A Comparison of Small Area Models Used in the Quality Indicator Program Sponsored by Agency for Healthcare Research and Quality[1]

Robert M. Baskin[1], Pamela L. Owens[1], Christopher J. Sroka[2] and Jeffrey J. Geppert[2]

1. Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville MD, 20850
2. Batelle Memorial Institute, 505 King Avenue, Columbus, OH 43201

Introduction

The Agency for Healthcare Research and Quality (AHRQ) has developed an array of measures called Quality Indicators™, which are health care decision making and research tools. The Quality Indicators can be used by program managers, researchers, and others at the Federal, State and local levels as measures of health care quality. The Prevention Quality Indicators (PQIs), one of four modules of the AHRQ Quality Indicators, identify hospital admissions in a geographic area that evidence suggests may have been potentially avoided through access to high-quality outpatient care. The PQI technical specifications and user-friendly software are used to calculate rates of hospitalization at the national, regional, and county-level. The current PQIs adjust to general populations in counties but not to condition specific populations. For example, diabetes or asthma PQIs are adjusted to the general population of the county instead of the condition-specific populations since those estimates are not currently available at the necessary level of geographic detail. For more information, visit the AHRQ Quality Indicator website: http://qualityndicators.ahrq.gov

The objective of this study is to evaluate domain specific models for estimating county-level counts of persons with specific conditions measured by the PQIs (e.g., diabetes or asthma). This information is only intended to inform the methodology of the AHRQ PQIs. The approach will be to build small area models for diabetes prevalence at the county level based on existing estimates of county level diabetes prevalence.

Data Sources

Two sources of data are available for model building but neither covers all geography consistently. The Behavioral Risk Factor Surveillance System (BRFSS), conducted by the Centers for Disease Control and Prevention, is a telephone survey of health risk factors and has been used to produce some county level estimates. BRFSS does collect information on conditions of interest such as diabetes but also has some issues. BRFSS does not provide public information at the county level if there is deemed to be too few observations in the county, i.e., observations can have county information suppressed. Some covariates in the BRFSS data have missing values. Because it is a telephone survey there are also concerns with population coverage. However, the BRFSS data are publicly available at the observation level and is therefore available for model building.

The Medical Expenditure Panel Survey – Household Component (MEPS-HC), conducted by the Agency for Healthcare Research and Quality, is a household level survey of medical conditions, usage, and expenditures that has also been used to produce some small area estimates. MEPS-HC has a large amount of medical information about each record but the geography is not publicly available below the Census region and the geographic coverage is very sparse. Therefore it was decided to use the BRFSS data for building the small area models.

Models

The small area models will be fit using a hierarchical Bayesian model and estimated using Markov chain Monte Carlo (MCMC) sampling and in some cases Laplace approximation of the posterior. As stated in Rao (2003) it is now generally accepted that when indirect estimators are required they should be based on explicit small area models and that the hierarchical Bayes method is extensively used because it is straightforward, inferences are 'exact', and complex problems can be handled with MCMC methods.

Because the PQIs adjust at the level of county by age, sex, and race/ethnicity the models will be fit at the level of county by age, sex and race/ethnicity categories (ASRE). The model will follow the work of Barker et. al. (2013) who built county level

[1] Disclaimer: The views expressed in this presentation are those of the authors and do not necessarily reflect those of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Bayesian models for diabetes but not at the level of categories of ASRE. Current PQI software uses 36 levels of ASRE categories, but because of sparseness of observations in some of the categories the categories were collapsed to 12 levels of ASRE with three levels of age, two levels of sex and two levels of Race-Ethnicity.

The hierarchical Bayesian model is

$$y_{ij} \sim \text{binomial}(p_{ij}, n_{ij})$$
$$\text{logit}(p_{ij}) = \alpha_j + \mu_{ij} + \nu_{s(i)j} + \omega_{s(i)j}$$

where

N = number of counties = 3143
M = number of age, sex, race/ethnicity categories = 12
n_{ij} = number of sampled persons in county i and ASRE j, weighted
($i = 1, 2, \ldots N; j = 1, 2, .. M$)
y_{ij} = number of persons with diabetes observed in county i and ASRE j, weighted
($i = 1, 2, \ldots N; j = 1, 2, .. M$)
p_{ij} = prevalence of diabetes in county i and ASRE j
($i = 1, 2, \ldots N; j = 1, 2, .. M$)
α_j = fixed effect of each ASRE, common across all counties
($j = 1, 2, .. M$)
μ_{ij} = random effect for each county-class combination
($i = 1, 2, \ldots N; j = 1, 2, .. M$)
$\nu_{s(i)j}$ = random effect for each state-class combination; s(i) is the state containing county i
($s(i) = 1, 2, \ldots, 51; j = 1, 2, \ldots, 12$)
$\omega_{s(i)j}$ = spatial effects for each state-class combination; s(i) is the state containing county i
($s(i) = 1, 2, \ldots, 51; j = 1, 2, \ldots, 12$)

The prior distributions for the model are as follows:

$$\alpha_j \sim \text{flat}$$
$$\mu_{ij} \sim \text{Normal}_M(0, T_\mu)$$
$$\nu_{s(i)j} \sim \text{Normal}_M(0, T_\nu)$$
$$\omega_{s(i)j} \sim \text{MVN CAR}(T_\omega)$$
$$T_\mu \sim \text{Wishart}(S, M)$$
$$T_\nu \sim \text{Wishart}(S, M)$$
$$T_\omega \sim \text{Wishart}(S, M)$$

where S is a M × M matrix with 1's along the diagonal and 0.001 for all other elements.

