

Small Area Estimates Reveal High Cigarette Smoking Prevalence in Low-Income Cities of Los Angeles County

Yan Cui, Susie B. Baldwin, Amy S. Lightstone, Margaret Shih, Hongjian Yu, and Steven Teutsch

ABSTRACT *Los Angeles County has among the lowest smoking rates of large urban counties in the USA. Nevertheless, concerning disparities persist as high smoking prevalence is found among certain subgroups. We calculated adult smoking prevalence in the incorporated cities of Los Angeles County in order to identify cities with high smoking prevalence. The prevalence was estimated by a model-based small area estimation method with utilization of three data sources, including the 2007 Los Angeles County Health Survey, the 2000 Census, and the 2007 Los Angeles County Population Estimates and Projection System. Smoking prevalence varied considerably across cities, with a more than fourfold difference between the lowest (5.3%) and the highest prevalence (21.7%). Higher smoking prevalence was generally found in socioeconomically disadvantaged cities. The disparities identified here add another layer of data to our knowledge of the health inequities experienced by low-income urban communities and provide much sought data for local tobacco control. Our study also demonstrates the feasibility of providing credible local estimates of smoking prevalence using the model-based small area estimation method.*

KEYWORDS *Cities, Small area estimation, Smoking prevalence, Los Angeles County*

INTRODUCTION

Over the past two decades, Los Angeles (LA) County has been a national leader in tobacco control and prevention. With a current smoking prevalence of 14%, the county has among the lowest smoking rates of large urban counties in the USA.¹ However, concerning disparities exist. African Americans and adults with lower socio-economic status smoke substantially more than other groups.² Also worrisome is the fact that the declining trend of smoking prevalence among county adults has stalled in recent years.

Strong anti-smoking policies have effectively reduced smoking prevalence.³ To date, the state of California has enacted policies including smoke-free workplaces

Cui, Baldwin, Lightstone, and Shih are with the Office of Health Assessment and Epidemiology, Los Angeles County Department of Public Health, Los Angeles, CA, USA; Yu is with the Center for Health Policy Research, University of California, Los Angeles, CA, USA; Teutsch is with the Los Angeles County Department of Public Health, Los Angeles, CA, USA.

Correspondence: Yan Cui, Los Angeles County Department of Public Health, Los Angeles, CA, USA. (E-mail: ycui@ph.lacounty.gov)

Electronic supplementary material The online version of this article (doi:10.1007/s11524-011-9615-0) contains supplementary material, which is available to authorized users.

(1995), smoke-free bars (1998), smoke-free playgrounds (2003), smoke-free doorways (2004), and smoke-free cars with children (2008). The experience of the last two decades indicates that efforts to further reduce tobacco use and exposure to secondhand smoke must include city- and county-level policies that reduce access to tobacco products, restrict smoking in public spaces, and create social norms that make smoking even less desirable. To this end, the LA County Department of Public Health's Tobacco Control and Prevention Program (TCPP) has shifted from a health education approach focusing on individual-level behavior change to a policy-based approach targeting community-level social norms. To further reduce smoking prevalence, the TCPP is facilitating the enactment of anti-smoking policies by local city governments.

LA County covers more than 4,000 mi² and consists of numerous distinct urban and suburban communities. Populations in these communities differ greatly in terms of racial-ethnic, socioeconomic, linguistic, and cultural backgrounds, and in their health status and behaviors. The locally based Los Angeles County Health Survey (LACHS) has shown that smoking prevalence varies substantially within the county, with high prevalence generally found in lower income areas.²

Baseline assessments of smoking status in target communities can form the basis for developing effective local tobacco control policies and programs. Since city governments have the jurisdiction to enact local anti-smoking ordinances, in this study, we aimed to calculate adult smoking prevalence within the 88 incorporated cities of LA County and to identify cities with high smoking prevalence.

