.	L.A.	Seattle
RDD sample size	7,000	12,300	7,500	22,600	8,400
Percentage in 1+ banks	47.8	61.7	50.7	51.7	44.6
Percentage in zero banks	32.0	19.5	23.1	26.4	23.7
Percentage in new exchanges	20.2	18.8	26.1	21.8	31.7

[1] The Atlanta site consisted of Fulton, DeKalb, and Cobb counties, GA; the Detroit site consisted of Wayne, Oakland, and Macomb counties, MI; the Philadelphia site consisted of Delaware and Philadelphia counties, PA; the Los Angeles site consisted of Los Angeles county, CA; the Seattle site consisted of King county, WA.

The sampling frame was created based on the Bellcore and Donnelly databases current in October 1994. Exchanges found on the Bellcore database with any listed residential numbers within the counties corresponding to the site (according to the Donnelly database) were eligible for inclusion on the sampling frame. Exchanges with a low concentration of phone numbers for the targeted counties representing very small percentages of all listed phone numbers in those counties (in aggregate less than 0.5 percent) were excluded from the site's frame.

These sampling frames for each site were also augmented with "new" exchanges: exchanges which are associated with the site counties but are not yet represented on the Donnelly database.

Table 2-1 presents the sample sizes and indicates the percentages of telephone numbers in each site's October 1994 sample within each bank type. As can be seen, the distribution of bank types differs considerably between sites. In Detroit, 61.7 percent of the frame is in 1+ banks. In Seattle only 44.6 percent of the frame is in 1+ banks. Under a list-assisted approach, the percentages of excluded banks would be considerably higher in Seattle than in Detroit.

2.1 Estimated Residential Rates by Bank type and Site

During the six month fielding period (November 1994 to April 1995), every telephone number in the sample was assigned a final status: residential, business, or nonworking. Table 2-2 below presents the sample percentages of residential numbers within bank types for each site. Banks from new exchanges are included with zero banks from established exchanges in these tables, since the differences between the zero banks and the new exchanges in terms of residency rates were negligible.

Table 2-2. Sample percentages of residential numbers within bank types for each site

% resid. numbers	Atlanta	Detroit	Phila.	L.A.	Seattle
1+ banks	46.83	51.84	53.01	48.96	52.42
Zero banks	2.04	1.73	1.20	1.28	2.41
All banks	23.58	32.70	27.47	25.99	24.63

The percentage of residential numbers in zero banks and new exchanges is considerably lower than that in 1+ banks and is consistent with results from national surveys. For example, in a national sample of 10,000 telephone numbers from zero banks, 1.4 percent were found to be residential (Brick, Waksberg, Kulp, and Starer (1995)). The results presented in Table 2-2, however, show the differences between sites.

A key question in using the list-assisted approach is the expected degree of undercoverage due to the exclusion of the zero and new exchange banks. This can be estimated for each study site by taking the estimated total number of residential numbers in such banks and dividing it by the estimated total number of residential numbers in all banks in the site.

Table 2-3 below presents estimates of the undercoverage rates that would be incurred by excluding the zero banks and new exchanges from the sampling frame. Confidence intervals are also given. These are approximate confidence intervals, as the undercoverage rate is a ratio of two estimates. We estimated the sampling variance of this ratio utilizing standard approximations. These results are also consistent with the Brick, Waksberg, Kulp, and Starer (1995) results, which estimated a residential undercoverage rate of 3.7 percent nationally when zero banks and new exchanges are excluded.

Table 2-3. Undercoverage (u.c.) rates from excluding zero banks and new exchanges

Confidence intervals for u.c. rates	Atlanta	Detroit	Phila.	L.A.	Seattle
Lower bound	3.40	1.55	1.46	1.94	4.38
Estimate of u.c. rate	4.49	2.03	2.15	2.37	5.44
Upper bound	5.58	2.51	2.84	2.80	6.50

As can be seen again, there are considerable differences between sites not explainable by sampling variability in the estimates. The estimated undercoverage rates range from 2 percent for the Detroit site to 5.4 percent for the Seattle site. Thus, using national rates in constructing the sample design for a local survey is imprecise due to such variation.

2.2 Cost Implications for the Sample Design

The losses in coverage from the list-assisted approach must be balanced in any real world sampling design by the extra costs incurred from calling these numbers. As can be seen from Table 2-1, there are as many zero bank and new exchange numbers in the RDD frame as there are 1+ bank numbers. The sample size necessary to obtain a targeted number of eligible households will need to be roughly twice the size if all banks are included in the sampling frame as compared to a sampling frame that includes only 1+ bank numbers.