In order to speed convergence the model without spatial effects is centered as described in Gilks and Roberts (1996). The centered model is reparameterized as

$$y_{ij} \sim \text{binomial}(p_{ij}, n_{ij})$$
$$\text{logit}(p_{ij}) = \alpha_j$$
$$\mu_{ij} \sim \text{Normal}_M(\nu_{s(i)j}, T_\mu)$$
$$\nu_{s(i)j} \sim \text{Normal}_M(\alpha_j, T_\nu)$$

Note that although we start with N = 3143 counties, we add "super counties" to each state to capture respondents with missing counties and we replace 29 Alaska counties with 5 regions because only one Alaska observation had a valid county code. As a result, N = 3170 for the model fitting.

BRFSS Data

Data on the number of persons with diabetes at the county and ASRE level was obtained from the BRFSS. Three years of BRFSS data, 2008, 2009, and 2010, were combined to provide estimates centered at 2009. For diabetes, we used the following question from the BRFSS core section:

Ever Told by Doctor You Have Diabetes

Section: 6.1 Diabetes **Type:** Num

Column: 87 **SAS Variable Name:** DIABETE2

Prologue:

Description: Have you ever been told by a doctor that you have diabetes (If "Yes" and respondent is female, ask "Was this only when you were pregnant?". If Respondent says pre-diabetes or borderline diabetes, use response code 4.)

Value	Value Label	Frequency	Percentage	Weighted Percentage
1	Yes	52,386	12.11	9.10
2	Yes, but female told only during pregnancy	3,163	0.73	0.91
3	No	369,824	85.49	88.66
4	No, pre-diabetes or borderline diabetes	6,809	1.57	1.23
7	Don't know/Not Sure	277	0.06	0.07
9	Refused	148	0.03	0.03

A new diabetes variable was created (DBNOGEST) to indicate whether the individual was ever diagnosed with diabetes that was not gestational.

$$DBNOGEST = \begin{cases} 1 \text{ if } DIABETE2 = 1 \\ 0 \text{ if } DIABETE2 = 2, 3, 4 \\ . \text{ otherwise} \end{cases}$$

A weighted version of DBNOGEST (called DBWTD) was created similarly:

$$DBWTD = \begin{cases} \_FINALWT \text{ if } DIABETE2 = 1 \\ 0 \text{ if } DIABETE2 = 2, 3, 4 \\ . \text{ otherwise} \end{cases}$$

where _FINALWT is the final weight applied to each observation that reflects both sampling weighting and weighting to adjust for nonresponse.

For demographic information, we use the BRFSS variables AGE (age in whole years), SEX, and RACE2. Note that RACE2 is a variable calculated by CDC based on responses to several questions regarding race and ethnicity.

Model Results

The model was initially fit as a generalized linear mixed model. The results from this model were then used as initial values for the Bayesian hierarchical model. As can be seen from the table comparing the Bayesian parameters to the GLMM parameters for ASRE the Bayesian model accounted for little change in the parameters. The output from the GLMM fit in R is given in Model Table 1 and the comparison of the Bayesian estimates with the GLMM estimates is given in Model Table 2 below.

Model Table 1.

- dbnogest is the binomial variable for non-gestational diabetes
- fipstco is the county FIPS code
- statecode is the state FIPS code
- asre_fmt*nn* is the *nn* level of the Age by Sex by Race indicator

Generalized linear mixed model fit by the Laplace approximation
Formula: cbind(dbnogest, n - dbnogest) ~ 0 + asre_fmt + (1 | fipstco) + (1 | statecode)
Data: model12
AIC BIC logLik deviance
27793 27906 -13882 27765

Random effects:
Groups Name Variance Std.Dev.
fipstco (Intercept) 0.030104 0.17351
statecode (Intercept) 0.036037 0.18984
Number of obs: 24079, groups: fipstco, 2245; statecode, 51

Fixed effects:
Estimate Std. Error z value Pr(>|z|)
asre_fmt01 -3.56435 0.03481 -102.39 <2e-16 ***
asre_fmt02 -3.01196 0.03930 -76.63 <2e-16 ***
asre_fmt03 -3.58433 0.03256 -110.10 <2e-16 ***
asre_fmt04 -3.01264 0.03441 -87.54 <2e-16 ***
asre_fmt05 -2.01426 0.02873 -70.10 <2e-16 ***
asre_fmt06 -1.41475 0.03117 -45.39 <2e-16 ***
asre_fmt07 -2.24930 0.02850 -78.91 <2e-16 ***
asre_fmt08 -1.42924 0.02979 -47.98 <2e-16 ***
asre_fmt09 -1.37377 0.02863 -47.98 <2e-16 ***
asre_fmt10 -0.81222 0.03266 -24.87 <2e-16 ***
asre_fmt11 -1.70399 0.02831 -60.18 <2e-16 ***
asre_fmt12 -0.85443 0.03045 -28.06 <2e-16 ***

