BMJ Open

Geospatial patterns of human papillomavirus vaccine uptake in Minnesota

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-008617
Article Type:	Research
Date Submitted by the Author:	28-Apr-2015
Complete List of Authors:	Nelson, Erik; Saint Louis University, Epidemiology Hughes, John; University of Minnesota, Division of Biostatistics Oakes, J.; University of Minnesota, Division of Epidemiology and Community Health Pankow, James; University of Minnesota, Division of Epidemiology and Community Health Kulasingam, Shalini ; University of Minnesota, Division of Epidemiology and Community Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Sexual health, Research methods, Infectious diseases
Keywords:	EPIDEMIOLOGY, Epidemiology < INFECTIOUS DISEASES, Epidemiology < ONCOLOGY, PUBLIC HEALTH

Title: Geospatial Patterns of Human Papillomavirus Vaccine Uptake in Minnesota

Author Names and Affiliations:

Erik J. Nelson, PhD, MPH (Corresponding Author)
Department of Epidemiology, College for Public Health and Social Justice, Saint Louis University, St. Louis, MO

John Hughes, PhD
Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN

J. Michael Oakes, PhD
Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

James S. Pankow, PhD, MPH
Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

Shalini L. Kulasingam, PhD
Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

Corresponding Author:
Erik J. Nelson, PhD, MPH
College for Public Health and Social Justice
Saint Louis University
3545 Lafayette Avenue, Room 472
Tel: (314) 977-4562
Email: nelsonej@slu.edu

Word Count: 2,818

ABSTRACT

Background: Research describing the geographic variability in human papillomavirus (HPV) vaccination uptake at the state or county level is limited, has relied on data collected from large national surveillance programs and has, to date, not accounted for spatial autocorrelation. This study aimed to determine geographic variation in vaccine uptake using ZIP code level data, and to identify factors associated with vaccination while accounting for spatial autocorrelation.

Methods: Data on HPV vaccination were collected for 760 individuals nested within 99 ZIP codes surrounding the downtown area of Minneapolis, Minnesota. Proper conditional autoregressive (CAR) models were used to identify factors associated with receipt of HPV vaccination.

Results: In all, 46.2% of participants had received ≥1 dose of HPV vaccine (67.7% of women and 13.0% of men). HPV vaccination was found to exhibit strong spatial dependence ($\hat{\rho} = 0.9951$). Accounting for spatial dependence, older age (OR = 0.76, 95% CI = 0.70-0.83) and male gender (OR=0.04, 95% CI = 0.03-0.07) were negatively associated with vaccination, while liberal political preferences (OR = 4.31, 95% CI = 2.32-8.01), and college education (OR = 2.58, 95% CI = 1.14-5.83) were found to be positively associated with HPV vaccination.

Conclusions: HPV vaccination exhibited strong spatial dependence, indicating that spatial statistical models are needed to accurately identify and estimate factors associated with HPV vaccine uptake across geographic regions. This study also underscores the need for more detailed data collected at local levels (e.g., ZIP code), as patterns of HPV vaccine receipt were found to differ significantly from aggregated state and national patterns.

INTRODUCTION

Human papillomavirus (HPV) is the most common sexually transmitted infection in the U.S.,[1] and is the necessary cause of cervical cancer.[2] HPV infections are also associated with other cancers (e.g. anogenital and oropharyngeal) as well as genital warts.[3, 4] Since mid-2006, the Advisory Committee on Immunization Practices (ACIP) has recommended routine vaccination of adolescent girls aged 11 or 12 years with the three-dose HPV vaccine series.[5] In October 2011, the ACIP extended their recommendation of the quadrivalent vaccine to include boys aged 11 or 12 years old.[6, 7] The ACIP also recommends catch-up vaccination for those aged 13 to 26 years. However, HPV vaccination uptake has been far lower than expected, with only 53.8% of girls and 20.8% of boys aged 13-17 years and 34.5% of women and 2.3% of men aged 19-26 years receiving at least one dose of the vaccine as of 2012.[8, 9] Despite lower than anticipated vaccine uptake, recently published HPV vaccine serosurvey results show significant reductions in HPV prevalence, and reductions in HPV-associated cancer incidence of approximately 70% are predicted in the coming decades.[10-14]

Initiation of the HPV vaccine (i.e. receiving at least one dose) has been shown to be higher among minority adolescent girls; however, completion of the three-dose series is substantially lower among blacks and Hispanics compared to whites.[15] Although male vaccination data are very limited, racial and income differences have also been observed among adolescent boys.[16] Disparities in receipt of the HPV vaccine have also been found to be associated with insurance covering the costs of the vaccine, clinical provider characteristics (e.g., age of physician, pediatricians, and physicians with a private medical practice), and parental perceptions of the HPV vaccine.[16-22]

Research on the geographic variability of HPV vaccination is limited, and has relied on data collected from large national surveillance programs to estimate uptake at the state or county

levels.[23-25] These national data on geographic variation in HPV vaccine uptake may mask a considerable amount of variability at the local (e.g., county, census tract, or ZIP code) level. Further, a major limitation of these geographic studies is that they do not account for the areal units from which geographically-defined data are collected, commonly referred to as the spatial structure of the data. Data collected in this manner typically exhibit spatial dependence (also referred to as spatial autocorrelation), with observations from areal units close together tending to have similar values.[26] Although a proportion of spatial dependence may be modeled by including known covariate risk factors (i.e., age, race, sex) in a traditional (non-spatial) regression model, it is common for spatial structure to not be accounted for and to remain in the residuals even after accounting for these covariate effects.[26] For example, one study noted several individual factors that were associated with receipt of HPV vaccination, including geographic region of residence, however they only used a categorical variable to account for geographic differences in uptake.[27] Another study that analyzed geographic variation in HPV vaccine uptake used a weighting scheme to account for dependence between study participants, but ignored the spatial dependence of respondents in neighboring geographic regions.[23] Thus, these studies inherently assume that factors associated with HPV vaccine uptake are homogeneous across areal units such as states or counties. Documenting geographic variation in vaccine disparities at local levels may help to identify specific areas with the largest disparities in HPV vaccine uptake (after accounting for spatial dependence) thereby informing outreach efforts, and may also provide new hypotheses regarding the underlying determinants of geographic patterns in uptake.

The objective of this study was to use HPV vaccination data measured at the ZIP code level to identify geographic variation in vaccine uptake, and to identify factors associated with the receipt of HPV vaccination while accounting for spatial dependence.

METHODS

Data

This study utilized data collected on 1,003 participants from the Survey of Minnesotans About Screening and HPV (SMASH) study, which is a cross-sectional study of English-speaking men and women aged 18-30 years from the Twin Cities Metropolitan Area of Minnesota, and has been described elsewhere.[28] Briefly, from November 2012 to January 2013, targeted advertisements were displayed on the social networking site Facebook™ to men and women who met the study eligibility criteria (as specified in their user profiles). Men and women who clicked on a study advertisement were redirected to the secured SMASH study website and invited to participate in an online survey. After providing consent, participants were asked to answer questions regarding HPV vaccination, cancer screening, and barriers/intentions regarding receipt of either.

The response to the question "*Have you ever received an HPV vaccine?"* was used as the current study's outcome variable for HPV vaccination status. Individuals (n=128) who responded *don't know*, refused, or who did not respond to this question were excluded from the study. Similarly, individuals who did not report their ZIP code (n=3), or who reported a ZIP code outside of the predetermined 25-mile radius of downtown Minneapolis, Minnesota (n=112) were excluded from the study in order to focus on this diverse metropolitan population. The resulting study sample consisted of 760 (75.8% of total enrolled) men and women nested within 99 ZIP codes within downtown Minneapolis (see Figure 1).

Spatial Data Analysis

We tested for spatial autocorrelation in the crude HPV vaccination uptake rates using choropleth maps and Moran's I.[29] Positive (negative) values of I indicate positive (negative) spatial correlation, meaning that nearby ZIP codes tend to exhibit similar (dissimilar) HPV vaccine uptake rates. The spatial adjacency of the data were defined in three different ways: rook contiguity, queen contiguity, and using the 5 nearest neighbors. Model results did not vary substantially by the neighborhood definition; therefore the queen contiguity structure was selected for the subsequent analyses.

Spatially dependent data violate the independence assumption required for generalized linear models. As such, ignoring the dependence of spatial data can lead to an underestimation of standard errors, resulting in overly narrow confidence interval estimates and, consequently, incorrect statistical inference.[30] To account for residual dependence the linear predictor can be augmented with a spatial random effect, as part of a Bayesian hierarchical model.[31] These random effects typically take the form of a conditional autoregression (CAR), which introduces spatial dependence through the adjacency structure of areal units.[31] CAR models are generally applied in a Bayesian setting, where inference is based on Markov-chain Monte Carlo (MCMC) simulation.[32]

To accommodate the potential spatial dependence of HPV vaccination, we implemented a spatial logistic regression model using ZIP code as the areal unit of analysis. To accomplish this, assume Y_i is the number of respondents who were vaccinated against HPV out of the total N_i sampled in each ZIP code j. The outcome can be modeled as a binomial response $Y_{ij} \sim bin(p_{ij}, N_{ij})$ such that p_{ij} is the true vaccine uptake proportion of individual i within a selected ZIP code j. The proportions were smoothed using the following model,

$$\text{logit}(p_{ij}) = \alpha + \beta X_{ij} + s_j \qquad (1)$$

where α is an intercept, which is interpreted as an overall log-odds coverage for all areas; β are the effects of the covariates X_{ij} in the model; and the s_j are spatially dependent random effects, such that neighboring areas have a similar vaccine uptake proportion. The parameter ρ (Rho) reflects the spatial dependence inherent in the data, measuring the average influence of a given ZIP code on neighboring ZIP codes.[31, 33, 34] Including information from neighboring ZIP codes to further inform the estimate for each ZIP code, even when the sample size is small, creates sufficient statistical power to generate reliable estimates.[35] This is achieved by assuming a proper CAR prior, defined as $N(s_{j|k}, 1/\tau_s m_j)$, where $s_{i|j}$ is the pooled mean of area *j*, based on the adjacent areas *k*, and m_j are the number of ZIP codes neighboring *j*, while τ_s is the precision that controls the amount of smoothing.[36, 37] By convention, the intercept and regression coefficients were assigned a conservative normal prior with a mean of 0 and a standard deviation of 1,000,000. Estimation of the model parameters was carried out with MCMC simulation techniques that were implemented in R version 3.0.1 (R Development Core Team, 2014). Model convergence was monitored using a Monte Carlo standard error threshold of 0.1.[38] For this analysis, a total of 1,000,000 posterior samples were generated.

