

Small-Area Estimation and Prioritizing Communities for Tobacco Control Efforts in Massachusetts

Wenjun Li, PhD, Thomas Land, PhD, Zi Zhang, MD, MPH, Lois Keithly, PhD, and Jennifer L. Kelsey, PhD

Tobacco use remains the leading cause of preventable chronic disease and death in the United States.¹ Each year, tobacco use is responsible for more deaths than automobile accidents, AIDS, homicides, suicides, and poisonings combined.² Although a decline in tobacco use in the United States is evident,³ cigarette smoking continues to be a major challenge to public health. According to national data from the 2005 Behavioral Risk Factor Surveillance System (BRFSS), approximately 95% of states, territories, metropolitan statistical areas, and counties do not meet the national goal of smoking prevalence rates of 12% or less among the adult population.³

The Massachusetts Tobacco Control Program, established in 1993 to coordinate and implement the state's tobacco control efforts, has developed an extensive local infrastructure, delivered comprehensive media campaigns, and led public policy on tobacco use.⁴ However, significant budget cuts after 2002 have made it difficult for the program to remain effective. One way to maintain the effectiveness of the program is to focus tobacco control efforts in communities with the greatest need. For this purpose, it is essential to be able to identify communities with high smoking prevalence and formulate health policies for the specific needs of those communities.

For Massachusetts, the BRFSS is the only source of population-based information on tobacco use. However, the current BRFSS does not provide health statistics at the community level (i.e., town, small city, or subdivision of a large city), in part, because the BRFSS is designed primarily for providing statewide or metropolitan area health statistics and also because the majority of communities do not have adequate sample sizes for directly calculating prevalence rates with reasonable precision. For example, in 2005, more than 58% of the communities had a sample size smaller than 15 people, and only 22% had a sample size of 30 people or more. Methods for producing reliable community-level statistics are needed.

Objectives. We developed a method to evaluate geographic and temporal variations in community-level risk factors and prevalence estimates, and used that method to identify communities in Massachusetts that should be considered high priority communities for smoking interventions.

Methods. We integrated individual-level data from the Behavioral Risk Factor Surveillance System from 1999 to 2005 with community-level data in Massachusetts. We used small-area estimation models to assess the associations of adults' smoking status with both individual- and community-level characteristics and to estimate community-specific smoking prevalence in 398 communities. We classified communities into 8 groups according to their prevalence estimates, the precision of the estimates, and temporal trends.

Results. Community-level prevalence of current cigarette smoking among adults ranged from 5% to 36% in 2005 and declined in all but 16 (4%) communities between 1999 and 2005. However, less than 15% of the communities met the national prevalence goal of 12% or less. High smoking prevalence remained in communities with lower income, higher percentage of blue-collar workers, and higher density of tobacco outlets.

Conclusions. Prioritizing communities for intervention can be accomplished through the use of small-area estimation models. In Massachusetts, socioeconomically disadvantaged communities have high smoking prevalence rates and should be of high priority to those working to control tobacco use. (*Am J Public Health*. 2009;99:470–479. doi:10.2105/AJPH.2007.130112)

With BRFSS data, we adapted mixed-effect logistic regression models to estimate and analyze geographic variations and temporal trends in smoking prevalence rates of 398 communities (including 339 towns and small cities, and 59 subdivisions of the 12 largest cities) in Massachusetts. Such information can assist in the planning of statewide tobacco control efforts and in the allocation of limited resources to those communities most in need.

METHODS

Definition of Communities and Geographic Unit of Analysis

We defined an adult's community as the town, city, or subdivision of a large city in which he or she resided at the time of interview. The city of Boston, which had more than 520 000 residents, was divided into 17 communities that were identical to the neighborhood planning districts defined by the Boston Redevelopment Authority. For Boston

residents, a person's community was determined by (in order of importance) self-reported Boston neighborhood, 5-digit zip code, and landline telephone exchange. The city of Worcester, which had more than 154 000 residents, was divided into 8 subdivisions; residents were sorted into communities based on their self-reported 5-digit zip code and landline telephone exchange. Each of the 8 subdivisions included 3 to 5 adjacent census tracts with comparable sociodemographic profiles. We used similar methods to divide other large cities, including Brockton, Cambridge, Fall River, Lawrence, Lowell, Lynn, New Bedford, Newton, Quincy, and Springfield, into subdivisions. In total, we defined 400 communities (geographic units). We excluded 2 communities (Monroe and Harbor Island in the city of Boston) because of insufficient data, leaving 398 communities for analysis. A unique community identification code was assigned to each of the defined communities, through which individual- and community-level

data were linked. The communities had wide variations in socioeconomic characteristics and demographic composition.

Source of Individual-Level Data

We obtained individual-level data from the BRFSS for 1999 to 2005. Since 1999, respondents' residences have been recorded at both the town or city and 5-digit zip code levels. Residents of Boston were also asked to give their neighborhood planning district. The BRFSS is a collaborative effort between the Centers for Disease Control and Prevention and health departments from 50 US states, the District of Columbia, and 3 territories. The BRFSS collects uniform state-specific data on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and infectious diseases in the adult population living in households. The BRFSS is one of the major sources of information for national, state, and local public health policymaking and evaluation.⁵ Data are collected annually from a random sample of adults via a telephone survey employing random-digit dialing. In the Massachusetts BRFSS, both cigarette smoking status and sociodemographic characteristics are assessed.

Source of Community-Level Data

Community-level sociodemographic data were obtained or derived from the 2000 Census.⁶ The data, including sociodemographic and cartographic boundary files, were obtained through either the US Census Bureau or MassGIS⁷ of the Commonwealth of Massachusetts. Population composition data at the town or city level were available in the 2000 Census. Population data at the city subdivision level were derived through aggregation of census tract-level data.

Community-level administrative and programmatic data, including substance abuse admissions, were obtained from the Tobacco Control Program of the Massachusetts Department of Public Health. The density of tobacco outlets was calculated based on data from the Massachusetts Department of Revenue's license registry of tobacco retailers.