Model Table 2. Comparison of Bayesian and GLMM parameters for the twelve ASRE categories

	Bayesian Parameter	**GLMM Estimate**	**Difference**	**% Difference from GLMM**
Age*Sex*Race[1]	-3.6139	-3.56435	-0.04955	1.39%
Age*Sex*Race[2]	-3.067	-3.01196	-0.05504	1.83%
Age*Sex*Race[3]	-3.6438	-3.58433	-0.05947	1.66%
Age*Sex*Race[4]	-3.0306	-3.01264	-0.01796	0.60%
Age*Sex*Race[5]	-2.048	-2.01426	-0.03374	1.68%
Age*Sex*Race[6]	-1.4499	-1.41475	-0.03515	2.48%
Age*Sex*Race[7]	-2.2825	-2.2493	-0.0332	1.48%
Age*Sex*Race[8]	-1.494	-1.42924	-0.06476	4.53%
Age*Sex*Race[9]	-1.3889	-1.37377	-0.01513	1.10%
Age*Sex*Race[10]	-0.8345	-0.81222	-0.02228	2.74%
Age*Sex*Race[11]	-1.7244	-1.70399	-0.02041	1.20%
Age*Sex*Race[12]	-0.9095	-0.85443	-0.05507	6.45%

Validation

Our model validation consisted of two stages. The first stage was to compare aggregated estimates to direct estimates from the BRFSS and the second stage involved cross-validation. Both types of validation were carried out in SAS.

Compare Aggregated Estimates

Using PROC SURVEYFREQ, direct estimates of diabetes and 95% confidence intervals were calculated from the BRFSS. Three sets of estimates were calculated:

- A national estimate
- Estimates for each state and DC
- Estimates for each of the 12 ASRE categories

PROC SURVEYFREQ was run using the BRFSS variable _STSTR for the stratum, _PSU for the cluster, and _FINALWT for the weight. For the ASRE estimates, a domain analysis was conducted by crossing the ASRE indicator with the DBNOGEST variable in the TABLES statement. For the state estimates, a BY statement was used instead since the data are already stratified by state.

The small area estimates were aggregated nationally, by state, and by ASRE category. We then compared these aggregated values to the 95% confidence intervals around the direct estimate. The results are presented in tables 1-3 below. In Tables 1 and 2, for the national and ASRE estimates, we see that each aggregated small area estimate falls within the corresponding 95% CI for the direct estimate. In Table 3 however, when aggregated to the state level, 39 of the aggregates fall within the 95% CI of the corresponding direct estimate, 11 of the aggregates fall above the 95% CI of the corresponding direct estimate, and the aggregated estimates for Oregon fall below the 95% CI for the direct estimate in Oregon.

Table 1: Comparison of Small Area Estimates to Direct BRFSS Estimates, National Level

Diabetes estimate, direct	Diabetes estimate, summing small area estimates	Population, sum of BRFSS weights	Population, Census 2009	Diabetes prevalence, direct	95% LCL, direct	95% UCL, direct	Diabetes prevalence, small area estimate	Comparison of SAE estimate to interval
20,867,439	21,235,459	230,436,745	232,358,081	9.06%	8.89%	9.22%	9.14%	Within CL

Table 2: Comparison of Small Area Estimates to Direct BRFSS Estimates, ASRE Level

Age category	Sex	Race category	Diabetes estimate, direct	Diabetes estimate, summing small area estimates	Population, sum of BRFSS weights	Population, Census 2009	Diabetes prevalence, direct	95% LCL, direct	95% UCL, direct	Diabetes prevalence, small area estimate	Comparison of SAE estimate to interval
18 to 44	Male	White	873,723	862,151	34,669,464	33,800,350	2.52%	2.23%	2.81%	2.55%	Within CL
18 to 44	Male	Other	919,076	874,963	22,362,108	22,776,100	4.11%	3.39%	4.83%	3.84%	Within CL
18 to 44	Female	White	785,344	824,299	34,314,935	33,186,818	2.29%	2.08%	2.49%	2.48%	Within CL
18 to 44	Female	Other	823,997	893,485	21,058,470	22,698,950	3.91%	3.41%	4.42%	3.94%	Within CL
45 to 64	Male	White	3,111,779	3,263,317	27,606,683	28,583,191	11.27%	10.86%	11.68%	11.42%	Within CL
45 to 64	Male	Other	1,844,385	1,909,470	10,061,438	10,562,102	18.33%	16.98%	19.68%	18.08%	Within CL
45 to 64	Female	White	2,560,963	2,611,015	28,928,368	29,414,451	8.85%	8.55%	9.16%	8.88%	Within CL
45 to 64	Female	Other	1,845,509	2,047,263	10,339,680	11,712,944	17.85%	16.82%	18.88%	17.48%	Within CL
65+	Male	White	2,695,987	2,838,237	13,273,311	13,817,786	20.31%	19.70%	20.93%	20.54%	Within CL
65+	Male	Other	1,068,316	1,004,825	3,278,729	3,207,638	32.58%	30.17%	34.99%	31.33%	Within CL
65+	Female	White	2,833,550	2,797,645	18,072,126	18,023,733	15.68%	15.25%	16.11%	15.52%	Within CL
65+	Female	Other	1,292,963	1,308,789	4,427,160	4,574,018	29.21%	27.51%	30.91%	28.61%	Within CL