All statistical models included *a priori* factors potentially associated with HPV vaccine uptake, including sex (categorized as male or female), age (mean-centered), race (categorized as white, African American, American Indian/Alaska native, Asian, or other), ethnicity (categorized as Hispanic or non-Hispanic), educational attainment (categorized as some high school, high school graduate, some college or technical school, college graduate, or graduate school), sexual orientation (categorized as heterosexual, homosexual/gay/lesbian, or bisexual), and political views (categorized as very conservative or conservative, moderate, liberal, or very liberal). Initially, the model was fit maintaining all the variables. The final model retained all covariates

that were statistically significant at $p < 0.05$. Odds ratios and the associated 95% credible intervals are presented. The random effect terms can be interpreted as the effect of ZIP code on HPV vaccination uptake for each individual.

RESULTS

Characteristics of the study sample are presented in Table 1. In all, 46.2% of participants had received at least one dose of HPV vaccine, with 67.7% of women reporting having been vaccinated compared to 13.0% of men. Of those who initiated the vaccine series, 71.1% completed the entire three-dose series (79.6% of women and 26.3% of men). Participants who had been vaccinated against HPV (i.e. received ≥1 dose of the vaccine) were younger (30.1% of those ≥25 years were vaccinated compared to 69.9% of those <25 years). Vaccine receipt was lower among those who identified themselves as politically "conservative" or "very conservative" as opposed to politically "liberal" (24.6% compared to 53.3%).

Table 1. Characteristics of study participants by HPV vaccination status.

	Vaccinated		**Not Vaccinated**		**Total**	
	(n=351)		**(n=409)**		**(n=760)**	
	N	%	N	%	N	%
Age, in years						
18-20	86	51.8	80	48.2	166	21.8
21-25	209	51.2	199	48.8	408	53.7
26-30	56	30.1	130	69.9	186	24.5
Gender						
Female	312	67.7	149	32.3	461	60.7
Male	39	13.0	260	87.0	299	39.3
Race						
White	298	46.3	346	53.7	644	84.7
Black	22	61.1	14	38.9	36	4.7
Am. Indian/AL native	4	50.0	4	50.0	8	1.1
Asian	15	34.1	29	65.9	44	5.8
Other	12	42.9	16	57.1	28	3.7

Hispanic/Latino						
Yes	13	48.1	14	51.9	27	3.6
No	336	46.2	391	53.8	727	96.4
Education						
Some High School	4	1.1	3	0.7	7	0.9
High School Graduate	19	5.4	25	6.1	44	5.8
Some College or Tech. School	135	38.5	151	36.9	286	37.6
College Graduate	152	43.3	175	42.8	327	43.0
Graduate School	41	11.7	55	13.4	96	12.6
Sexual Orientation						
Heterosexual	311	45.7	370	54.3	681	89.7
Homosexual, gay, or lesbian	12	36.4	21	63.6	33	4.3
Bisexual	21	65.6	11	34.4	32	4.2
Don't know/Refused	7	53.8	6	46.2	13	1.7
Political Views						
Very Conservative	1	4.5	21	95.5	22	2.9
Conservative	28	29.2	68	70.8	96	12.6
Moderate	103	44.6	128	55.4	231	30.4
Liberal	154	52.0	142	48.0	296	38.9
Very Liberal	65	56.5	50	43.5	115	15.1

HPV indicates human papillomavirus

[a]Other indicates Native Hawaiian or Pacific Islander, more than one race, or a response of "other."

HPV vaccination was found to exhibit strong spatial dependence ($\hat{\rho} = 0.9951$). The CAR model also successfully converged, as the maximum Monte Carlo standard error was 0.028 (which was below our threshold of 0.1), indicating that a sufficient number of posterior samples were generated for the estimates to stabilize. Estimates for the best-fitting CAR model are shown in Table 2. After accounting for spatial dependence using the CAR model, age, sex, education, and political preferences remained significantly associated with HPV vaccine receipt. Specifically, older age (OR = 0.77 per year, 95% CI = 0.72-0.83) and being male (OR=0.03, 95% CI = 0.02-0.06) were associated with a decreased odds of HPV vaccine receipt. Higher educational attainment (referent to receiving some high school or high school graduates) was associated with an increased odds of HPV vaccine receipt (some college OR = 2.58, 95% CI =

1.14-5.83; college graduate OR = 3.93, 95% CI = 1.66-9.30; graduate degree OR = 4.74, 95% CI = 1.71-13.17). Moderate and liberal political preferences (referent to very conservative and conservative preferences) were also associated with an increased odds of HPV vaccine receipt (moderate OR = 3.24, 95% CI = 1.62-6.49; liberal OR = 5.32, 95% CI = 2.68-10.58). Race was not found to be significantly associated with HPV vaccine uptake. For comparison, odds ratios (and corresponding 95% confidence intervals) from a traditional logistic regression model that does not account for spatial dependence were also estimated and are also presented in Table 2. Compared to the traditional logistic model, estimates from the CAR model were greater in magnitude for all covariates. Of note, in the traditional logistic regression model, having received some college education was not a statistically significant factor but became significant in the CAR model (traditional OR = 1.88, 95% CI = 0.90-3.93; spatial CAR OR = 2.58, 95% CI = 1.14-5.83).

Table 2. Odds ratio estimates for factors associated with HPV vaccination from traditional logistic regression and spatial CAR models.

	Traditional Logistic Model		**Spatial CAR Model**	
	Odds Ratio	95% CI[a]	Odds Ratio	95% CI[b]
Age[c]	0.84	(0.78-0.90)	0.76	(0.70-0.83)
Sex				
Female	1	(referent)	1	(referent)
Male	0.07	(0.05-0.11)	0.04	(0.03-0.07)
Political Views				
Conservative[d]	1	(referent)	1	(referent)
Moderate	2.34	(1.30-4.19)	3.06	(1.61-5.81)
Liberal	2.76	(1.57-4.85)	4.31	(2.32-8.01)
Very Liberal	3.42	(1.76-6.62)	4.82	(2.34-9.94)
Education				
High School[e]	1	(referent)	1	(referent)
Some College	1.88	(0.90-3.93)	2.58	(1.14-5.83)
College Graduate	2.51	(1.15-5.45)	3.93	(1.66-9.30)
Graduate Degree	2.59	(1.03-6.52)	4.74	(1.71-13.17)

HPV indicates human papillomavirus
[a]95% Confidence Interval
[b]95% Credible Interval
[c]Age is centered at the mean (23.24 years old)
[d]Referent group consists of "conservative" and "very conservative" responses
[e]Referent group consists of "some high school" and "high school graduate"

DISCUSSION

In this study, HPV vaccination was found to exhibit strong spatial dependence, indicating that spatial statistical models are needed to accurately identify and estimate factors associated with HPV vaccine uptake. As a result, ignoring this spatial dependence can lead to biased point estimates and overly narrow credible intervals. Consistent with other studies, younger age, female gender, higher education, and political views were found to be significantly associated with HPV vaccination (after accounting for spatial dependence).[21, 27, 39, 40] The associations of age and sex with HPV vaccine receipt can be attributed, in part, to the evolving ACIP recommendations, as they were first recommended for use in young girls and were later expanded to include young boys. Conservative political views have also been found to be associated with decreased knowledge of HPV, lower perceived risk of infection with HPV, and stronger views against premarital sex.[41]

However, contrary to other studies that have not accounted for spatial dependence, this study found that race was not significantly associated with HPV vaccination.[21, 27, 39, 40, 42, 43] Racial disparities (and other disparities) have been shown to be pronounced in some areas, while less evident (or absent) in other areas.[44-46] Although the existence of these disparities is well documented, the overall average effects (i.e., national level data) can mask variation across local areas.[47, 48] For example, in a traditional regression analysis where minority girls live in regions with systematically different rates of HPV vaccine uptake, and the region is not

controlled for, one could erroneously conclude that racial "disparities" exist when in fact *where* people live is the significant factor associated with vaccination. Thus, ignoring geography (i.e., the spatial dependence of the data) may lead to incorrect inference. Previous studies that have attempted to describe geographic variation in HPV vaccine uptake have either ignored spatial dependence completely or have not correctly accounted for it using spatial statistical models. These studies may have incorrectly concluded that covariates such as race are significantly associated with HPV vaccine receipt when, in fact, these conclusions are likely to be erroneous because they are based on models that did not account for spatial dependence. As our analysis shows, using models that account for spatial dependence may greatly improve the identification of independent factors that are truly associated with HPV vaccination (as opposed to spatially confounded covariates), particularly when analyzing data from varying geographic locations.

Previous studies have shown that HPV vaccination uptake exhibits significant geographic variability.[23-25, 27] HPV vaccine policies, availability, costs, financial assistance, and availability of education materials to promote uptake collectively contribute to this variability, as they vary widely across and within states.[49] As a result, variation at state levels may not reflect the variation in HPV vaccine uptake occurring at a more local level. However, a more refined level of analysis was not possible in these studies because of the sparseness of data at the county and ZIP code level, which is in part attributable to national surveys aggregating or suppressing responses due to participant identification concerns. One strength of this study is that ZIP code level data were available to conduct a more detailed spatial analysis.

