Journal of Exposure Science and Environmental Epidemiology (2006) 16, 147–155
© 2006 Nature Publishing Group All rights reserved 1559-0631/06/$30.00
www.nature.com/jes

Validation of self-reported proximity to agricultural crops in a case–control study of neural tube defects

RUDOLPH P. RULL,[a,b] BEATE RITZ[b,c] AND GARY M. SHAW[d]

[a]*Center for Health Policy Research, University of California, Los Angeles, USA*
[b]*Department of Epidemiology, University of California, Los Angeles, USA*
[c]*Department of Environmental Health Sciences, University of California, Los Angeles, USA*
[d]*March of Dimes Birth Defects Foundation, California Birth Defects Monitoring Program, Berkeley, California, USA*

Self-reported perinatal exposures to chemicals or pollutant sources in case–control studies of birth defects may be inaccurate due to misreporting among mothers. In a case–control study of neural tube defects delivered in California in 1987–1988, mothers of case and control infants were asked whether they lived within 0.25 mile (400 m) of agricultural crops. Responses were compared against a gold standard derived from historical agricultural land-use survey maps. The odds ratio for self-reported proximity to any crops (1.62, 95% confidence interval: 1.08, 2.43) appeared to be positively biased compared with the estimate for map-based proximity (1.17, 95% confidence interval: 0.79, 1.71). This pattern was also observed for proximity to specific crops such as nonpermanent and orchard crops. For vineyards, however, we observed an increased risk associated with map-based proximity (odds ratio = 2.45, 95% confidence interval: 1.08, 5.58) but not with self-reported proximity (1.09, 95% confidence interval: 0.51, 2.34). The sensitivity of self-reported proximity to any crops was greater for case (65.7%) than control mothers (50.0%) while specificity was about the same for case and control mothers (87.5 vs. 89.3%), suggesting that control mothers under-reported proximity to crops. Differential reporting was also observed between geographic regions, urban and rural residents, and across levels of maternal employment and education. These results suggest differential reporting between case and control mothers as well as an influence from maternal demographic characteristics on reporting accuracy.
Journal of Exposure Science and Environmental Epidemiology (2006) **16,** 147–155. doi:10.1038/sj.jea.7500444; published online 27 July 2005

Keywords: *agricultural crops, epidemiologic bias, geographic information systems, mental recall, neural tube defects, validation studies.*

Introduction

Population-based case–control studies have been used as a feasible and efficient alternative to cohort designs in the study of causes of birth defects in humans, but exposure assessment using this approach has inherent limitations. In the absence of records of environmental contaminants that can be used to ascertain gestational exposures during the relevant period of organogenesis, these studies usually rely on interview or questionnaire data from parents that is often obtained several months or even years after a pregnancy ended and are thus susceptible to reporting errors (Werler et al., 1989; Coughlin, 1990; Little, 1992; Neutra et al., 1992; Swan et al., 1992). Furthermore, reporting error may be differential between cases and controls, and may possibly be affected by subjective differences in perception or reporting related to demographic characteristics, including maternal age, race/ethnicity, geographic region, or educational level (Delgado-Rodriguez et al., 1995; Schieve et al., 1999; Maynard et al., 2003). It has been speculated that parents of affected infants might be more likely than those born healthy to report accurately exposure when they were truly exposed (higher sensitivity) or exaggerate and thus over-report exposure when they were truly unexposed (lower specificity). While differential recall is often alleged and cited as a potential study limitation, rarely does the opportunity exist to assess the presence and magnitude of bias in effect estimation since a gold standard to validate self-reported exposures is usually lacking (Raphael, 1987; Swan et al., 1992). Furthermore, while previous validation studies of environmental exposures focused on identifying differential recall between case and control parents, these studies did not evaluate potential exposure over-reporting and under-reporting associated with geographic and demographic characteristics of study participants (Feldman et al., 1989; Mackenzie and Lippman, 1989; Little, 1992; Infante-Rivard and Jacques, 2000).

Recent advances in geographic information systems (GIS) technology and the availability of historical land-use survey maps of agricultural crops in California offered us a unique

1. Abbreviations: CBDMP, California Birth Defects Monitoring Program; CI, confidence interval; GIS, geographic information systems; NTD, neural tube defect.

2. Address all correspondence to: Dr. R.P. Rull, UCLA Center for Health Policy Research, 10911 Weyburn Ave, Suite 300, Los Angeles, CA 90024, USA. Tel.: +1 310 794 0589. Fax: +1 310 794 2686.
E-mail: rude@ucla.edu
Received 26 March 2005; accepted 6 June 2005; published online 27 July 2005

opportunity to create a gold standard for a self-reported exposure measure used in a population-based case–control study of neural tube defects (NTDs) conducted by the California Birth Defects Monitoring Program (CBDMP) (Shaw et al., 1999). The study observed that infants born to mothers who reported living within 0.25 mile (400 m) of agricultural crops had increased risks of having NTDs. The study relied on maternal reports of residential proximity to crops as a surrogate for pesticide exposures and the authors discussed whether the observed risk increase may have been an artifact of differential recall. Here, we evaluate the validity of these self-reports of proximity to agricultural crops by comparing responses to land-use survey maps as a gold standard. Furthermore, we assess the impact of differential recall on effect estimates and examine maternal factors that may be associated with under- or over-reporting.

