

Sources of Variation in Response Rates to the Behavioral Risk Factor Surveillance System

David W. Smith, Ph.D., M.P.H.
University of Texas School of Public Health

Abstract

The BRFSS is an RDD telephone survey conducted by state and territorial health departments under the sponsorship of the Centers for Disease Control. Sample designs, sample sizes, and questionnaires are similar but not identical in all states in any year. In recent years the number of respondents per state has been about 2000-6000, totalling 200,000 to 350,000 from all states.

The BRFSS reports several indicators of the quality of survey execution for each state, including cooperation rates, response rates, and call disposition codes. A public use data file includes item specific missing data codes for refused and don't know/not sure (DK/NS). Respondents who complete partial surveys can, largely, be distinguished in the public use data file.

The variation among the indicators of survey quality exceeds chance across states and years with much more variation among years than among states. Variation in item nonresponse depends on questionnaire length and the type of question. Patterns of don't know/not sure (DK/NS) rates and refusal rates are quite different, depending on the question asked.

Keywords: missing data, survey, Behavioral Risk Factor Surveillance System

1. Introduction

The Behavioral Risk Factor Surveillance System (BRFSS) surveys are random telephone surveys conducted by states (including territories and the District of Columbia) under the sponsorship of the Centers for Disease Control and Prevention (CDC). State health departments collect data, through telephone interviews of adults, on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and infectious diseases. Each state manages the conduct of its own survey within mutually agreed standards (CDC 2005a; Holtzman 2003). These standards specify such characteristics as the sample design and sample selection, dates of data collection each month, the number of callbacks to contact a household, making appointments for interviews, and the questionnaire structure and wording. In recent years the number of respondents has been similar in each state in any single year, though a few states are

unusually larger in each year in order to conduct special studies.

BRFSS content and questionnaire length vary from year to year and by state, in a controlled and structured way. The core content (the first part) changes annually and is used by all states. This includes such information as general health, quality of life, smoking, access to health care, presence of chronic conditions such as diabetes and hypertension, and sociodemographic information. Optional modules, the second part, are selected by states from a menu which changes annually. As the third part, states often add questions they write themselves or obtain from other sources but the number is usually a small part of the total.

About half the states contract with a small number of vendors for data collection. Several states contract with vendors who conduct the BRFSS for one or two states. Some states hire and train their own staff to collect data. States request proposals and contract with vendors according to state purchasing laws, typically every few years.

States and vendors have several opportunities to influence the quality of the data collection process within the shared standards. It is possible that variation in response rates among the states is due to variation in local conditions.

To identify patterns in variation in response rates, I used reports of response rates for 1994 through 2004 to measure variation across states (here including the District of Columbia and territories with surveys done every year). To illustrate patterns of variation in item nonresponse I used the public use file for 2005.

2. Methods

2.1 Data Source

I used annual reports from 1998 through 2004 that reported response rates for each state for each year from 1994 through 2004 (CDC 1999; CDC 2000; CDC 2001; CDC 2002; CDC 2003; CDC 2004; CDC 2005b, Source: http://www.cdc.gov/brfss/technical_infodata/quality.htm, obtained June 20, 2006). Fifty-two states (including Puerto Rico and the District of Columbia) had reported response rates in the source tables. Three (Puerto Rico, the District of Columbia, and Hawaii)

had missing rates for one to three years and were deleted, leaving 49 states for analysis.

I used the BRFSS public use file for 2005, which included 356,112 respondents, to compute nonresponse rates for selected items. (Source, June 20, 2006: http://www.cdc.gov/brfss/technical_infodata/surveydata/2005.htm.)

2.2 BRFSS Questionnaire Structure

The questionnaire is organized in the following sequence:

1. Opening script
2. Core sections
3. Closing or transitional statement
4. Optional modules
5. Closing or transitional statement
6. State-added questions
7. Closing

There were 20 core sections in 2005, identical in every state. There were 26 optional modules in 2005. The core sections, in their 2005 sequence with the number of primary questions and conditional questions, were:

- | | |
|-------------------------|------|
| 1. Health Status | 1 |
| 2. Healthy Days | 3 |
| 3. Health Care Access | 4 |
| 4. Exercise | 1 |
| 5. Diabetes | 1+1 |
| 6. Hypertension | 1+1 |
| 7. Cholesterol Screen | 1+2 |
| 8. Cardiovascular | 3 |
| 9. Asthma | 1+1 |
| 10. Immunizations | 3 |
| 11. Tobacco Use | 1+2 |
| 12. Alcohol | 1+4 |
| 13. Demography | 16+2 |
| 14. Veteran Status | 1 |
| 15. Disability | 2 |
| 16. Arthritis Burden | 1+4 |
| 17. Fruits & Vegetables | 6 |
| 18. Physical Activity | 3+4 |
| 19. HIV/AIDS | 1+3 |
| 20. Emotional Support | 2 |