METHODS

Since 1997, the LACHS has provided local agencies with invaluable smoking data by Service Planning Area (SPA) and Health District (HD),* the geographic units used by the county for health care planning and delivery.⁴ However, due to its limited sample size, the survey cannot provide reliable design-based estimates of smoking prevalence for most of the county's 88 incorporated cities. Here, we used a model-based small area estimation approach to calculate smoking prevalence at the city level.^{5,6} This approach starts with survey data designed for estimating statistics at larger geographic areas (in our case, LA County and its SPAs and HDs) and then uses associations established at the larger geographic level to derive estimates for smaller geographic areas (in our case, cities in LA County). Since the City of Los Angeles itself is large and diverse, we also estimated smoking prevalence for its 15 Council Districts.

Data Sources

We used three main data sources in the estimation: the adult component of the 2007 LACHS, the 2000 Census Summary File 3 (SF3), and the 2007 LA County Population Estimates and Projection System (PEPS).

The adult component of the 2007 LACHS was designed to represent the non-institutionalized adult population (18+ years) residing in LA County, CA, USA. A random sample of LA County households was selected using random digit dialing; one adult was randomly selected from each sampled household and queried using a

*LA County has 26 HDs and 8 SPAs. Each SPA consists of 1–5 HDs. HDs and SPAs have coincident boundaries.

computer-assisted telephone interviewing system. To account for linguistic diversity among the county residents, interviews were conducted in English, Spanish, Cantonese, Mandarin, Korean, and Vietnamese. A sample of 7,200 adults completed the survey. The overall response rate (AAPOR RR3) for the survey was 18%, calculated as the ratio of households interviewed to the entire number of eligible households. The cooperation rate (AAPOR COOP3) was 40%, reflecting the percent of households reached in which a respondent successfully completed the survey.⁷ Survey weights were developed to account for differences in the probability of selection of households and adults into the sample, to adjust for households without telephone service, and to align survey estimates to known geographic and demographic characteristics of the county population. Approval of the 2007 LACHS was obtained from the LA County Department of Public Health Institutional Review Board.

The 2007 LACHS provided data on smoking status and individual-level explanatory variables including age, sex, race/ethnicity and ratio of household income to federal poverty level. We assigned 19 participants with missing data for age to one of the age group categories using hot deck imputation method. Approximately 20% of survey participants did not provide sufficient income information to be assigned to a specific category of income-to-poverty ratio. A Markov chain Monte Carlo method was used to impute these missing values. All other variables had limited missing data, and no imputation was performed.

To retain small area variability and to prevent small area estimates from shrinking towards the global mean, we accounted for neighborhood characteristics; neighborhoods were defined using zip codes. Since data on neighborhood characteristics were not collected in the 2007 LACHS, a broad range of contextual variables regarding population makeup, citizenship, English language proficiency, income, educational attainment, and housing occupancy (see [Supplemental data](#)) were extracted from the 2000 Census SF3.

We also used mid-year population and poverty estimates at the census tract level from the 2007 PEPS, which were created using population counts from the 2000 Census as the base and projecting population changes due to fertility, mortality, and migration.⁸ Detailed population counts were available for subgroups jointly defined by age, sex, race/ethnicity and poverty status.

Principal Components Analysis

After extracting the contextual variables from the SF3 (see [Supplemental data](#)), we performed principal components analysis within the categories of population makeup, citizenship, English language proficiency, income, and educational attainment to eliminate multi-collinearity among variables and to reduce the number of variables to be included in regression models discussed below. Ten principal components were derived from the initial contextual variables. These principal components, together with the housing occupancy variables, were then merged with the LACHS and PEPS data using zip codes as the common identifiers.

Model Building at County Level

We used logistic regression models to assess the probability of being a current cigarette smoker based on individual and contextual characteristics. A current cigarette smoker was defined as having smoked at least 100 cigarettes in a lifetime and currently smoking. The individual-level variables included age, sex, race ethnicity, and household income. Contextual variables included the aforementioned

principal components and housing occupancy. Including contextual variables in the model allowed adults with same individual demographics to have different probabilities of being a smoker if they lived in different neighborhoods, thereby retaining local variations. To account for the LACHS design effects, we adjusted for sampling weights.