Methods

Study Population

The study population for this case–control study was previously described (Shaw et al., 1999). Briefly, 315 cases of NTDs (anencephaly, spina bifida cystica, craniorrhachischisis, or iniencephaly), including elective terminations, were ascertained from 344,214 singleton infants and fetal deaths delivered between January 1, 1987 and December 31, 1988 to women residing in most California counties, excluding metropolitan Los Angeles and San Francisco and sparsely populated counties not monitored by the CBDMP during this period. A total of 652 control subjects were randomly selected from all live-born infants (without reportable congenital anomalies before the first birthday) delivered in the same geographic area and time period as the cases. For both cases and controls, the month of conception was estimated from the gestational age.

After excluding case or control mothers who spoke languages other than English or Spanish, 265 (84% of eligible) NTD case and 481 (76%) control mothers were interviewed over the telephone. These interviews occurred, on average, 3.7 and 3.8 years after the date of delivery for cases and controls, respectively, and focused on the 4-month periconceptional period ranging between one month before and 3 months after conception. During the course of the interview, mothers were asked to list all residential addresses they had lived at for 2 weeks or more during the periconceptional period. For each of these residences, mothers were asked: "…were agricultural crops or commercial flowers grown within 0.25 mile of this home?" If the mother answered yes, she was asked to name the crops and indicate whether she knew if any products were applied on these crops to control insects or weeds. Reported residential addresses were geocoded (mapped to latitude and longitude coordinates) using the Dynamap/2000 street centerline database (GDT Inc., Lebanon, NH, USA) derived from 1990 US Census data (Croen and Shaw, 1996; Shaw et al., 1999).

We restricted the study population to those case and control mothers whose entire periconceptional residential histories could be geocoded and fell within counties or regions of California where land-use surveys were conducted within 9 years of the year of conception (89% of interviewed cases and 90% of interviewed controls). We further excluded four cases and five controls with "don't know" responses for residential proximity to crops and five cases and six controls with missing information on maternal characteristics, leaving for analyses 227 (86% of interviewed) cases and 424 (88% of interviewed) controls. The distributions of geographic and demographic characteristics among case and control mothers are presented in Table 1.

Land-use Survey-based Residential Proximity

The California Department of Water Resources performs countywide surveys (1:24,000 scale) of land use and crop cover every 7–10 years (California Department of Water Resources, 2005). Most surveys were obtained as paper maps and digitized for use with ArcView GIS software (ESRI, Redlands, CA, USA), while a few surveys were available in digital format. We linked mothers' residential addresses to the appropriate county surveys conducted in the survey year closest to the year of conception. The observed mean absolute difference between the years of the survey and conception was 1.7 years (SD = 1.5 years), with a range of 9 years.

Depending on the type of crop, land-use maps identifying crops grown at a specific point in time may be inaccurate when surveys are conducted only during the summer every 7–10 years. Orchards and vineyards tend to stand for several years or decades and will not substantially differ between surveys. However, seasonal rotations used for nonpermanent crops such as cotton and tomatoes lead to uncertainty regarding which specific crop was planted in a specific location or point in time (Mitchell et al., 2001). Owing to this uncertainty, we combined these crops into a single class of nonpermanent crops (Rull and Ritz, 2003). Furthermore, we collapsed orchard crops into deciduous and citrus/subtropical classes due to the low frequency of specific crops reported by mothers and because some orchards identified in the land-use survey codes were only labeled as deciduous or citrus/subtropical.

We created exposure estimates for the period between conception and the closure of the neural tube (up to the 30th day of gestation) (Elwood et al., 1992). In order to cover the period of neurulation, we selected the residence reported for the calendar month of conception and the month after conception. If a mother lived at more than one address during this period (9% of cases and 6.5% of controls), we selected the longest-lived residence. If a subject had two

Table 1. Characteristics of women with NTD-affected pregnancies (cases) and those without such pregnancies (controls), California, 1987–1988.

