COMPARISON OF TRADITIONAL AND MODIFIED WAKSBERG

Charlene Smith and Emma L. Frazier, Centers for Disease Control
Charlene Smith, 1600 Clifton Road, MS K30, Atlanta, GA 30033

KEY WORDS: Waksberg sampling design, Behavioral Risk Factor Surveillance System

1. Introduction

Because of limited budgets, many health care survey researchers are considering telephone surveys instead of in-person interviews. Random digit dialing (RDD) is the preferred method for conducting telephone surveys. Unlike telephone directory listings, RDD gives every working number a chance to be included in the sample regardless of whether the number is a listed number. However, a shortcoming of this procedure is that a generated telephone number may not be connected to a household. Additionally, efforts to determine whether a number reaches a household, may increase the cost of the survey.

To identify households in RDD surveys more efficiently, a two stage sampling procedure was developed by Mitofsky and Waksberg (Waksberg, 1978). Clusters or blocks of telephone numbers comprise the primary sampling frame. If the generated number reaches a household, the cluster of numbers is retained in the sample. However, if a business or nonworking number is reached, the block is rejected from the sample. The rationale behind this procedure is if one working household number is found in a cluster, then it is likely that other working household numbers will appear in the cluster. Usually, the telephone company assigns new numbers consecutively in blocks. The advantage of this approach is the number of calls to businesses or not-in-service numbers are limited. The cost of survey may decrease.

To obtain equal selection probability for this random digit dialing procedure, a constant number of households per cluster out of all accepted clusters is required. There are several disadvantages to this method. It is necessary to monitor how many telephone numbers in a cluster resulted into interviews. If an insuffient number of interviews are obtained from a cluster, more telephone numbers cannot be added to achieve the desired cluster size (Waksberg, 1984). In addition, it can take a substantial amount of time to determine the residential status of the cluster, because several call backs may need to be made. This procedure is inconvenient if a limit time is required for data collection. Hereafter, this method will be referred to as the traditional version of Waksberg.

To overcome the inconveniences of obtaining the required constant number of household interviews for each cluster, an alternative strategy of completing a cluster was introduced (Waksberg, 1984). This method will be referred to as the modified version of Waksberg. Instead of designating a constant number of interviewed households per cluster, the modified Waksberg specifies a constant number of telephone numbers to be dialed per cluster. The primary sampling frame as is the same as in the traditional version, with cluster or blocks of numbers. Residential status of telephone stills has to be determine for the block of numbers. There is no tracking of how many interviewed households. The idea is to dial all available telephone numbers in each cluster. The modified procedure is believed to accelerate the data collection phase of a study (Waksberg, 1984). Along with the increase in speed, however, there may be a loss in precision of estimates because of the resulting variation in cluster sizes.

The Centers for Disease Control (CDC) has recommended that states conducting surveys for the Behavioral Risk Factor Surveillance System (BRFSS) use the traditional Waksberg method. With the desire to increase sample sizes for the states and increase the efficiency of the work flow, the modified version of the Waksberg method is appealing. This paper presents results from a comparsion of both operational and statistical aspects of the two versions of Waksberg, using 1989 BRFSS data collected in California.

2. Background and Methods

2.1 Behavioral Risk Factor Surveillance System

In 1984, the Centers for Disease Control in collaboration with state health departments,

initiated the Behavioral Risk Factor Surveillance System (BRFSS). In this year, 15 states collected behavioral risk factor data to monitor health behaviors of civilian, non-institutionalized adults (aged 18 and older) through monthly telephone surveys. By 1989, 39 states and the District of Columbia were participating in the BRFSS. Staff from each state conducted from 100 to more than 200 interviews each month during an 8- to 14-day period. Along with demographics, information about behaviors related to alcohol use, hypertension, physical activity, seat belt use, and smoking is collected in the BRFSS interviews (Frazier et al, 1992).

