Estimating Health Conditions for Small Areas: Asthma Symptom Prevalence for State Legislative Districts

Carolyn A. Mendez-Luck, Hongjian Yu, Ying-Ying Meng, Mona Jhawar, and Steven P. Wallace

---

Research Objective. To create prevalence estimates of asthma symptoms for California legislative districts.

Data Sources. Three main data sources were used for this study: 2001 California Health Interview Survey, 2000 Census, and 2000–2002 March Current Population Surveys.

Study Design. Secondary data analyses were conducted from cross-sectional data to distribute the joint probability of ever having an asthma diagnosis and symptoms in the last 12 months within an Assembly district. We applied hierarchical logistic regressions to estimate the parameters for selected survey and census data that predicted the probabilities of diagnosed asthmatics with asthma symptoms. Predictors included individual-level variables and contextual variables at zip code levels.

Principal Findings. Asthma symptom prevalence geographically varied by age within and across Assembly districts throughout California.

Conclusions. With modest investments in establishing analytic data files and estimating regression parameters for target conditions, small area estimation (SAE) procedures can create health data estimates not otherwise available at the sub-county level. Applying SAE procedures to asthma symptom prevalence suggest that these data can become essential reference tools for advocates and policy makers currently addressing this and other public health concerns in the state.

Key Words. Small-area estimation, asthma, health policy

Asthma is a national and state health concern that has generated much research and policy activity. In California, an estimated 3.9 million children and adults, 11.9 percent of the state's population, reported that they had been diagnosed with asthma at some point in their lives (Meng et al. 2003). About 3 million children and adults who had ever been diagnosed with asthma also experienced asthma symptoms, such as wheezing, breathlessness, chest

tightness, and nighttime or early morning coughing (Sears 1997), at least once during the previous 12 months (Mendez-Luck, Yu, Meng, Chia, Jhawar et al. 2004).

The 2003–2004 California legislative session considered 28 bills focusing explicitly on asthma risk factors, seven on asthma treatment, and five that used asthma as a rationale for other action. Thirteen bills passed and were signed into law. The legislative activity related to asthma leads many elected representatives to want to understand the salience of the problem and consider the data for their jurisdictions. California also has a number of state programs and agencies that are responsible for tracking asthma and related risk factors; they also require local area data for their planning and policy work. The California Department of Health Services published a strategic plan for asthma that outlined research, education, treatment, prevention, and policy goals for the coming years. The California Asthma Public Health Initiative (see <http://www.dhs.ca.gov/cdic/caphi/>) funded 20 local community agencies throughout the state for programs to reduce asthma morbidity for school-aged children.

It is reasonable that asthma symptom rates would vary across locales, given variations in access to care, quality of care and the triggers of asthma exacerbation in different regions (National Asthma Education and Prevention Program 2003). However, local-level information on this ambulatory-sensitive chronic disease is virtually nonexistent. Most national surveys do not typically have identifiers for smaller geographic areas such as counties. Even the California Health Interview Survey (CHIS), the largest state health survey in the country—designed to make county-level estimates of health status and health services utilization—is limited in its power to make subcounty estimates.

California policy makers and advocates continually request data on asthma and other health issues for policy activities. However, California's Assembly, Senate, or Congressional districts do not follow any other administrative boundaries, making it difficult to assess the health of legislators' con-

---

Address correspondence to Carolyn A. Mendez-Luck, Ph.D., M.P.H., UCLA School of Public Health, Department of Community Health Sciences, UCLA Center for Health Improvement in Minority Elders/Resource Centers for Minority Aging Research, Box 951772, 21-275B CHS, Los Angeles, CA 90095-1772. Hongjian Yu, Ph.D., and Ying-Ying Meng, Dr. P.H., are with the UCLA Center for Health Policy Research, Los Angeles, CA. Mona Jhawar, M.P.H., is with the Los Angeles County Department of Public Health, Los Angeles, CA. Steven P. Wallace, Ph.D., is with the UCLA School of Public Health, Department of Community Health Sciences, UCLA Center for Health Improvement in Minority Elders/Resource Centers for Minority Aging, UCLA Center for Health Policy Research, Los Angeles, CA.

stituencies. District-specific information on the burden of disease—in this case, asthma—is especially important for establishing the need for and impact of policies. In response to ongoing requests for such data, we conducted a study to estimate the prevalence of various ambulatory care-sensitive chronic diseases, including asthma symptoms, for California’s Assembly, Senate, and Congressional districts. The purpose of the study was to develop a small-area estimation (SAE) methodology to create such estimates and to produce a series of policy briefs for a broad audience of legislators and advocates. The policy briefs can be found at <http://www.healthpolicy.ucla.edu/pubs/pubList.asp?topicID=12&subTopicID=42>.

The primary purpose of this article is to describe the SAE method we used to determine asthma symptom prevalence among California residents. We defined “asthma symptom prevalence” as ever having had a prior diagnosis of asthma and having asthma symptoms in the previous 12 months. The secondary purpose is to illustrate the utility of SAE methods for examining health issues at local levels, using asthma symptom prevalence at the Assembly district level as an example.