Definition of Cigarette Smoking Status

In each year of the survey, interviewees were asked whether they had smoked at least 100 cigarettes in their lifetime and whether they currently smoked cigarettes. We classified respondents

as either current smokers (being a smoker at the time of interview; n=11 241, 20.0%) or noncurrent smokers (never smoked or having quit smoking; n=44 711, 79.3%); we excluded unknowns (n=376, 0.7%) from the analysis according to the established BRFSS coding methods.

Statistical Analysis

Statistical models. Following Malec et al.,^{8,9} we adapted mixed effects logistic regression models to estimate community-level smoking prevalence. The models estimated community-level smoking prevalence based on the association between respondents' smoking status and individual- and community-level socioeconomic characteristics, taking into account community demographics. Such models have been successfully applied to estimate undercount for small geographic areas by the US Census,^{8,9} county-level health statistics by Centers for Disease Control and Prevention,¹⁰ and smoking prevalence in the United Kingdom.¹¹

The small-area estimation model includes both individual- and community-level predictors of cigarette smoking. Odds ratios (ORs) for each predictor are calculated by applying exponential transformation of the unstandardized coefficient estimates in the model. This 2-level model is specified as

$$(1) \text{ logit}(y_{ij}) = \beta_0 + \sum_{p=1}^P x_{p,ij}\beta_p + t_{ij}\beta_t + \gamma_{0i} + \sum_{q=1}^Q z_{q,i}\gamma_q + t_{ij}\gamma_{t,i} + \varepsilon_{ij},$$

where y_{ij} is the smoking status of the j th respondent in the i th community, with $i = 1, \dots, 398$, $y_{ij} = 1$ for current smoker, and $y_{ij} = 0$ for noncurrent smoker; x_p is the p th individual-level covariate, with $p = 1, \dots, P$; β_p is the coefficient corresponding to the p th individual-level covariate; t_{ij} is the year of interview (used as a quantitative variable); β_t is the state average slope (temporal trend) over interview year; γ_0 is the community-level random intercept that follows a normal distribution with zero mean; z_q is the q th community-level covariate, $q = 1, \dots, Q$; γ_q is the coefficient corresponding to the q th community-level covariate; $\gamma_{t,i}$ is the community-level random effect of slope over interview year and follows a normal distribution with zero mean; and ε_{ij} is the random error, which follows a normal distribution with zero mean.

Individual- and community-level predictors.

We included several individual-level characteristics as fixed effects in the models: gender, age group, race/ethnicity, marital status, education level, employment status, and annual household income level. These variables were derived from Massachusetts BRFSS data and have been documented as personal risk factors for cigarette smoking.¹² Missing values for these variables were replaced with either medians or means or the most frequent categories of these variables when appropriate. These individual-level categorizations are similar to those in the 2000 Census, so the community-level census data can be incorporated in the small-area estimations to account for the demographic compositions of the communities.

The models also included selected community-level characteristics as fixed effects: median per capita income, percentage of owner-occupied housing units, percentage of blue-collar jobs in the total employed labor force, racial diversity, percentage of vacant housing units, percentage of population in rural area, the crude rate of admission to Department of Public Health-funded substance abuse treatment programs, and density of tobacco outlets (number per mile of road).^{13,14} We determined functional forms of these variables in relation to smoking with fractional polynomial regression models or locally averaged smoothing curves controlled for individual-level variables. We categorized variables with nonlinear associations by quartiles or thresholds if present.

Assessment of temporal trend. Year of interview (t_{ij}) was included in the model to estimate the community-specific slopes of change in smoking prevalence over time and to examine the variations in smoking prevalence changes among communities. Additional community-level variations were modeled as random effects and included in the model as random intercepts.

We examined potential spatial autocorrelations in the community-level random effects between adjacent communities, and observed no strong autocorrelations (Moran Index= 0.04; $P = .17$).

Estimation of community-level smoking prevalence. With the 2-level model and the data on the sociodemographic characteristics of the communities, the logit value of i th community in year T was estimated as

$$(2) \bar{\eta}_{i,T} = \sum_{p=1}^P \bar{x}_{p,i} \hat{\beta}_p + \sum_{q=1}^Q z_{q,i} \hat{\gamma}_q + (\hat{\beta}_0 + \gamma_{0i}) + T(\hat{\beta}_t + \hat{\gamma}_{t,i}),$$

where $\bar{x}_{p,i}$ is the mean value of the p th individual-level covariate according to the 2000 Census. The confidence interval (CI) of $\bar{\eta}_{i,T}$ was calculated with the variances of the random effects. Community-level smoking prevalence rates and associated CIs were then calculated by applying exponential transformation of the community-level logit values and confidence intervals.

Classification of communities. We classified communities into 8 priority classes according to the smoking prevalence estimates and their precision and temporal trends and then provided specific recommendations on tobacco control efforts.

We conducted all analyses with Stata MP version 9.1 (StataCorp, College Station, TX) and mapped the smoking prevalence estimates with ESRI ArcGIS version 9.2 (ESRI, Redlands, CA).

RESULTS

Our study sample included 56 328 adult respondents, of whom 55 467 (98.5%) from 398 communities were included in our analyses. Individuals were excluded if their towns or city subdivisions of residence could not be determined ($n=639$, 1.1%) or if smoking status was unknown ($n=376$, 0.7%); some individuals fell into both categories. The study sample was 40% men and 60% women and the mean age of participants was 47 years.

With BRFSS data, we estimated Massachusetts's average prevalence of current cigarette smoking between 1999 and 2005 to be 19.1% (95% CI=18.7%, 19.6%), which was lower than the national average. The smoking prevalence rates of the state's 3 largest cities, Boston, Worcester, and Springfield, were 18.1% (95% CI=17.0%, 19.2%), 25.4% (95% CI=23.6%, 27.3%), and 25.6% (95% CI=23.8%, 27.5%), respectively.