The proportion of all adults in this study who had been vaccinated against HPV (i.e. received at least one dose of the HPV vaccine) was 46.2% (67.7% for women and 13.0% for men). These estimates are much higher than the HPV vaccine coverage estimates from the 2012

National Health Interview Survey (NHIS) for women (34.5%) and men (2.3%) aged 19 to 26.[38] Although the results for women are more similar to those obtained from the National Immunization Survey – Teen for girls (53.8%), the estimate for men is much lower than the NIS-Teen estimate for boys (20.8%) aged 13 to 17 who received at least one dose of HPV vaccine in 2012.[39] Although the differences in the observed rates may be partially explained by the sampling frame, response rates, or the small number of eligible respondents who received the HPV vaccine question series in the national surveys, the estimates of HPV vaccine uptake are noticeably different from the current study.

There are several limitations to this study. First, all study measures were self-reported by persons over the Internet and may be subject to under or over-reporting. However, recent studies have shown recall of HPV vaccination status to be accurate.[50] In addition, Internet-based studies have shown increased self-disclosure and reporting with online surveys, which may reduce potential response biases (e.g. interviewer bias or social desirability).[51, 52] Second, analyses by race may have been underpowered due to small numbers, however, the distribution of racial groups was proportionate to estimates from the U.S. Census for the study area.[28] Third, the spatial analyses were conducted at the ZIP code level and assume a common ZIP code level effect, so within-ZIP code differences may be masked. However, to our knowledge, this is only the second study to examine HPV vaccination at such a small areal unit.[48] Finally, this study utilizes cross-sectional data and temporal effects cannot be established.

In conclusion, the results from this study demonstrate that more detailed and local assessments of HPV vaccine uptake that account for spatial dependence are necessary as ZIP code level patterns differ significantly from aggregated state and national patterns. Future work

is needed to further pinpoint areas with the greatest disparities and how to then access these populations to improve vaccine uptake.

Figure 1 Caption

The spatial distribution of 760 survey responses across the Twin Cities Metropolitan Area of Minnesota.

Abbreviations

HPV: human papillomavirus; CAR: conditional autoregression.

Author Contributions

SK and EJN conceived and designed the study. EJN also conducted the data collection for this study and drafted the manuscript. JH, JP, and JMO assisted in the survey design, supervised the statistical analysis, and assisted in reviewing/revising the manuscript. SK also provided contributions to the concept and analytic approach for the article, and oversaw the analysis, interpretation, and reviewing/revising of manuscript. All authors read and approved the final manuscript.

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was supported by the J.B. Hawley Student Research Award from the University of Minnesota School of Public Health and by the Minnesota Medical Foundation through Grant 4120-9227-12.

Data Sharing Statement

SMASH study data have not been published but may be made available for reasonable requests.

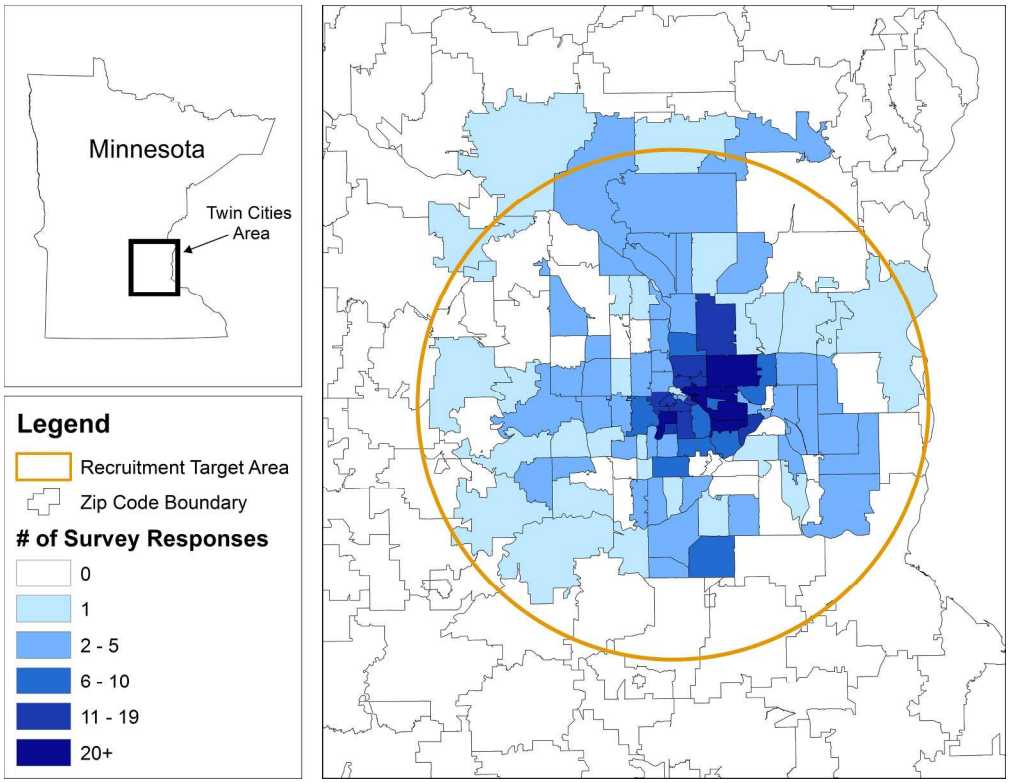
REFERENCES

1. Weinstock, H., S. Berman, and W. Cates, Jr., *Sexually Transmitted disease among American youth: incidence and prevalence estimates, 2000.* Perspect Sex Reprod Health, 2004. **36**(1): p. 6-10. PMID:
2. Walboomers, J.M., et al., *Human papillomavirus is a necessary cause of invasive cervical cancer worldwide.* J Pathol, 1999. **189**(1): p. 12-9. PMID: 10451482.
3. Centers for Disease Control and Prevention, *Human papillomavirus-associated cancers - United States, 2004-2008.* MMWR Morb Mortal Wkly Rep, 2012. **61**: p. 258-61. PMID: 22513527.
4. Saraiya, M., et al., *Toward using National Cancer Surveillance data for preventing and controlling cervical and other human papillomavirus-associated cancers in the US.* Cancer, 2008. **113**(10 Suppl): p. 2837-40. PMID: 18980202.
5. Markowitz, L.E., et al., *Quadrivalent Human Papillomavirus Vaccine: Recommendations of the Advisory Committee on Immunization Practices (ACIP).* MMWR Recomm Rep, 2007. **56**(RR-2): p. 1-24. PMID: 17380109.
6. Centers for Disease Control and Prevention, *FDA licensure of bivalent human papillomavirus vaccine (HPV2, Cervarix) for use in females and updated HPV vaccination recommendations from the Advisory Committee on Immunization Practices (ACIP).* MMWR Morb Mortal Wkly Rep, 2010. **59**(20): p. 626-9. PMID: 20508593.
7. Centers for Disease Control and Prevention, *Recommended adult immunization schedule - United States, 2012.* MMWR Morb Mortal Wkly Rep, 2012. **61**(4). PMID:
8. Williams, W.W., et al., *Noninfluenza vaccination coverage among adults - United States, 2012.* MMWR Morb Mortal Wkly Rep, 2014. **63**(5): p. 95-102. PMID: 24500288.
9. *Human papillomavirus vaccination coverage among adolescent girls, 2007-2012, and postlicensure vaccine safety monitoring, 2006-2013 - United States.* MMWR Morb Mortal Wkly Rep, 2013. **62**(29): p. 591-5. PMID: 23884346.
10. Giuliano, A.R., et al., *Efficacy of quadrivalent HPV vaccine against HPV Infection and disease in males.* N Engl J Med, 2011. **364**(5): p. 401-11. PMID: 21288094.

11. Joura, E.A., et al., *Efficacy of a quadrivalent prophylactic human papillomavirus (types 6, 11, 16, and 18) L1 virus-like-particle vaccine against high-grade vulval and vaginal lesions: a combined analysis of three randomised clinical trials.* Lancet, 2007. **369**(9574): p. 1693-702. PMID: 17512854.
12. Paavonen, J., et al., *Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women.* Lancet, 2009. **374**(9686): p. 301-14. PMID: 19586656.
13. Markowitz, L.E., et al., *Reduction in human papillomavirus (HPV) prevalence among young women following HPV vaccine introduction in the United States, National Health and Nutrition Examination Surveys, 2003-2010.* J Infect Dis, 2013. **208**(3): p. 385-93. PMID: 23785124.
14. Kim, J.J. and S.J. Goldie, *Health and economic implications of HPV vaccination in the United States.* N Engl J Med, 2008. **359**(8): p. 821-32. PMID: 18716299.
15. Niccolai, L.M., N.R. Mehta, and J.L. Hadler, *Racial/Ethnic and poverty disparities in human papillomavirus vaccination completion.* Am J Prev Med, 2011. **41**(4): p. 428-33. PMID: 21961471.
16. Jeudin, P., et al., *Race, Ethnicity, and Income Factors Impacting Human Papillomavirus Vaccination rates.* Clin Ther, 2014. **36**(1): p. 24-37. PMID: 24417783.
17. Vadaparampil, S.T., et al., *Provider factors associated with disparities in human papillomavirus vaccination among low-income 9- to 17-year-old girls.* Cancer, 2013. **119**(3): p. 621-8. PMID: 23341308.
18. Bednarczyk, R.A., et al., *Health disparities in human papillomavirus vaccine coverage: trends analysis from the National Immunization Survey-Teen, 2008-2011.* Clin Infect Dis, 2014. **58**(2): p. 238-41. PMID: 24162745.
19. Bednarczyk, R.A., et al., *Human papillomavirus vaccine uptake and barriers: association with perceived risk, actual risk and race/ethnicity among female students at a New York State university, 2010.* Vaccine, 2011. **29**(17): p. 3138-43. PMID: 21376797.
20. Chao, C., et al., *Correlates for human papillomavirus vaccination of adolescent girls and young women in a managed care organization.* Am J Epidemiol, 2010. **171**(3): p. 357-67. PMID: 20047978.
21. Liddon, N.C., J.S. Leichliter, and L.E. Markowitz, *Human papillomavirus vaccine and sexual behavior among adolescent and young women.* Am J Prev Med, 2012. **42**(1): p. 44-52. PMID: 22176845.
22. Hertweck, S.P., et al., *Health care decision making by mothers for their adolescent daughters regarding the quadrivalent HPV vaccine.* J Pediatr Adolesc Gynecol, 2013. **26**(2): p. 96-101. PMID: 23518189.
23. Wei, F., P.C. Moore, and A.L. Green, *Geographic variability in human papillomavirus vaccination among U.S. young women.* Am J Prev Med, 2013. **44**(2): p. 154-7. PMID: 23332332.
24. Pruitt, S.L. and M. Schootman, *Geographic disparity, area poverty, and human papillomavirus vaccination.* Am J Prev Med, 2010. **38**(5): p. 525-33. PMID: 20409501.
25. Eberth, J.M., et al., *County-level estimates of human papillomavirus vaccine coverage among young adult women in Texas, 2008.* Tex Public Health J, 2013. **65**(1): p. 37-40. PMID: 24466565.