Rao (2003) and Pfeiffermann (2002) provide comprehensive reviews of the methodologies and applications of SAE in health-related research. Few studies have applied SAE to asthma symptom prevalence (Gottlieb et al. 1995; Salmond et al. 1999; Cesaroni et al. 2003). We used SAE methods to identify asthma patterns throughout the state for the purposes of estimating asthma symptom prevalence among the state’s residents. We specifically sought to provide policy audiences with data that aided their efforts to identify the scope of asthma prevalence in their areas. We improved upon a methodology previously developed to study health insurance coverage in California (Yu et al. 2007). The method is similar to the one independently developed by Twigg et al. (2000), with key differences that are elaborated elsewhere (Yu et al. 2007).

DATA AND METHODS

Data Sources

Three main data sources were used for this study: the CHIS 2001, the 2000 Census, and the 2000–2002 Current Population Surveys (CPS; see Figure 1). Other data sources used in this study included legislative boundary files from the California State Senate Office of Demographics, which reflected redistricting based on the 2000 Census.

Figure 1: Overview of Study's Small-Area Methodology.

The outcome variable, asthma symptom prevalence, and other individual-level data for respondents, came from the CHIS 2001. CHIS is a statewide random-digit dial (RDD) telephone survey of adults, adolescents, and children. Sampling was stratified at the county and county-group level to provide

stable direct estimates for most counties. Respondents were asked their zip codes to identify their subcounty locations. CHIS 2001 interviewed more than 55,000 households drawn from every county in California, providing a representative sample of the state's noninstitutionalized population. CHIS 2001 was weighted to the 2000 Census at the county or county-group level by race, Hispanic ethnicity, sex, and age. A detailed methodology of CHIS can be found at <http://www.chis.ucla.edu/methods.html>.

We used the 2000 Census for population data at the block, block group, and zip code levels. SF1 data files provided population counts for each census block in California by age, sex, and race/ethnicity. SF3 and SF4 files provided data at the block group and zip code levels on contextual characteristics.

We used the March CPS to derive a set of poverty ratios that could be used to impute a wider range of poverty ratios than was available in the Census SF data. The CPS, conducted by the U.S. Census Bureau, interviews approximately 60,000 households nationwide annually and obtains information on more than 150,000 individuals, including more than 14,000 in California. The March supplemental survey uses a comprehensive questionnaire that includes information on income and provides detailed data on poverty ratios in the public use files.

Methods

The estimates produced in this study were considered "synthetic" because we started with survey data designed for estimating rates at larger geographic areas—state and counties—and used the pattern of associations for these larger areas to derive estimates for smaller geographical areas—Assembly districts (Gonzalez 1973). Conventional SAE uses a blend of synthetic and direct estimates in statistical modeling to achieve an optimal trade-off between bias, minimized by direct estimates, and variance, minimized by synthetic estimates (Rao 1999). However, we were not able to include direct estimates in our SAE methods because the CHIS 2001 sampling design did not ask adequate geographic information to assign respondents to unique Assembly districts.

Figure 1 presents an overview of the major steps of our SAE methodology. The steps outlined in Figure 1 were built on a SAE approach we previously developed for studying health insurance coverage. Our original methodology involved three major steps—modeling, merging, and calibrating—which are detailed elsewhere (Yu et al. 2007). In the current study, we improved our methodology to replace the manual calibration step and include a new step—principal components analyses—to add contextual data. The

modeling and merging steps remained the same. In this article we elaborate on our revised methodology.

Creating a Consolidated Dataset. Our small-area methodology started with creating a consolidated dataset, using data from the CPS and 2000 Census SF1 file (steps 1–3, Figure 1). Although the 2000 Census contained poverty data, the SF1 did not have detailed income-to-poverty ratios (IPRs) at the block level by age by sex by race/ethnicity. Therefore, we statistically modeled the distribution of IPRs (below 100, 100–199, 200+ percent) using 3 years of combined CPS (2000–2002), concatenated to increase sample size (step 1, Figure 1). We then applied the predicted probabilities of the IPRs to the 2000 Census SF1 by age by race/ethnicity by sex, and aggregated the data to block groups (steps 2 and 3, Figure 1).

Principal Components Analyses. Our next step (step 4, Figure 1) was to conduct principal component analyses (Afifi and Clark 1996) of sociodemographic Census data at the ZCTA level (the Census' approximation of zip codes). The purposes of the principal components analyses were to (1) identify unique factors that would increase small area variability in our models to partially compensate for the lack of direct estimates and (2) to eliminate collinearity problems among many of these variables if used individually in the models. We chose data from the 2000 Census SF3 and SF4 files that approximated neighborhood characteristics, such as racial/ethnic composition, income, immigration status, language proficiency, educational attainment, and housing occupancy. Table 1 shows the original number and names of zip code data and the number of composite indices that resulted from the principal components analyses. We then merged the principal components into the individual-level CHIS survey data to create a multilevel dataset of person-level data and zip code-level data (step 5, Figure 1).