2.2 Study Design

Between July and September, 1989, the state of California, in collaboration with the Centers for Disease Control, conducted simultaneous BRFSS surveys using both the traditional and modified versions of Waksberg methods. The goal of the traditional survey was to obtain three completed interviews per cluster, while the goal of the modified method was to call nine telephone numbers per cluster. Based on previous studies, nine numbers were estimated to be the number of calls needed to obtain an average of three completed interviews per cluster in California. Interviewers were randomly assigned to each of the methods during each of the interviewing time periods. The length of the interviewing period was 73 and 68 days, for the traditional and modified methods, respectively.

2.3 Methods

Considering operational issues for the two methods, the outcome of calls and several quality controls measures are compared. Quality control measures include response rate, survey efficiency, and refusal rate. The survey response rate for BRFSS is based on the Council of American Survey Research Organization (CASRO) definition. CASRO (1983) suggests that the calls of unknown status be divided into eligible and ineligible categories in the same proportions as calls with known status. A call is classified as unknown if after five attempts to reach a number there is no answer or busy. The CASRO formula including the deposition codes for calculating the response rate is shown in Appendix A.

Survey efficiency, and refusal rate are quality control measures that indicate how successful the interviewers were in obtaining completed interviews. Survey efficiency is the percentage of all telephone numbers dialed which resulted in a completed interview. The refusal rate is defined as the proportion of eligible respondents who were contacted but declined to be interviewed or terminated the interview. The formulas for these indicators are also provided in Appendix A.

Weighted prevalence estimates for various risk factors computed from BRFSS were computed and compared between the two methodologies. (Descriptions of the weighting procedures for both methods are provided in the next section.) To examine the precision of the estimates for the two methods, standard errors, design effects, and effective sample sizes were compared.

The design effect is the variance of a measure taking into account the survey sample design divided by the variance of a measure if the survey had been conducted by simple random sample. The effective sample size represents the number of observations that contribute to the precision of the prevalence estimates. Effective sample size is computed as the ratio of actual sample size divided by the design effect.

Statistical tests based on the chi-square and t-test were computed by the Statistical Analysis System (SAS). All standard errors and design effects were calculated using Standard Errors Program for Computing of Standardized Rates from Sample Survey Data (SESUDAAN), a program designed for complex survey analysis.

2.4 Weighting Procedures

All responses from the BRFSS are adjusted by weighting to compensate for the variation in selection probabilities and the disproportionate representation of individuals in the survey sample. Weighting procedures are necessary for both tradtional and modified Waksberg methods. Weighting is performed to adjust for the multiple telephone numbers in the household, number of adults, the number of interviews completed per cluster, and the demographic distribution of the survey sample.

Weights for BRFSS are derived as the product of three terms. The first two terms reflect

the actual selection probabilities for each respondent. The first term adjusts for homes with more than one telephone number and the number of adults residing in the household. This is computed by calculating the ratio of the number of adults divided by the number of different telephone numbers that reach the household.

The second term, called the cluster size adjustment, compensates for unequal cluster sizes. Because equal cluster sizes are needed to achieve equal-probability samples of households, a ratio of the expected cluster size divided by actual cluster size is computed to compensate for missing interviews.

The last term of the weighting computation for BRFSS is to adjust for undercoverage and nonresponse of the target population. This adjustment, called the post-stratification factor, is calculated by computing the ratio of age, race, and sex distribution of the California 1989 intercensal population divided by that of the survey sample. This factor is then multiplied by the first two terms to compute a final weight. The final weights are used to generate weighted prevalence estimates for selected behaviors in BRFSS.

The weighting procedure used in the modified Waksberg method is similar to that used in the traditional method. Since the modified version specifies dialing a constant number of telephone numbers instead of obtaining a constant number of interviews, the weighting procedure was modified for the second factor, the cluster size adjustment. In the modified Waksberg version, the cluster size adjustment is defined as the ratio of the expected average cluster size divided by the actual cluster size (Westat, 1984). In this study, the cluster size ranged from one to nine, with an average cluster size of 4.9.