BRFSS questions typically allow respondents to refuse to answer a question or to state that they don't know or are not sure of the answer. Refusals are usually coded in one category and respondents who don't know or are not sure are coded in another category. Some responses in the final data set are blank. Blanks either indicate questions that were skipped as part of the structure of the questions, such as number of cigarettes

for nonsmokers, or else indicate that the interview was terminated partway through.

2.3 Response Rates

Analysis of response rates focuses on two summary measures for each year and state: the cooperation rates and the CASRO (Council of American Survey Research Organizations) rates. The former is an indicator of the willingness of individuals to give an interview. The latter is an industry-wide standard formula for response rates.

The cooperation rate is the ratio of the number of respondents who agree to an interview over the number of households where the respondent is both selected and contacted, that is, the proportion of people who, after being contacted, agree to be interviewed. This measures the willingness of individuals to give an interview when asked (CDC 2005a, page 2).

The CASRO response rate is calculated assuming that telephone numbers that were sampled but not contacted or otherwise resolved have the same proportion of eligible households as the telephone numbers whose eligibility to respond is clearly determined. Since the BRFSS makes 15 call attempts to each sampled number, this rate is thought to be an underestimate of the response rate (CDC 2005a, page 3).

I used graphical methods for summary statistics of response rates. I computed annual correlations of the two rates, CASRO and cooperation. I analyzed the response rates by state and year for their components of variation, that is, state, year, and all other (combining residual and interaction) using analysis of variance. I estimate components of variance for state, year, and all other sources of variation and their percentage of the total variation in response rates.

2.4 Item Nonresponse Rates

I analyze the rates (per 100,000 respondents) for three categories, refused, DK/NS, and blank, for selected items that applied to all respondents.

One group of items is demographic characteristics. Besides income these include weight, county of residence and zip code of residence. Income is well-known to have high nonresponse rates and the other items provide a context for these rates. Another group of items includes self-reported health characteristics and medical screening and prevention status. These include general health status, days in the last month that physical health was not good, days in the last month that mental health was not good, having health insurance, having

high blood pressure, having had a cholesterol check, having had a flu shot in the last 12 months, and ever having had a pneumonia vaccination.

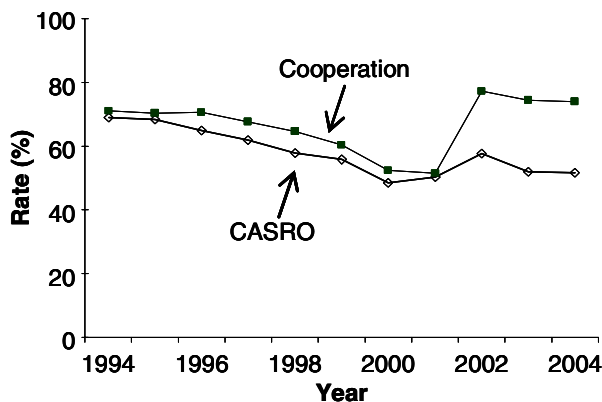
A third group of items shows the changes in rates over the length of the questionnaire. All of these items are from the core questionnaire and were asked of all respondents. Two demographic items are included: county and zip code of residence, both from near the end of section 13. For each of the following sections except section 19 (14, 15, 16, 17, 18, and 20), the first item was selected, except for section 18. Item 18.2 was used since item 18.1 pertains to the workplace and is asked only of employed respondents. Items that were asked before the demographic section (section 13) were not included, since there were nearly no blank items. This is a result of the policy of not including respondents who did not answer a specified number of the demographic items.

3. Results

3.1 Relationship of Cooperation and CASRO Rates over Time

The average cooperation rates and CASRO rates for each year from 1994 through 2004 are shown in Figure 1. These rates followed a similar pattern from 1994 through 2001, decreasing slowly, but since 2002 the average cooperation rate has been substantially higher than the CASRO rate. The cooperation and CASRO rates have a moderately strong relationship. The correlations for each of 11 years ranged from .848 to 0.996.