Modeling Asthma Symptom Prevalence. The next major step was to model asthma symptom prevalence. We conducted hierarchical logistic regressions to predict the probability of having asthma symptoms using the newly created multilevel CHIS 2001 dataset (step 6, Figure 1). Our outcome variable was asthma symptom prevalence. We defined asthma symptom prevalence as the joint probability of having asthma symptoms in the previous 12 months and having a prior diagnosis of asthma. We dichotomized the outcome variable to 1 = one or more symptoms and 0 = no symptoms. We ran separate models

Table 1: Composite Indices from Principal Components Analyses of Zip Code Data from 2000 Census SF3 and SF4 File

<i>Original Variables</i>	<i>Number of Principal Components</i>
Income (total = 9)	7
Percent of <100% FPL	
Percent of 100–199% FPL	
Percent of 200+% FPL	
Median HH income in 1999	
Average HH income in 1999	
Percent of HH ≥ \$75,000	
Percent of HH ≥ \$100,000	
Percent of not employed for ages 16–64	
Percent of public assistance and no SSI	
Language (total = 15)	8
Percent of English limited for all children ages 5–17	
Percent of English limited for Spanish children ages 5–17	
Percent of English limited for Asian children ages 5–17	
Percent of English limited for all adults ages 18–64	
Percent of English limited for Spanish adults ages 18–64	
Percent of English limited for Asian adults ages 18–64	
Percent of English limited for all older adults ages 65+	
Percent of English limited for Spanish older adults ages 65+	
Percent of English limited for Asian older adults ages 65+	
Percent of English limited for all persons ages 5+	
Percent of English limited for all Spanish ages 5+	
Percent of English limited for all Asian ages 5+	
Percent of foreign language linguistically isolated	
Percent of Spanish linguistically isolated	
percent of Asian linguistically isolated	
Citizenship (total = 5)	4
Percent of foreign-born	
Percent of naturalized citizens	
Percent of noncitizens	
Immigrant concentration	
Percent of immigrants who entered after 1990	
Population (total = 10)	7
Percent of ages 0–14	
Percent of ages 15–64	
Percent of ages 65+	
Percent of female	
Percent of population in urban areas	
Percent of housing units in urban areas	
Population density (# of persons per square mile)	
Percent of single moms with own children under 18 years old	
Population from Census 2000 SF1	
Population from Census 2000 SF3	

for children (ages 0–17), nonelderly adults (ages 18–64), and elderly adults (age 65 and older). We included in our models predictors at both the individual and contextual levels. Predictors at the individual level included age (0–4, 5–17, 18–29, 30–44, 45–64, 65+), sex (male, female), race/ethnicity (Latino, non-Latino white, African American, Asian/other) and poverty levels (below 100 percent, 100–199 percent, 200+ percent). Other individual-level predictors included two-way interactions of age, sex, race/ethnicity, and poverty levels. Contextual variables included the principal components described above. We also included 48 dummy geographic variables of California counties.

We ordered the models as follows: first, the individual-level interaction terms were selected while forcing the four main effects (the individual-level predictors) to stay in the model, using backward selection. The criterion for backward selection was p -value $< .05$. Second, the zip code-level predictors were selected while forcing the main effects and their two-way interactions (selected in the previous step) to stay in the model. Finally, the 48 geographic variables were added as fixed effects.

Table 2 shows the specific individual and contextual level predictors of the final models for each age group. For each age group, we were able to achieve a parsimonious model with specific neighborhood information. The final models were fitted using *SUDAAN* to take into account the design effect that resulted from the CHIS sampling scheme, incorporating that into the calculation of the variance–covariance matrix of regression coefficients.

Merging. The last steps in our methodology were to estimate the number of persons with asthma symptoms per Assembly district (steps 7 and 8, Figure 1). We merged the estimated model coefficients and their variance–covariance matrix to our consolidated dataset by age by sex by race/ethnicity for block groups (step 7, Figure 1). We then calculated the numbers of diagnosed asthmatics with asthma symptoms by block group and aggregated the block groups to the Assembly district level (step 8, Figure 1). We estimated asthma symptom prevalence by dividing these numbers by the appropriate age population in each Assembly district.

Finally, the estimates and their respective coefficients of variation (CVs) were calculated for all legislative districts to assess the stability of the estimates (step 9, Figure 1). The variances were calculated by applying the Delta method, as provided by Hosmer and Lemeshow, scenario two (26). Further details are available request from the first author. We designated 30 percent as