The cost, however, of using all banks for the sampling frame will not be double that of using only the 1+ banks, as the highest interviewing cost is incurred from calling residential households: the cost from calling a nonworking or business number is smaller as the interviewer time necessary to confirm a number as nonworking or business is smaller. Our experience with the Women's CARE study has been that the interviewer time necessary to finalize, for example, 1,000 zero bank and new exchange numbers is roughly half that necessary to finalize 1,000 1+ bank numbers. This difference is due to the much larger percentages of nonworking and business numbers among the zero bank and new exchange numbers.

A method for reducing interviewing costs further is to screen telephone numbers with a tritone screening system, described in Sections 3 and 4.

3. The Screening of Nonworking Numbers Using the Tritone Technology

The tritone screening system (GENESYS-ID) used for the Women's CARE study utilizes a combination of both database and computer telephony technology. The software, which employs telephony technology similar to a predictive dialer (used extensively for high-volume outbound telemarketing applications), was developed by GENESYS Sampling Systems.

Before being screened for nonworking status using the tritone technology, each number was checked for business or residential status against a listing of known business or residential numbers. Numbers found on these listings were not screened using the tritone technology and were called by interviewers. (An alternative approach would have been to drop all business numbers identified, but, to help maximize coverage, all such numbers were called.)

After deletion of those numbers listed as working, the remaining sample numbers were processed through the "dialer." This software is specifically designed to detect tritone, FAX, and modem tones. The software was modified to control ringing in the event a tone is not detected initially. To limit intrusiveness, most calls are terminated so that a household might hear at most one ring. In terminating the call so quickly, however, the technology is unable to identify some nonworking numbers (the nonworking intercept may not cut in until the call is allowed to ring multiple times).

The tritone screening technology may identify working numbers as nonworking in some situations. Telephone line problems due to busy circuits, bad weather, trunk outages, or other related problems can result in a tritone-like signal for an otherwise working number.

The dialing system classifies nonworking numbers as tritone signal numbers and no ringback numbers. The tritone signal numbers are numbers for which the three tone tritone signal is "heard" by the technology. The no ringback numbers fall into two categories. The first category consists of numbers for which there is a nonworking result based on receipt of a 40 second interval of silence. That is, once the number has been dialed and a connection made, the software receives no return signal of any kind. The second class of no ringback calls consist of numbers for which a nonworking result is assigned after receiving six seconds of continuous "noise." An example would be a connection followed by a return to a dial tone. Telephone line problems mentioned in the previous paragraph can occasionally generate a tritone signal or no ringback result for working numbers.

The only other significant situation we know of where misidentification may result would occur in screening those households with a dual-use residential number with a modem or facsimile machine attached. In this situation, if the modem or facsimile machine was connected, but not in control of the line, and the dialer tested the number, that number could be identified as nonworking if the modem or facsimile answered the call within the first ring.

4. Experience with the Tritone Screening System in the Women's CARE Study

The tritone technology discussed in Section 3 was applied to every telephone number in the study RDD samples, except those that were residential or business according to the white and yellow page listings. Multiple (generally two) passes were made for each of these telephone numbers. Three possible assignments were made to each telephone number after these passes: (1) nontritone: all passes indicate working status; (2) pure tritone: all passes indicate nonworking status, and (3) mixed tritone. Some passes indicate working status and some nonworking status (in the case of two passes, one identification as working and one as nonworking).

There were a substantial number of telephone numbers that fell into the third category. Table 4-1 below indicates the percentages of telephone numbers in each site sample that were assigned to each category. These percentages are from the second RDD sample, selected in October 1994. Table 4-1 also includes estimated percentages of nonworking numbers in each site.

Table 4-1. Percentages of tritone and nonworking numbers in the October 1994 sample[1]

	Atlanta	Detroit	Phila.	L.A.	Seattle
Pure tritones	19.9	22.8	24.8	1.9	2.0
Mixed tritones	7.3	3.8	3.5	5.4	7.4
Nontritones	72.9	73.4	71.7	92.7	90.7
Nonworking	49.8	44.0	48.9	51.2	50.1

[1]The Atlanta site consised of Fulton, DeKalb, and Cobb counties, GA; the Detroit site consisted of Wayne, Oakland, and Macomb counties, MI; the Philadelphia site consisted of Delaware and Philadelphia counties, PA; the Los Angeles site consisted of Los Angeles county, CA; the Seattle site consisted of King county, WA.

The percentages of numbers with at least one tritone result were much smaller in Los Angeles and Seattle. The percentages of nonworking numbers were actually larger in the two Western sites, so the tritone technology was less successful in these sites (the tritone technology was implemented from the Philadelphia metropolitan area: the effectiveness of the tritone screening may vary by distance). In addition, the two passes for each number were more consistent in the three Eastern sites and the ratio of pure tritones to mixed tritones was much higher in the Eastern sites: it ranged from .27 to .35 in the Western sites, but from 2.7 to 7.1 in the Eastern sites.