We first modeled the probability of being a current cigarette smoker in relation to individual-level variables and their two-way interactions. We employed backward selection procedure to select interaction terms using a $P < 0.05$ as the criterion, while forcing main effects of the individual-level variables to remain in the model. We then added contextual variables to the model and used the same criterion to perform backward selection. Finally, we added SPAs of residence to the model and modeled the geographic information as fixed effects. The final model included all individual-level variables, two-way interactions between race/ethnicity and gender and between race/ethnicity and household income, principal components accounting for community population makeup and educational attainment, and SPAs of residence. Testing for goodness of fit showed acceptable model fit ($P = 0.32$).

Calculation of Number of Smokers and Smoking Prevalence

By applying the parameter estimates and the variance–covariance matrix of regression coefficients to the PEPS dataset, we calculated smoking prevalence and confidence intervals at the city or city council district level. The formula used to calculate variances was based on the Delta method, as provided by Hosmer and Lemeshow, scenario two.⁹ To evaluate the stability of the small area estimates, we calculated coefficients of variation (CVs) and designated 30% as the acceptable cutoff for CVs, consistent with the standard practice of the National Center for Health Statistics.¹⁰ All the estimates had CVs below 30%. P values were two sided. All statistical analyses were performed in SAS 9.1 (SAS Institute, Cary, NC, USA).

RESULTS

Parameter estimates and statistical testing of the logistic regression model are summarized in Table 1. Middle-aged adults, African American men, Asian/Pacific Islander men, American Indians, and low-income adults smoked substantially more than their counterparts. In addition, the probability of being a smoker varied by population makeup and overall educational attainment in the communities where an individual lived. Adults living in SPA 1 were more likely to smoke cigarettes than all the other SPAs.

Smoking prevalence was calculated for 81 incorporated cities with a population size of 5,000 or more. Smoking prevalence varied considerably across cities, with the lowest in San Marino (5.3%) and the highest in Lancaster (21.7%), a more than fourfold difference (Table 2). Furthermore, more densely populated cities such as those in SPAs 4 and 6 had higher smoking prevalence overall (Figure 1).

Although cities with high smoking prevalence appeared to concentrate in certain geographic areas (Figure 1), smoking prevalence varied considerably among cities within some SPAs. For example, in SPA 8, the smoking prevalence for the City of Long Beach was 15.3%, while in Palos Verdes Estates, an adjacent city, the smoking prevalence was 7.4%. Smoking prevalence also varied considerably across the 15 Council Districts in the City of Los Angeles (data not shown).

TABLE 1 Parameter estimates for individual demographics and contextual factors in association with smoking status