26. Lee, D., *CARBayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors.* Journal of Statistical Software, 2013. **55**(13). PMID:
27. Rahman, M., C.J. McGrath, and A.B. Berenson, *Geographic variation in human papillomavirus vaccination uptake among 13-17 year old adolescent girls in the United States.* Vaccine, 2014. **32**(21): p. 2394-8. PMID: 24637175.
28. Nelson, E.J., et al., *Estimation of geographic variation in human papillomavirus vaccine uptake in men and women: an online survey using facebook recruitment.* J Med Internet Res, 2014. **16**(9): p. e198. PMID: 25231937.
29. Moran, P.A., *Notes on continuous stochastic phenomena.* Biometrika, 1950. **37**(1-2): p. 17-23. PMID: 15420245.
30. Waller, L.A. and C.A. Gotway, *Applied Spatial Statistics for Public Health Data*. 2004, Hoboken, NJ: John Wiley & Sons.
31. Besag J, Y.J., Mollie A, *Bayesian Image Restoration with Two Applications in Spatial Statistics.* The Annals of the Institute of Statistics and Mathematics, 1991. **43**(1): p. 1-59. PMID:
32. Lee, D., *A comparison of conditional autoregressive models used in Bayesian disease mapping.* Spat Spatiotemporal Epidemiol, 2011. **2**(2): p. 79-89. PMID: 22749587.
33. Leroux, B., X. Lei, and N. Breslow, *Estimation of disease rates in small areas: a new mixed model for spatial dependence*, in *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, M. Halloran and D. Berry, Editors. 1999, Springer-Verlag: New York. p. 135-178.
34. Stern, H. and N. Cressie, *Inference for extremes in disease mapping*, in *Disease Mapping and Risk Assessment for Public Health*, A. Lawson, et al., Editors. 1999, John Wiley & Sons. p. 63-84.
35. Carlin, B.P. and T.A. Louis, *Bayes and Empirical Bayes methods for data analysis*. 2nd Edition ed. 2000, Boca Raton: Chapman & Hall/CRC.
36. Assun\c cão, R. and E. Krainski, *Neighborhood dependence in Bayesian spatial models.* Biometrical Journal, 2009. **51**(5): p. 851-869. PMID:
37. Besag, J. and C. Kooperberg, *On conditional and intrinsic autoregression.* Biometrika, 1995. **82**(4): p. 733-746. PMID:
38. Flegal, J.M., M. Haran, and G.L. Jones, *Markov chain Monte Carlo: Can we trust the third significant figure?* Statistical Science, 2008. **23**(2): p. 250-260. PMID: ISI:000259275400006.
39. Laz, T.H., M. Rahman, and A.B. Berenson, *Human papillomavirus vaccine uptake among 18- to 26-year-old women in the United States: National Health Interview Survey, 2010.* Cancer, 2013. **119**(7): p. 1386-92. PMID: 23508594.
40. Anhang Price, R., et al., *Use of human papillomavirus vaccines among young adult women in the United States: an analysis of the 2008 National Health Interview Survey.* Cancer, 2011. **117**(24): p. 5560-8. PMID: 21732336.
41. Gerend, M.A. and J.E. Shepherd, *Correlates of HPV knowledge in the era of HPV vaccination: a study of unvaccinated young adult women.* Women Health, 2011. **51**(1): p. 25-40. PMID: 21391159.
42. Sadry, S.A., L.R. De Souza, and M.H. Yudin, *The impact of ethnicity on awareness and knowledge of and attitudes towards the human papillomavirus and vaccine among adult women.* J Obstet Gynaecol Can, 2013. **35**(11): p. 995-1003. PMID: 24246399.

43. Miller, M.K., et al., *Views on Human Papillomavirus Vaccination: A Mixed-Methods Study of Urban Youth.* J Community Health, 2014. PMID: 24664875.
44. Skinner, J., et al., *Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients.* N Engl J Med, 2003. **349**(14): p. 1350-9. PMID: 14523144.
45. Baicker, K., A. Chandra, and J.S. Skinner, *Geographic variation in health care and the problem of measuring racial disparities.* Perspect Biol Med, 2005. **48**(1 Suppl): p. S42-53. PMID: 15842086.
46. Chen, J., et al., *Racial differences in the use of cardiac catheterization after acute myocardial infarction.* N Engl J Med, 2001. **344**(19): p. 1443-9. PMID: 11346810.
47. Baicker, K., et al., *Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries.* Health Aff (Millwood), 2004. **Suppl Variation**: p. VAR33-44. PMID: 15471775.
48. Trogdon, J.G. and T. Ahn, *Geospatial patterns in human papillomavirus vaccination uptake: evidence from uninsured and publicly insured children in north Carolina.* Cancer Epidemiol Biomarkers Prev, 2015. **24**(3): p. 595-602. PMID: 25576528.
49. Katz, M.L., et al., *Human papillomavirus (HPV) vaccine availability, recommendations, cost, and policies among health departments in seven Appalachian states.* Vaccine, 2009. **27**(24): p. 3195-200. PMID: 19446191.
50. Ojha, R.P., et al., *The accuracy of human papillomavirus vaccination status based on adult proxy recall or household immunization records for adolescent females in the United States: results from the National Immunization Survey-Teen.* Ann Epidemiol, 2013. **23**(5): p. 281-5. PMID: 23453240.
51. Cantrell, M.A. and P. Lupinacci, *Methodological issues in online data collection.* Journal of Advanced Nursing, 2007. **60**(5): p. 544-549. PMID: ISI:000250498900010.
52. Rhodes, S.D., D.A. Bowie, and K.C. Hergenrather, *Collecting behavioural data using the world wide web: considerations for researchers.* Journal of Epidemiology and Community Health, 2003. **57**(1): p. 68-73. PMID: ISI:000180078500017.



The spatial distribution of 760 survey responses across the Twin Cities Metropolitan Area of Minnesota.
215x166mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
Outcome data	15*	Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses

Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Geospatial patterns of human papillomavirus vaccine uptake in Minnesota

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-008617.R1
Article Type:	Research
Date Submitted by the Author:	24-Jul-2015
Complete List of Authors:	Nelson, Erik; Saint Louis University, Epidemiology Hughes, John; University of Minnesota, Division of Biostatistics Oakes, J.; University of Minnesota, Division of Epidemiology and Community Health Pankow, James; University of Minnesota, Division of Epidemiology and Community Health Kulasingam, Shalini ; University of Minnesota, Division of Epidemiology and Community Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Sexual health, Research methods, Infectious diseases
Keywords:	EPIDEMIOLOGY, Epidemiology < INFECTIOUS DISEASES, Epidemiology < ONCOLOGY, PUBLIC HEALTH

Title: Geospatial Patterns of Human Papillomavirus Vaccine Uptake in Minnesota

Author Names and Affiliations:

Erik J. Nelson, PhD, MPH (Corresponding Author)
Department of Epidemiology, College for Public Health and Social Justice, Saint Louis University, St. Louis, MO

John Hughes, PhD
Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN

J. Michael Oakes, PhD
Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

James S. Pankow, PhD, MPH
Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

Shalini L. Kulasingam, PhD
Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

Corresponding Author:
Erik J. Nelson, PhD, MPH
College for Public Health and Social Justice
Saint Louis University
3545 Lafayette Avenue, Room 472
Tel: (314) 977-4562
Email: nelsonej@slu.edu

Word Count: 3,434

ABSTRACT

Objectives: To identify factors associated with human papillomavirus (HPV) vaccination and to determine the geographic distribution of vaccine uptake while accounting for spatial autocorrelation.

Design: This study is cross-sectional in design using data collected via the Internet from the Survey of Minnesotans About Screening and HPV study.

Setting and participants: The sample consists of 760 individuals aged 18-30 years nested within 99 ZIP codes surrounding the downtown area of Minneapolis, Minnesota.

Results: In all, 46.2% of participants had received ≥1 dose of HPV vaccine (67.7% of women and 13.0% of men). Prevalence of HPV vaccination was found to exhibit strong spatial dependence ($\hat{\rho} = 0.9951$) across ZIP codes. Accounting for spatial dependence, age (OR = 0.76, 95% CI = 0.70-0.83) and male gender (OR=0.04, 95% CI = 0.03-0.07) were negatively associated with vaccination, while liberal political preferences (OR = 4.31, 95% CI = 2.32-8.01), and college education (OR = 2.58, 95% CI = 1.14-5.83) were found to be positively associated with HPV vaccination.