Table 2: Final Stratified Asthma Models

Final Asthma Models for Children, Adults, and Elderly

<i>Children, Ages 0–17</i>	<i>Adults, Ages 18–64</i>	<i>Adults, Ages 65+</i>	<i>Variable Name</i>	<i>Categories</i>	
Individual level					
X			Age	0–4	
X				05–11	
X				12–17	
	X			18–29	
	X			30–44	
	X			45–64	
		X		65+	
X	X	X	Sex	Male	
X	X	X		Female	
X	X	X	Race	Latino	
X	X	X		Non-Latino white	
X	X	X		African American	
X	X	X		Asian American and other	
X	X	X	Poverty	< 100%	
X	X	X		100–199%	
X	X	X		200%+	
X	X	X	California Regions	Northern California/ Sierras	
X	X	X		Greater San Francisco Bay area	
X	X	X		Sacramento	
X	X	X		San Joaquin Valley	
X	X	X		Central Coast	
X	X	X		Los Angeles	
X	X	X		Greater Southern California Region	
Contextual level					
Single variables					
	X	X	% College graduates		
X			% High school graduates		
	X		% Crowded occupancy (tenure occupancy per room)		
		X	% Owner-occupied housing units out of all housing units		
X	X	X	Strata	48 geographic categories	

continued

Table 2. *Continued*

<i>Final Asthma Models for Children, Adults, and Elderly</i>				
<i>Children, Ages 0–17</i>	<i>Adults, Ages 18–64</i>	<i>Adults, Ages 65+</i>	<i>Variable Name</i>	<i>Categories</i>
Principal components				
	X		Income	
	X		zinc1	
			zinc2	
			Language	
X			zlang1	
X			zlang3	
		X	zlang4	
			Citizenship	
		X	zcitshp1	
			Population	
	X		zpop1	
	X		zpop2	
Interaction terms				
X	X		race1 × pov1	Latino X <200% poverty
X	X		race1 × pov2	Latino X 200%+ poverty
X	X		race2 × pov1	NLW X <200% poverty
X	X		race2 × pov2	NLW X 200%+ poverty
X	X		race3 × pov1	African American X <200% poverty
X	X		race3 × pov2	African American X 200% poverty

Note. “X” represents the variables that were included in the final statistical models, by age.

the cutoff for acceptable CVs, consistent with the standard practice of the National Center for Health Statistics (27). For CVs larger than 30 percent, the entire process was repeated from the beginning with new models that had fewer predictors or new, usually larger, geographic groupings for “calibration” or both until all the estimates had CVs below 30 percent.

RESULTS

Table 3 shows the highest and lowest estimates of asthma symptom prevalence by age group in California’s 80 Assembly districts. The estimates show that

Table 3: Highest and Lowest Asthma Symptom Prevalence Rates in California Assembly Districts by Age

Percent of California Residents Reporting Asthma Symptoms at Least Once in the Last 12 Months									
Children Ages 0–17*			Nonelderly Adults Ages 18–64			Elderly Adults Age 65 and Older			
District	County Location	Rate (%)	District	County Location	Rate (%)	District	County Location	Rate (%)	
29	Fresno, Madera, Tulare	16	29	Fresno, Madera, Tulare	13	18	Alameda	12	
08	Solano, Yolo	15	36	Los Angeles, San Bernardino	13	16	Alameda	11	
31	Fresno, Tulare	14	08	Solano, Yolo	13	02	Butte, Colusa, Glenn, Modoc, Shasta, Siskiyou, Sutter, Tehama, Yolo	11	
63	Riverside, San Diego	14	04	Alpine, El Dorado, Placer, Sacramento	12	:	:	:	
06	Marin, Sonoma	14	05	Placer, Sacramento	12	:	:	:	
59	Los Angeles, San Bernardino	14	01	Del Norte, Humboldt, Lake, Mendocino, Sonoma, Trinity	12	:	:	:	
62	San Bernardino	13	07	Napa, Solano, Sonoma	12	:	:	:	
07	Napa, Solano, Sonoma	13	09	Sacramento	12	:	:	:	
:	:	:	11	Contra Costa	12	:	:	:	
:	:	:	:	:	:	:	:	:	
:	:	:	:	:	:	:	:	:	
40	Los Angeles	7	:	:	:	:	:	:	
49	Los Angeles	7	49	Los Angeles	6	44	Los Angeles	5	
47	Los Angeles	7	58	Los Angeles	6	49	Los Angeles	5	
45	Los Angeles	7	45	Los Angeles	6	71	Orange, Riverside	5	
28	Monterey, San Benito, Santa Clara, Santa Cruz	7	79	San Diego	6	69	Orange	5	
39	Los Angeles	7	28	Monterey, San Benito, Santa Clara, Santa Cruz	6	55	Los Angeles	5	
50	Los Angeles	7	19	San Mateo	6	06	Marin, Sonoma	5	

continued

Table 3: Continued

Percent of California Residents Reporting Asthma Symptoms at Least Once in the Last 12 Months								
Children Ages 0-17*			Nonelderly Adults Ages 18-64			Elderly Adults Age 65 and Older		
District	County Location	Rate (%)	District	County Location	Rate (%)	District	County Location	Rate (%)
69	Orange	6	39	Los Angeles	6	51	Los Angeles	5
48	Los Angeles	6	50	Los Angeles	5	53	Los Angeles	5
52	Los Angeles	5	69	Orange	5	64	Orange	4
46	Los Angeles	5	46	Los Angeles	5	54	Los Angeles	4

*CHHS 2001 did not ask about asthma diagnosis and asthma symptoms among children ages 0–1 because asthma is commonly not clinically diagnosed for that age group. However, this age group is included to be consistent with census data. Asthma prevalence among the overall 0–17 age group may actually be higher than reported here.