Characteristic	Cases (total = 227) No. (%)[a]	Controls (total = 424) No. (%)[a]
Rural/urban residence[b]		
Urban	151 (66.5)	287 (67.7)
Rural	76 (33.5)	137 (32.3)
Geographic region[c]		
Southern (South Coast and Desert)	96 (42.3)	146 (34.4)
Central (San Joaquin Valley and Central Coast)	79 (34.8)	154 (36.3)
Northern (Sacramento Valley and North Coast)	52 (22.9)	124 (29.3)
Maternal ethnicity/race		
Caucasian	137 (60.4)	250 (59.0)
US-born Latina	28 (12.3)	64 (15.1)
Foreign-born Latina	42 (18.5)	50 (11.8)
Other	20 (8.8)	60 (14.2)
Age at conception		
14–19	21 (9.3)	41 (9.7)
20–24	56 (24.7)	124 (29.2)
25–29	82 (36.1)	137 (32.3)
30–34	51 (22.5)	88 (20.8)
≥35	17 (7.5)	32 (7.5)
Periconceptional employment[d]		
Unemployed	101 (44.5)	191 (45.0)
Employed	126 (55.5)	233 (55.0)
Education completed		
College graduate	29 (12.8)	70 (16.5)
Some college	68 (30.0)	126 (29.7)
High school graduate	73 (32.2)	129 (30.4)
<High school graduate	57 (25.1)	99 (23.3)
Cigarette smoking[d]		
None	187 (82.4)	316 (74.5)
1–19/day	27 (11.9)	84 (19.8)
≥20/day	13 (5.7)	24 (5.7)
Vitamin use[d]		
No vitamins	87 (38.3)	132 (31.1)
Used vitamins	140 (61.7)	292 (68.9)

[a]Percentages may not equal 100 due to missing data or rounding.
[b]Based on location of month-of-conception address and 1990 Census Urbanized Areas.
[c]Based on location of month-of-conception address within the California Agriculture Bulletin's geographic groupings of counties.
[d]During the period between the month before conception and 3 months after conception.

longest-lived addresses, the address closest to agricultural crops listed on land-use surveys was selected.

We overlaid subjects' residential addresses as latitude/longitude coordinates onto the land-use maps and plotted buffers with a 0.25-mile (400 m) radius around each residence using ArcView GIS. A residence was defined as within proximity of crops if a field, orchard, or vineyard was located inside or on the edge of the buffer. In addition, we assessed residential proximity to crops within distances of 0.31 and 0.50 miles (500 and 800 m, respectively). We also considered validating whether mothers who reported living near crops knew if any products were applied to control insects or weeds, but "don't know" or "maybe" accounted for 40% of all responses. We also plotted addresses over a map of urbanized areas based on the 1990 US Census to distinguish urban and rural residences. To categorize residences by geographic region, we grouped residences into six geographic regions according to the California Agricultural Bulletin (California Employment Development Department, 2005). Owing to the small numbers of cases and control mothers within certain geographic regions, we collapsed the six regions into three categories: Southern (South Coast and Desert), Central (San Joaquin Valley and Central Coast), and Northern (Sacramento Valley and North Coast). Throughout the proximity assessment, we were blinded to the case–control status of the residences.

Statistical Analysis

We employed logistic regression to estimate odds ratios and 95% confidence intervals (CIs) for delivering an infant with an NTD based on self-reported proximity or land-use map-based proximity within 0.25, 0.31, and 0.50 miles (400, 500, and 800 m, respectively) of any and specific crop classes. Each model controlled for maternal race/ethnicity (Caucasian, US-born Latina, foreign-born Latina, and other), educational level (college graduate, some college, high school graduate, <high school graduate), cigarette smoking (none, <20 per day, ≥20 per day during the periconceptional period), and vitamin use (any use vs. none during the periconceptional period) (Shaw et al., 1999).

Using land-use map-based proximity within 0.25 mile of any or specific crop types as the gold standard, we estimated sensitivities and specificities of self-reported proximity for cases and controls as well as case–control differences with 95% CIs. In addition, we stratified the case and control groups by various geographic and demographic characteristics, including geographic region, urban or rural residence, maternal race/ethnicity, age, educational level, and periconceptional employment, cigarette smoking, and vitamin use and estimated stratum-specific sensitivity and specificity to assess whether there were any differences in reporting behavior within strata of case and control groups. In this analysis, we further dichotomized maternal education as having any college education (college graduate or some college) vs. not having a college education (high school graduate or <high school graduate).

Logistic regression was used to estimate odds ratios for maternal misreporting of (1) residential proximity within 0.25 mile of any or specific crops (i.e., under-reporting) and (2) no

residential proximity to any or specific crops (i.e., over-reporting) (Maynard et al., 2003). To examine factors that may influence under-reporting, we selected case and control mothers who lived within 0.25 mile of any or specific crop types according to land-use surveys and defined the dependent variable as not reporting (vs. reporting) of proximity to the crop type. Similarly, to examine over-reporting, we selected case and control mothers not living within 0.25 mile of any or specific crop types and defined the dependent variable as reporting (vs. not reporting) of proximity to the crop type. In both models, we evaluated the following variables as potential predictors: case–control status, geographic region, urban or rural residence, maternal race/ethnicity, education completed, age at conception, periconceptional employment, and periconceptional cigarette smoking. All analyses were conducted using SAS version 8.2 software (SAS Institute, Cary, NC, USA).