Figure 1. Cooperation Rates and CASRO Rates of the BRFSS for 1994-2004

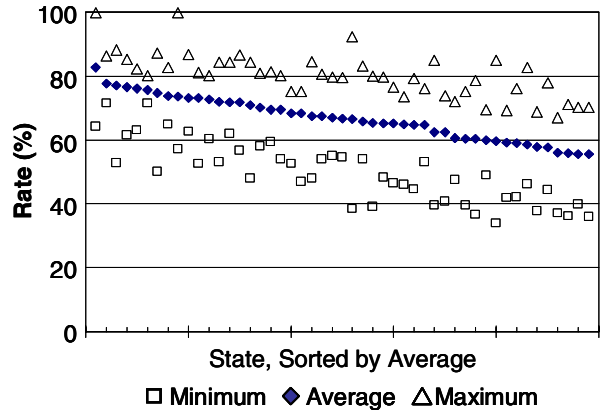


3.2 Cooperation Rates

The cooperation rates ranged from 34.0% to 99.8% over 49 states and 11 years (539 values). The average

cooperation rate for each state, with its minimum and maximum over 11 years, is shown in Figure 2.

Figure 2. Cooperation Rates by State of the BRFSS for 1994-2004



An analysis of variance of cooperation rates (in percent) by year and state is shown in Table 1 with estimates of variance components. The F-statistics for both year ($F=70.9$, 10, 480 df) and state ($F=6.3$, 48, 480 df) are quite large with significance levels below 1 in 10^{30} . Variation among states contributes 1.0% of the variation in the response rates, while that among years contributes 90.0% and the residual contributes 9.0%.

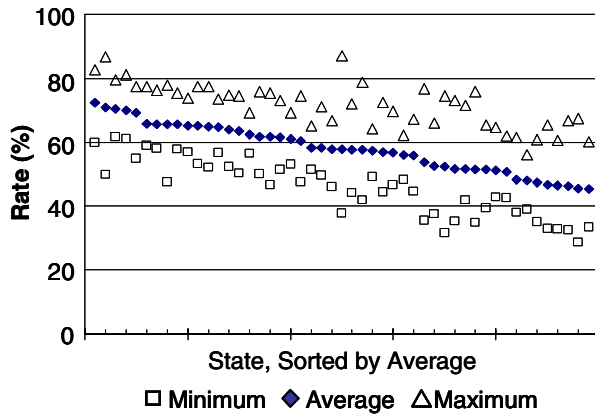
Table 1. ANOVA and Variance Components of Cooperation Rates by State and Year

Analysis of Variance					Variance Components	
Source	SS	df	MS	F	Est	Pct
Year	36,861.8	10	3,686.2	70.9	519.2	90.0
State	15,720.9	48	327.5	6.3	5.6	1.0
Other	24,963.6	480	52.0		52.0	9.0
Total	77,546.3	538			576.8	

3.3 CASRO Rates

The CASRO rates ranged from 28.8% to 87.1% over 49 states and 11 years (539 values). The average CASRO rate for each state, with its minimum and maximum over 11 years, is shown in Figure 3.

Figure 3. CASRO Rates by State of the BRFSS for 1994-2004



An analysis of variance of CASRO rates (in percent) by year and state is shown in Table 4 with estimates of variance components. The F-statistics for both year ($F=50.8, 10, 480$ df) and state ($F=7.9, 48, 480$ df) are quite large with significance levels below 1 in 10^{30} . Variation among states contributes 1.7% of the variation in the response rates, while that among years contributes 86.2% and the residual contributes 12.1%.

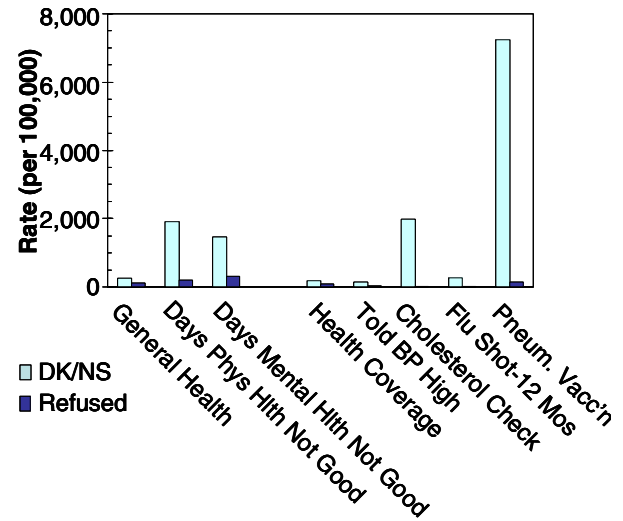
Table 2. ANOVA and Variance Components of CASRO Rates by State and Year

