Mean Square Error Analysis of Health Estimates from the Behavioral Risk Factor Surveillance System for Counties along the United States-Mexico Border Region

Joe Fred Gonzalez, Jr.[1], Machell Town[2], and Jay J. Kim[1]
Centers for Disease Control, National Center for Health Statistics, [1]
National Center for Chronic Disease Prevention and Health Promotion[2]

Summary

The Behavioral Risk Factor Surveillance System (BRFSS) is a State telephone based survey of the civilian non-institutionalized adult (18 years and over) population residing in the United States. Consequently, the BRFSS final weights that are currently available in the data files are designed to produce unbiased estimates of socio-demographic and health characteristics for adults at the State level. In addition to State and national level BRFSS estimates, there is another geographical subpopulation of interest, that is, the border counties within the four United States-Mexico border States: Arizona, California, New Mexico, and Texas. The focus of this paper will be on the 44 counties which are within 100 kilometers (62 miles) of the border, as defined by the "Healthy Border 2010 Program," United States-Mexico Border Health Commission. The purpose of this paper is to investigate alternative ways of arriving at post-stratification factors (ratio adjustments) by collapsing rows or columns by age-sex-ethnicity/race for producing final weights/estimates for this border region. A conditional mean square analysis was used to observe the effect of cell collapsing on bias and variance estimators for several BRFSS socio-demographic and health characteristics.

Keywords: Collapsing, Post-stratification factors, Bias, Variance

1. Introduction

The BRFSS is a State telephone based survey of the civilian non-institutionalized adult (18 years and over) population residing in the United States. However, there is interest in another geographical subpopulation, the 44 counties which are within 100 kilometers (62 miles) of the border, as defined by the "Healthy Border 2010 Program," United States-Mexico Border Health Commission. Healthy Border 2010 aims to improve health in the United States - Mexico border region, an area defined as 100 kilometers (62 miles) north and south of the United States – Mexico border (Figure 1.) This area includes 80 municipios (equivalent of counties) in 6 Mexican states and 48 counties in 4 U.S. states. Since there are 44 counties of interest for the "Healthy Border 2010 Program," United States-Mexico Border Health Commission, this investigation will also focus on U.S. data which are limited to the same 44 border counties, excluding Maricopa, Pinal, and La Paz counties in Arizona and Riverside County in California (Figure 2.).

First, it was determined that it would be worthwhile to produce BRFSS percent estimates for the adult population in the border region by certain age-sex-ethnicity/race cells. The desired six age groups were: 18-24, 25-34, 35-44, 45-54, 55-64, and 65 and over. The desired three ethnicity/race groups were: Hispanic, White Non-Hispanic, and Black/Multiracial and others. Initially, BRFSS sample counts were tabulated by age-sex-ethnicity/race within each border county. However, the sample counts were insufficient for many cells within each border county. Secondly, cell counts were investigated across all border counties within each of the border States. However, again cell sizes were insufficient within each State. Finally, sample sizes were aggregated by the desired age-sex-ethnicity/race cells for all 44 border counties across all four border States. Fortunately, cell sizes were sufficiently large for the desired age-sex-ethnicity/race cells for both Hispanics and White Non-Hispanics, and in a few instances for Black/Multiracial and others. This level of geographical aggregation was defined as the United States-Mexico Border Region for the purpose of our paper. Hereafter, the United States-Mexico Border Region will be simply referred to as the "border region" and is similarly defined for each of the three years 2001-2003. In addition, the same age-sex-ethnicity/race crosstabulation that was used for determining sample size sufficiency was also used as the weighting matrix for this investigation.

2. Sample Weighting Procedures in the Border Region

As mentioned earlier, the purpose of this paper is to investigate alternative ways of producing ratio adjustment factors (modified post-stratification factors) by collapsing adjacent age-sex cells within three ethnicity/race groups (Hispanics, White-Non-Hispanic, and Blacks/Multiracial and Others) within the 44 counties which are within 100 kilometers (62 miles) of the U.S.-Mexico Border.

Post-stratification is used for incorporating population distributions of key socio-demographic variables into survey estimates. Some references about post-stratification follow.

- Holt and Smith (1979): If a sample is badly balanced for some characteristic, post-stratified estimator automatically corrects for this.
- Little (1993): Post-stratification can improve the accuracy of estimates both by reducing bias and by increasing precision.
- Kim (2004): Effect of Collapsing Rows/Columns of Weighting Matrix on Weights.