Determinants of Smoking Status

Table 1 presents ORs for smoking according to each individual- and community-level factor. Individual-level factors associated with higher probability of smoking include male gender,

TABLE 1—Multivariable Adjusted Odds Ratios (AORs) of Individual- and Community-Level Factors for Cigarette Smoking: Behavioral Risk Factor Surveillance System, Massachusetts, 1999–2005

Risk Factor	AOR (95% CI)
Change over any 5-y period ^a	0.96 (0.91, 1.02)
Gender	
Women (Ref)	1.00
Men	1.07 (1.02, 1.12)
Age at time of interview	
18–44 y	3.96 (3.64, 4.32)
45–65 y	3.03 (2.79, 3.29)
≥65 y (Ref)	1.00
Race/ethnicity	
Non-Hispanic White (Ref)	1.00
Non-Hispanic Black	0.62 (0.56, 0.69)
Hispanic	0.42 (0.39, 0.46)
Other	0.70 (0.62, 0.79)
Marital status	
Married and living together (Ref)	1.00
Divorced	1.97 (1.84, 2.10)
Widowed	1.24 (1.12, 1.37)
Separated	1.78 (1.59, 1.99)
Never married	1.45 (1.37, 1.54)
Education	
College degree (Ref)	1.00
Some college	2.03 (1.91, 2.16)
High school diploma	2.64 (2.48, 2.80)
Less than high school	3.27 (3.01, 3.56)
Employment status	
Employed or not in labor force (Ref)	1.00
Unemployed	1.54 (1.41, 1.67)
Household income, \$	
< 15 000	1.80 (1.66, 1.96)
15 000–24 999	1.50 (1.39, 1.61)
25 000–34 999	1.50 (1.39, 1.62)
35 000–49 999	1.38 (1.29, 1.48)
50 000–75 000	1.13 (1.05, 1.21)
> 75 000 (Ref)	1.00
Average per capita income, \$	
< 25 000	1.66 (1.29, 2.13)
25 000–34 999	1.52 (1.19, 1.94)
35 000–50 000	1.21 (0.94, 1.56)
> 50 000 (Ref)	1.00
% of blue-collar job workers	1.06 (1.02, 1.10)
Racial diversity ^b (range=0–1)	1.03 (1.00, 1.05)
% of Population living in urban area	
< 25 (Ref)	1.00
25–89	1.19 (1.10, 1.29)
≥ 90	1.32 (1.12, 1.55)

Continued

TABLE 1—Continued

Housing units vacant	
< 5% (Ref)	1.00
≥ 5%	1.10 (1.03, 1.18)
Substance abuse hospitalization rate increasing by 1 per 1000 adults	1.03 (1.00, 1.06)
Tobacco retailer density increase of 1 additional retailer per 10 miles of road	1.13 (1.00, 1.27)

Note. CI = confidence interval.

^aInterview years were 1999 to 2005.

^bA value of 0 indicates a single-race community (e.g., 100% non-Hispanic White or 100% non-Hispanic Black). A value of 1 indicates that equal proportions of 2 or more races are present (e.g., 25% non-Hispanic White, 25% non-Hispanic Black, 25% Hispanic, and 25% other).

younger age, White race, living alone, less education, unemployment at the time of interview, and lower household income.

Independent of individual-level variables, community-level factors associated with higher odds of smoking in individuals include lower per capita income, higher percentage of blue-collar workers in the employed labor force, higher vacancy rate for housing units, higher racial diversity, and higher density of tobacco outlets.

Community-Level Smoking Prevalence Estimation and Precision

We used the demographic compositions of the communities and estimated associations to estimate community-level prevalence of smoking. For the most recent year with available data (2005), smoking prevalence ranged from less than 5% in communities that were most socioeconomically advanced to 36.3% in those that were most disadvantaged. The majority (85%) of communities did not meet the national goal of reducing smoking prevalence to 12% or less of the adult population.

Table 2 shows the 20 communities with the highest and lowest smoking prevalence rates, along with their key sociodemographic characteristics. Compared with the communities with the lowest smoking prevalence rates, communities with the highest smoking prevalence rates had, on average, one third the per capita income, 2.5 times the unemployment rate, 7 times the rate of admission to substance abuse treatment programs, and 6 times the density of tobacco retailers.

The margins of error (or precision, defined as half the width of the 95% CI) of prevalence

estimates ranged from 1 to 5 percentage points, with more than 85% of the communities having margins of error of fewer than 4 percentage points. For example, prevalence estimates of 2005 had margins of error of fewer than 4 percentage points in 80% of the communities.

Temporal Trend in Smoking Prevalence

The prevalence of smoking in all except 16 (4%) communities declined between 1999 and 2005. There were wide variations in the rates of decline. The ORs of change over any 5-year period ranged from 0.90 (declining) to 1.08 (increasing), with a median of 0.96.

Classification of Communities and Recommended Actions

We sorted the communities into 8 priority classes according to prevalence of smoking, precision of estimates, and temporal trends. Precision criteria vary according to prevalence, which is necessary because precision is proportional to prevalence and the use of a single strict precision criterion (e.g., fewer than 3 percentage points) would result in a large number of communities with high prevalence rates being classified into lower-priority groups. The 8 classes, with recommended specific tobacco control actions for each class of community, are:

Class 1. Definite hotspots: communities with smoking prevalence estimates above the national average (>21%) and of good precision ($\leq 4\%$), or those that had an increasing trend over time. These communities should be a high priority for intervention;

we classified 49 communities as definite hotspot communities.

Class 2. Probable hotspots: communities with smoking prevalence estimates above the national average (>21%) and of limited precision (>4%). These communities should be a high priority for both intervention and enhanced surveillance; we classified 72 communities as probable hotspot communities.