Results

Exposure prevalences and effect estimates for having an NTD-affected pregnancy based on crop proximity derived from self-reports or land-use maps are listed in Table 2. For proximity within 0.25 mile of any crops, we observed a positively biased effect estimate for self-reported proximity (odds ratio = 1.62, 95% CI: 1.08, 2.43) compared with the estimate for land-use map-based proximity (odds ratio = 1.17, 95% CI: 0.79, 1.71). Similar effect estimates were observed for exposures defined as map-based proximity within 0.31 and 0.50 miles. The same pattern was observed for effect estimates for proximity to any nonpermanent and any orchard crop types, as well as for deciduous orchard and subtropical/citrus orchard crop types (results not shown). The exception was proximity to vineyards, for which the self-report-based effect estimate (odds ratio = 1.09, 95% CI: 0.51, 2.34) was considerably smaller than that estimate derived from map-based proximity within 0.25 mile (odds ratio = 2.45, 95% CI: 1.08, 5.58).

Sensitivity and specificity of self-reported proximity, compared with survey-based proximity within 0.25 mile, are presented in Table 3 by case–control status. Overall, sensitivity was poor to moderate (maximum of 65.7% for cases within 0.25 mile of any crops) and tended to be considerably higher for cases. Large differences in sensitivity

Table 2. Odds ratios of NTDs associated with maternal self-reported or land-use survey-based crop proximity, California, 1987–1988.

Crop proximity method (distance: n miles)[a]	Cases ($n = 227$) No. (%)	Controls ($n = 424$) No. (%)	Odds ratio[b]	95% CI[c]
Proximity to any agricultural crops				
Self-reported (0.25)	64 (28.2)	91 (21.5)	1.62	1.08, 2.43
Land-use survey (0.25)	67 (29.5)	116 (27.4)	1.17	0.79, 1.71
Land-use survey (0.31)	84 (37.0)	148 (34.9)	1.17	0.81, 1.68
Land-use survey (0.50)	117 (51.5)	209 (49.3)	1.17	0.83, 1.66
Proximity to specific crop types[d]				
Any nonpermanent crops[e]				
Self-reported (0.25)	37 (16.4)	52 (12.4)	1.45	0.90, 2.34
Land-use survey (0.25)	49 (21.7)	85 (20.2)	1.09	0.72, 1.65
Land-use survey (0.31)	64 (28.3)	117 (27.9)	1.07	0.73, 1.57
Land-use survey (0.50)	100 (44.3)	175 (41.7)	1.17	0.82, 1.67
Any orchards[f]				
Self-reported (0.25)	31 (13.7)	36 (8.6)	1.95	1.14, 3.34
Land-use survey (0.25)	25 (11.1)	42 (10.0)	1.12	0.65, 1.93
Land-use survey (0.31)	40 (17.7)	60 (14.3)	1.31	0.83, 2.07
Land-use survey (0.50)	62 (27.4)	108 (25.7)	1.14	0.77, 1.68
Vineyards				
Self-reported (0.25)	12 (5.3)	22 (5.2)	1.09	0.51, 2.34
Land-use survey (0.25)	15 (6.6)	12 (2.9)	2.45	1.08, 5.58
Land-use survey (0.31)	19 (8.4)	19 (4.5)	2.04	1.01, 4.13
Land-use survey (0.50)	32 (14.2)	49 (11.7)	1.40	0.81, 2.41

[a]0.25 mile = 400 m; 0.31 mile = 500 m; 0.50 mile = 800 m.
[b]Adjusted for maternal ethnicity, education, and periconceptional cigarette smoking and vitamin use.
[c]CI: confidence interval.
[d]Percentages exclude one case and four controls who reported proximity to any crops but did not list any names of crops.
[e]Examples include cotton, tomato, beans, alfalfa, and lettuce.
[f]Examples include deciduous crops such as almonds, apples, cherries, and peaches, and subtropical/citrus crops such as lemons, oranges, avocados, and olives.

Table 3. Sensitivity and specificity for self-reported proximity to agricultural crops[a] among mothers of NTD cases and controls, California, 1987–1988.

Crop type	Land-use-based proximity = yes No.	Sensitivity %	Sensitivity difference % (95% CI[b])	Land-use-based proximity = no No.	Specificity %	Specificity difference % (95% CI[b])
Any agricultural crops						
Cases	67	65.7	15.7 (1.1, 30.2)	160	87.5	−1.8 (−8.0, 4.4)
Controls	116	50.0		308	89.3	
Specific crop types						
Any nonpermanent crops						
Cases	49	44.9	10.8 (−6.4, 28.0)	177	91.5	−1.6 (−6.5, 3.3)
Controls	85	34.1		335	93.1	
Any orchards						
Cases	25	64.0	28.3 (4.5, 52.0)	201	92.5	−1.9 (−6.2, 2.4)
Controls	42	35.7		378	94.4	
Vineyards						
Cases	15	46.7	5.0 (−32.6, 42.6)	211	97.6	1.8 (−1.0, 4.6)
Controls	12	41.7		408	95.8	

[a]Compared with land-use survey-based proximity within 0.25 mile (400 m).
[b]CI: confidence interval.

between case and control mothers were observed for 0.25-mile proximity to any crops (15.7%, 95% CI: 1.1, 30.2%) and any orchards (28.3%, 95% CI: 4.5, 52.0%). Specificity, conversely, was quite high. Comparing case and control mothers, we observed a slightly lower specificity among cases for proximity to any or specific crop types, with the exception of vineyards (97.6 vs. 95.8%).