For this analysis, the variable _WT2, which is available in the 2001-2003 BRFSS data sets is the initial sample weight as follows:

_WT2 =_STRWT * NAD / NPH

where,

STRWT = within State stratum weight,
NAD = number of adults in household, and
NPH = number of phones in the household.

For purposes of this investigation, the initial sample weight (_WT2) was used to create the "initial poststratification factors (PSF)" which were calculated in the usual manner by age (6 groups)-sex (2)-ethnicity/race (Hispanic, White Non-Hispanic, and Black/Multiracial and Others) as follows:

PSF = Census pop. count within an i-th cell / sum of _WT2 within same i-th cell.

Table 1 shows the initial 2003 poststratification factors (PSF) that were multiplied by the basic sampling weights (_WT2) for adults to produce unbiased estimates of adult health characteristics in the border region. Similarly, "initial poststratification factors (PSF)" were calculated for the years 2001-2002, but are not shown in this paper due to space limitations.

The "initial poststratified Final Weights" used in this investigation were calculated in same fashion for each of the three years 2001-2003 as follows:

"Final_ Weight" = _WT2*PSF.

where PSF is as previously defined.

The "Final_ Weights" were used to produce BRFSS percent estimates of adult characteristics for all three years (2001- 2003) using the following 11 binary health variables for this research:

- Ever had Asthma
- Asthma now
- Ever had high blood pressure
- High cholesterol
- Diabetes
- Taking insulin
- Taking diabetic pills
- Having health insurance
- Current smoker
- Former smoker
- Any exercise.

Additionally, there were three binary health variables (sigmoidoscopy, Pap smear, and Mammogram) that were available for only the years 2002 and 2003.

3. Conditional Mean Square Error Analysis

First, we will introduce the usual notation involved in doing a mean square error (MSE) analysis as follows:

$$MSE(p) = [Bias(p)]^2 + [se(p)]^2$$

where p= percent estimator of a health characteristic, and
se (p)= standard error estimator for the same health characteristic.

To get a sense of the effect of the bias on the MSE, the root mean square error (RMSE) was calculated by simply taking the square root of the MSE, that is,

$$RMSE(p) = \sqrt{MSE(p)} = \sqrt{[Bias(p)]^2 + [se(p)]^2}.$$

Therefore, if the RMSE (p) was close to the standard error (se) of a percent estimator, the bias was small. Alternatively, if the RMSE (p) was much larger than the se (p), then the bias was large.

The percent estimates of health characteristics using the "initial poststratified Final_Weights" are unbiased estimates and are referred to as "**P**" in the first column of Table 5. The entries in the "**P**" column of Table 5 are treated as "*parameters,*" that is, as true values of health characteristics for the adult population in the border region for this mean square error (MSE) analysis . So, in reality, the bias, variance, and the mean square error (MSE) analysis is *conditiona*l. The MSE analysis was performed by comparing these "*parameters (***P***)*" of health characteristics with corresponding percent estimates of health characteristics generated by using the following two approaches:

- ♦ *standard cell collapsing (S)* the usual approach which is driven by the usual sample size considerations (minimum cell count = 20), maximum ratio criteria by domains (Table 2), and row adjacency, and

- ♦ *truncation (or censoring of ratios) and collapsing (T)* approach.

"New" PSF, corresponding Final Weights, and corresponding percent estimates were produced by using each of the two approaches above.

Table 3 (Kim, 2004) defines the quantities that are involved for producing PSF using *standard cell collapsing (S)*..

A collapsing example for 2003 BRFSS Black/Multiracial/Others follows. The collapsing ratio criterion for 2003 Black is 6. The 18-24 group has the ratio of 8.06 and 25-34 has 10.18. Both rows have to be collapsed. Even if they are combined, the ratio should be much bigger than 6, so the combined cell is further collapsed with 35-44. The results of collapsing are shown in Table 4 . [Note in Table 4 that f_i is the PSF for cell i and $c= N_2 / N_1 = 0.53$.]