CHHS, California Health Interview Survey.

asthma symptom prevalence varied by age across Assembly districts. The proportions of all children, nonelderly adults, and elderly adults who reported having asthma symptoms in the previous year ranged from 5 to 16, 5 to 13, and 4 to 12 percent, respectively. The majority of Assembly districts had rates similar to the statewide averages for each age group, although some districts across the state fared better or worse than these averages.

The patterns of asthma symptom prevalence were slightly different for children and nonelderly adults. The Assembly districts with the highest asthma symptom prevalence rates for children were located primarily in the state's interior regions, including parts of Fresno, Madera, Tulare, Solano, Yolo, San Bernardino, Riverside, and San Diego counties. Assembly districts with the lowest rates of asthma symptom prevalence were located primarily in Southern California and along the central coast. The rates of asthma symptom prevalence for nonelderly adults were more broadly distributed across the state, compared with children. Assembly districts with the highest asthma symptom prevalence for nonelderly adults were primarily in the central, northern and interior regions of the state. Districts with the lowest asthma symptom prevalence were primarily located in the greater Southern California region and along the state's central coast. However, low rates still translated to large numbers because the nonelderly adult population represented a significant proportion of the state's total population. Therefore, even Assembly districts with asthma symptom prevalence of 5 or 6 percent had approximately 13,000–17,000 nonelderly adults who reported having asthma symptoms in the last 12 months (data not shown). The Assembly districts with the highest asthma symptom prevalence among California's older population were located in the northern and central valley regions of the state. Similar to our findings for nonelderly adults, the Assembly districts with the lowest asthma symptom prevalence for the 65 and older population were located in Southern California, specifically in Orange, Los Angeles, and Riverside counties.

Table 4 shows the Assembly districts with the highest asthma symptom prevalence rates by age group, compared to their corresponding county rates and the California statewide average. The Assembly district with the highest rate for both children ages 0–17 and adults ages 18–64 was Assembly District 29 (AD29), located in California's interior region of the Central Valley, one of the state's prime agricultural regions. The small-area estimate for asthma symptom prevalence for children in this district was 16 percent and 13 percent for nonelderly adults. However, no single county associated with this district had a direct estimate as high as 16 percent; the CHIS 2001 direct estimates

Table 4: Asthma Symptom Prevalence Rates for California, Selected Counties and Corresponding Assembly Districts, All Ages

		<i>California Residents Reporting Asthma Symptoms at Least Once in Last 12 Months</i>		
		<i>Children (%)*</i>	<i>Nonelderly Adults (%)</i>	<i>Elderly Adults (%)</i>
		<i>Ages 0–17</i>	<i>Ages 18–64</i>	<i>Age 65 and older</i>
California		10	9	8
County				
	Alameda	10	9	8
	Fresno	15	13	9
	Los Angeles	8	7	8
	Madera	8	10	14
	San Bernardino	13	9	8
	Solano	16	14	10
	Tulare	12	9	6
	Yolo	14	10	7
Assembly	Corresponding County			
District				
16	Alameda	11	11	11
18	Alameda	10	10	12
29	Fresno, Madera, Tulare	16	13	10
31	Fresno, Tulare	14	11	9
36	Los Angeles, San Bernardino	12	13	9
08	Solano, Yolo	15	13	10
02	Butte, Colusa, Glenn, Modoc, Shasta, Siskiyou, Sutter, Tehama, Yolo	10	11	11

*CHIS 2001 did not ask about asthma diagnosis and asthma symptoms among children ages 0–1 because asthma is commonly not clinically diagnosed for that age group. However, this age group is included to be consistent with census data. Asthma prevalence among the overall 0–17 age group may actually be higher than reported here.
CHIS, California Health Interview Survey.

for children in these counties ranged from 8 percent in Madera County to 15 percent in Fresno County. A similar result was found for Assembly District 36 (AD36) for adults ages 18–64. AD36 had the second highest asthma symptom prevalence rate for this age group, compared with the rest of the state. AD36 is comprised of parts of San Bernardino and Los Angeles counties and is geographically considered part of the Southern California region referred to as the Inland Empire. The CHIS 2001 direct estimates of asthma symptom prevalence in Los Angeles and San Bernardino counties were 7 and 9 percent, respectively, for this age group. However, the small-area

estimate of asthma symptom prevalence in AD36 for this age group was 13 percent.

DISCUSSION

Our methodology is a first attempt to use SAE methods to develop asthma symptom prevalence at a geographic area smaller than county. The ability to produce credible estimates for legislative districts is particularly important for front burner policy issues because legislators are particularly interested in information that helps them make or justify decisions on those topics (Oliver and Singer 2006). The strength of this new methodology is its applicability for a wide variety of health conditions for geographic areas where existing data cannot support direct estimates. By applying predictive models to population data and integrating person-level characteristics and zip code-level contextual data, this method opens up the possibility of generating estimates for many different irregular geographic units, such as legislative districts, cities, and school districts.