Table 4 lists sensitivity and specificity of self-reported 0.25-mile proximity to any crops stratified according to maternal and geographic characteristics and case–control status. No particular covariate stratum appeared to contribute disproportionately to differences in sensitivity. We observed some differences in sensitivity and specificity within strata of case and control mothers, especially within categories of geographic region, urban or rural residence, maternal race/ethnicity, periconceptional employment, and completed education.

In Table 5, we present odds ratios obtained from multivariate logistic regression models for predictors of under-reporting and over-reporting of residential proximity to any crops. Case mothers appeared to be less likely to under-report than control mothers (odds ratio = 0.53, 95% CI: 0.26, 1.06). Under-reporting was also less likely for Central California residents (odds ratio = 0.39, 95% CI: 0.16, 1.00). In addition, maternal employment during the periconceptional period was positively associated with under-reporting (odds ratio = 2.27, 95% CI: 1.16, 4.44). Over-reporting was more likely among rural residents (odds ratio = 2.05, 95% CI: 1.02, 4.13) and Central California residents (odds ratio = 2.28, 95% CI: 1.09, 4.74), while women without any college education were less likely to over-report (odds ratio = 0.41, 95% CI: 0.19, 0.88). Due to the small numbers of cases and controls reporting proximity to specific crop types, we were only able to evaluate predictors of over- and under-reporting for nonpermanent field crops. Effect estimates for under- and over-reporting were similar in size and direction to those for residential proximity to any crops (results not shown).

Discussion

These findings quantify potential errors associated with using self-reported residential proximities as exposure metrics in case–control studies. Effect estimates based on self-reports suggest an increased risk of NTD-affected pregnancies associated with proximity to any crops and to orchard crops. Estimates based on land-use survey maps, however, do not suggest such an association and indicate that odds ratios from self-reports were positively biased away from the null due to differential reporting. We observed that case and control mothers did not differ with respect to over-reporting residential proximity to crops, which contradicts our hypothesis that case mothers would be more likely to over-report. However, case mothers were more likely to report accurately residential proximity to crops than control mothers, that is, control mothers under-reported their proximity to crops, which resulted in artificially elevated

Table 4. Sensitivity and specificity for self-reported proximity to any agricultural crops[a] among mothers of NTD cases and controls, stratified by geographic and maternal characteristics, California, 1987–1988.

	Land-use survey-based proximity = yes				Land-use survey-based proximity = no			
	Cases ($n = 67$)		Controls ($n = 116$)		Cases ($n = 160$)		Controls ($n = 308$)	
	Sensitivity		Sensitivity		Specificity		Specificity	
Characteristic	No.	%	No.	%	No.	%	No.	%
Urban or rural residence								
Urban	23	70	48	31	128	88	239	92
Rural	44	64	68	63	32	88	69	80
Geographic region								
Southern	15	80	25	20	81	88	121	93
Central	40	68	62	65	39	85	92	79
Northern	12	42	29	45	40	90	95	94
Ethnicity								
Caucasian	38	63	67	52	99	86	183	89
US-born Latina	13	92	19	58	15	87	45	87
Foreign-born Latina	13	46	16	50	29	93	34	94
Other	3	67	14	29	17	88	46	89
Employment								
Unemployed	34	68	53	64	67	90	138	91
Employed	33	64	63	38	93	86	170	88
Education completed								
Any college	25	72	42	45	72	83	154	86
No college	42	62	74	53	88	91	154	92
Age at conception								
14–19	6	50	12	50	15	87	29	86
20–24	13	62	37	62	43	91	87	93
25–29	23	65	35	51	59	88	102	90
30–34	18	89	23	35	33	82	65	85
≥ 35	7	29	9	33	10	90	23	91
Cigarette smoking								
None	55	64	87	49	132	88	229	90
1–19/day	10	70	21	52	17	88	63	87
≥ 20/day	2	100	8	50	11	82	16	81
Vitamin use								
No vitamins used	25	68	40	50	62	86	92	91
Vitamins used	42	64	76	50	98	89	216	88

[a]Compared with land-use survey-based proximity ≤ 0.25 mile (400 m).

effect estimates. This observation contradicts a speculation in the original report, which stated that the observed association between self-reported residential proximity to crops and NTDs was not likely explained by under-reporting among control mothers because the proportion of control mothers reporting proximity to crops was similar to the proportion observed in another California-based study of stillbirths (Pastore et al., 1997; Shaw et al., 1999). It seems possible that in both studies, control mothers may have been prone to under-reporting.