Conclusions: Strong spatial dependence and heterogeneity of HPV vaccination prevalence was found across ZIP codes, indicating that spatial statistical models are needed to accurately identify and estimate factors associated with vaccine uptake across geographic units. This study also underscores the need for more detailed data collected at local levels (e.g., ZIP code), as patterns of HPV vaccine receipt were found to differ significantly from aggregated state and national patterns. Future work is needed to further pinpoint areas with the greatest disparities in HPV vaccination and how to then access these populations to improve vaccine uptake.

Strengths and limitations of this study

- This is the first study to identify factors associated with human papillomavirus (HPV) vaccination at the ZIP code level using statistical models that account for spatial dependence.
- Study strengths include the large representative sample of 18-30 year olds in Minneapolis, Minnesota, adjustment for factors known to be associated with HPV vaccination, and the use of robust spatial statistical models.
- This study reveals a gap between local estimation of HPV vaccination and estimates from large national surveillance programs.
- Potential limitations include the reliance on self-reported data collected via the Internet, selection bias, and the absence of information regarding study participants' age at vaccination and income.

INTRODUCTION

Human papillomavirus (HPV) is the most common sexually transmitted infection in the U.S.,[1] and is the necessary cause of cervical cancer.[2] HPV infections are also associated with other cancers (e.g. anogenital and oropharyngeal) as well as genital warts.[3, 4] Since mid-2006, the Advisory Committee on Immunization Practices (ACIP) has recommended routine vaccination of adolescent girls aged 11 or 12 years with the three-dose HPV vaccine series.[5] In October 2011, the ACIP extended their recommendation of the quadrivalent vaccine to include boys aged 11 or 12 years old.[6, 7] The ACIP also recommends catch-up vaccination for those aged 13 to 26 years. However, HPV vaccination uptake has been far lower than expected, with only 57.3% of girls and 34.6% of boys aged 13-17 years and 36.9% of women and 5.9% of men aged 19-26 years receiving at least one dose of the vaccine as of 2013.[8, 9] Despite lower than anticipated vaccine uptake, recently published HPV vaccine serosurvey results show significant reductions in HPV prevalence,[10-12] and reductions in HPV-associated cancer incidence of approximately 70% are predicted in the coming decades.[13, 14]

Initiation of the HPV vaccine (i.e., receiving at least one dose) has been shown to be higher among minority adolescent girls; however, completion of the three-dose series is substantially lower among blacks and Hispanics compared to whites.[15] Although male vaccination data are very limited, racial and income differences have also been observed among adolescent boys.[16] Disparities in receipt of the HPV vaccine have also been found to be associated with insurance covering the costs of the vaccine, clinical provider characteristics (e.g., age of physician, pediatricians, and physicians with a private medical practice), poverty, and parental perceptions of the HPV vaccine.[16-22]

Research on the geographic variability of HPV vaccination is limited, and has relied on data collected from large national surveillance programs to estimate uptake at the state or county

levels.[23-25] These national data on geographic variation in HPV vaccine uptake may mask a considerable amount of variability at the local (e.g., county, census tract, or ZIP code) level. Further, a major limitation of these geographic studies is that they do not account for the areal units from which geographically-defined data are collected, commonly referred to as the spatial structure of the data. Data collected in this manner typically exhibit spatial dependence (also referred to as spatial autocorrelation), with observations from areal units close together tending to have similar values.[26] Although a proportion of spatial dependence may be modeled by including known covariate risk factors (i.e., age, race, sex) in a traditional (non-spatial) regression model, it is common for spatial structure to not be accounted for and to remain in the residuals even after accounting for these covariate effects.[26] For example, one study noted several individual factors that were associated with receipt of HPV vaccination, including geographic region of residence, however they only used a categorical variable to account for geographic differences in uptake.[27] Another study that analyzed geographic variation in HPV vaccine uptake used a weighting scheme to account for dependence between study participants, but ignored the spatial dependence of respondents in neighboring geographic regions.[23] Thus, these studies inherently assume that factors associated with HPV vaccine uptake are homogeneous across areal units such as states or counties. Documenting geographic variation in vaccine disparities at local levels may help to identify specific areas with the largest disparities in HPV vaccine uptake (after accounting for spatial dependence) thereby informing outreach efforts, and may also provide new hypotheses regarding the underlying determinants of geographic patterns in uptake.

The objective of this study was to use HPV vaccination data measured at the ZIP code level to identify geographic variation in vaccine uptake, and to identify factors associated with the receipt of HPV vaccination while accounting for spatial dependence.

METHODS

Data

This study utilized data collected on 1,003 participants from the Survey of Minnesotans About Screening and HPV (SMASH) study, which is a cross-sectional study of English-speaking men and women aged 18-30 years from the Twin Cities Metropolitan Area of Minnesota, and has been described elsewhere.[28] Briefly, from November 2012 to January 2013, targeted advertisements were displayed on the social networking site Facebook™ to men and women who met the study eligibility criteria (as specified in their user profiles). Men and women who clicked on a study advertisement were redirected to the secured SMASH study website and invited to participate in an online survey. After providing consent, participants were asked to answer questions regarding HPV vaccination, cancer screening, and barriers/intentions regarding receipt of either.

The response to the question "*Have you ever received an HPV vaccine?"* was used as the current study's outcome variable for HPV vaccination status. Individuals (n=128) who responded *don't know*, refused, or who did not respond to this question were excluded from the study. Similarly, individuals who did not report their ZIP code (n=3), or who reported a ZIP code outside of the predetermined 25-mile radius of downtown Minneapolis, Minnesota (n=112) were excluded from the study in order to focus on this diverse metropolitan population. The resulting study sample consisted of 760 (75.8% of total enrolled) men and women nested within 99 ZIP codes within downtown Minneapolis (see Figure 1).

Spatial Data Analysis

We tested for spatial autocorrelation in the crude HPV vaccination uptake rates using choropleth maps and Moran's I.[29] Positive (negative) values of I indicate positive (negative) spatial correlation, meaning that nearby ZIP codes tend to exhibit similar (dissimilar) HPV vaccine uptake rates. The spatial adjacency of the data were defined in three different ways: rook contiguity, queen contiguity, and using the 5 nearest neighbors. Model results did not vary substantially by the neighborhood definition; therefore the queen contiguity structure was selected for the subsequent analyses.

Spatially dependent data violate the independence assumption required for generalized linear models. As such, ignoring the dependence of spatial data can lead to an underestimation of standard errors, resulting in overly narrow confidence interval estimates and, consequently, incorrect statistical inference.[30] To account for residual dependence the linear predictor can be augmented with a spatial random effect, as part of a Bayesian hierarchical model.[31] These random effects typically take the form of a conditional autoregression (CAR), which introduces spatial dependence through the adjacency structure of areal units.[31] CAR models are generally applied in a Bayesian setting, where inference is based on Markov-chain Monte Carlo (MCMC) simulation.[32]

To accommodate the potential spatial dependence of HPV vaccination, we implemented a spatial logistic regression model using ZIP code as the areal unit of analysis. To accomplish this, assume Y_i is the number of respondents who were vaccinated against HPV out of the total N_i sampled in each ZIP code j. The outcome can be modeled as a binomial response $Y_{ij} \sim bin(p_{ij}, N_{ij})$ such that p_{ij} is the true vaccine uptake proportion of individual i within a selected ZIP code j. The proportions were smoothed using the following model,

$$\text{logit}(p_{ij}) = \alpha + \beta X_{ij} + s_j \qquad (1)$$

where α is an intercept, which is interpreted as an overall log-odds coverage for all areas; β are the effects of the covariates X_{ij} in the model; and the s_j are spatially dependent random effects, such that neighboring areas have a similar vaccine uptake proportion. The parameter ρ (Rho) reflects the spatial dependence inherent in the data, measuring the average influence of a given ZIP code on neighboring ZIP codes.[31, 33, 34] Including information from neighboring ZIP codes to further inform the estimate for each ZIP code, even when the sample size is small, creates sufficient statistical power to generate reliable estimates.[35] This is achieved by assuming a proper CAR prior, defined as $N(s_{jlk}, 1/\tau_s m_j)$, where s_{ilj} is the pooled mean of area *j*, based on the adjacent areas *k*, and m_j are the number of ZIP codes neighboring *j*, while τ_s is the precision that controls the amount of smoothing.[36, 37] By convention, the intercept and regression coefficients were assigned a conservative normal prior with a mean of 0 and a standard deviation of 1,000,000. Estimation of the model parameters was carried out with MCMC simulation techniques that were implemented in R version 3.0.1 (R Development Core Team, 2014). Model convergence was monitored using a Monte Carlo standard error threshold of 0.1.[38] For this analysis, a total of 1,000,000 posterior samples were generated.

All statistical models included *a priori* factors potentially associated with HPV vaccine uptake, including sex (categorized as male or female), age (mean-centered), race (categorized as white, African American, American Indian/Alaska native, Asian, or other), ethnicity (categorized as Hispanic or non-Hispanic), educational attainment (categorized as some high school, high school graduate, some college or technical school, college graduate, or graduate school), sexual orientation (categorized as heterosexual, homosexual/gay/lesbian, or bisexual), and political views (categorized as very conservative or conservative, moderate, liberal, or very liberal). Initially, the model was fit maintaining all the variables. The final model retained all covariates

that were statistically significant at $p < 0.05$. Odds ratios and the associated 95% credible intervals are presented. The random effect terms can be interpreted as the effect of ZIP code on HPV vaccination uptake for each individual.