Class 3. Moderate priority: communities with (1) smoking prevalence estimates around the state average (16%–21%) but below the national average (21%), (2) good estimate precision ($\leq 3.5\%$), and (3) prevalence estimates that decline slower than the average state rate. These communities should be a moderate priority for intervention; we classified 30 communities as class 3 moderate priority communities.

Class 4. Moderate priority: communities with (1) smoking prevalence estimates around the state average (16%–21%) but below the national average (21%), (2) good estimate precision ($\leq 3.5\%$), and (3) prevalence estimates that decline faster than the average state rate. These communities require continued surveillance; we classified 85 communities as class 4 moderate priority communities.

Class 5. Moderate priority: communities with smoking prevalence estimates around the state average (16%–21%) but below the national average (21%) and of limited precision (>3.5%). These communities are a priority for enhanced surveillance; we classified 33 communities as class 5 moderate priority communities.

Class 6. Lower priority: communities with smoking prevalence estimates (12%–16%) lower than the state average (18%) but above the national goal (<12%) and with limited precision (>3%). These communities require enhanced surveillance and interventions that sustain the low prevalence rates; we classified 34 communities as class 6 lower priority communities.

Class 7. Lower priority: communities with prevalence estimates lower than the state average (18%) but above the national goal (<12%) and with good precision ($\leq 3\%$). These communities should maintain their current level of surveillance and sustain the

TABLE 2—Characteristics of the Massachusetts Communities with the 20 Highest and the 20 Lowest Cigarette Smoking Prevalence Rates: Behavioral Risk Factor Surveillance System, 2005

Rank ^a	Community (5-digit zip code ^b)	Total Sample Size 1999–2005, No.	Smoking Prevalence, % (95% CI)	Per capita Income, \$1000	Unemployment Rate, %	Racial Diversity Index ^c	Rural Population, ^d %	Substance Abuse Hospitalizations per 1000 Adults	Tobacco Outlets per 10 Miles of Road, No.
Communities with the 20 highest smoking prevalence rates									
398	New Bedford (02744)	169	36.3 (31.4, 41.5)	11.7	6.5	0.63	0.00	4.52	12.3
397	New Bedford (02746)	172	34.7 (29.9, 39.7)	11.4	6.9	0.55	0.00	4.52	8.1
396	Fall River (02724)	285	34.6 (30.2, 39.3)	14.3	4.3	0.86	0.00	4.66	9.4
395	Fall River (02721)	370	34.5 (30.3, 39.0)	15.6	3.9	0.83	0.00	4.66	7.0
394	Fall River (02723)	238	34.2 (29.7, 39.0)	12.1	4.4	0.85	0.00	4.66	7.0
393	Worcester (01610)	330	33.8 (29.6, 38.3)	11.6	5.4	0.42	0.00	4.46	13.4
392	Worcester (01607)	143	31.3 (26.7, 36.3)	12.2	10.1	0.68	0.00	4.46	5.2
391	North Adams	94	30.3 (25.6, 35.4)	16.4	3.4	0.91	8.30	2.84	2.5
390	Fall River (02720)	462	29.9 (26.1, 33.9)	20.4	3.6	0.89	2.30	4.66	3.2
389	New Bedford (02745)	369	28.5 (24.6, 32.6)	20.2	3.6	0.84	1.30	4.52	2.4
388	Springfield (01103, 01105, 01107)	305	28.2 (24.4, 32.3)	17.4	6.2	0.39	0.00	5.24	14.2
387	Winchendon	63	28.1 (23.5, 33.2)	18.8	3.8	0.92	40.9	1.43	0.9
386	Revere	272	28.1 (24.1, 32.5)	19.7	3.4	0.69	0.00	3.15	7.3
385	Lowell (01850)	262	28.0 (24.0, 32.4)	16.7	5.4	0.45	0.00	3.51	7.0
384	Webster	157	27.8 (23.6, 32.6)	20.4	2.9	0.89	6.90	2.37	2.7
383	Warren	30	27.5 (22.8, 32.7)	17.2	3.3	0.95	53.8	1.80	0.8
382	Salisbury	40	27.0 (22.5, 32.2)	21.6	2.5	0.95	11.9	2.37	2.4
381	Worcester (01603)	306	27.0 (23.3, 31.2)	18.3	3.3	0.61	0.00	4.46	6.6
380	Gardner	140	27.0 (22.8, 31.6)	18.6	2.8	0.85	6.1	2.28	2.6
379	Fitchburg	445	26.8 (22.9, 31.1)	17.3	5.3	0.67	5.5	3.08	3.0
	Average	233	30.2	16.6	4.6	0.74	6.9	3.68	5.9
Communities with the 20 lowest smoking prevalence rates^e									
1	Carlisle Town	31	4.7 (3.7, 6.0)	59.6	0.5	0.88	85.6	0.62	0.16
2	Weston Town	63	5.5 (4.4, 7.0)	79.6	1.0	0.80	0.0	0.30	1.10
3	Dover Town	41	5.6 (4.4, 7.1)	64.9	0.9	0.90	41.3	0.38	0.26
4	Sherborn Town	23	5.9 (4.7, 7.5)	58.1	2.9	0.92	66.8	0.70	0.89
5	Concord Town	118	6.1 (4.8, 7.7)	51.5	1.1	0.82	11.8	0.78	1.60
6	Wellesley Town	130	6.3 (5.0, 7.9)	52.9	2.2	0.78	0.0	0.28	1.25
7	Wayland Town	78	6.7 (5.3, 8.4)	52.7	1.3	0.85	9.9	0.59	0.80
8	Winchester Town	124	7.0 (5.6, 8.8)	50.4	1.6	0.87	0.0	0.55	1.00
9	Lincoln Town	38	7.1 (5.6, 8.9)	49.1	1.1	0.77	22.6	0.26	0.26
10	Lexington Town	220	7.2 (5.7, 8.9)	46.1	1.6	0.77	0.0	0.38	1.02
11	Sudbury Town	113	7.4 (5.9, 9.3)	53.3	1.3	0.88	10.0	0.53	1.50
12	Newton (02461, 02464, 02468)	88	7.7 (6.1, 9.6)	38.0	1.6	0.78	0.0	0.45	2.09
13	Harvard Town	39	7.7 (6.1, 9.7)	40.9	2.0	0.73	98.8	0.45	0.06
14	Newton (02459, 02467)	183	7.7 (6.2, 9.6)	37.6	2.3	0.74	0.0	0.45	1.14
15	Sharon Town	101	7.9 (6.3, 9.8)	41.3	1.7	0.82	15.5	0.43	1.63
16	Boxford Town	44	8.3 (6.6, 10.4)	48.8	1.1	0.95	48.6	0.76	0.16
17	Newton (02462, 02466)	49	8.6 (6.8, 10.6)	37.2	3.6	0.75	0.0	0.45	1.90
18	Westwood Town	85	8.7 (6.9, 10.8)	41.6	1.6	0.93	2.9	0.47	2.07
19	Longmeadow Town	93	8.8 (7.0, 10.9)	38.9	2.1	0.91	0.0	0.60	2.01