Certain maternal characteristics were associated with erroneous reporting of residential proximity. Under-reporting was positively associated with periconceptional employment and over-reporting with education level. These are novel findings because very few studies have investigated the association of these characteristics with over- and under-reporting of environmental exposures (Feldman et al., 1989; Mackenzie and Lippman, 1989; Little, 1992). Misreporting of residential proximity was also associated with geographic region. Mothers in rural areas were more likely than those in

Table 5. Odds ratios from multivariate logistic regression models for under-reporting[a] or over-reporting[b] residential proximity[c] to any agricultural crops among mothers of NTD cases and controls, California, 1987–1988.

	Under-reporting (81 of 183 mothers (44.2%))[a]		Over-reporting (52 of 466 mothers (11.1%))[b]	
Characteristic	Odds ratio	95% CI[d]	Odds ratio	95% CI[d]
Case or control				
Control	Reference		Reference	
Case	0.53	0.26, 1.06	1.34	0.71, 2.55
Rural or urban residence				
Urban	Reference		Reference	
Rural	0.59	0.28, 1.25	2.05	1.02, 4.13
Geographic region				
Southern	Reference		Reference	
Central	0.39	0.16, 1.00	2.28	1.09, 4.74
Northern	1.14	0.42, 3.15	0.67	0.28, 1.61
Ethnicity				
Caucasian	Reference		Reference	
US-born Latina	0.59	0.21, 1.61	1.01	0.39, 2.64
Foreign-born Latina	2.12	0.75, 6.01	0.87	0.26, 2.93
Other	2.46	0.74, 8.15	1.13	0.45, 2.88
Employment				
Unemployed	Reference		Reference	
Employed	2.27	1.16, 4.44	1.19	0.62, 2.30
Educational completed				
Any college	Reference		Reference	
No college	1.34	0.61, 2.95	0.41	0.19, 0.88
Age at conception				
14–19	2.39	0.72, 7.89	2.09	0.66, 6.60
20–24	0.94	0.38, 2.32	0.77	0.32, 1.82
25–29	Reference		Reference	
30–34	0.90	0.37, 2.19	1.52	0.70, 3.27
≥35	2.73	0.73, 10.14	0.78	0.20, 3.00
Cigarette smoking				
None	Reference		Reference	
1–19/day	0.84	0.35, 2.02	1.11	0.48, 2.57
≥20/day	0.56	0.13, 2.37	2.56	0.82, 7.94

[a]Mothers who reported proximity = no among mothers whose land-use survey-based proximity = yes.
[b]Mothers who reported proximity = yes among mothers whose land-use survey-based proximity = no.
[c]0.25 mile (400 m).
[d]CI: confidence interval.

urban areas to over-report proximity to crops. A resident in a rural area with intensive agricultural activity such as Central California may underestimate the distance of crops from her home, while a resident of an urban area far from agricultural land such as metropolitan Southern California is more likely to report accurately that she does not live near crops.

The effect estimates for land-use survey-based proximity to crops (Table 2) do not suggest an association between NTDs and general proximity to any crops or specific crop types, except for vineyards. Proximity to agriculture has been used as an indicator for environmental exposure to agricultural pesticides in previous studies (White et al., 1988; Pastore et al., 1997; Ward et al., 2000; Xiang et al., 2000; Schreinemachers, 2003). Residence near crops, however, fails to capture regional and seasonal variations in pesticide use. This is especially problematic when investigating outcomes such as birth defects in which the critical exposure period may only be a few months (Hertz-Picciotto et al., 1996; Bell et al., 2001; Teitelbaum, 2002; Rull and Ritz, 2003). Pesticide use is associated with regional distributions of specific insects, plant diseases, and other pests and coincides with growing seasons as well as seasonal infestation

(Bell et al., 2001; Rull and Ritz, 2003). Furthermore, the fate and drift of pesticides in the environment are determined by several factors, including the persistence (i.e., half-life) of the pesticide in the environment, the application method, wind speed and direction, and other weather conditions (Menzie, 1972).

We observed an association between land-use map-based residential proximity within 0.25 and 0.31 miles of vineyards and NTDs. This association may be attributable to exposure to applications of pesticides with known teratogenic properties such as benomyl on vineyards. Benomyl, a benzimidazole fungicide, has been associated with congenital malformations, including NTDs, in animals (Seiler, 1975; Kavlock et al., 1982; Ellis et al., 1987). Benomyl would have also been applied on several other crops, including celery, strawberries, and almonds (Wales, 1999), but the necessity to group these crops into nonpermanent and orchard categories prevented us from estimating effects of proximity to these specific crops on NTDs.