RESULTS

Characteristics of the study sample are presented in Table 1. In all, 46.2% of participants had received at least one dose of HPV vaccine, with 67.7% of women reporting having been vaccinated compared to 13.0% of men. Of those who initiated the vaccine series, 71.1% completed the entire three-dose series (79.6% of women and 26.3% of men). Participants who had been vaccinated against HPV (i.e. received ≥1 dose of the vaccine) were younger (30.1% of those ≥25 years were vaccinated compared to 69.9% of those <25 years). Vaccine receipt was lower among those who identified themselves as politically "conservative" or "very conservative" as opposed to politically "liberal" (24.6% compared to 53.3%).

Table 1. Characteristics of study participants by HPV vaccination status.

	Vaccinated		**Not Vaccinated**		**Total**	
	(n=351)		**(n=409)**		**(n=760)**	
	N	%	N	%	N	%
Age, in years						
18-20	86	24.5	80	19.6	166	21.8
21-25	209	59.5	199	48.7	408	53.7
26-30	56	16.0	130	31.8	186	24.5
Gender						
Female	312	88.9	149	36.4	461	60.7
Male	39	11.1	260	63.6	299	39.3
Race						
White	298	84.9	346	84.6	644	84.7
Black	22	6.3	14	3.4	36	4.7
Am. Indian/AL native	4	1.1	4	1.0	8	1.1
Asian	15	4.3	29	7.1	44	5.8
Other	12	3.4	16	3.9	28	3.7

Hispanic/Latino						
Yes	13	3.7	14	3.5	27	3.6
No	336	96.3	391	96.5	727	96.4
Education						
Some High School	4	1.1	3	0.7	7	0.9
High School Graduate	19	5.4	25	6.1	44	5.8
Some College or Tech. School	135	38.5	151	36.9	286	37.6
College Graduate	152	43.3	175	42.8	327	43.0
Graduate School	41	11.7	55	13.4	96	12.6
Sexual Orientation						
Heterosexual	311	88.6	370	90.7	681	89.7
Homosexual, gay, or lesbian	12	3.4	21	5.1	33	4.3
Bisexual	21	6.0	11	2.7	32	4.2
Don't know/Refused	7	2.0	6	1.5	13	1.7
Political Views						
Very Conservative	1	0.3	21	5.1	22	2.9
Conservative	28	8.0	68	16.6	96	12.6
Moderate	103	29.3	128	31.3	231	30.4
Liberal	154	43.9	142	34.7	296	38.9
Very Liberal	65	18.5	50	12.2	115	15.1

HPV indicates human papillomavirus

[a]Other indicates Native Hawaiian or Pacific Islander, more than one race, or a response of "other."

HPV vaccination was found to exhibit strong spatial dependence ($\hat{\rho} = 0.9951$). The CAR model also successfully converged, as the maximum Monte Carlo standard error was 0.028 (which was below our threshold of 0.1), indicating that a sufficient number of posterior samples were generated for the estimates to stabilize. Estimates for the best-fitting CAR model are shown in Table 2. After accounting for spatial dependence using the CAR model, age, sex, education, and political preferences remained significantly associated with HPV vaccine receipt. Specifically, older age (OR = 0.77 per year, 95% CI = 0.72-0.83) and being male (OR=0.03, 95% CI = 0.02-0.06) were associated with a decreased odds of HPV vaccine receipt. Higher educational attainment (referent to receiving some high school or high school graduates) was associated with an increased odds of HPV vaccine receipt (some college OR = 2.58, 95% CI =

1.14-5.83; college graduate OR = 3.93, 95% CI = 1.66-9.30; graduate degree OR = 4.74, 95% CI = 1.71-13.17). Moderate and liberal political preferences (referent to very conservative and conservative preferences) were also associated with an increased odds of HPV vaccine receipt (moderate OR = 3.24, 95% CI = 1.62-6.49; liberal OR = 5.32, 95% CI = 2.68-10.58). Race was not found to be significantly associated with HPV vaccine uptake. For comparison, odds ratios (and corresponding 95% confidence intervals) from a traditional logistic regression model that does not account for spatial dependence were also estimated and are also presented in Table 2. Compared to the traditional logistic model, estimates from the CAR model were greater in magnitude for all covariates. Of note, in the traditional logistic regression model, having received some college education was not a statistically significant factor but became significant in the CAR model (traditional OR = 1.88, 95% CI = 0.90-3.93; spatial CAR OR = 2.58, 95% CI = 1.14-5.83).

Table 2. Odds ratio estimates for factors associated with HPV vaccination from traditional logistic regression and spatial CAR models.

	Traditional Logistic Model		**Spatial CAR Model**	
	Odds Ratio	95% CI[a]	Odds Ratio	95% CI[b]
Age[c]	0.84	(0.78-0.90)	0.76	(0.70-0.83)
Sex				
Female	1	(referent)	1	(referent)
Male	0.07	(0.05-0.11)	0.04	(0.03-0.07)
Political Views				
Conservative[d]	1	(referent)	1	(referent)
Moderate	2.34	(1.30-4.19)	3.06	(1.61-5.81)
Liberal	2.76	(1.57-4.85)	4.31	(2.32-8.01)
Very Liberal	3.42	(1.76-6.62)	4.82	(2.34-9.94)
Education				
High School[e]	1	(referent)	1	(referent)
Some College	1.88	(0.90-3.93)	2.58	(1.14-5.83)
College Graduate	2.51	(1.15-5.45)	3.93	(1.66-9.30)
Graduate Degree	2.59	(1.03-6.52)	4.74	(1.71-13.17)

HPV indicates human papillomavirus
[a]95% Confidence Interval
[b]95% Credible Interval
[c]Age is centered at the mean (23.24 years old)
[d]Referent group consists of "conservative" and "very conservative" responses
[e]Referent group consists of "some high school" and "high school graduate"

Figure 2 shows a choropleth map of HPV vaccine uptake attributable to the conditional autoregressive random effects in the CAR model. These values represent the spatial heterogeneity of HPV vaccine uptake conditional on population size and the factors included in the model. Heterogeneous HPV vaccine uptake is evident, in that a cluster of ZIP codes with lower uptake are concentrated in the downtown area (shown in light blue), with uptake increasing as distance from city center increases (dark blue ZIP codes).

DISCUSSION

In this study, HPV vaccination was found to exhibit strong spatial dependence, indicating that spatial statistical models are needed to accurately identify and estimate factors associated with HPV vaccine uptake. The spatial analysis also revealed that ZIP codes tend to have HPV vaccine uptake rates that were similar to their neighbors. Ignoring this spatial dependence can lead to biased point estimates and overly narrow credible intervals. Consistent with other studies, younger age, female gender, higher education, and political views were found to be significantly associated with HPV vaccination (after accounting for spatial dependence).[21, 27, 39, 40] The associations of age and sex with HPV vaccine receipt can be attributed, in part, to the evolving ACIP recommendations, as they were first recommended for use in young girls and were later expanded to include young boys. Conservative political views have also been found to be associated with decreased knowledge of HPV, lower perceived risk of infection with HPV, and stronger views against premarital sex.[41]

However, contrary to other studies that have not accounted for spatial dependence, this study found that race was not significantly associated with HPV vaccination.[21, 27, 39, 40, 42, 43] Racial disparities (and other disparities) have been shown to be pronounced in some areas, while less evident (or absent) in other areas.[44-46] Although the existence of these disparities is well documented, the overall average effects (i.e., national level data) can mask variation across local areas.[47, 48] For example, in a traditional regression analysis where minority girls live in regions with systematically different rates of HPV vaccine uptake, and the region is not controlled for, one could erroneously conclude that racial "disparities" exist when in fact *where* people live (e.g., the context of their neighborhood) is the significant factor associated with vaccination. Thus, ignoring geography (i.e., the spatial dependence of the data) may lead to incorrect inference. Previous studies that have attempted to describe geographic variation in HPV vaccine uptake have either ignored spatial dependence completely or have not correctly accounted for it using spatial statistical models.[24, 49] These studies may have incorrectly concluded that covariates such as race are significantly associated with HPV vaccine receipt when, in fact, these conclusions are likely to be erroneous because they are based on models that did not account for spatial dependence. As our analysis shows, using models that account for spatial dependence may greatly improve the identification of independent factors that are truly associated with HPV vaccination (as opposed to spatially confounded covariates), particularly when analyzing data from varying geographic locations.

Previous studies have shown that HPV vaccination uptake exhibits significant geographic variability.[23-25, 27] HPV vaccine policies, availability, costs, poverty, financial assistance, and availability of education materials to promote uptake collectively contribute to this variability, as they vary widely across and within states.[18, 50] As a result, variation at state

levels may not reflect the variation in HPV vaccine uptake occurring at a more local level. However, a more refined level of analysis was not possible in these studies because of the sparseness of data at the county and ZIP code level, which is in part attributable to national surveys aggregating or suppressing responses due to participant identification concerns. One strength of this study is that ZIP code level data were available to conduct a more detailed spatial analysis.

The proportion of all adults in this study who had been vaccinated against HPV (i.e. received at least one dose of the HPV vaccine) was 46.2% (67.7% for women and 13.0% for men). These estimates are much higher than the HPV vaccine coverage estimates from the 2012 National Health Interview Survey (NHIS) for women (34.5%) and men (2.3%) aged 19 to 26.[38] Although the results for women are more similar to those obtained from the National Immunization Survey – Teen for girls (53.8%), the estimate for men is much lower than the NIS-Teen estimate for boys (20.8%) aged 13 to 17 who received at least one dose of HPV vaccine in 2012.[39] Although the differences in the observed rates may be partially explained by the sampling frame, response rates, or the small number of eligible respondents who received the HPV vaccine question series in the national surveys, the estimates of HPV vaccine uptake are noticeably different from the current study.