Analysis of Variance					Variance Components	
Source	SS	df	MS	F	Est	Pct
Year	25,716.9	10	2,571.7	50.8	360.2	86.2
State	19,208.3	48	400.2	7.9	7.1	1.7
Other	24,280.5	480	50.6		50.6	12.1
Total	69,205.7	538			417.9	

3.4 Missing Data Rates for Personal Health

The unweighted rates (per 100,000) of respondents who report DK/NS and refused are shown for selected health items in Figure 4. All items had very low refusal rates. There was substantial variation in the rates of DK/NS among items. Items requiring recall of facts had higher rates than items requiring personal judgement. The rates for Days Physical and Days Mental, requiring recall for the last 30 days, though high, were much lower than the rate for pneumonia vaccination, which is pertinent only to patients with chronic diseases or people over age 65.

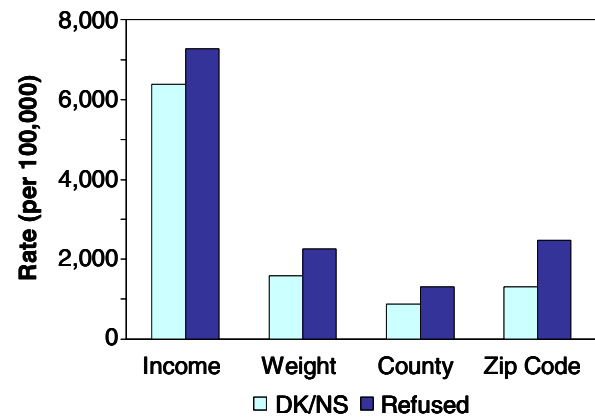
Figure 4. Rates (per 100,000) of DK/NS and Refused for Selected Health Characteristics



3.5 Missing Data Rates for Demographic Items

The rates of DK/NS and refused for several demographic items are shown in Figure 5. Both rates are highest for income but they are also moderately high for weight, county of residence, and zip code. The corresponding rates for weight and zip code are similar, with both exceeding the rates for county of residence.

Figure 5. Rates (per 100,000) of DK/NS and Refused for Selected Demographic Items



3.6 Missing Data Rates Across the Questionnaire

The rates of blank responses for selected items across the first part of the questionnaire, common to all states, are shown in Figure 6. The rates of DK/NS for these items are shown in Figure 7 and the rates of refusal are shown in Figure 8. Questions asked later in the interview had higher rates of blank responses. There

was no regular pattern of DK/NS or refused by questionnaire sequence.

Figure 6. Rates (per 100,000) of Blank Responses for Selected Items in Sequence

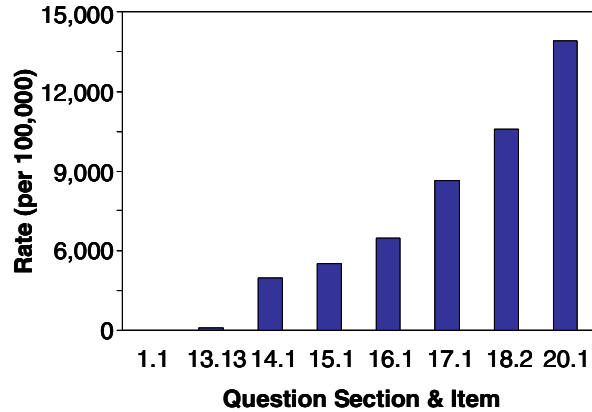


Figure 7. Rates (per 100,000) of DK/NS Responses for Selected Items in Sequence

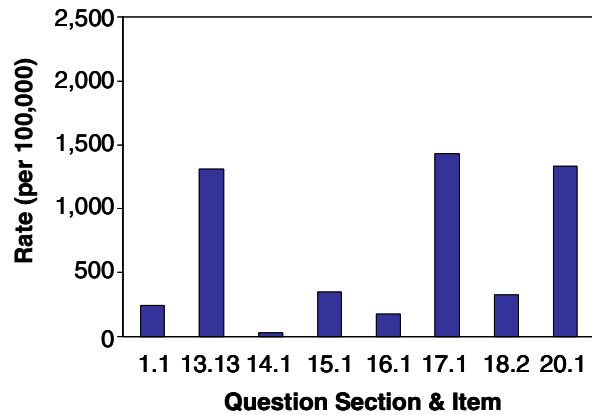
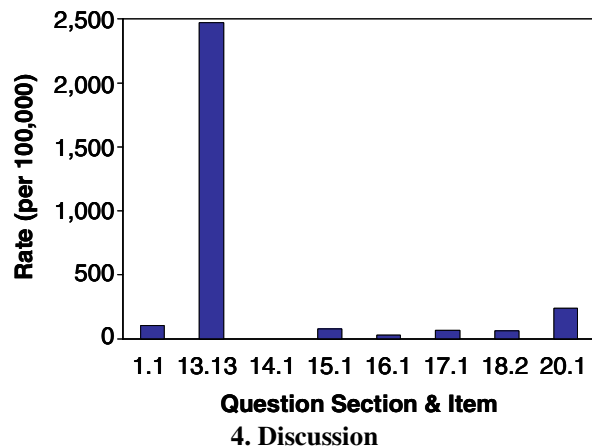