Model covariates	Parameter estimate (β)	Standard error	P value
Individual-level variables			
Age group (Ref., 40–49)			
18–24	–0.40	0.20	0.05
25–29	–0.19	0.19	0.31
30–39	–0.02	0.14	0.90
50–59	0.05	0.12	0.70
60–64	–0.17	0.16	0.27
65 or over	–0.90	0.14	<0.01
Gender (Ref., male)			
Female	–0.24	0.14	0.09
Race/ethnicity (Ref., white)			
Latino	–0.07	0.19	0.70
African American	0.39	0.28	0.17
Asian/Pacific Islander	0.32	0.24	0.18
American Indian and white/American Indian	1.39	0.53	<0.01
Federal poverty level (FPL) (Ref., 200% or above FPL)			
0–99% FPL	0.97	0.26	<0.01
100–199% FPL	0.91	0.19	<0.01
Interactions of individual-level variables			
Gender \times race/ethnicity			
Female \times Latino	–0.84	0.22	<0.01
Female \times African American	–0.43	0.31	0.16
Female \times Asian/Pacific Islander	–1.87	0.38	<0.01
Female \times American Indian and white/American Indian	–0.24	0.70	0.74
Race/ethnicity \times FPL			
Latino \times 0–99% FPL	–0.81	0.32	0.01
Latino \times 100–199% FPL	–0.81	0.27	<0.01
African American \times 0–99% FPL	–0.03	0.39	0.95
African American \times 100–199% FPL	–0.16	0.42	0.69
Asian/Pacific Islander \times 0–99% FPL	–0.81	0.43	0.06
Asian/Pacific Islander \times 100–199% FPL	–0.90	0.40	0.03
American Indian and white/American Indian \times 0–99% FPL	–1.34	1.15	0.24
American Indian and white/American Indian \times 100–199% FPL	–1.24	0.85	0.15
Principal components of contextual variables			
Population makeup			
1st principal component	–0.07	0.06	0.24
2nd principal component	0.15	0.06	0.02
3rd principal component	0.17	0.08	0.05
Educational attainment			
1st principal component	0.19	0.07	<0.01
2nd principal component	0.05	0.06	0.41
Geographic areas			
Service Planning Areas (SPA) (Ref., SPA 2)			
SPA 1	0.56	0.24	0.02

TABLE 1 (Continued)

Model covariates	Parameter estimate (β)	Standard error	P value
SPA 3	-0.04	0.15	0.81
SPA 4	-0.11	0.18	0.55
SPA 5	-0.42	0.27	0.11
SPA 6	0.18	0.20	0.37
SPA 7	-0.01	0.18	0.95
SPA 8	0.11	0.15	0.46

A fixed-effects logistic regression model was fitted to assess the associations between smoking status (yes vs. no) and the variables listed in the table

DISCUSSION

Our study reveals that smoking prevalence varies considerably across cities in LA County and that impoverished cities have higher smoking rates overall. The geographic disparities identified by our small area estimation add another layer of data to our knowledge of the health disparities experienced by low-income urban communities. These communities are also deeply affected by issues such as high prevalence of obesity, food deserts, and lack of safe places to be physically active.¹¹⁻¹⁵ Taken collectively, these health disparities call for comprehensive approaches to inform and mobilize local residents, and to foster built environments and social norms conducive to healthy lifestyles. Given that cigarette smoking remains the leading preventable cause of death and disability,¹⁶ the creation of healthy environments and social norms requires tobacco control. Our estimates provide data sought by city government agencies and local tobacco control advocates to inform implementation of anti-smoking policies at the city level. Specifically, small area estimates of tobacco use can help policymakers identify cities in greatest need of tobacco control efforts and deploy resources accordingly.

Model-based small area estimation methods have gained popularity in recent years and have provided urban jurisdictions a tool for calculating data useful for local program planning and policymaking. Model-based approaches may provide more valid and

TABLE 2 Highest and lowest adult smoking prevalence among 81 incorporated cities of Los Angeles County, CA, USA, 2007

City	Prevalence	
	Percent	95% CI
Lancaster	21.7	(16.4%, 27.1%)
West Hollywood	19.6	(14.0%, 25.1%)
Palmdale	18.5	(13.7%, 23.3%)
Hawthorne	18.3	(13.7%, 22.9%)
Lawndale	17.7	(14.2%, 21.2%)
Palos Verdes Estates	7.4	(4.1%, 10.6%)
Calabasas	7.3	(4.5%, 10.1%)
La Canada Flintridge	6.4	(3.9%, 9.0%)
Malibu	5.8	(2.7%, 9.0%)
San Marino	5.3	(2.7%, 7.9%)

Smoking prevalence was not calculated for seven incorporated cities with population sizes <5,000