Continued

TABLE 2—Continued

20	Needham	172	8.8 (7.1, 10.9)	44.5	1.9	0.90	2.9	0.28	1.20
	Average	92	7.2	49.4	1.7	0.80	20.8	0.49	0.98

Note. CI = confidence interval.

^aRank of the community-level smoking prevalence estimates (highest = 398, lowest = 1).

^bZip codes are only listed for communities located within a larger city whose multiple communities are identifiable only by 5-digit zip code.

^cDiversity index of racial composition in the community with a range of 0 to 1. A value of 0 indicates a single-race community (e.g., 100% non-Hispanic White or 100% non-Hispanic Black). A value of 1 indicates that equal proportions of 2 or more races are present (e.g., 25% non-Hispanic White, 25% non-Hispanic Black, 25% Hispanic, and 25% other).

^dPer the 2000 US Census.

^eThese communities met the national goal of smoking prevalence rates of 12% or less of the adult population.

low smoking prevalence rates; we classified 37 communities as class 7 lower priority communities.

Class 8. Lowest priority: communities with smoking prevalence estimates that are lower than the national goal (<12%) and have good precision ($\leq 3\%$). These communities require continued surveillance and should maintain their tobacco control efforts; we classified 58 communities as lowest priority communities.

Mapping of Prevalence and the Priority Classification of Communities

Smoking prevalence estimates of each community in Massachusetts for 2005 (the most recent year for which data were available) are shown in Figure 1. Geographic disparities in smoking prevalence are evident, with high prevalence rates in socioeconomically disadvantaged communities and rural communities in central, western, and southeastern Massachusetts. These communities are characterized by low community-level income, lack of white-collar and professional job opportunities, and large travel distances from major job centers. The geographic disparity is largely but not entirely explained by the high correlation between smoking prevalence and community-level per-capita income (Spearman rank correlation = -0.83 ; $P < .001$).

The priority classifications of communities are illustrated in Figure 2. The figure shows which communities are a high priority for intervention, for enhanced surveillance, or for both, and which communities have met the national goal of reducing smoking prevalence to 12% of the population or less.

DISCUSSION

We used BRFSS data to demonstrate a promising method for analyzing geographic variations and temporal trends in community-level prevalence of risk behaviors. We applied the method to the analysis of smoking prevalence in Massachusetts communities and showed how the results can assist in the planning of tobacco control efforts.

Our method was adapted from statistical methods that have been used for identifying individual- and community-level risk factors. Our approach is straightforward and cost-effective, can be easily translated into routine practice, and can be implemented with existing data and via readily available statistical software packages. Proper adaptation of these methods may expand the scope of existing national health surveillance and data collection systems and provide health information and specific recommendations that are directly relevant to local communities and governments.

We defined the operating concept of community as the town, city, or subdivision of a large city of the respondent's residence. Although community-level political actions and legislation are critical to the overall success of the state's tobacco control efforts, little is known about whether associations between smoking and community-level factors exist at such small geopolitical levels and how such associations can be applied to improve local prevalence estimation, policymaking, and evaluation. In Massachusetts, the lowest level of government at which civil services operate, including land use planning, property taxation, elections, administration, social services, and police and emergency responses, is the town or city level. With the existing government

infrastructure, community-based interventions are likely to be more effective at this level than at the county or state level. Methodologically, community-level ecologic inference is less susceptible to aggregation (effect) bias and specification bias (loss of variability) than inference at the county level.¹⁵

Although the overall smoking prevalence in Massachusetts is lower than the national average, we found that most communities (85%) in Massachusetts did not meet the national goal of smoking prevalence rates of 12% or less. High smoking prevalence rates persist in socioeconomically disadvantaged communities; sixteen communities that had increases in smoking prevalence were predominantly poor urban areas with a high density of tobacco outlets and a high concentration of minorities. It is important to understand why smoking continues to be common in these communities, to assess whether tobacco control efforts have been adequately delivered to them, and to evaluate whether current tobacco control programs are effective.

The geographic disparities in smoking prevalence are consistent with the analysis of Sheer et al. on local restaurant smoking regulations in Massachusetts communities.¹⁶ Sheer et al. reported that socioeconomically disadvantaged communities were less likely to adopt stronger restaurant smoking regulations and that the pattern of smoke-free restaurant policy enactment contributes to disparities in tobacco control.