The collapsing adjustment factor (CAF) (Kim, 2004) for the 18-34 age group = $\frac{(1+c)\, f_2}{cf_1 + f_2} = 0.55.$ That is, the sampling weight of the 18-34 age group decreases by 45% due to collapsing. Similarly, the CAF for the 35-44 age group = $\frac{(1+c)\, f_1}{cf_1 + f_2} = 1.85.$ That is, the sum of the poststratified sampling weights for the 35-44 group increases by 85% due to collapsing. Therefore, there is too much of a shift in weights in Table 4 due to collapsing, In the above example, the PSF for the age group 18-34 is 9.11, but that for 35-44 is 2.72. When they are combined, the newly computed PSF for combined group can go down below 6. Note that any cell whose PSF is 5.9 or close to it will keep the same PSF, but the cell whose PSF is greater than 6 can become smaller than 6 after collapsing, which causes underestimation for the poorly covered groups. To avoid this situation Kim, Tompkins, Li, and Valiant (2005) proposed and implemented in their simulation study two approaches which involve censoring the PSF.

The approach we adopted for this paper called "truncation (censoring)/collapsing" is slightly different from the approaches mentioned above. The current approach is based on the CAF. Any CAF less than 0.8 suggests that the associated cell would lose at least 20 percent of its own weights. Thus to mitigate this loss, the approach that was taken so that the final CAF would be around 0.8, required pre-multiplication of each weight in the severely under-covered cells by 1.15, 1.25, etc.

Table 5 shows some 2002 MSE (RMSE) analysis results for comparing "standard collapsing (S)" vs. "truncation(censoring)/collapsing (T)." Table 6 shows a summary comparison of MSE (RMSE) analysis for ethnic/racial groups for several years.

Using the normal approximation, a hypothesis test at the 1% significance level for one population proportion was performed to see if the observed proportion ($\hat{p} = \frac{x}{n} = \frac{76}{105}$. See Table 6) of cases when the Truncation (censoring)/collapsing approach was superior to that of the Standard Collapsing (S) approach the majority of the time. The right-tail test of hypothesis for one proportion follows:

H_0: P = 0.5
Ha: P > 0.5.
The value of the test statistic is

$$Z = \frac{\hat{p} - 0.5}{\sqrt{\frac{0.5(1-0.5)}{105}}} = \frac{\frac{76}{105} - 0.5}{0.0488} = 4.59.$$

The critical value for Z at the 1% level of significance is 2.33 < 4.59, therefore, we conclude that the MSE (RMSE) of health estimates based on "Truncation/Collapsing" is significantly lower the majority of the time.
NOTE: As mentioned earlier, the observed proportion of "Truncation/Collapsing Wins" is $\hat{p} = \frac{x}{n} = \frac{76}{105}$, see last "total "row in Table 6. Note from Table 6 that n = 76 + 29 =105 (ties excluded). The sample proportion ($\hat{p}$) proportion of "Truncation/Collapsing Wins" could be determined in two ways: ties included and ties excluded. The reason why ties occur is that all the persons with certain characteristics belong to the age groups whose weights were not impacted by weight truncations. What this suggests is that it is fair to ignore the ties in statistical tests.

4. Conclusion and Further Research

Results from this investigation indicate that in general when calculating PSF the truncation (censoring)/collapsing (T) of cells/rows/columns approach is superior to standard collapsing (S) in terms of lower RMSE. These findings were based on a heuristic approach, but were also confirmed by a simulation study (Kim, et al, 2005). Ideally, it would be extremely useful if an optimal strategy could be developed for creating the "best poststratification factors" in terms of producing a *minimum* mean square error for the health estimates of interest. This is an area for further research.

References

1.) Kim, Jay J. (2004) Effect of Collapsing Rows/Columns of Weighting Matrix on Weights, Proceedings of the ASA Survey Research Methods Section, 2004.

2.) Kim, Jay J., Tompkins, Linda, Li, Jianzhu, and Valliant, Richard, A Simulation Study of Cell Collapsing in Poststratification (to appear in the Proceedings of the American Statistical Association, Survey Research Methods Section, 2005)

3.) Holt, D. and Smith, T.M.F. (1979). Post Stratification, Journal of Royal Statistical Society, Series A, Vol. 142, Part I, pp. 33-46.

4.) Little, R.J.A. (1993). Post-Stratification: A Modeler's Perspective, Journal of the American Statistical Association, Vol. 88, No. 23, pp. 1001-1012.

Figure 1.