There are several limitations to this study. First, all study measures were self-reported by persons over the Internet and may be subject to under or over-reporting. However, recent studies have shown recall of HPV vaccination status to be accurate.[51] In addition, Internet-based studies have shown increased self-disclosure and reporting with online surveys, which may reduce potential response biases (e.g. interviewer bias or social desirability).[52, 53] Second, analyses by race may have been underpowered due to small numbers, however, the distribution

of racial groups was proportionate to estimates from the U.S. Census for the study area.[28] Similarly, we cannot rule out selection bias although several procedures were utilized to obtain a representative sample.[28] Third, this study used the age of participants at the time of the survey, not the age of participants at the time of vaccination, to assess differences by age. It should be noted that our objective was to estimate factors associated with the overall prevalence of vaccine uptake among young adults, not to estimate prevalence by age. Fourth, the spatial analyses were conducted at the ZIP code level and assume a common ZIP code level effect, so within-ZIP code differences may be masked. However, to our knowledge, this is only the second study to examine HPV vaccination at such a small areal unit.[48] Another limitation is that this study did not directly adjust for the income of participants, as this information was not available. However, accounting for spatial dependence in this study sample likely incorporates some of the variability for unmeasured factors such as income.[54] Finally, this study utilizes cross-sectional data and temporal effects cannot be established.

In conclusion, the results from this study demonstrate that more detailed and local assessments of HPV vaccine uptake that account for spatial dependence are necessary as ZIP code level patterns differ significantly from aggregated state and national patterns. Future work is needed to further pinpoint areas with the greatest disparities and how to then access these populations to improve vaccine uptake.

Figure 1 Caption

The spatial distribution of 760 survey responses across the Twin Cities Metropolitan Area of Minnesota.

Figure 2 Caption

Uptake of the human papillomavirus vaccine that is attributable to the conditional autoregressive random effects in the spatial CAR model.

Abbreviations

HPV: human papillomavirus; CAR: conditional autoregression.

Author Contributions

SK and EJN conceived and designed the study. EJN also conducted the data collection for this study and drafted the manuscript. JH, JP, and JMO assisted in the survey design, supervised the statistical analysis, and assisted in reviewing/revising the manuscript. SK also provided contributions to the concept and analytic approach for the article, and oversaw the analysis, interpretation, and reviewing/revising of manuscript. All authors read and approved the final manuscript.

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was supported by the J.B. Hawley Student Research Award from the University of Minnesota School of Public Health and by the Minnesota Medical Foundation through Grant 4120-9227-12.

Data Sharing Statement

Only SMASH study team members have full access to the raw data. Researchers interested in using SMASH study data may request permission directly from the authors and will be considered on a case-by-case basis.

REFERENCES

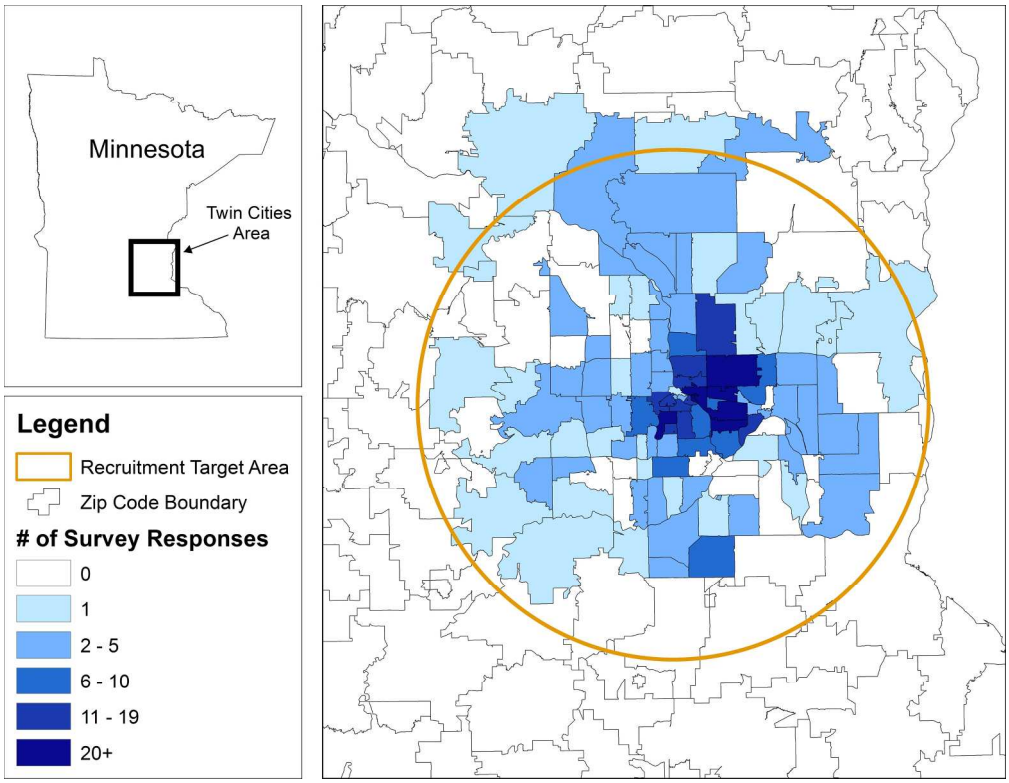
1. Weinstock, H., S. Berman, and W. Cates, Jr., *Sexually Transmitted disease among American youth: incidence and prevalence estimates, 2000.* Perspect Sex Reprod Health, 2004. **36**(1): p. 6-10. PMID:
2. Walboomers, J.M., et al., *Human papillomavirus is a necessary cause of invasive cervical cancer worldwide.* J Pathol, 1999. **189**(1): p. 12-9. PMID: 10451482.
3. CDC, *Human papillomavirus-associated cancers - United States, 2004-2008.* MMWR Morb Mortal Wkly Rep, 2012. **61**: p. 258-61. PMID: 22513527.
4. Saraiya, M., et al., *Toward using National Cancer Surveillance data for preventing and controlling cervical and other human papillomavirus-associated cancers in the US.* Cancer, 2008. **113**(10 Suppl): p. 2837-40. PMID: 18980202.
5. Markowitz, L.E., et al., *Quadrivalent Human Papillomavirus Vaccine: Recommendations of the Advisory Committee on Immunization Practices (ACIP).* MMWR Recomm Rep, 2007. **56**(RR-2): p. 1-24. PMID: 17380109.
6. CDC, *FDA licensure of bivalent human papillomavirus vaccine (HPV2, Cervarix) for use in females and updated HPV vaccination recommendations from the Advisory Committee on Immunization Practices (ACIP).* MMWR Morb Mortal Wkly Rep, 2010. **59**(20): p. 626-9. PMID: 20508593.
7. CDC, *Recommended adult immunization schedule - United States, 2012.* MMWR Morb Mortal Wkly Rep, 2012. **61**(4). PMID:
8. Elam-Evans, L.D., et al., *National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years--United States, 2013.* MMWR Morb Mortal Wkly Rep, 2014. **63**(29): p. 625-33. PMID: 25055186.
9. Williams, W.W., et al., *Vaccination coverage among adults, excluding influenza vaccination - United States, 2013.* MMWR Morb Mortal Wkly Rep, 2015. **64**(4): p. 95-102. PMID: 25654611.
10. Giuliano, A.R., et al., *Efficacy of quadrivalent HPV vaccine against HPV Infection and disease in males.* N Engl J Med, 2011. **364**(5): p. 401-11. PMID: 21288094.
11. Joura, E.A., et al., *Efficacy of a quadrivalent prophylactic human papillomavirus (types 6, 11, 16, and 18) L1 virus-like-particle vaccine against high-grade vulval and vaginal lesions: a combined analysis of three randomised clinical trials.* Lancet, 2007. **369**(9574): p. 1693-702. PMID: 17512854.
12. Paavonen, J., et al., *Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types*

(PATRICIA): final analysis of a double-blind, randomised study in young women. Lancet, 2009. **374**(9686): p. 301-14. PMID: 19586656.