3. RESULTS

3.1 Quality Control Measures

Table 1 shows the outcome of calls, survey efficiency, and response rate for both versions of the Waksberg design. A total of 1647 and 1809 telephone numbers were dialed to obtain 600 and 772 completed interviews, respectively, using the traditional and modified versions of Waksberg. The survey efficiency (or the percentage of completes) was 36.4 for traditional Waksberg and 42.7 for modified Waksberg. This suggests that the modified version of Waksberg was more successful in obtaining completed interviews.

The percentage of outcome of business, ring-no-answer, and nonworking numbers was 45.3 for traditional Waksberg and 42 for modified Waksberg. This indicates these methods are similar. Using a chi-square test of independence, no significant differences for the outcome groups of the 2 methods are found at $p < 0.05$ level of significance.

Table 1 also shows the response and refusal rates for the study. The response rates for this study are 63.9 for traditional and 71.0 for modified procedures. Both rates are below the CASRO recommended guideline of 75 percent. The refusal rates are 16.2 for traditional and 13.2 for modified.

3.2 Statistical Measures

In Table 2, the age, race, and sex distributions are compared for the traditional and modified methods, as well as the 1989 intercensal estimates. Intercensal estimates are not significantly different from the sample distributions, except for the race groups and the age group 18-24. To correct for any possible biases from the age, race, and sex distributions, the data from the survey was adjusted bythe post-stratification factor in the weighting procedure.

Table 3 presents weighted prevalence estimates and related statistics for seven BRFSS self reported behaviors for traditional and modified Waksberg. For 4 out of 7 behaviors presented, the estimates are larger for the modified version of Waksberg. However, the differences between the prevalence estimates between traditional and modified Waksberg are not statistically significant.

The design effects increased as often as they decreased for the behaviors presented in Table 3. Drinking and driving had the largest change in the design effects with 1.09. The average of design effect for the traditional is 1.41 and 1.51 for modified.

The computations for the effective sample sizes produced similar results to the design effects. For example, drinking and driving, the large increase in the design effect caused a large reduction in effective sample size from the actual sample size. Effective sample sizes for the average

design effect are 426 for traditional and 511 for the modified.

4. Conclusions and Recommendations

The modified version of Waksberg does speed up the phase of data collection. The modified version was able to produce 172 more sample observations than the traditional in approximately same amount of time. The desired sample size can be reached faster with the modified version.

No consistent increase or decrease is identified from the individual design effects. This finding was surprising. Since the cluster size varies in the modified version, an increase in the design effects from traditional was expected in this study. The variability among the clusters in the selection of inteviewed households causes the variances of the estimates to larger than those in the traditional Waksberg method where the each household has the same probability of selection. The overall average of design effects shows the expected increase from the traditional version of Waksberg. The effective sample sizes for the average design effect demonstrates that more observations contributing to precision of estimates are lost when the design effects or the variability is larger. This study made attempt to adjust for the expected increase in design effect by increasing the sample size for the modified Waksberg. But still the effective sample size for the average of design effect for modified Waksberg is below the desired sample size of 600.

This study lacks some of the cost data to adequately compare operational aspects of both designs. The cost information is vital to actually comparing the methods because more telephone numbers will have to be dialed to obtain a large sample for the modified Waksberg. A large sample is needed to compensate for loss in precision so that the effective sample size will be equal to the desired sample size. One recommendation is that the Centers for Disease Control conduct another study using BRFSS that will more closely track the costs, sample administration, supervisor and interviewer work effort required to adequately conduct both methods. Only after an assessment of the relative differences in cost has been made, can a strong recommendation be made for one method over the other.

APPENDIX A

The following are quality control indicators used by BRFSS:

$$\text{CASRO response rate} = \frac{01}{(01+02+07+09) + \frac{(01+02+07+09)}{(01+02+07+09)+(03+05+06+08+11)} \times (04+10)}$$

$$\text{refusal rate} = \frac{02+09}{01+02+09}$$

where

- 01 = Completed interview
- 02 = Refused interview
- 03 = Nonworking interview
- 04 = Ring-no-answer
- 05 = Business phone
- 06 = No eligible respondent at this number
- 07 = No eligible respondent available during interviewing period
- 08 = Language barrier
- 09 = Interview terminated
- 10 = Busy
- 11 = Respondent unable to communicate due to physical or mental impairment.