We developed a new series of steps from our original SAE methodology (Yu et al. 2007). These new steps represented improvements to our original methodology in two key ways. First, we incorporated contextual data into our models. This approach allowed us to add variables that approximated community characteristics to better geographically distribute the observations of our outcome variable based on the CHIS 2001 data. This level of data helped preserve the specificity of local areas. Principal component analyses captured most of the information in those zip code-level variables as well as addressing the problem of multicollinearity. The contextual variables added predictive power to our models and helped the small area estimates from shrinking toward our global mean, the statewide average.

The second key improvement was including 48 unique geographic variables (mostly counties) in our models to retain regional variations within the state. Utilizing a fixed-effects model at the county level constrained the estimates to sum to the direct county and statewide prevalence rates. Including geographic variables ensured that the modeled estimates were consistent with a “gold standard,” the direct CHIS estimates. Thus, the model estimates matched the direct estimates at the county level, eliminating the calibration step from our original methodology. The calibration step we had previously conducted was a manual process to ensure that the aggregation of small-area estimates equaled the direct estimates for larger areas (Yu et al. 2007). This methodological change greatly streamlined the analytic process.

Calculating the variances was also a critical step because it provided a measure of the stability of the estimates. Because policy makers were primarily interested only in the point estimates, we insured that no point estimate had a CV over 30 percent, in addition to providing the 95 percent confidence intervals along with the point estimates.

Our synthetic estimate-only approach differs from conventional SAE methods that use a combination of direct and synthetic estimates. Both approaches provide more stable estimates of an outcome variable across much smaller subpopulation groups than would be possible from only using direct estimates from survey data. However, in conventional SAE shrinkage usually occurs and is justifiable to balance the trade-off between variance and bias. Direct estimates were not an option for us because of the limited geocoding of the CHIS 2001 data. Most public use data files, such as CHIS 2001, do not provide sufficient geocoding for making direct estimates of unusual geographies such as legislative districts. By using only synthetic estimates, our SAE method has a basic inability to determine whether the differences we found between Assembly districts reflected true geographic clustering or bias from shrinkage to the county-level means. In other words, some of the observed differences in prevalence across the state could have been artifacts of our estimation method due to the less than optimal levels of trade-off between variance and bias customary in SAE. Nevertheless, the contribution of our SAE approach is in providing more stable estimates of the distribution of the asthma symptom prevalence in smaller and irregular geographic areas when sample size and geocoding constraints make direct estimates impossible.

We must note the limitations of our study. First, the estimates we created were based on the patterns derived from predictors of asthma across larger geographic areas (county and statewide). As a result, the estimates do not reflect the efforts of particular local-level interventions and cannot be used for evaluating the effectiveness of such programs. For example, our estimates did not account for changes in asthma symptom prevalence due to specific school-based disease management programs or community-based asthma awareness activities.

Second, our SAE methodology relies on the availability and timeliness of census data for population counts and contextual data. Decennial census data become more outdated as we move farther away from its most current release date. However, we offer an approach to address this issue (Yu et al. 2007). Updated census tract data from the Census Bureau's American Community Survey should also improve the precision of intercensal estimates (Holden 2002).

Third, our methodology relies on the availability of population-based survey data. We utilized data from CHIS 2001. Although comprehensive population-based health surveys are being conducted in only a handful of states, the CDC sponsors a number of population-based data collection activities that offers states the opportunity to examine emerging public health concerns in their jurisdictions. The largest activity is the Behavioral Risk Factor Surveillance System (BRFSS, see <http://www.cdc.gov/brfss/index.htm>). BRFSS is a state-based system of health surveys that collects information on health risk behaviors, preventive health practices, and health care access primarily related to chronic disease and injury. Each year, all 50 states collect data using a standard core questionnaire and optional modules that cover additional health topics or more detailed questions on a health topic included in the core. The CDC also supports the Pregnancy Risk Assessment Monitoring System (PRAMS, see <http://www.cdc.gov/prams/>), a surveillance project that collects state-specific, population-based data on maternal attitudes and experiences before, during, and after pregnancy. Twenty-nine states are currently participating in PRAMS. Another option is the State and Local Area Integrated Telephone Survey (SLAITS, see <http://www.cdc.gov/nchs/slaits.htm>), developed by National Center for Health Statistics. SLAITS collects health care data at state and local levels to supplement national data collection strategies. SLAITS uses the same RDD telephone design approach and sampling frame as the ongoing National Immunization Survey and uses standardized questions to produce comparative data across states and for the nation. Researchers, policy advocates, and state administrators can adapt our SAE methodology to examine a wide range of health issues at various geographic levels, with modest investments in establishing analytic data files from these population-based data sources and estimating the regression parameters for their variables of interest.

