Historical Pesticide Exposure in California Using Pesticide Use Reports and Land-Use Surveys: An Assessment of Misclassification Error and Bias

Rudolph P. Rull[1] and Beate Ritz[1,2]

[1]Department of Epidemiology and [2]Center for Occupational and Environmental Health, School of Public Health, University of California-Los Angeles, Los Angeles, California, USA

We used California's Pesticide Use Report (PUR) and land-use survey data to conduct a simulation study evaluating the potential consequences of misclassifying residential exposure from proximity to agricultural pesticide application in health effect studies. We developed a geographic model linking the PUR with crop location data from land-use surveys to assess the impact of exposure misclassification from simpler exposure models based solely on PUR or land-use data. We simulated the random selection of population controls recruited into a hypothetical case–control study within an agricultural region. Using residential parcel data, we derived annual exposure prevalences, sensitivity, and specificity for five pesticides and relied on the PUR plus land-use model as the "gold standard." Based on these estimates, we calculated the attenuation of prespecified true odds ratios (ORs), assuming nondifferential exposure misclassification. True ORs were severely attenuated *a*) when residential exposure status was based on a larger geographic area yielding higher sensitivity but low specificity for exposure, in contrast to relying on a smaller area and increasing specificity; *b*) for less frequently applied pesticides; and *c*) with increasing mobility of residents among the study population. Considerable effect estimate attenuation also occurred when we used residential distance to crops as a proxy for pesticide exposure. Finally, exposure classifications based on annual instead of seasonal summaries of PUR resulted in highly attenuated ORs, especially during seasons when applications of specific pesticides were unlikely to occur. These results underscore the importance of increasing the spatiotemporal resolution of pesticide exposure models to minimize misclassification. *Key words:* agriculture, bias, California, epidemiology, exposure assessment, geographic information systems, land use, misclassification, pesticide use, residential exposure. *Environ Health Perspect* 111:1582–1589 (2003). doi:10.1289/ehp.6118 available via *http://dx.doi.org/* [Online 20 May 2003]

Agricultural pesticides are the largest group of poisonous substances intentionally disseminated throughout the environment for the purpose of combating animal pests and diseases that devastate crops. Many are known to be acutely toxic to nontargeted organisms, including humans (Ecobichon and Joy 1994). Most epidemiologic studies investigating acute or chronic health effects from human pesticide exposure have been conducted in heavily exposed occupational groups such as pesticide applicators or manufacturers (Zahm et al. 1997). The number of workers occupationally exposed to a specific pesticide formulation, however, is often relatively small or not representative of certain susceptible populations (e.g., pregnant women), thus hampering investigations of less common chronic diseases suspected to be caused by some pesticides, including specific cancers (Zahm et al. 1997), Parkinson's disease (Engel et al. 2001), and birth defects (Shaw et al. 1999). Furthermore, it may not be appropriate to use results from studies of acute exposures to predict chronic health effects, especially when low-level and long-term exposures are more widespread in the general population.

Residential proximity to agricultural pesticide applications may be an important source of ambient environmental exposure in rural communities throughout the United States. Pesticides applied from the air or ground have been observed to drift from their intended treatment sites, with measurable concentrations detected in the air and in plants and animals several hundred meters away (Chester and Ward 1984; Currier et al. 1982; Frost and Ware 1970; MacCollom et al. 1986; Woods et al. 2001). Herbicides transported downwind can cause unintended damage to crops (Byass and Lake 1977), and acute pesticide poisonings have been observed in communities downwind from agricultural fields after applications (Ames et al. 1993). Children residing near agricultural fields tend to have higher urinary levels of dimethylthiophosphate (DMTP), a metabolite of organophosphorus pesticides commonly used in agriculture (Loewenherz et al. 1997). Ward et al. (2000) recently evaluated the feasibility of assessing pesticide exposures due to residential proximity to agricultural pesticide applications in rural populations. Using geographic information system (GIS) technology, they created estimates of pesticide exposure in rural Nebraska by combining land-use and crop-cover information obtained from satellite images with statewide annual estimates of pesticide application rates and acres treated. The spatial and temporal validity of this exposure assessment approach, however, is limited because Nebraska pesticide-application data report only regional and annual summaries.