Table 3: Comparison of Small Area Estimates to Direct BRFSS Estimates, State Level

State	Diabetes estimate, direct	Diabetes estimate, summing small area estimates	Population, sum of BRFSS weights	Population, Census 2009	Diabetes prevalence, direct	95% LCL, direct	95% UCL, direct	Diabetes prevalence, small area estimate	Comparison of SAE estimate to interval
Alabama	431,459	427,417	3,529,430	3,617,440	12.22%	11.17%	13.28%	11.82%	Within CL
Alaska	29,607	31,694	505,020	511,285	5.86%	4.61%	7.12%	6.20%	Within CL
Arizona	407,339	407,888	4,825,594	4,710,373	8.44%	7.27%	9.61%	8.66%	Within CL
Arkansas	218,692	229,229	2,158,067	2,185,263	10.13%	9.02%	11.25%	10.49%	Within CL
California	2,550,223	2,418,109	27,946,732	27,630,961	9.13%	8.56%	9.69%	8.75%	Within CL
Colorado	215,598	225,366	3,746,519	3,752,946	5.75%	5.24%	6.27%	6.01%	Within CL
Connecticut	176,551	201,275	2,663,737	2,738,555	6.63%	5.95%	7.31%	7.35%	Above UCL
Delaware	53,717	60,694	665,134	684,726	8.08%	7.08%	9.07%	8.86%	Within CL
District of Columbia	35,306	38,199	471,517	488,907	7.49%	6.53%	8.45%	7.81%	Within CL
Florida	1,558,407	1,482,590	14,569,452	14,628,829	10.70%	9.67%	11.72%	10.13%	Within CL
Georgia	689,100	685,872	7,199,254	7,128,347	9.57%	8.53%	10.61%	9.62%	Within CL
Hawaii	84,866	87,192	1,002,100	1,043,274	8.47%	7.60%	9.33%	8.36%	Within CL
Idaho	89,273	101,335	1,123,541	1,127,907	7.95%	7.15%	8.74%	8.98%	Above UCL
Illinois	782,785	795,181	9,690,472	9,652,155	8.08%	7.25%	8.90%	8.24%	Within CL
Indiana	448,637	471,147	4,789,941	4,845,940	9.37%	8.68%	10.06%	9.72%	Within CL
Iowa	173,067	184,619	2,277,133	2,303,256	7.60%	6.89%	8.31%	8.02%	Within CL
Kansas	178,246	206,660	2,087,770	2,109,742	8.54%	8.08%	8.99%	9.80%	Above UCL
Kentucky	376,381	356,396	3,261,475	3,291,899	11.54%	10.55%	12.53%	10.83%	Within CL
Louisiana	367,682	367,763	3,331,568	3,373,172	11.04%	10.24%	11.83%	10.90%	Within CL
Maine	86,784	89,154	1,039,308	1,049,724	8.35%	7.70%	9.00%	8.49%	Within CL
Maryland	394,569	398,773	4,248,679	4,368,450	9.29%	8.46%	10.11%	9.13%	Within CL
Massachusetts	393,882	385,364	4,978,771	5,085,358	7.91%	7.32%	8.50%	7.58%	Within CL
Michigan	711,038	718,915	7,612,963	7,514,998	9.34%	8.67%	10.01%	9.57%	Within CL
Minnesota	252,937	258,842	3,972,835	3,994,103	6.37%	5.66%	7.07%	6.48%	Within CL