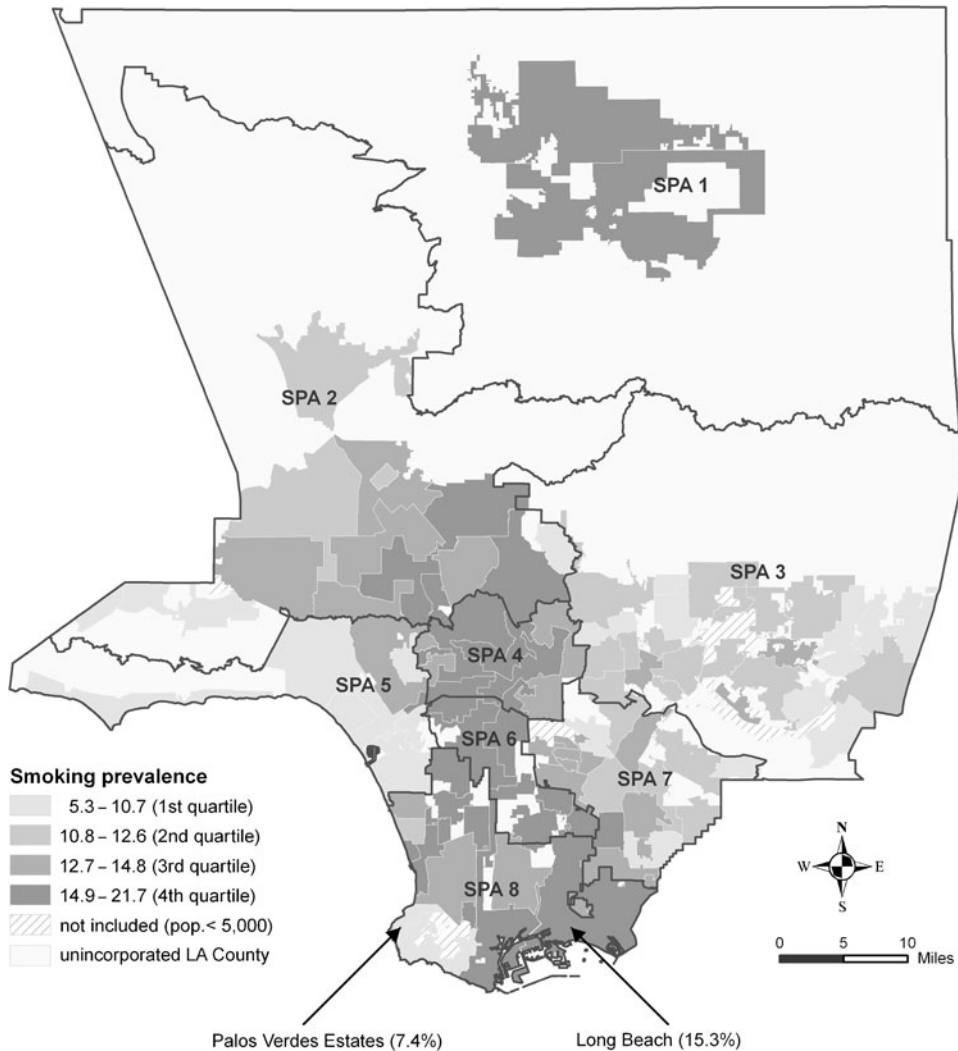


FIGURE 1. Variations of adult smoking prevalence by city within Service Planning Areas (SPAs) of Los Angeles County, California.

precise estimates than synthetic method and spatial data smoothing, two other commonly used small area estimation techniques.^{17,18} Model-based methods have been used to provide small area estimates on such topics as health behaviors and chronic disease prevalence, although modeling approaches vary.^{5,6,19-24} The model-based approach used in this study allowed flexibility in aggregating data to different geographic levels. The small area estimates compared favorably to direct estimates from the LACHS at the SPA level (correlation coefficient=0.98, $P<0.0001$).