If both pure and mixed tritones are excluded as nonworking, then 25 to 30 percent of the numbers in the original sample can be excluded in the Eastern sites. This can save considerable costs in interviewer calls to establish status. The reduction in the two Western sites are much more limited: 5 to 10 percent.

The RDD samples in the study were equal probability samples from the sampling frame. Under a list-assisted approach, samples would be drawn only from the 1+ banks in the sampling frame. Table 4-2 indicates the percentages of numbers assigned to each status by the tritone technology among the 1+ bank numbers from the October 1994 study sample.

If pure and mixed tritone numbers are excluded as nonworking in a list-assisted approach where tritone screening is also used, then 20 to 25 percent of the numbers would be excluded in the Eastern sites, and 4 to 6 percent in the Western sites. These are less than the percentages in Table 4-1 because there is a smaller percentage of nonworking numbers in the 1+ banks compared to the zero banks and new exchange banks.

Table 4-2. Percentages of tritone numbers among 1+ bank numbers in October 1994 CARE RDD sample

1+ bank only	Atlanta	Detroit	Phila.	L.A.	Seattle
Pure tritones	15.1	20.0	19.5	1.6	0.8
Mixed tritones	5.1	2.4	1.5	3.6	3.9
Non tritones	79.8	77.6	79.0	94.8	95.3

4.1 Percentages of Residential Numbers Among the Tritone Numbers

The tritone technology will identify only some of the nonworking numbers as nonworking. The small cost of using this technology compared to interviewer costs makes it cost efficient to use even if the percentage of nonworking numbers identified is relatively small. A different issue is that of numbers which are assigned as nonworking by the tritone technology that are in fact working residential numbers. The misclassification of these numbers can result in potential bias due to undercoverage.

The tritone screening technology is relatively new, so there have been a limited number of studies evaluating its accuracy. One such study of GENESYS-ID occurred in its application to the State and Local Immunization Coverage and Health Survey (SLICHS), as presented in Battaglia, Starer, Oberkofler, and Zell (1995).

In order to evaluate the extent to which this false identification of nonworking status occurred in the CARE study, Westat performed two quality control experiments. The first experiment consisted of a random drawing of 1,000 numbers which had a mixed tritone result in the April 1994 sample, and the second consisted of a sample of 1,000 numbers with a pure tritone result.

These numbers were then called by our interviewers to verify nonworking status. These calls were made up to a month following the tritone trials for the mixed tritones, and three months for the pure tritones, so that some numbers became working during the interim period. To account for this, the interviewers asked respondents when the telephone number was first in service. If service was initiated after the tritone trials were done, we assumed that the number was in fact nonworking during the tritone trials.

Tables 4-4 and 4-5 below present the counts of confirmed residential numbers among the mixed and pure tritones in the quality control sample. These numbers do not include those that were residential but were indicated as being initiated into service after the date of the tritone trials. Numbers that were confirmed to be residential but could not be confirmed as being in service or not in service at the date of the tritone trial were imputed as in service or not in service. In addition, numbers which were not confirmed as residential were imputed as residential or not residential. The fractional values reflect the imputations, which were a limited part of the Table 4-4 and 4-5 counts.

Table 4-4. Residential numbers in the mixed tritone quality control samples

Site	Mixed tritone sample size	Estimated resid. count	Estimated resid. percentage
Atlanta	84	2.1	2.5
Detroit	118	1.8	1.5
Philadelphia	65	0.0	0.0
L.A.	544	38.7	7.1
Seattle	189	1.1	0.6
Total	1,000	43.7	4.4

The percentage of residential numbers among the mixed tritones was nonnegligible, particularly in Los Angeles. Only in Philadelphia were there no false positives among the mixed tritones. There were fewer residential numbers misclassified among the pure tritones, as indicated in Table 4-5 below.

Table 4-5. Estimated in-service residential numbers in the pure tritone quality control samples, after imputation

Site	Pure tritone sample size	In service residential count	Residential percentage
Atlanta	169	1.1	0.6
Detroit	363	4.9	1.4
Phil.	238	0.0	0.0
L.A.	200	2.0	2.0
Seattle	30	0.0	0.0
Total	1,000	7.0	0.7

4.2 Undercoverage from Excluding Tritone Numbers

Exclusion of mixed or pure tritone numbers as nonworking will result in some undercoverage of residential numbers in most of the sites. This undercoverage can be estimated by comparing the residential counts for the tritone sets to the residential counts for the overall sample.

Table 4-6 below presents estimated undercoverage rates from excluding all of the numbers designated as nonworking by the tritone technology (the mixed and pure tritones). The columns are as follows:

1. Estimated tritone residentials. The sum of the mixed and pure residential estimates for the October 1994 sample if all tritone numbers had been called.
2. Total fielded residentials. The total number of telephone numbers that received final dispositions as residential during the fielding of the October 1994 sample.
3. Estimated undercoverage rate. The tritone residentials divided by the total residentials.