State	Diabetes estimate, direct	Diabetes estimate, summing small area estimates	Population, sum of BRFSS weights	Population, Census 2009	Diabetes prevalence, direct	95% LCL, direct	95% UCL, direct	Diabetes prevalence, small area estimate	Comparison of SAE estimate to interval
Mississippi	249,716	267,300	2,156,680	2,198,543	11.58%	10.87%	12.28%	12.16%	Within CL
Missouri	357,585	402,208	4,495,628	4,528,528	7.95%	7.07%	8.84%	8.88%	Above UCL
Montana	50,781	71,121	750,086	759,038	6.77%	6.14%	7.40%	9.37%	Above UCL
Nebraska	99,800	141,087	1,333,960	1,354,725	7.48%	6.89%	8.07%	10.41%	Above UCL
Nevada	153,206	176,658	1,955,464	2,016,556	7.83%	6.59%	9.08%	8.76%	Within CL
New Hampshire	72,662	76,695	1,024,486	1,023,564	7.09%	6.33%	7.85%	7.49%	Within CL
New Jersey	573,665	587,323	6,583,475	6,680,230	8.71%	7.99%	9.44%	8.79%	Within CL
New Mexico	127,795	138,301	1,487,436	1,519,500	8.59%	7.85%	9.34%	9.10%	Within CL
New York	1,309,923	1,296,952	14,773,788	14,938,438	8.87%	8.05%	9.69%	8.68%	Within CL
North Carolina	672,074	703,263	6,979,965	7,170,869	9.63%	8.85%	10.41%	9.81%	Within CL
North Dakota	37,058	51,116	494,204	515,389	7.50%	6.71%	8.29%	9.92%	Above UCL
Ohio	876,043	905,592	8,688,138	8,770,904	10.08%	9.33%	10.83%	10.32%	Within CL
Oklahoma	302,549	298,319	2,746,333	2,792,174	11.02%	10.22%	11.82%	10.68%	Within CL
Oregon	241,767	212,278	2,920,170	2,937,494	8.28%	7.24%	9.32%	7.23%	Below LCL
Pennsylvania	869,938	920,210	9,615,276	9,844,790	9.05%	8.32%	9.77%	9.35%	Within CL
Rhode Island	56,827	64,060	817,195	826,092	6.95%	6.29%	7.62%	7.75%	Above UCL
South Carolina	353,190	360,633	3,402,873	3,502,997	10.38%	9.50%	11.26%	10.29%	Within CL
South Dakota	44,043	54,076	604,762	605,260	7.28%	6.60%	7.96%	8.93%	Above UCL
Tennessee	489,618	513,503	4,759,872	4,804,768	10.29%	9.22%	11.35%	10.69%	Within CL
Texas	1,649,711	1,734,539	17,697,952	18,005,487	9.32%	8.47%	10.17%	9.63%	Within CL
Utah	115,755	125,314	1,900,021	1,865,790	6.09%	5.58%	6.61%	6.72%	Above UCL
Vermont	30,356	33,178	489,648	492,927	6.20%	5.62%	6.78%	6.73%	Within CL
Virginia	481,315	512,586	5,919,125	6,075,298	8.13%	7.25%	9.01%	8.44%	Within CL
Washington	386,515	394,268	5,011,026	5,086,749	7.71%	7.27%	8.15%	7.75%	Within CL
West Virginia	176,174	185,301	1,426,006	1,456,914	12.35%	11.33%	13.38%	12.72%	Within CL

State	Diabetes estimate, direct	Diabetes estimate, summing small area estimates	Population, sum of BRFSS weights	Population, Census 2009	Diabetes prevalence, direct	95% LCL, direct	95% UCL, direct	Diabetes prevalence, small area estimate	Comparison of SAE estimate to interval
Wisconsin	355,064	341,470	4,303,481	4,323,228	8.25%	7.23%	9.27%	7.90%	Within CL
Wyoming	28,161	38,438	402,683	424,219	6.99%	6.31%	7.68%	9.06%	Above UCL

Cross Validation

Counties were randomly assigned to groups A, B, and C. We then cross-validated the model by fitting the data only in groups A and B, only in groups A and C, and only in groups B and C. For each of these 3 model runs, we compared the estimated value from the model to the direct value in the data and calculated a percent error. Thus, every county has a percentage error derived from the model that excluded the group to which the county belongs.

The three tables below provide the mean error, squared error, and percent error using the direct as the denominator for each group – A, B, and C.

Group A

Variable	Label	Mean
error	Error	6667.65
sqderror	Squared error	676479378
pcterror	Percent error	0.9417582

Group B

Variable	Label	Mean
error	Error	6224.04
sqderror	Squared error	391472939
pcterror	Percent error	0.8797789