Lastly, we found that the burden of asthma was not equally distributed across California. However, our results did not explain the causes of variations among Assembly districts. Variations in rates may have been due to one or more of the following factors: (1) population characteristics in local areas; (2) the levels of local environmental triggers, such as air pollution and seasonal weather; (3) access to asthma-related health services; and (4) characteristics of our outcome variable. First, the demographic composition of the population can affect asthma rates, such as having a large number of Latino children who are less likely to report having asthma or asthma symptoms (Holtby et al. 2006). Second, asthma is exacerbated by air pollutants, such as ozone, small matter particulates, and nitrogen dioxide, especially in children (Byrd and Joad 2006), and intrastate variations in these pollutants are well documented

by the California Air Resources Board's air quality monitoring program (State of California Air Resources Board October 2003). Another contributory factor may be the range in environmental climates within California; many counties have multiple climactic areas where the geography in a single county can include coastal, high mountain, and inland valley systems. Third, because asthma symptoms can be controlled through proper management (National Asthma Education and Prevention Program 2003), variations across California in access to health care, such as insurance coverage and having usual sources of care, may have contributed to our findings (Mendez et al. 2003; Jhavar et al. 2004; Mendez-Luck et al. 2004). Fourth, our findings may be driven in part by our outcome variable, which measures the joint probability of both having a prior asthma diagnosis and having symptoms in the last 12 months given a prior diagnosis. Therefore, our observed variations in asthma symptom prevalence may reflect underlying differences in asthma diagnosis rates, variations in asthma severity, geographic variations in persistent exposure to environmental triggers, frequency of symptoms due to inadequate medical control, or a combination of these factors. Our methodology does not separately identify these issues, leaving the policy debate of disease course versus disease management for other analyses.

Asthma costs California approximately \$1.26 billion yearly, \$721 million in direct medical expenditures, and \$545 million in indirect costs (Asthma & Allergy Foundation of America 1998). Asthma is on the front burner of policy activity in California, especially in many of the local communities highlighted in our study, such as in the Central Valley and in the Inland Empire. Advocacy groups such as the American Lung Association of California and the Community Action to Fight Asthma (CAFA), a statewide network of asthma coalitions, work with local communities to build support for asthma awareness, treatment, and management services for California's children.

The California legislature is actively pursuing legislation to address indoor air in homes and in schools, outdoor air, and clinical policies. While some of these bills have been signed by Governor Arnold Schwarzenegger, others were either narrowly defeated in committee or vetoed by the governor (CAFA 2006).

Our study particularly informs policy efforts by representatives of districts with high rates of asthma symptoms. As a direct result of this study, our research team was asked to give testimony on the subject at a Fresno County Air Quality Senate Committee Hearing in 2004. Several bills have since been proposed to improve the Central Valley's air quality because the Central

Valley suffers from some of the worst air pollution (American Lung Association 2006) and asthma symptoms in the state (see Tables 3 and 4). The proposed legislation included incentives for owners of stationary sources of air pollution to invest in air pollution control equipment, monetary fines for violations in excessive emissions of particulate matter, and removal of gross-polluting vehicles from the road.

Faced with decision making on bills such as those proposed for California's Central Valley, partisans on both sides of the asthma policy debate want data to contextualize their positions on proposed legislation. The SAE data we provide are essential references for state politicians and advocates to gauge their constituencies' needs against others around the state. Our policy brief—on which this article is based—has been downloaded from the UCLA Center for Health Policy Research website over 6,000 times since October 2004; our other SAE policy briefs on health insurance and chronic conditions have collectively been downloaded more than 15,000 times since May 2003. The high numbers of downloads for these policy briefs demonstrate the ongoing demand for local-level health data in California. The data are useful for decision making and planning efforts for both new and long-standing health problems that face California and our nation. The strength of this new methodology is its applicability to estimate prevalence data for a wide variety of health conditions for geographic areas where existing data cannot support direct estimates.

ACKNOWLEDGMENTS

Financial support for this research was provided by a grant from The California Endowment. Drs. Mendez-Luck and Wallace received support from the University of California, Los Angeles, Center for Health Improvement in Minority Elderly/Resource Centers for Minority Aging Research, NIH/NIA/NCMHD, under Grant P30-AG-021684. The authors gratefully thank Y. Jenny Chia and Alek Sripipatana for their contributions during the initial development of the research. The authors also thank Jose Calderon, Honghu Liu, and Jeff Luck for their feedback on earlier versions of this manuscript.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Affi, A. A., and V. Clark. 1996. *Computer-Aided Multivariate Analysis*. London: Chapman & Hall.