Our study is subject to several limitations. This model-based approach relies heavily on the quality of survey and population data. Census data from 2000 may not accurately reflect neighborhood characteristics in 2007. The 2007 LACHS was a landline telephone interview survey that excluded residents who only used cellular telephones, a growing population including many young and low-income people.²⁵ The 2007 LACHS had a relatively low response rate, reflecting a decline in telephone survey response rates nationwide. Nevertheless, studies have demonstrated

that non-response does not necessarily introduce substantial biases into survey estimates.^{26,27} We assessed our survey sample demographics and found the sample closely reflected the population makeup of the county's non-institutionalized adult population. Another limitation is that self-reported data collected in the LACHS are subject to reporting bias. For example, current smokers might not want to report their smoking status or might classify themselves as ex-smokers due to perceived social norms that negativize smoking. This could lead to underestimated smoking prevalence in the survey.^{28,29} In addition, we chose not to provide estimates for cities with a population size <5,000 since population estimates for cities with small population sizes might be less accurate.

The model-based estimation method reported here uses shared characteristics in a large area to derive estimates for small areas of interest. Due to its "borrowing strength," the method yields relatively stable small area estimates. Nevertheless, the range of small area estimates may be artificially narrower than direct estimates from local surveys.⁵ To mitigate the shrinkage effect, we accounted for neighborhood characteristics in the regression model. Finally, not all factors that could predict smoking status were included in the model. For example, local tobacco control ordinances or differences in community norms may affect smoking behavior among local residents, but we could not account for these in the model. As a result, the small area estimates presented here are not suitable for assessing the effectiveness of local tobacco control interventions.

In conclusion, our study demonstrates the feasibility of providing credible local estimates of smoking prevalence using the model-based small area estimation method. These local estimates can help guide tobacco control and prevention efforts in local communities and inform targeted interventions to reduce smoking in urban areas at highest risk.

ACKNOWLEDGMENTS

We thank Yajun Du for helping with the 2000 Census data, Aida Angelescu and Alex Ho for helping with PEPS data and creating the map, and as Linda Aragon and Mark Weber for providing expertise on local tobacco control efforts. We are also indebted to the staff of the Health Assessment Unit for their ongoing work on the LACHS, and to Paul Simon who provided guidance on the project.

REFERENCES

1. Los Angeles County Department of Public Health. *Cigarette Smoking in Los Angeles County: Local Data to Inform Tobacco Policy*. Los Angeles, CA: Los Angeles County Department of Public Health, Office of Health Assessment and Epidemiology; June 2010.
2. Los Angeles County Department of Public Health. *LA Health Trends: Cigarette Smoking among Los Angeles County Adults*. Los Angeles, CA: Los Angeles County Department of Public Health, Office of Health Assessment and Epidemiology; 2006.
3. Warner KE, Mendez D, Alshanteqy O. Tobacco control success versus demographic destiny: examining the causes of the low smoking prevalence in California. *Am J Public Health*. 2008; 98(2): 268–9.
4. Simon PA, Wold CM, Cousineau MR, Fielding JE. Meeting the data needs of a local health department: the Los Angeles County Health Survey. *Am J Public Health*. 2001; 91(12): 1950–2.