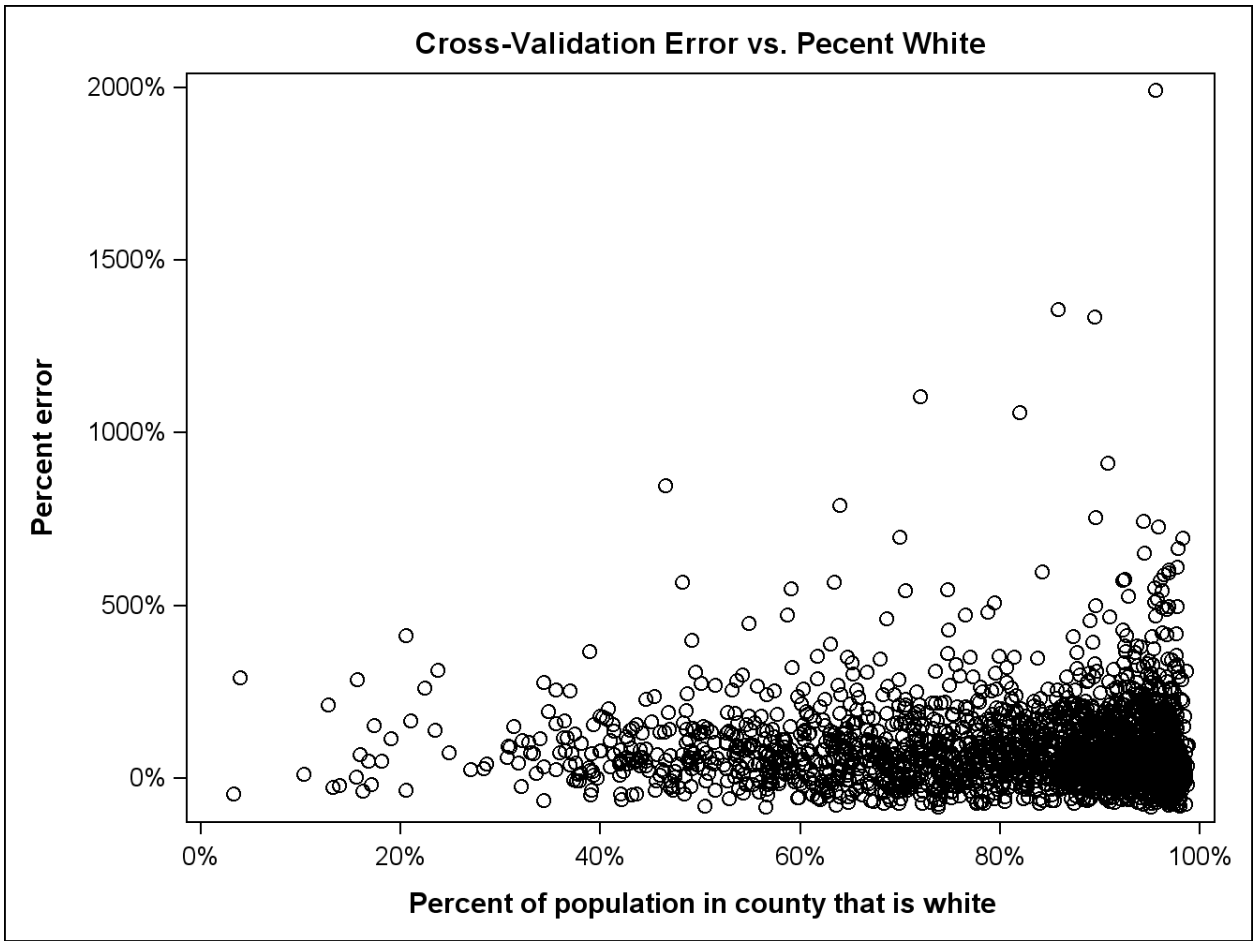
Group C

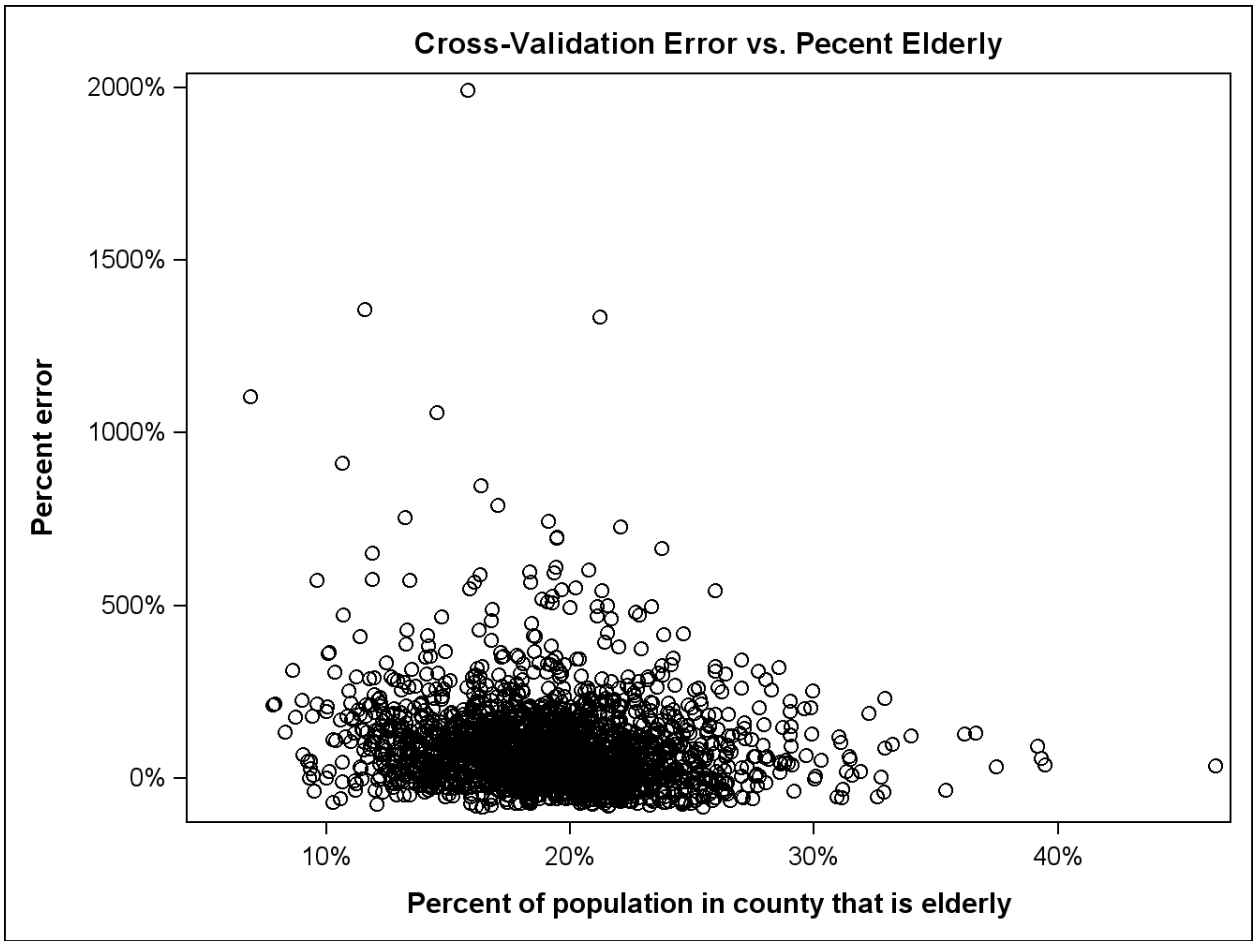
Variable	Label	Mean
error	Error	5437.72
sqderror	Squared error	281457552
pcterror	Percent error	0.8250096