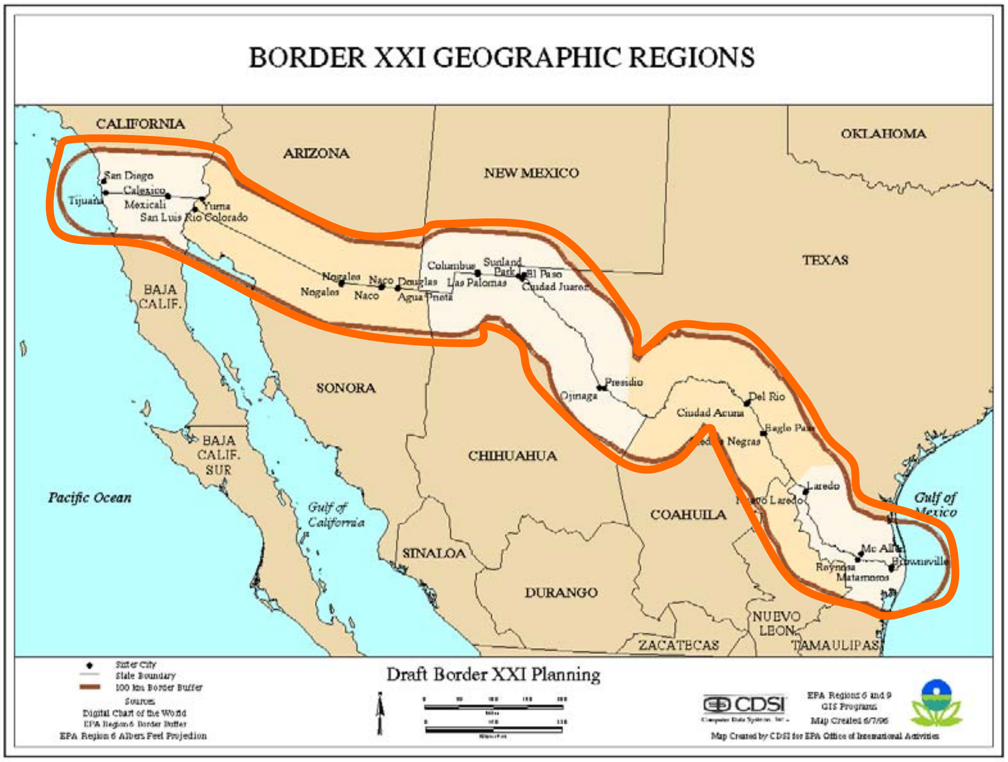

Figure 2.