- American Lung Association. 2006. "State of the Air 2006" [accessed on November 3, 2006]. Available at http://lungaction.org/reports/SOTA06_stateozone.html
- Asthma & Allergy Foundation of America. 1998. "State Costs at a Glance - 1998" [accessed on November 3, 2006]. Available at <http://www.aafa.org/display.cfm?id=6&sub=63&cont=142>
- Byrd, R. S., and J. P. Joad. 2006. "Urban Asthma." *Current Opinion in Pulmonary Medicine* 12 (1): 68-74.
- Cesaroni, G., S. Farchi, M. Davoli, F. Forastiere, and C. A. Perucci. 2003. "Individual and Area-Based Indicators of Socioeconomic Status and Childhood Asthma." *European Respiratory Journal* 22 (4): 619-24.
- Community Action to Fight Asthma (CAFA). 2006. "California State Policies Related to Asthma—October 5, 2006" [accessed on November 3, 2006]. Available at http://www.calasthma.org/uploads/resources/legislative_updates1.doc
- Gonzalez, M. E. 1973. "Use and Evaluation of Synthetic Estimates." In *Proceedings of the Social Statistics Section, American Statistical Association*, pp. 33-6, Washington, DC.
- Gottlieb, D. J., A. S. Beiser, and G. T. O'Connor. 1995. "Poverty, Race, and Medication Use are Correlates of Asthma Hospitalization Rates: A Small Area Analysis in Boston." *Chest* 108: 28-35.
- Holden, C. 2002. "U.S. Demographics. New Annual Survey Brings Census into 21st Century." *Science* 295 (5563): 2202-3.
- Holtby, S., E. Zahnd, N. Lordi, C. McCain, Y. J. Chia, and J. H. Kurata. 2006. "Health of California's Adults, Adolescents and Children: Findings from CHIS 2003 and CHIS 2001." Report, UCLA Center for Health Policy Research.
- Jhawar, M., C. A. Mendez-Luck, H. Yu, Y. Y. Meng, J. Chia, A. Sripipatana, and S. P. Wallace. 2004. "Many Children Remain Uninsured and Not Eligible for Medi-Cal and Health Families." Policy Brief, UCLA Center for Health Policy Research (PB2004-6), pp 1-4.
- Mendez, C. A., H. Yu, Y. Y. Meng, J. Chia, and E. R. Brown. 2003. "California's New Assembly and Senate Districts: Geographic Disparities in Health Insurance Coverage." Policy Brief, UCLA Center for Health Policy Research.
- Mendez-Luck, C. A., H. Yu, Y. Y. Meng, J. Chia, M. Jhawar, A. Sripipatana, and S. P. Wallace. 2004. "Asthma Among California's Children, Adults and the Elderly: A Geographic Look by Legislative Districts." Policy Brief, UCLA Center for Health Policy Research (PB2004-5), pp. 1-8.
- Mendez-Luck, C. A., H. Yu, Y. Y. Meng, J. Chia, B. L. Newman, A. Sripipatana, and S. P. Wallace. 2004. "Many Uninsured Children Qualify for Medi-Cal or Health Families." Policy Brief, UCLA Center for Health Policy Research (PB2004-3), pp. 1-8.
- Meng, Y., S. Babey, E. Malcolm, E. R. Brown, and N. Chawla. 2003. "Asthma in California: Findings from the 2001 California Health Interview Survey." Policy Brief, UCLA Center for Health Policy Research.
- National Asthma Education and Prevention Program. 2003. *Expert Panel Report: Guidelines for the Diagnosis and Management of Asthma—Update on Selected Topics 2002*. Bethesda, MD: National Institutes of Health.

- Oliver, T. R., and R. F. Singer. 2006. "Health Services Research as a Source of Legislative Analysis and Input: The Role of the California Health Benefits Review Program." *Health Services Research* 41 (3, part 2): 1124–58.
- Pfeffermann, D. 2002. "Small Area Estimation—New Developments and Directions." *International Statistical Review* 70: 125–43.
- Rao, J. N. K. 1999. "Some Recent Advances in Model-Based Small Area Estimation." *Survey Methodology* 25 (2): 175–86.
- . 2003. *Small Area Estimation*. Indianapolis, IN: Wiley Publishing Inc.
- Salmond, C., P. Crampton, S. Hales, S. Lewis, and N. Pearce. 1999. "Asthma Prevalence and Deprivation: A Small Area Analysis." *Journal of Epidemiology and Community Health* 53: 476–80.
- Sears, M. R. 1997. "Descriptive Epidemiology of Asthma." *Lancet* 350 (suppl 2): SII1–4.
- State of California Air Resources Board. 2003. "Assessment of California's State-wide Air Monitoring Network for the Children's Environmental Health Protection Act (SB 25)" [accessed on July 14, 2006]. Available at <http://www.arb.ca.gov/ch/programs/sb25/adequacyreport.pdf>
- Twigg, L., G. Moon, and K. Jones. 2000. "Predicting Small-Area Health-Related Behaviour: A Comparison of Smoking and Drinking Indicators." *Social Science and Medicine* 50 (7–8): 1109–20.
- Yu, H., Y. Meng, C. A. Mendez-Luck, M. Jhavar, and S. P. Wallace. 2007. "Meeting Local Data Needs: Small Area Estimation of Health Insurance Coverage for California Legislative Districts." *American Journal of Public Health* 97 (3): 1–7.