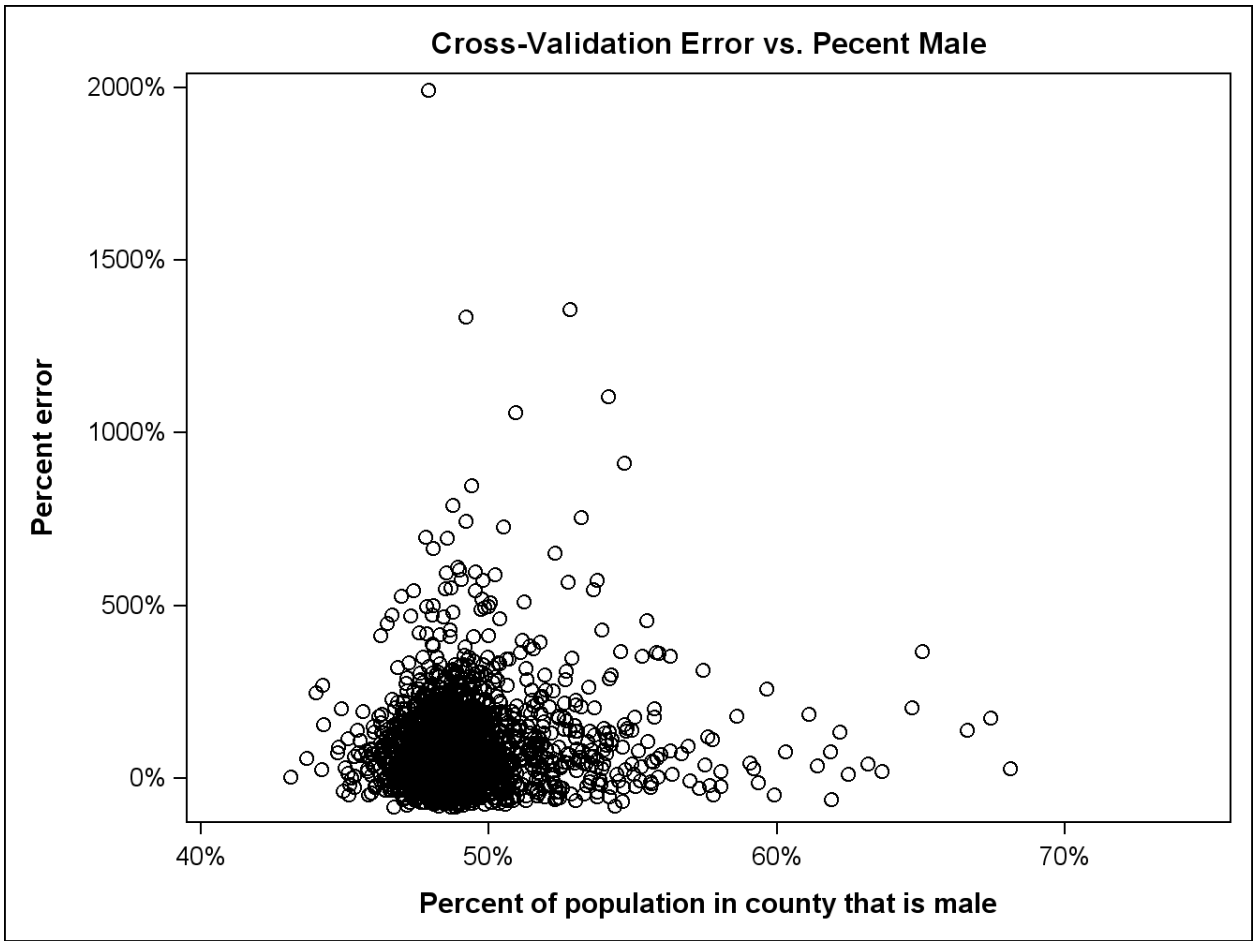
We looked for patterns in the percent error to county demographics – percent of population that is white, percent that is elderly (age 65 and over), and percent that is male. We also looked at the percent error relative to the sample size for the county, as a percent of the population. Finally, we compared the percent error to the direct estimate. The only patterns to emerge are that