Certain limitations of land-use survey and geocoded address data affect their validity as a gold standard for assessing residential proximity to crops. While the surveys accurately reflect land-use during the summer of the year the survey was conducted, changes in land-use may have occurred between county survey years as a result of urban or suburban development expanding into rural areas. Another potential factor affecting the validity of the land-use surveys is short-term market-driven change in the production of crop types (e.g., changing an orchard to a vegetable crop). In addition, crop or orchards may enlarge, shrink, or change location during years between surveys (Rull and Ritz, 2003). As a result, we may have incorrectly classified a correct self-report of proximity to a crop type as over-reporting or a correct self-report of no proximity as under-reporting. Furthermore, the rotation of nonpermanent crops prevented us from validating self-reports of proximity to specific crops such as cotton, tomatoes, and corn.

Errors in residential-address geocoding may have occurred as a result of geocoding procedures that interpolate addresses within a range of street numbers and potentially place residences in imprecise locations. This interpolation method, which was used by the CBDMP to geocode maternal residences and is the default geocoding procedure commonly used in GIS software packages, may have limited utility in rural areas (Ratcliffe, 2001; Durr and Froggatt, 2002; Cayo and Talbot, 2003). Rural residences may lie between 90 and 200 feet away from the street curb location where the address is geocoded using GIS software and may even be obstructed from the street by a crop field, vineyard, or orchard (Ward et al., 2000). As a result, for a residence where the true distance to crops is close to or just beyond 0.25 mile, geocoding error will lead to misclassification of survey-based proximity. While these land-use survey and geocoding errors would affect the sensitivity and specificity of our results, we did not expect the error to be differential between cases and controls. Furthermore, when we compared self-reported proximity within 0.25 mile of crops to survey-based proximity, we still observed differential reporting between cases and controls within larger distances of 0.31 and 0.50 miles.

The subjective recall of residential proximity to locations of interest within an arbitrary distance may be prone to error that can lead to biased effect estimates when used as a proxy for exposure (Infante-Rivard and Jacques, 2000). The recent emergence of GIS technology in epidemiology and the availability of geo-referenced environmental data allow for the objective measurement of distances and assessment of proximity to exposure sources. Using land-use maps as a gold standard, we observed differential reporting associated with maternal case–control status and geographic and demographic characteristics. Nevertheless, distance to a pollutant source is a proxy for true exposure to the pollutant. Self-reported and map-based residential proximity to any or specific crops as a proxy measure of environmental exposure to pesticides cannot sufficiently capture spatial and temporal variations of pesticide use.

Acknowledgements

We thank the California Department of Water Resources for supplying us with their land-use survey data and assisting our application of these data to this project. We also thank Drs. Jess Kraus, Barbara Visscher, and Wendie Robbins for comments that improved this paper, and Marlena Kane and Jenesha Narayanan for reviewing drafts.

References

Bell E.M., Hertz-Picciotto I., and Beaumont J.J. A case–control study of pesticides and fetal death due to congenital anomalies. *Epidemiology* 2001: 12: 148–156.

California Department of Water Resources. Land Use Surveys. Available at http://www.landwateruse.water.ca.gov/basicdata/landuse/surveys.cfm. Accessed March 25, 2005.

California Employment Development Department. Labor Market Information: Agricultural Regions. Available at http://www.calmis.ca.gov/file/agric/regcos.htm. Accessed March 25, 2005.

Cayo M.R., and Talbot T.O. Positional error in automated geocoding of residential addresses. *Int J Health Geogr*, [serial online] 2003: 2, Available at http://www.ij-healthgeographics.com/content/2/1/10. Accessed March 25, 2005.

Coughlin S.S. Recall bias in epidemiologic studies. *J Clin Epidemiol* 1990: 43: 87–91.

Croen L.A., and Shaw G.M. *The Occurrence of Neural Tube, Heart, and Oral Cleft Defects in Areas with National Priorities List Sites: A Case–Control Study*. California Birth Defects Monitoring Program, March of Dimes Birth Defects Program, Berkeley, 1996.

Delgado-Rodriguez M., Gomez-Olmedo M., Bueno-Cavanillas A., Garcia-Martin M., and Galvez-Vargas R. Recall bias in a case–control study of low birth weight. *J Clin Epidemiol* 1995: 48: 1133–1140.

Durr P.A., and Froggatt A.E. How best to geo-reference farms? A case study from Cornwall, England. *Prev Vet Med* 2002: 56: 51–62.

Ellis W.G., Semple J.L., Hoogenboom E.R., Kavlock R.J., and Zeman F.J. Benomyl-induced craniocerebral anomalies in fetuses of adequately nourished and protein-deprived rats. *Teratog Carcinog Mutagen* 1987: 7: 357–375.

Elwood J.M., Little J., and Elwood J.H. Classification, anatomy, and embryology. In: Elwood J.M., Little J., and Elwood J.H. (Eds.). *Epidemiology and Control of Neural Tube Defects*. Oxford University Press, Oxford, 1992, pp. 10–36.