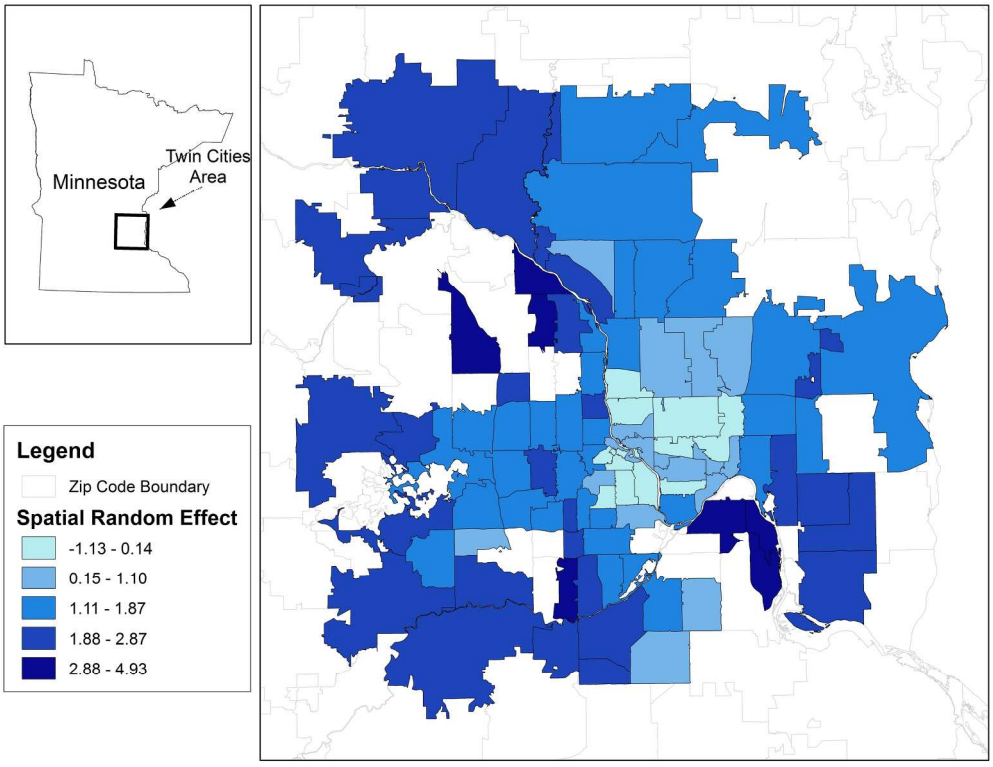
13. Markowitz, L.E., et al., *Reduction in human papillomavirus (HPV) prevalence among young women following HPV vaccine introduction in the United States, National Health and Nutrition Examination Surveys, 2003-2010.* J Infect Dis, 2013. **208**(3): p. 385-93. PMID: 23785124.
14. Kim, J.J. and S.J. Goldie, *Health and economic implications of HPV vaccination in the United States.* N Engl J Med, 2008. **359**(8): p. 821-32. PMID: 18716299.
15. Niccolai, L.M., N.R. Mehta, and J.L. Hadler, *Racial/Ethnic and poverty disparities in human papillomavirus vaccination completion.* Am J Prev Med, 2011. **41**(4): p. 428-33. PMID: 21961471.
16. Jeudin, P., et al., *Race, Ethnicity, and Income Factors Impacting Human Papillomavirus Vaccination rates.* Clin Ther, 2014. **36**(1): p. 24-37. PMID: 24417783.
17. Vadaparampil, S.T., et al., *Provider factors associated with disparities in human papillomavirus vaccination among low-income 9- to 17-year-old girls.* Cancer, 2013. **119**(3): p. 621-8. PMID: 23341308.
18. Bednarczyk, R.A., et al., *Health disparities in human papillomavirus vaccine coverage: trends analysis from the National Immunization Survey-Teen, 2008-2011.* Clin Infect Dis, 2014. **58**(2): p. 238-41. PMID: 24162745.
19. Bednarczyk, R.A., et al., *Human papillomavirus vaccine uptake and barriers: association with perceived risk, actual risk and race/ethnicity among female students at a New York State university, 2010.* Vaccine, 2011. **29**(17): p. 3138-43. PMID: 21376797.
20. Chao, C., et al., *Correlates for human papillomavirus vaccination of adolescent girls and young women in a managed care organization.* Am J Epidemiol, 2010. **171**(3): p. 357-67. PMID: 20047978.
21. Liddon, N.C., J.S. Leichliter, and L.E. Markowitz, *Human papillomavirus vaccine and sexual behavior among adolescent and young women.* Am J Prev Med, 2012. **42**(1): p. 44-52. PMID: 22176845.
22. Hertweck, S.P., et al., *Health care decision making by mothers for their adolescent daughters regarding the quadrivalent HPV vaccine.* J Pediatr Adolesc Gynecol, 2013. **26**(2): p. 96-101. PMID: 23518189.
23. Wei, F., P.C. Moore, and A.L. Green, *Geographic variability in human papillomavirus vaccination among U.S. young women.* Am J Prev Med, 2013. **44**(2): p. 154-7. PMID: 23332332.
24. Pruitt, S.L. and M. Schootman, *Geographic disparity, area poverty, and human papillomavirus vaccination.* Am J Prev Med, 2010. **38**(5): p. 525-33. PMID: 20409501.
25. Eberth, J.M., et al., *County-level estimates of human papillomavirus vaccine coverage among young adult women in Texas, 2008.* Tex Public Health J, 2013. **65**(1): p. 37-40. PMID: 24466565.
26. Lee, D., *CARBayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors.* Journal of Statistical Software, 2013. **55**(13). PMID:
27. Rahman, M., C.J. McGrath, and A.B. Berenson, *Geographic variation in human papillomavirus vaccination uptake among 13-17 year old adolescent girls in the United States.* Vaccine, 2014. **32**(21): p. 2394-8. PMID: 24637175.

28. Nelson, E.J., et al., *Estimation of geographic variation in human papillomavirus vaccine uptake in men and women: an online survey using facebook recruitment.* J Med Internet Res, 2014. **16**(9): p. e198. PMID: 25231937.
29. Moran, P.A., *Notes on continuous stochastic phenomena.* Biometrika, 1950. **37**(1-2): p. 17-23. PMID: 15420245.
30. Waller, L.A. and C.A. Gotway, *Applied Spatial Statistics for Public Health Data*. 2004, Hoboken, NJ: John Wiley & Sons.
31. Besag J, Y.J., Mollie A, *Bayesian Image Restoration with Two Applications in Spatial Statistics.* The Annals of the Institute of Statistics and Mathematics, 1991. **43**(1): p. 1-59. PMID:
32. Lee, D., *A comparison of conditional autoregressive models used in Bayesian disease mapping.* Spat Spatiotemporal Epidemiol, 2011. **2**(2): p. 79-89. PMID: 22749587.
33. Leroux, B., X. Lei, and N. Breslow, *Estimation of disease rates in small areas: a new mixed model for spatial dependence*, in *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, M. Halloran and D. Berry, Editors. 1999, Springer-Verlag: New York. p. 135-178.
34. Stern, H. and N. Cressie, *Inference for extremes in disease mapping*, in *Disease Mapping and Risk Assessment for Public Health*, A. Lawson, et al., Editors. 1999, John Wiley & Sons. p. 63-84.
35. Carlin, B.P. and T.A. Louis, *Bayes and Empirical Bayes methods for data analysis*. 2nd Edition ed. 2000, Boca Raton: Chapman & Hall/CRC.
36. Assun\c cão, R. and E. Krainski, *Neighborhood dependence in Bayesian spatial models.* Biometrical Journal, 2009. **51**(5): p. 851-869. PMID:
37. Besag, J. and C. Kooperberg, *On conditional and intrinsic autoregression.* Biometrika, 1995. **82**(4): p. 733-746. PMID:
38. Flegal, J.M., M. Haran, and G.L. Jones, *Markov chain Monte Carlo: Can we trust the third significant figure?* Statistical Science, 2008. **23**(2): p. 250-260. PMID: ISI:000259275400006.
39. Laz, T.H., M. Rahman, and A.B. Berenson, *Human papillomavirus vaccine uptake among 18- to 26-year-old women in the United States: National Health Interview Survey, 2010.* Cancer, 2013. **119**(7): p. 1386-92. PMID: 23508594.
40. Anhang Price, R., et al., *Use of human papillomavirus vaccines among young adult women in the United States: an analysis of the 2008 National Health Interview Survey.* Cancer, 2011. **117**(24): p. 5560-8. PMID: 21732336.
41. Gerend, M.A. and J.E. Shepherd, *Correlates of HPV knowledge in the era of HPV vaccination: a study of unvaccinated young adult women.* Women Health, 2011. **51**(1): p. 25-40. PMID: 21391159.
42. Sadry, S.A., L.R. De Souza, and M.H. Yudin, *The impact of ethnicity on awareness and knowledge of and attitudes towards the human papillomavirus and vaccine among adult women.* J Obstet Gynaecol Can, 2013. **35**(11): p. 995-1003. PMID: 24246399.
43. Miller, M.K., et al., *Views on Human Papillomavirus Vaccination: A Mixed-Methods Study of Urban Youth.* J Community Health, 2014. PMID: 24664875.
44. Skinner, J., et al., *Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients.* N Engl J Med, 2003. **349**(14): p. 1350-9. PMID: 14523144.

45. Baicker, K., A. Chandra, and J.S. Skinner, *Geographic variation in health care and the problem of measuring racial disparities.* Perspect Biol Med, 2005. **48**(1 Suppl): p. S42-53. PMID: 15842086.
46. Chen, J., et al., *Racial differences in the use of cardiac catheterization after acute myocardial infarction.* N Engl J Med, 2001. **344**(19): p. 1443-9. PMID: 11346810.
47. Baicker, K., et al., *Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries.* Health Aff (Millwood), 2004. **Suppl Variation**: p. VAR33-44. PMID: 15471775.
48. Trogdon, J.G. and T. Ahn, *Geospatial patterns in human papillomavirus vaccination uptake: evidence from uninsured and publicly insured children in north Carolina.* Cancer Epidemiol Biomarkers Prev, 2015. **24**(3): p. 595-602. PMID: 25576528.
49. Rahman, M., T.H. Laz, and A.B. Berenson, *Geographic variation in human papillomavirus vaccination uptake among young adult women in the United States during 2008-2010.* Vaccine, 2013. PMID: 24071591.
50. Katz, M.L., et al., *Human papillomavirus (HPV) vaccine availability, recommendations, cost, and policies among health departments in seven Appalachian states.* Vaccine, 2009. **27**(24): p. 3195-200. PMID: 19446191.
51. Ojha, R.P., et al., *The accuracy of human papillomavirus vaccination status based on adult proxy recall or household immunization records for adolescent females in the United States: results from the National Immunization Survey-Teen.* Ann Epidemiol, 2013. **23**(5): p. 281-5. PMID: 23453240.
52. Cantrell, M.A. and P. Lupinacci, *Methodological issues in online data collection.* Journal of Advanced Nursing, 2007. **60**(5): p. 544-549. PMID: ISI:000250498900010.
53. Rhodes, S.D., D.A. Bowie, and K.C. Hergenrather, *Collecting behavioural data using the world wide web: considerations for researchers.* Journal of Epidemiology and Community Health, 2003. **57**(1): p. 68-73. PMID: ISI:000180078500017.
54. Paciorek, C.J., *The importance of scale for spatial-confounding bias and precision of spatial regression estimators.* Stat Sci, 2010. **25**(1): p. 107-125. PMID: 21528104.



The spatial distribution of 760 survey responses across the Twin Cities Metropolitan Area of Minnesota.
215x166mm (300 x 300 DPI)



Uptake of the human papillomavirus vaccine that is attributable to the conditional autoregressive random effects in the spatial CAR model.
215x166mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
Outcome data	15*	Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses

Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.