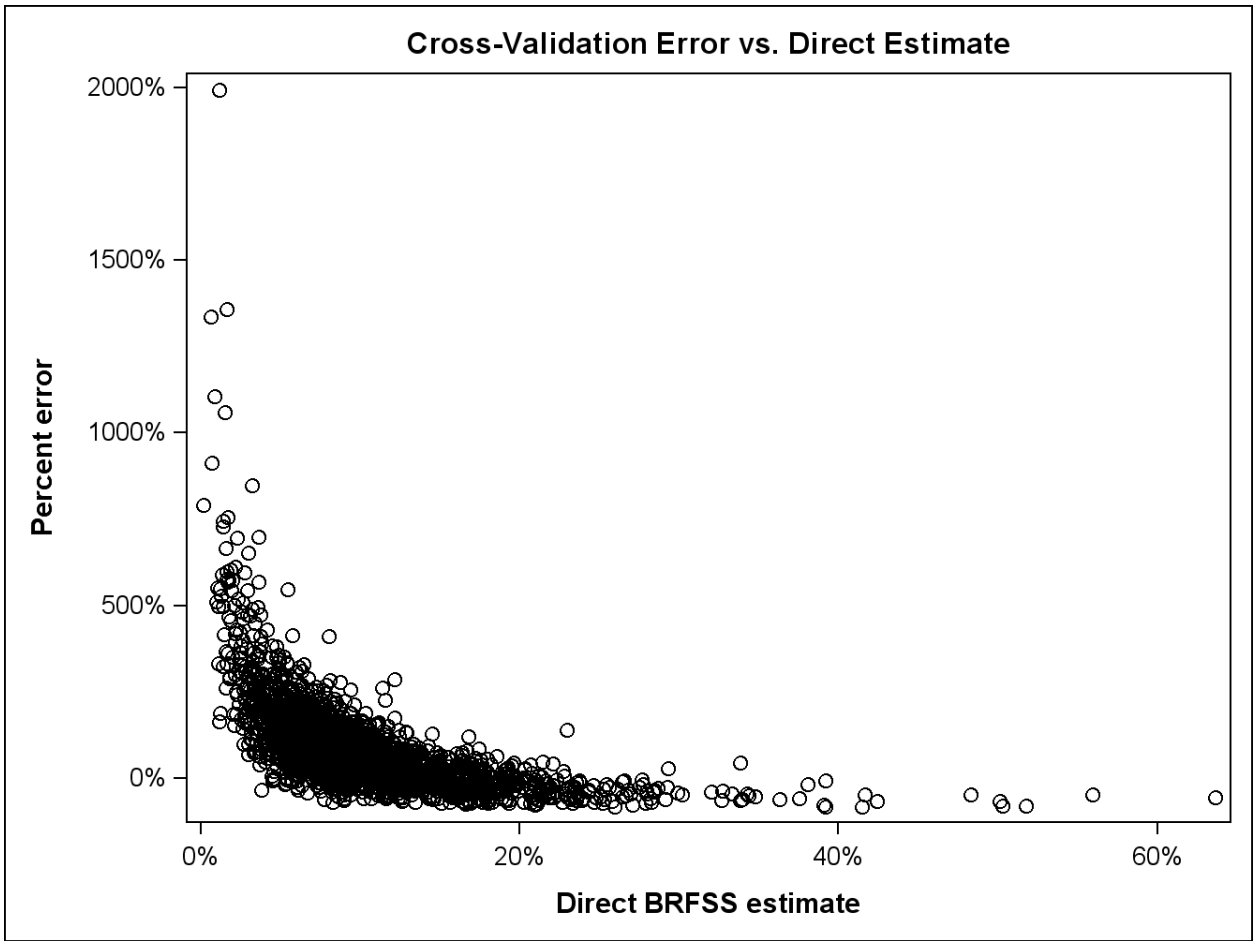
- The percent error becomes more variable as the percentage of whites increases. For some counties, the percent error increases as the percentage of whites increases, though there are still many counties with a high percentage of whites that have small errors.
- The percent error becomes less variable as the sample size (as a percentage of county population) increases.
- Errors decrease as the direct BRFSS estimate increases.

These figures are shown on the following pages.


Cross-Validation Error vs. Pecent White
Percent error
2000%
1500%
1000%
500%
0%
0%
20%
40%
60%
80%
100%
Percent of population in county that is white


Cross-Validation Error vs. Pecent Elderly
Percent error
2000%
1500%
1000%
500%
0%
10%
20%
30%
40%
Percent of population in county that is elderly


Cross-Validation Error vs. Pecent Male
Percent error
2000%
1500%
1000%
500%
0%
40%
50%
60%
70%
Percent of population in county that is male



Conclusion

Small area models for diabetes prevalence at the level of county by ASRE were built using Bayesian hierarchical models based on BRFSS data. The validations of the small area models for diabetes at the level of county by ASRE are proof of concept that small area models can provide a more accurate model than the current de facto model used in the PQIs which assumes that adjusting at the level of the general population is equivalent to adjusting at the level of the diabetes specific population.

References

Barker, LE, Thompson, TJ, Kirtland, KA, Boyle, JP, Geiss, LS, McCauley. MM, Albright, AL, 2013, "Bayesian Small Area Estimates of Diabetes Incidence by United States County, 2009", Journal of Data Science 11(2013), 249-267

Gilks, WR, Roberts, GO, 1996, "Strategies for Improving MCMC", in Markov Chain Monte Carlo in Practise, Gilks et. al. eds, (1996) Chapman and Hall/CRC

Rao, JNK, 2003, "Small Area Estimation", John Wiley and Sons