Feldman Y., Koren G., Mattice K., Shear H., Pellegrini E., and MacLeod S.M. Determinants of recall and recall bias in studying drug and chemical exposure in pregnancy. *Teratology* 1989: 40: 37–45.

Hertz-Picciotto I., Pastore L.M., and Beaumont J.J. Timing and patterns of exposures during pregnancy and their implications for study methods. *Am J Epidemiol* 1996: 143: 597–607.

Infante-Rivard C., and Jacques L. Empirical study of parental recall bias. *Am J Epidemiol* 2000: 152: 480–486.

Kavlock R.J., Chernoff N., Gray Jr L.E., Gray J.A., and Whitehouse D. Teratogenic effects of benomyl in the Wistar rat and CD-1 mouse, with emphasis on the route of administration. *Toxicol Appl Pharmacol* 1982: 62: 44–54.

Little J. Ascertainment, registration, and assessment of exposure. In: Elwood J.M., Little J., and Elwood J.H. (Eds.). *Epidemiology and Control of Neural Tube Defects*. Oxford University Press, Oxford, 1992, pp. 37–95.

Mackenzie S.G., and Lippman A. An investigation of report bias in a case–control study of pregnancy outcome. *Am J Epidemiol* 1989: 129: 65–75.

Maynard L.M., Galuska D.A., Blanck H.M., and Serdula M.K. Maternal perceptions of weight status of children. *Pediatrics* 2003: 111: 1226–1231.

Menzie C.M. Fate of pesticides in the environment. *Annu Rev Entomol* 1972: 17: 199–222.

Mitchell J.P., Lanini W.T., Miyao E.M., et al. Growing processing tomatoes with less tillage in California. In: Hartz T.K. (Ed.). *Acta Hort 542: VII International Symposium on the Processing of Tomato*. ISHS, Leuven, Belgium, 2001 pp. 347–354.

Neutra R.R., Swan S.H., Hertz-Picciotto I., et al. Potential sources of bias and confounding in environmental epidemiologic studies of pregnancy outcomes. *Epidemiology* 1992: 3: 134–142.

Pastore L.M., Hertz-Picciotto I., and Beaumont J.J. Risk of stillbirth from occupational and residential exposures. *Occup Environ Med* 1997: 54: 511–518.

Raphael K. Recall bias: a proposal for assessment and control. *Int J Epidemiol* 1987: 16: 167–170.

Ratcliffe J.H. On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units. *Int J Geogr Inform Sci* 2001: 15: 473–485.

Rull R.P., and Ritz B. Historical pesticide exposure in California using pesticide use reports and land-use surveys: an assessment of misclassification error and bias. *Environ Health Perspect* 2003: 111: 1582–1589.

Schieve L.A., Perry G.S., Cogswell M.E., et al. Validity of self-reported pregnancy delivery weight: an analysis of the 1988 National Maternal and Infant Health Survey. NMIHS Collaborative Working Group. *Am J Epidemiol* 1999: 150: 947–956.

Schreinemachers D.M. Birth malformations and other adverse perinatal outcomes in four US Wheat-producing states. *Environ Health Perspect* 2003: 111: 1259–1264.

Seiler J.P. Toxicology and genetic effects of benzimidazole compounds. *Mutat Res* 1975: 32: 151–168.

Shaw G.M., Wasserman C.R., O'Malley C.D., Nelson V., and Jackson R.J. Maternal pesticide exposure from multiple sources and selected congenital anomalies. *Epidemiology* 1999: 10: 60–66.

Swan S.H., Shaw G.M., and Schulman J. Reporting and selection bias in case–control studies of congenital malformations. *Epidemiology* 1992: 3: 356–363.

Teitelbaum S.L. Questionnaire assessment of nonoccupational pesticide exposure in epidemiologic studies of cancer. *J Expo Anal Environ Epidemiol* 2002: 12: 373–380.

Wales P. *Use Information and Air Monitoring Recommendation for the Pesticide Active Ingredient Benomyl*. California Department of Pesticide Regulation, Sacramento, November 1999.

Ward M.H., Nuckols J.R., Weigel S.J., Maxwell S.K., Cantor K.P., and Miller R.S. Identifying populations potentially exposed to agricultural pesticides using remote sensing and a Geographic Information System. *Environ Health Perspect* 2000: 108: 5–12.

Werler M.M., Pober B.R., Nelson K., and Holmes L.B. Reporting accuracy among mothers of malformed and nonmalformed infants. *Am J Epidemiol* 1989: 129: 415–421.

White F.M., Cohen F.G., Sherman G., and McCurdy R. Chemicals, birth defects and stillbirths in New Brunswick: associations with agricultural activity. *Can Med Assoc J* 1988: 138: 117–124.

Xiang H., Nuckols J.R., and Stallones L. A geographic information assessment of birth weight and crop production patterns around mother's residence. *Environ Res* 2000: 82: 160–167.