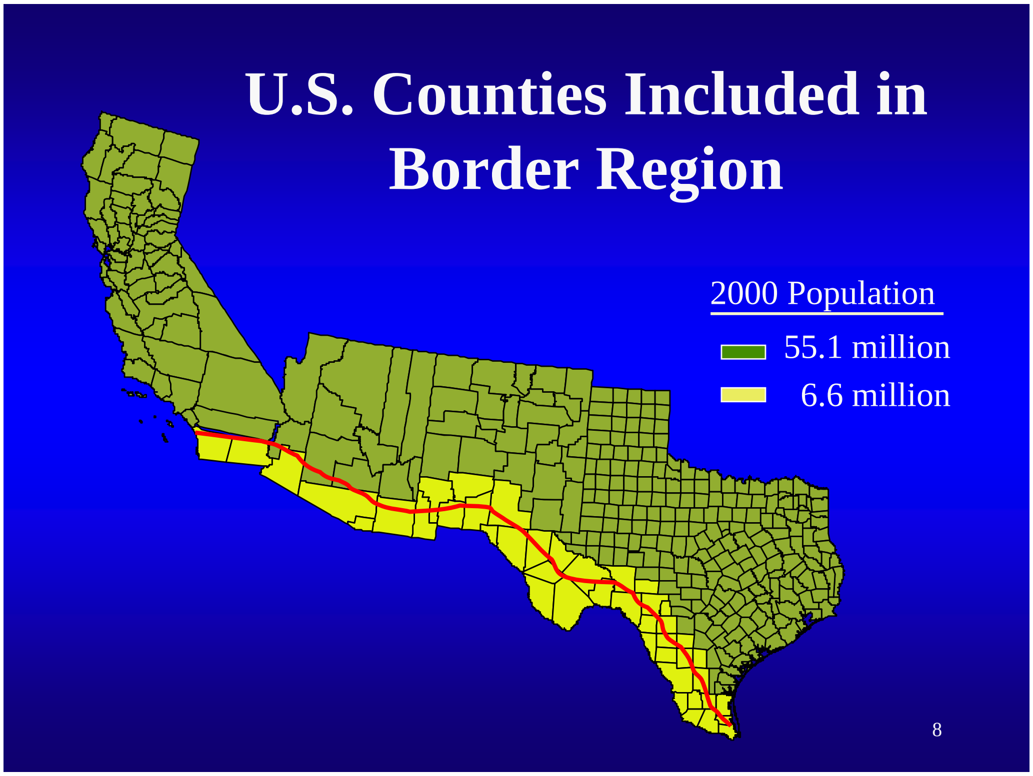

Table 1. Initial 2003 Post-stratification Factors (PSF) by Age-Sex-Ethnicity/Race for the Border Region.

Age and Sex	Ethnicity/Race		
	Hispanic	White Non-Hispanic	Black/Multiracial/Other
Male			
18-24	6.5	5.4	8.1
25-34	5.3	5.0	10.2
35-44	5.7	3.5	2.7
45-54	4.6	4.2	3.0
55-64	3.8	4.2	2.4
65+	9.2	3.1	3.9
Female			
18-24	4.5	3.9	4.1
25-34	2.6	3.2	6.4
35-44	3.3	3.3	4.0
45-54	2.6	2.8	7.5
55-64	3.6	2.9	4.2
65+	4.3	3.5	5.0

Table 2. Maximum Ratio Criteria Used for Standard Collapsing (S) by Year

Ethnicity/race	Year		
	2001	2002	2003
Hispanic	4	5	6
White	4	4	4
Black/Multiracial/Others	5	6	6

Table 3. Weighting Matrix for Calculating Usual PSF.

Rows	Raw Sample Count	Initially Weighted Sample Count	Control Count
Row 1	n_1	W_1	N_1
Row 2	n_2	W_2	N_2
Row 3	n_3	W_3	N_3

Table 4. Standard Collapsing (S) Example.

Age Group	N_i	W_i	f_i
18-34	113255	1243.273	9.11
35-44	60012	22073.83	2.72

Table 5[†]. Some Results (in percent) of MSE (RMSE) Analysis for Several Health Variables, Assuming "P" as the Population Parameter for each Health Variable, and a Comparison of "Standard Collapsing (S)" vs. "Truncation (Censoring)/Collapsing (T)," 2002 Black/Multiracial/Other Males, 45-54.

Variable (Yes)	P	P-se	S-est	S-se	S-RMSE	T-est	T-se	T-RMSE
Asthma-ever	20.9	16.5	24.1	18.3	18.6	23.2	17.8	17.9*
Asthma-now	6.7	8.8	6.7	8.8	8.8	6.7	8.8	tie
Diabetes	11.8	7.1	12.4	7.9	7.95	12.3	7.68	7.69*
Taking insulin	31.7	26.5	34.6	28.7	28.8	33.8	28.1	28.13*
Diabetes Pills	23.0	17.8	15.8	13.4	15.2	17.8	14.7	15.6
Current smoker	20.0	11.5	17.7	10.0	10.2	18.4	10.4	10.5
Former smoker	27.2	11.5	21.6	9.9	11.4	23.2	10.4	11.1*

[†] NOTE: **P** = estimated percent without collapsing which is assumed to be the true value of a health characteristic of adults in the population based on this study; **P-se** = standard error of **P**; **S-est** = percent estimate using Standard Collapsing (S); **S-se** = standard error of percent estimator using Standard Collapsing (S); **S-RMSE** =root mean square error of percent estimator using Standard Collapsing (S); **T-est** = percent estimate using Truncation (Censoring)/Collapsing (T); **T-se** = standard error of percent estimator using Truncation (Censoring)/Collapsing (T); **T-RMSE** = root mean square error of percent estimator using Truncation (Censoring)/Collapsing (T).

Table 6. Some MSE Analysis Findings when Comparing "Truncation(censoring)/ Collapsing (T)," vs. "Standard Collapsing (S)" for Ethnic/Racial Groups for Several Years.

Group	Truncation(censor)/ Collapsing (T). Wins	Standard Collapsing (S) Wins	Tie
2001 Hispanics	11	3	0
2001 Black/Multiracial/Other	29	15	3
2002 Black	15	3	10
2003 Hispanics	8	3	0
2003 Black/Multiracial/Other	13	5	10
Total	76	29	23