The individual- and community-level determinants of smoking behavior observed in our study are consistent with those documented by others. Higher probability of smoking has been associated with male gender, younger age, White race, living alone, less education, unemployment, and lower income. Independent

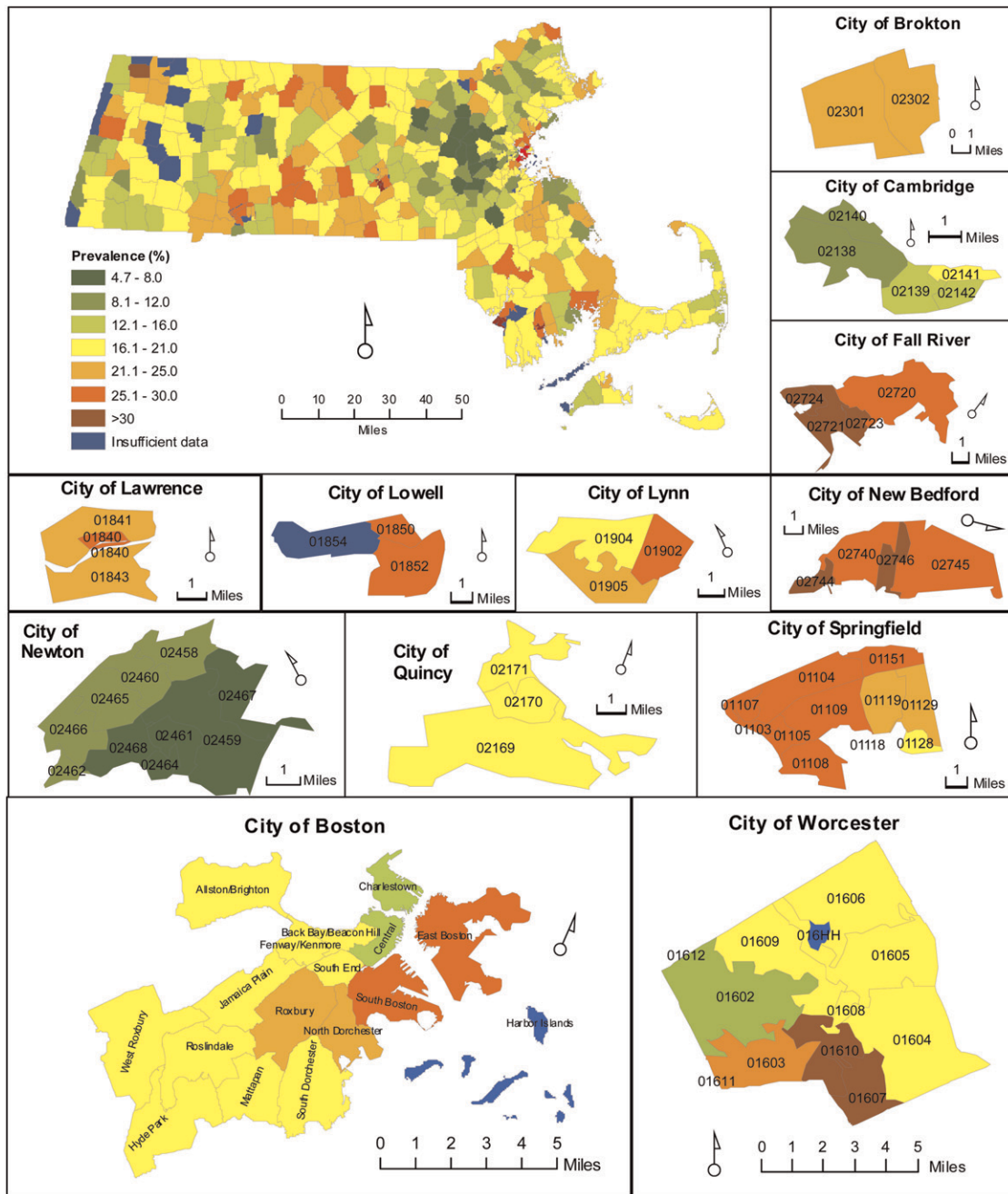


FIGURE 1—Town- and city-level prevalence of cigarette smoking: Massachusetts, 2005.

of individual-level factors, adults in socioeconomically disadvantaged communities—which are characterized by lower income,^{17,18} lower social class,¹¹ less education,¹⁸ higher unemployment rate,¹⁸ and higher residential deprivation^{19–24}—have a higher likelihood of being a smoker.^{25,26}

We observed that adult residents in communities with a higher density of tobacco retailers were more likely to be smokers after adjusting for both individual- and community-level risk factors (Table 1). The result is consistent with several recent studies that associated tobacco outlet density with smoking

behaviors among both adults^{26,27} and adolescents.^{28,29} Retail tobacco outlets tend to be disproportionately located in communities that are socioeconomically disadvantaged^{27,28} or have a higher percentage of minorities.³⁰ Because smokers have been shown to be concerned with the costs of smoking, lowering access costs