California is the most agriculturally productive state in the United States, accounting for approximately 13% ($957 million) of all agricultural chemical expenditures, including pesticides [U.S. Department of Agriculture (USDA) 1997]. In 2000, approximately 172 million pounds of pesticide active ingredients were applied for production agriculture in California [California Department of Pesticide Regulation (CDPR) 2000a]. In 1972, California mandated by law the filing of pesticide-use reports (PUR) for commercial applications of restricted-use pesticides (i.e., agents with harmful environmental or toxicologic effects). The law was extended to cover all pesticides in 1990 (CDPR 2000b). The locations of agricultural pesticide applications are reported according to the Public Land Survey System (PLSS), a grid that parcels land into sections with an area of approximately 1 mi^2 (640 acres or 259 ha) and is used in the 30 westernmost states formed from lands in the public domain.

Recently, California PUR data have been used to identify population groups residing in high pesticide use areas at the county, zip code, census block, or PLSS section level and to examine links between use patterns and outcomes of interest, such as adult and childhood cancers (Clary and Ritz 2003; Gunier et al. 2001; Mills 1998; Reynolds et al. 2002), Parkinson's disease (Ritz and Yu 2000), and fetal deaths (Bell et al. 2001).

Although the California PUR system records specific pesticide use, the spatial resolution of the PUR data alone does not allow for the assessment of exposures from residential

Address correspondence to B. Ritz, Department of Epidemiology, UCLA School of Public Health, PO Box 951772, Los Angeles, CA 90095-1772 USA. Telephone: (310) 206-7458. Fax: (310) 206-7371. E-mail: britz@ucla.edu

We thank T. Quach, A. Krishnadasan, and G. Maglinte for their assistance in initiating this study and K. Hoggatt for reviewing drafts.

This study was funded in part by grant R01 ES010544 from the National Institute of Environmental Health Sciences and pilot grants from the American Parkinson's Disease Association, the Southern California Environmental Health Sciences Center, and the University of California Toxic Substances Research and Training Program.

The authors declare they have no conflict of interest.

Received 20 November 2002; accepted 20 May 2003.

proximity to pesticides at distances < 1 mi (1,609 m). Yet some dispersion studies suggest that pesticides are measurable only at considerably smaller distances of ≤ 500 m from the application site (Chester and Ward 1984; Frost and Ware 1970; MacCollom et al. 1986; Woods et al. 2001). Thus, exposure misclassification may occur if the drift range of an applied pesticide is considerably less than one 1 mi. One way to increase the spatial resolution of the PUR beyond the square-mile PLSS section is by using land-use survey data available from the California Department of Water Resources (CDWR 2002; Miller et al. 2002b).

Exposure misclassification may result from residential mobility when an individual changes residences during the exposure period of interest (Khoury et al. 1988). In studies of residential exposures where detailed residential histories are lacking, only one residential address (usually the most recent) is used as a proxy for all addresses. Another potential source of misclassification is the use of an exposure model based on aggregated annual data. This type of data may be adequate for diseases such as Parkinson's disease and adult cancers in which the relevant exposure period may consist of several years or even decades. For diseases in which the critical exposure period may be considerably less than 1 year (i.e., the first trimester for birth defects), however, aggregated annual data may be a poor proxy for detailed temporal data of pesticide applications, especially when the frequency of pesticide applications fluctuates seasonally.

In the present analysis, we used California's unique PUR and land-use survey databases to conduct a simulation study relying on actual data of historical pesticide use and crop cover in an agricultural region (Western Kern County). We evaluated the potential consequences of misclassifying residential exposure due to proximity to agricultural pesticide application in health effect studies that *a*) use land-use information only (e.g., as proposed for states or regions where historical pesticide use data is unavailable); *b*) rely on PUR data without land-use information; *c*) assume long-term residential stability when a population is relatively mobile; or *d*) employ annual use averages for seasonally applied pesticides.

Materials and Methods

For our simulation study, we selected Kern County, the second most agriculturally productive county in the United States (by market value of production) (USDA 1997), which is located in the southern end of the Central Valley region in California. Because one of our objectives was to explore effects of long-term pesticide exposure on chronic diseases with several decades of latency, we used restricted-use PUR data collected between 1972 and 1989, rather than relying on the subsequent full-use reporting system and selected PUR data from 1988. We used the land-use survey closest in time, conducted in 1990, to map the most likely land use during 1988. Because this survey predominantly covered the agricultural western half of Kern County, we restricted our study to this area of the county.