Figure 8. Rates (per 100,000) of Refusals for Selected Items in Sequence



4. Discussion

The analysis of the components of variance for both cooperation rates and CASRO rates shows that variation in states is quite small compared with variation in years. Moreover, the variance among states is rather small compared with the residual variation, which is the total of state-year interactions and any other sources.

The large variation among years suggests that much of the variation is due to shared influences, both external influences that affect all the states and internal practices that affect the design and conduct of the BRFSS: sample design, questionnaire design, and interview procedures.

The size of the residual component suggests that some of the variation may be due to sources that are either not under the control of states or sources that might be controlled but are not at the present time. That is, no particular state, or its vendor, is able to maintain a uniformly better response rate over several years.

Consistent with other reports, the rates of missing values for income are very high, but not the highest. The rate of DK/NS for pneumonia vaccination is much higher but the rate of refusal is much lower. Moreover, the rates for DK/NS and refused to state income are fairly similar. The rates for county of residence are lowest and the rates for zip code of residence are somewhat higher. This pattern suggests that large numbers of respondents do not know their income rather than that they know their income and will not state it. The difference between rates for county and zip code suggests that as the requested information verges into the more personal and more detailed that larger numbers both do not know the requested information and are not willing to report it.

The refusal rates for all the health characteristics are, by contrast, quite low. The primary variation in these items is in the rates of DK/NS. One important contrast is in general health status, the first question, which calls for a personal judgement, and the next two items, days of poor physical and mental health in the last month. The DK/NS rate for both these items is much higher than for general health, suggesting that the request for explicit factual information presents a substantial cognitive challenge to many more respondents that is a barrier to obtaining an answer.

Three factual health characteristics had very low rates for DK/NS: health care coverage, high blood pressure, and a recent flu shot. The rates for having had one's cholesterol checked and having had a pneumonia immunization are much higher. Pneumonia immunization

has no relevance to most respondents, it is recommended for people with chronic disease and people over age 65. A cholesterol check can be done with little notice to a patient and, if negative, no further discussion with a physician is necessary.

The rates of blank fields increase steadily for later sections of the interview, reflecting cumulative dropping out as the interview progresses. These rates appear to have nothing to do with the rates of DK/NS or refused. While they are an alternative, permanent form of refusal, the processes determining them may be quite different from item nonresponse.

5. Conclusion

Variation in the CASRO and cooperation rates is much greater across years than among states. This suggests that uniform standards for 50 or more different surveys are effective in maintaining response rates and that reduced response rates over time are due to external conditions faced by all the states.

The processes giving rise to these three categories of missing data, DK/NS, refused, and a blank field, may be quite different. This suggests that different methods of compensating for missing data, depending on its type, may be useful in overall analysis, particularly imputation.

References

- Centers for Disease Control and Prevention (1999). 1998 BRFSS Summary Quality Control Report.
- Centers for Disease Control and Prevention (2000). 1999 BRFSS Summary Quality Control Report.
- Centers for Disease Control and Prevention (2001). 2000 Summary Data Quality Report.
- Centers for Disease Control and Prevention (2002). 2001 Summary Data Quality Report.
- Centers for Disease Control and Prevention (2003). 2002 Summary Data Quality Report.
- Centers for Disease Control and Prevention (2004). 2003 Summary Data Quality Report.
- Centers for Disease Control and Prevention (2005a). Behavioral Risk Factor Surveillance System Operational and User's Guide, Version 3.0, March 4, 2005.
- Centers for Disease Control and Prevention (2005b). 2004 Summary Data Quality Report.
- Centers for Disease Control and Prevention (2000). BRFSS Summary Quality Control Report. Atlanta (GA): Centers for Disease Control and Prevention, 2001.
- Holtzman D. Analysis and interpretations of data from the US Behavioral Risk Factor Surveillance Sys-

tem. In: McQueen DV, Puska P, Editors. Global Behavioral Risk Factor Surveillance. New York: Kluwer Academic/Plenum Publishers; 2003, p. 35-46.