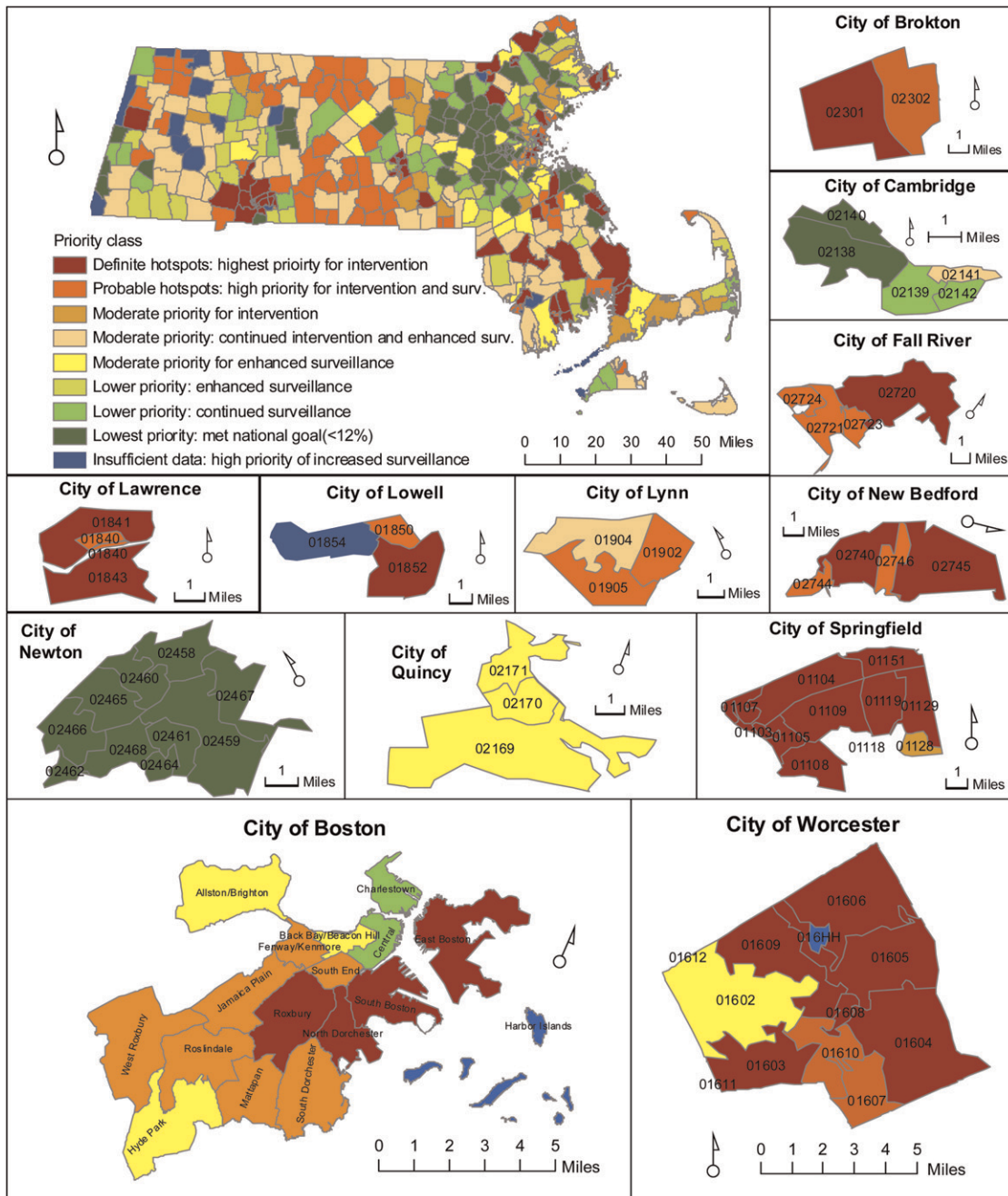


FIGURE 2—Smoking intervention priority classes of communities: Massachusetts, 2005.

(e.g., reducing the travel time to a retail tobacco outlet) is likely to increase consumption in these communities.²⁷

The small-area estimation method we used could be improved in several aspects. First, there were wide variations in town- or city-specific sample sizes. Prevalence estimates of

smoking for communities with small sample sizes had limited precision and thus may not be reliable. The best solution to this problem would be to create a new sampling design for the BRFSS to produce reliable community-level risk factor prevalence estimates without compromising the precision of state- and

county-level risk factor prevalence estimates. In Massachusetts, this would be feasible by increasing sample sizes in rural and smaller communities, especially in the central and western part of the state. The margin of error depends on both the town-specific sample size and the risk factor's prevalence rate; as sample

size increases and prevalence decreases, the margin of error decreases. Therefore, to achieve similar precision, communities with high prevalence rate estimates would require enhanced surveillance efforts (i.e., a larger sample size) than would communities with low prevalence rate estimates. Less optimally, the variation in sample sizes could be lessened by developing complex models that incorporate temporal and spatial correlations among the community-level random effects.

Second, the sampling weights were not directly included in the development of the predictive models. The weights that were available in the Massachusetts BRFSS data were derived for the purpose of producing state- and county-level estimates and were not appropriate for producing community-level estimates. The BRFSS program and associated state public health agencies should develop new weighting methods that reflect the sampling scheme and demographic composition of the communities under analysis.

Finally, a study carefully designed for validating small-area estimation models would be of great importance. We have begun to develop a study that to find effective approaches for validating these models and evaluating the progress of tobacco control. The methods we have developed and the results we have obtained have already had a direct impact on the surveillance and planning of tobacco control efforts in Massachusetts. Previously, the tobacco control program efforts focused on statewide initiatives; however, since local smoking prevalence estimates were developed, 30 new programs have been launched in communities with high prevalence rates. The large variations in community-level smoking prevalence that we uncovered can be used to motivate, formulate, and implement necessary tobacco control policies at the community level. We encourage others to explore further the potential use of our small-area estimation method. ■

About the Authors

At the time of the study, Wenjun Li and Jennifer L. Kelsey were with the Division of Preventive and Behavioral Medicine, Department of Medicine, University of Massachusetts Medical School, Worcester. Thomas Land, Zi Zhang, and Lois Keithly are with the Massachusetts Department of Public Health, Boston.

Requests for reprints should be sent to Wenjun Li, Division of Preventive and Behavioral Medicine, University of Massachusetts Medical School, Shaw Building, SH2-230, 55 Lake Ave N, Worcester, MA 01655 (e-mail: wenjun.li@umassmed.edu).

This article was accepted May 1, 2008.

Contributors

W. Li, T. Land, Z. Zhang, and L. Keithly originated this project. W. Li conducted the data analysis. T. Land and L. Keithly provided programmatic application of the study results. Z. Zhang assisted with preparation of the analytic dataset from the Behavioral Risk Factor Surveillance System. All authors helped interpret findings and write and review drafts of the article.

Acknowledgments

The authors thank Massachusetts Behavioral Risk Factor Surveillance System and Massachusetts Tobacco Control Program for providing survey and administrative data. Several components of the town- or city-level data were obtained from a study supported by the National Institute of Environmental Health Sciences (grant 1R21ES014195).

Human Participant Protection

The study was approved by the institutional review board of the University of Massachusetts Medical School.