Pesticide use reports. For agricultural applications in California, each PUR record documents the name of each pesticide's active ingredient, the pounds applied, the crop and acreage of the field, the application method, and the date and location of the application. [A current PUR data sheet is available online (CDPR 1999).] The spatial resolution of a PUR is one PLSS, or approximately 1 mi^2. For our simulation exercise, we selected a diverse set of five pesticides representing different physicochemical properties, use specifications (e.g., herbicides, fungicide, insecticide), and application frequencies: methomyl, a carbamate insecticide; parathion, an organothiophosphate insecticide; paraquat, a pyridine defoliant herbicide; endosulfan, an organochlorine insecticide; and maneb, a dithiocarbamate fungicide. Of these, methomyl, parathion, and paraquat were among the most frequently applied agents, while endosulfan and maneb were more scarcely applied in 1988. Records of application of these chemical agents on agricultural land in Kern County were queried from the PUR database for 1988. Reports of other uses not reported at the PLSS section level were excluded (i.e., nonagricultural and structural or indoor treatments).

We linked the PUR to a database of California PLSS sections to remove PUR records erroneously reporting nonexisting sections and to identify records with potential data entry errors, including extremely high application rates (applied pounds ÷ treated acres; CDPR 2000b). For these later records we imputed a new value for applied pounds based on the statewide median application rate for that pesticide. To make the PUR compatible with land-use survey data, we collapsed all nonpermanent field crops, including cotton, tomatoes, potatoes, grains, and alfalfa, into a single class of "field crops."

Land-use surveys. The CDWR (2002) performs countywide, large-scale surveys (1:24,000, or 1 in. = 2,000 ft) of land use and crop cover every 7–10 years. As previously stated, for 1988, the Kern County land-use survey closest in time was conducted in 1990 (Figure 1) and was readily available in digital format (i.e., ArcView GIS shapefile; ESRI, Redlands, CA). In the shapefile, fields, vineyards, orchards, and other land-use types exist as contiguous polygons that are individually linked to their respective attribute information (e.g., land use type, acreage) in a database table. PLSS section boundaries, however, were not included in the data set and were added by merging the land-use data with a shapefile of PLSS sections.

Depending on the type of crop, land-use data 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 field, truck, grain, and pasture crops (e.g., cotton and tomatoes) lead to uncertainty regarding which specific crop was planted in a specific location or point in time (Mitchell et al. 2001). Because of this

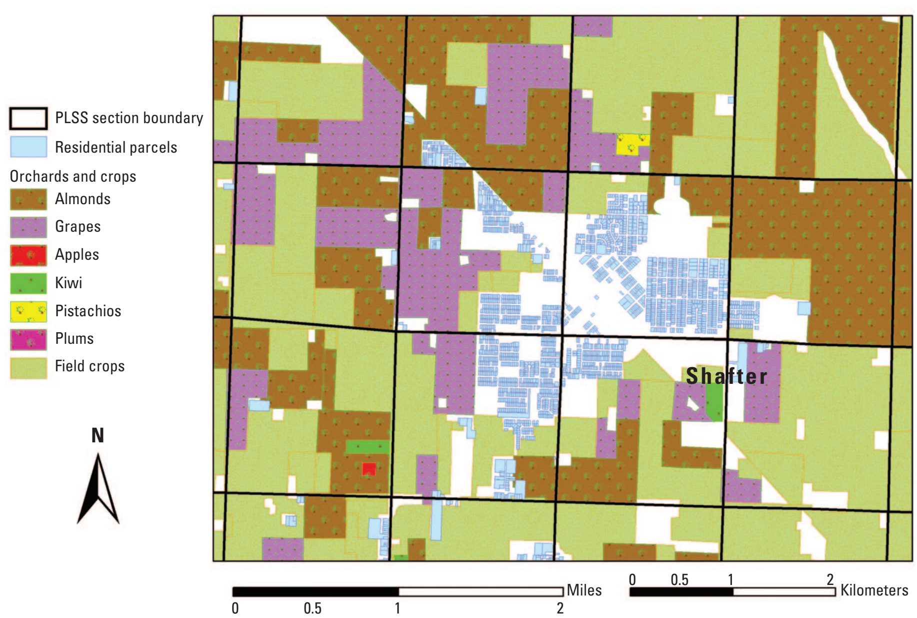

Figure 1. Agricultural land use, Shafter, Kern County, California, 1990.

uncertainty, we decided to combine these crop categories into a class of field crops, assuming that for a reported pesticide application on a specific field, truck, grain, or pasture crop in a PLSS section, all areas with these crop designations were equally likely application sites. After conducting this reclassification, we recalculated acreage estimates for each field, vineyard, or orchard polygon and summed them to obtain PLSS-section-specific estimates of total field crop acreage.