REFERENCES

Centers for Disease Control. BRFSS Operations Manual (1989), Atlanta: Centers for Disease Control, 1989.

Frazier, Emma L., A. L. Franks, L. M. Sanderson, Behavioral Risk Factor Data. Using Chronic Disease Data: A Handbook for Public Health Practitioners, Atlanta: Centers for Disease Control, 1992.

Groves, Robert M., R. L. Kahn (1979), Surveys By Telephone, Academic Press, New York, NY.

Gentry, E. M., W. D. Kalsbeek, G. C. Hogelin, J. T. Jones, K. L. Gaines, M. R. Forman, J. S.

Marks, F. L. Trowbridge, "The Behavioral Risk Factor Surveys: II. Design, Methods, and Estimates from Combined State Data (1985)," American Journal of Preventive Medicine, vol. 1, no. 6, pp. 9 - 14.

Rossi, P. H., J. D. Wright, A. B. Anderson (1983), Handbook of Survey Research, Academic Press, Orlando, FL.

Shah, B. V., SESUDAAN: Standard Errors Program for Computing of Standardized Rates from Sample Data. Research Triangle Park, NC: Research Triangle Institute, 1981.

Waksberg, J. (March 1978), "Sampling Methods for Random Digit Dialing," Journal of the American Statistical Association, vol. 73, no. 361, pp. 40-46.

Waksberg, J. (December 1984), "Efficiency of Alternative Methods of Establishing Cluster Sizes in RDD Sampling," WESTAT memorandum.

****

Table 1. Summary of Calls for Traditional and Modified Waksberg Methods

	Traditional	Modified
Outcome of Call	N (%)	N (%)
Complete interview	600 (36.4)	772 (42.7)
Refusal	116 (7.0)	118 (6.5)
Ineligible respondents	183 (11.1)	157 (8.7)
Business	170 (10.3)	209 (11.5)
Ring-No-Answer/Busy	223 (13.5)	203 (11.2)
Nonworking	355 (21.5)	350 (19.3)
Total	1647	1809
Refusal Rate	16.2	13.3
CASRO Response Rate	63.9	71.0

Table 2. Unweighted Distribution of Sex, Race, and Age Groups for Traditional and Modified Waksberg and 1989 Intercensal Estimates

Distribution Groups		Traditional	Modified	Intercensal
		%	%	%
Sex				
	Male	43.7	46.0	49.1
	Female	56.3	54.0	50.9
Race				
	White	74.8	75.0	83.5
	Nonwhite	25.2	25.0	16.5
Age				
	18 - 24	12.2	10.6	14.9
	25 - 34	28.7	26.1	25.8
	35 - 44	22.7	22.8	20.3
	45 - 54	12.8	12.8	13.3
	55 - 64	11.3	13.2	11.1
	65+	12.3	14.5	14.6

****

Table 3 . Prevalence Estimates and Effective Sample Sizes for Traditional and Modified Waksberg

	Traditional(N=599)				Modified(N=771)			
Risk Behaviors	%	s.e	design effect	effective sample size	%	s.e	design effect	effective sample size
Chronic Drinking	4.02	0.89	1.21	496	4.99	0.91	1.35	572
Drinking/Driving	3.92	0.91	1.31	458	5.34	1.25	2.40	322
Hypertension	16.42	1.65	1.81	331	19.91	1.90	1.74	444
Sedentary Lifestyle	47.21	2.47	1.47	409	46.67	2.14	1.41	548
Lack of Seatbelt	9.38	1.55	1.69	355	7.35	1.01	1.14	677
Currently Smoking	22.06	2.04	1.45	414	20.50	1.89	1.69	457
Overweight	15.80	1.74	1.36	442	18.00	1.61	1.36	567
Mean		1.41		426			1.51	511