References

1. *The Health Consequences of Smoking: A Report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention; 2004.
2. *Massachusetts Deaths 2005*. Boston: Massachusetts Department of Public Health; 2007. Available at: http://www.mass.gov/Eoehhs2/docs/dph/research_epi/death_report_05.pdf. Accessed September 24, 2008.
3. Chowdhury PP, Balluz L, Murphy W, et al. Surveillance of certain health behaviors among states and selected local areas—United States, 2005. *MMWR Surveill Summ*. 2007;56(4):1–160.
4. Hamilton WL, Rodger CN, Chen X, Njobe TK, Kling R, Norton G. *Independent Evaluation of the Massachusetts Tobacco Control Program. Eighth Annual Report: January 1994–June 2001*. Cambridge, MA: Abt Associates Inc; 2003. Available at: http://www.mass.gov/Eoehhs2/docs/dph/tobacco_control/abt_8th_report.pdf. Accessed September 24, 2008.
5. *Behavioral Risk Factor Surveillance System Survey Data*. Atlanta, GA: Centers for Disease Control and Prevention; 2002.
6. Census 2000. Your Gateway to Census 2000: U.S. Census Bureau. Available at <http://www.census.gov/main/www/cen2000.html>. Accessed December 24, 2008.
7. MassGIS. The Massachusetts Geographic Information System (Mass GIS): Office of Geographic and Environmental Information, Commonwealth of Massachusetts. Available at: <http://www.mass.gov/mgis>. Accessed December 24, 2008.
8. Malec D, Sedransk J, Moriarity CL, LeClere FB. Small area inference for binary variables in the National Health Interview Survey. *J Am Stat Assoc*. 1997;92(439):815–826.
9. Malec D, Davis WW, Cao X. Model-based small area estimates of overweight prevalence using sample selection adjustment. *Stat Med*. 1999;18(23):3189–3200.
10. Jia H, Link M, Holt J, Mokdad AH, Li L, Levy PS. Monitoring county-level vaccination coverage during the 2004–2005 influenza season. *Am J Prev Med*. 2006;31(4):275–280.
11. 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):931–937.
12. Giovino G. The tobacco epidemic in the United States. *Am J Prev Med*. 2007 Dec;33(6 Suppl):S318–26.
13. Peterson NA, Lowe JB, Reid RJ. Tobacco outlet density, cigarette smoking prevalence, and demographics at the county level of analysis. *Subst Use Misuse*. 2005;40(11):1627–1635.
14. Schneider JE, Reid RJ, Peterson NA, Lowe JB, Hughey J. Tobacco outlet density and demographics at the tract level of analysis in Iowa: implications for environmentally based prevention initiatives. *Prev Sci*. 2005;6(4):319–325.
15. Crawford CAG, Young LJ. A spatial view of the ecological inference problem. In: King G, Rosen O, Tanner MA, eds. *Ecological Inference*. 1st ed. New York, NY: Cambridge University Press; 2004:233–244.
16. Skeer M, George S, Hamilton WL, Cheng DM, Siegel M. Town-level characteristics and smoking policy adoption in Massachusetts: are local restaurant smoking regulations fostering disparities in health protection? *Am J Public Health*. 2004;94:286–292.
17. Ross CE. Walking, exercising, and smoking: does neighborhood matter? *Soc Sci Med*. 2000;51(2):265–274.
18. Tseng M, Yeatts K, Millikan R, Newman B. Area-level characteristics and smoking in women. *Am J Public Health*. 2001;91:1847–1850.
19. Reijneveld SA. The impact of individual and area characteristics on urban socioeconomic differences in health and smoking. *Int J Epidemiol*. 1998;27(1):33–40.
20. Kleinschmidt I, Hills M, Elliott P. Smoking behaviour can be predicted by neighbourhood deprivation measures. *J Epidemiol Community Health*. 1995;49(Suppl 2):S72–S77.
21. Duncan C, Jones K, Moon G. Smoking and deprivation: are there neighbourhood effects? *Soc Sci Med*. 1999;48(4):497–505.
22. Sundquist J, Malmstrom M, Johansson SE. Cardiovascular risk factors and the neighbourhood environment: a multilevel analysis. *Int J Epidemiol*. 1999;28(5):841–845.
23. Cubbin C, Hadden WC, Winkleby MA. Neighbourhood context and cardiovascular disease risk factors: the contribution of material deprivation. *Ethn Dis*. 2001;11(4):687–700.
24. Diez Roux AV, Merkin SS, Hannan P, Jacobs DR, Kiefe CI. Area characteristics, individual-level socioeconomic indicators, and smoking in young adults: the coronary artery disease risk development in

young adults study. *Am J Epidemiol*. 2003;157(4):315–326.

25. Shohaimi S, Luben R, Wareham N, et al. Residential area deprivation predicts smoking habit independently of individual educational level and occupational social class. A cross sectional study in the Norfolk cohort of the European Investigation into Cancer (EPIC-Norfolk). *J Epidemiol Community Health*. 2003;57(4):270–276.

26. Chuang YC, Cubbin C, Ahn D, Winkleby MA. Effects of neighbourhood socioeconomic status and convenience store concentration on individual level smoking. *J Epidemiol Community Health*. 2005;59(7):568–573.

27. Hyland A, Travers MJ, Cummings KM, Bauer J, Alford T, Wieczorek WF. Tobacco outlet density and demographics in Erie County, New York. *Am J Public Health*. 2003;93:1075–1076.

28. Leatherdale ST, Strath JM. Tobacco retailer density surrounding schools and cigarette access behaviors among underage smoking students. *Ann Behav Med*. 2007;33(1):105–111.

29. Novak SP, Reardon SF, Raudenbush SW, Buka SL. Retail tobacco outlet density and youth cigarette smoking: a propensity-modeling approach. *Am J Public Health*. 2006;96:670–676.

30. Reid RJ, Peterson NA, Lowe JB, Hughey J. Tobacco outlet density and smoking prevalence: does racial concentration matter? *Drugs Educ Prev Policy*. 2005;12(3):233–238.