5. Yu H, Meng YY, Mendez-Luck CA, Jhawar M, Wallace SP. Small-area estimation of health insurance coverage for California legislative districts. *Am J Public Health*. 2007; 97(4): 731–7.
6. Mendez-Luck CA, Yu H, Meng YY, Jhawar M, Wallace SP. Estimating health conditions for small areas: asthma symptom prevalence for state legislative districts. *Health Serv Res*. 2007; 42(6 Pt 2): 2389–409.
7. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 5th ed. Lenexa, KS: AAPOR: The American Association for Public Opinion Research; 2008.
8. County of Los Angeles. *Estimated Census Tract-City Split Population by Age, Sex, and Race/Ethnicity*. Los Angeles, CA: County of Los Angeles, Chief Executive Office, Service Integration Branch; 2008.
9. Hosmer DW, Lemeshow S. Confidence interval estimates of an index of quality performance based on logistic regression models. *Stat Med*. 1995; 14(19): 2161–72.
10. Cohen RA, Bloom B. *Trends in Health Insurance and Access to Medical Care for Children under Age 19 Years: United States, 1998–2003*. Advance data from Vital and Health Statistics. Hyattsville, MD: National Center for Health Statistics; 2005.
11. Lewis LB, Sloane DC, Nascimento LM, et al. African Americans' access to healthy food options in South Los Angeles restaurants. *Am J Public Health*. 2005; 95(4): 668–73.
12. Kipke MD, Iverson E, Moore D, et al. Food and park environments: neighborhood-level risks for childhood obesity in East Los Angeles. *J Adolesc Health*. 2007; 40(4): 325–33.
13. Simon PA, Kwan D, Angelescu A, Shih M, Fielding JE. Proximity of fast food restaurants to schools: do neighborhood income and type of school matter? *Prev Med*. 2008; 47(3): 284–8.
14. Sloane DC, Diamant AL, Lewis LB, et al. Improving the nutritional resource environment for healthy living through community-based participatory research. *J Gen Intern Med*. 2003; 18(7): 568–75.
15. Jetter KM, Cassady DL. The availability and cost of healthier food alternatives. *Am J Prev Med*. 2006; 30(1): 38–44.
16. Smoking-attributable mortality, years of potential life lost, and productivity losses—United States, 2000–2004. *MMWR Morb Mortal Wkly Rep*. 2008; 57(45): 1226–1228.
17. Goodman MS. Comparison of small-area analysis techniques for estimating prevalence by race. *Prev Chronic Dis*. 2010; 7(2): A33.
18. Jia H, Muennig P, Borawski E. Comparison of small-area analysis techniques for estimating county-level outcomes. *Am J Prev Med*. 2004; 26(5): 453–60.
19. Malec D, Davis WW, Cao X. Model-based small area estimates of overweight prevalence using sample selection adjustment. *Stat Med*. 1999; 18(23): 3189–200.
20. Li W, Kelsey JL, Zhang Z, et al. Small-area estimation and prioritizing communities for obesity control in Massachusetts. *Am J Public Health*. 2009; 99(3): 511–9.
21. Li W, Land T, Zhang Z, Keithly L, Kelsey JL. Small-area estimation and prioritizing communities for tobacco control efforts in Massachusetts. *Am J Public Health*. 2009; 99(3): 470–9.
22. Knutson K, Zhang W, Tabnak F. Applying the small-area estimation method to estimate a population eligible for breast cancer detection services. *Prev Chronic Dis*. 2008; 5(1): A10.
23. Twigg L, Moon G. Predicting small area health-related behaviour: a comparison of multilevel synthetic estimation and local survey data. *Soc Sci Med*. 2002; 54(6): 931–7.
24. Twigg L, Moon G, Jones K. Predicting small-area health-related behaviour: a comparison of smoking and drinking indicators. *Soc Sci Med*. 2000; 50(7–8): 1109–20.
25. Blumberg SJ, Luke JV. *Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2009*. Hyattsville, MD: Division of Health Interview Statistics, National Center for Health Statistics; 2010.
26. Keeter S, Kennedy C, Dimock M, Best J, Graighill P. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. *Public Opin Q*. 2006; 70(5): 759–79.

27. Keeter S, Miller C, Kohut A, Groves RM, Presser S. Consequences of reducing nonresponse in a national telephone survey. *Public Opin Q.* 2000; 64(2): 125–48.
28. Murray RP, Connett JE, Istvan JA, Nides MA, Rempel-Rossum S. Relations of cotinine and carbon monoxide to self-reported smoking in a cohort of smokers and ex-smokers followed over 5 years. *Nicotine Tob Res.* 2002; 4(3): 287–94.
29. Pell JP, Haw SJ, Cobbe SM, et al. Validity of self-reported smoking status: comparison of patients admitted to hospital with acute coronary syndrome and the general population. *Nicotine Tob Res.* 2008; 10(5): 861–6.