4/5. Approximate lower and upper bounds for 95 percent confidence interval. These bounds are computed using standard asymptotic results. The "0+" indicates that the lower bound was negative.

Table 4-6. Estimated undercoverage rates from excluding all tritone numbers

Site	Estimated tritone resid.	Total resid.	Estimated under-coverage rate	Confi-dence interval lower bound	Confi-dence interval upper bound
Atlanta	21.5	1332	1.6	0.1	3.6
Detroit	45.0	3265	1.4	0.4	2.6
Phila.	0	1679	0	*	*
L.A.	233.4	4690	5.0	3.6	6.5
Seattle	6.5	1666	0.4	0+	1.3

The estimated undercoverage rate was highest in Los Angeles (5.0%), and varied significantly between sites. In Philadelphia and Seattle, the undercoverage rate appears to be negligible. This source of undercoverage needs to be considered in any sample design where tritone screening is planned. Under a list-assisted approach, the undercoverage rates will be quite close to these rates, as most of the residential numbers among both the tritone and total sample numbers come from 1+ banks.

The undercoverage rates in Table 4-6 assume all mixed and pure tritone numbers are excluded as nonworking. An alternative approach would be to exclude only the pure tritone numbers as nonworking, and to call the mixed tritone numbers. Table 4-7 presents estimated undercoverage rates in each site using this sample design.

Table 4-7. Estimated undercoverage rates from excluding pure tritone numbers

Site	Ex-cluded tritone resid.	Total resid.	Under-coverage rate	Confi-dence interval lower bound	Confi-dence interval upper bound
Atlanta	8.8	1332	0.7	0.0+	2.2
Detroit	38.0	3265	1.2	0.2	2.3
Phila.	0.0	1679	0.0	*	*
L.A.	3.3	4690	0.1	0.0+	0.2
Seattle	0.0	1666	0.0	*	*

The undercoverage rates were considerably lower for Los Angeles and Seattle, as compared to the sample design that excludes all tritone numbers. Some improvement in coverage occured for Atlanta and Detroit. Of course, if this approach is used, considerably more numbers would have to be worked by interviewers, increasing survey costs.

5. Conclusions

The sample design for the Women's CARE study allowed us to explore two issues of current relevance to survey practitioners. The first issue is the variation in undercoverage across different metropolitan areas that can result from carrying out a list-assisted RDD survey. The second issue is the effectiveness of the tritone screening technology both in terms of the elimination of nonworking numbers and the error rate in assigning nonworking status to residential numbers.

With respect to the first issue, we estimated the undercoverage rates that would have been incurred from a list-assisted sampling design. These estimates are based on the residential numbers obtained from banks on the sampling frame containing no listed residential numbers. (Under a list-assisted approach, these banks are excluded from the frame.) These estimated undercoverage rates roughly correspond to estimates derived from similar studies, but we saw considerable variation between the sites (ranging from 2.0 to 5.4 percent across the sites).

As for the second issue, the evidence of this study is that the effectiveness of the tritone screening technology in assigning nonworking status varies considerably by site. A smaller percentage of nonworking numbers in the RDD sample were identified in the two Western sites (Los Angeles and Seattle) compared to the Eastern sites. In this screening process residential numbers were occasionally identified as nonworking. The residential undercoverage rates induced by this false identification were estimated to be in the range of 0 to 2 percent for four of the five sites, and 5 percent for Los Angeles (if both mixed and pure tritones are excluded from the sample as nonworking). Depending on the requirements of a study, the magnitude of such undercoverage rates may be important. As previously mentioned, the tritone technology was implemented from the metropolitan Philadelphia area, and the effectiveness of the technology may vary by distance.

In surveys where full coverage of the residential numbers in the population of interest is of great importance, one could consider requiring tritone outcomes to be experienced for multiple trials. In addition, it would likely prove useful to call a random set of the numbers designated as nonworking by the tritone technology as a matter of course in any RDD surveys that utilize this technology. This would help ensure that any undercoverage resulting from false positives in the tritone screening process is in accord with expectations for the study.

References

Battaglia, M. P., Starer, A., Oberkofler, J., and Zell, E. R. (1995). "Pre-identification of nonworking and business telephone numbers in list-assisted random-digit-dialing samples", unpublished technical note.

Brick, M., Waksberg, J., Kulp, D., and Starer, A. (1995). "Bias in list-assisted telephone samples", *Public Opinion Quarterly, 59*, 218-235.

Cochran, W. G. (1977). *Sampling Techniques*, 3rd ed. New York: John Wiley & Sons.

Tucker, C., Casady, R., and Lepkowski, J. (1993). "A hierarchy of list-assisted stratified telephone sample design options." Paper presented at the Annual Conference of the American Association for Public Opinion Research.