We used a three-tiered approach to link the PUR information to our reclassified land-use survey data. First, when a PUR matched exactly to land-use polygons in a PLSS section by crop type, both records were directly linked. The highest percentages of matches between aggregate PUR and land-use polygons were obtained for maneb (93% of 152 PLSS sections) and paraquat (92% of 2,323), followed by methomyl (90% of 1,713), endosulfan (88% of 257), and parathion (79% of 903). Second, if pesticide use was reported on a crop that did not match any of the crops listed in the land-use survey in a PLSS section, yet the section contained other field, vineyard, or orchard crops, we assumed that these crop locations were possible sites where the reported crop was grown during the years between the available land-use surveys. For example, a PUR for applications on apples may be recorded for a PLSS section, while the land-use survey specifically reported only field, vineyard, and/or other orchard crops (excluding apples); in this case, the PUR was linked to these other crops. This procedure was necessary for 6% of all applications for maneb, 7% for paraquat, 9% for methomyl, 11% for endosulfan, and 20% for parathion.

Third, if we found a PUR for a given PLSS section, but according to the land-use survey no field, vineyard, or orchard crops were present in the section, we assumed that any area within the entire section could have been treated. Although such linkage to an entire section will decrease specificity when assessing exposure proximity, third-tier matches were rare and only necessary for 0.7% (1 polygon) of all maneb applications, 0.8% of endosulfan, 1.1% of paraquat, and 1.3% of methomyl and parathion applications. We calculated annual application rates (total applied pounds ÷ total crop acres in a PLSS section) for all polygons linked to the PUR data.

Residential parcels. We obtained a GIS-shapefile of real-estate properties, or parcels, from the Kern County Assessor (2002) to identify residential locations. This data set maps the locations of all parcels in 1998 (the earliest year parcel data was available in shapefile format) as polygons. We selected all parcels with a residential-use code (e.g., family residences, apartment complexes, convalescent homes). Because residential exposure to even the most commonly applied agricultural pesticides is highly unlikely in an urban area, we intentionally excluded parcels in highly urbanized areas of the county not located near agricultural land (i.e., central Bakersfield). As a result, we restricted the parcels in this simulation to those located in rural western Kern County by selecting those whose geometric centroids were within the area of the available land-use survey for 1990 and within or adjacent to a PLSS section containing an agricultural land-use polygon.

Simulation exercise. We simulated the random selection of population controls for a case–control study by randomly selecting residences from the parcel database described above. We drew 1,000 random samples of 200 addresses and calculated the distribution of exposure to each of the five pesticides selected for the year 1988 based on *a*) both PUR data and land-use maps (the "gold standard"), *b*) PUR data only, and *c*) land-use maps only (i.e., using locations of specific crops as a proxy indicator of application sites for specific pesticides).

Using our gold standard (i.e., the PUR plus land-use model), we called a residential parcel exposed to a pesticide if its geometric centroid was within 500 or 1,000 m of the edge of the nearest field, vineyard, or orchard potentially treated with this pesticide (Figure 2). For comparison (PUR only), we used a zonal exposure model developed by Bell et al. (2001) in which PUR data per PLSS section (without land-use information) is used to determine whether a residence is exposed; i.e., whether it is located within a section ("narrow" exposure definition) with an application reported, or within or adjacent to a section with an application reported ("broad" exposure definition). For each of the samples, we calculated the control exposure prevalences for the gold standard as well as Bell's narrow and broad zonal definitions. We specified three moderate "true" effect sizes [odds ratios (ORs) of 1.5, 2.0, and 3.0] to be observed when using our gold standard and calculated the respective case exposure prevalences as a function of the "true" OR and the gold-standard control prevalence (Appendix).

For each sample, we calculated the respective sensitivities and specificities of the broad and narrow zonal exposure definitions in comparison to the exposure classifications derived from the PUR plus land-use model (Appendix). We then estimated the case exposure prevalences under the broad and narrow definitions as a function of the sensitivity, specificity, and gold-standard case exposure prevalence and calculated the observed ORs using the broad- and narrow-definition case and control exposure prevalences. In addition, we measured the degree of attenuation of the true OR due to nondifferential misclassification (Thompson and Walter 1988).

To simulate residential mobility (i.e., changes in address), fixed proportions (10, 25, and 40%) of residential parcels selected in the first round were resampled from all eligible parcels, and these newly sampled residential addresses were substituted for the original addresses. We assessed exposures for these newly selected parcels using the PUR plus land-use and the PUR-only exposure models and estimated the degree of misclassification

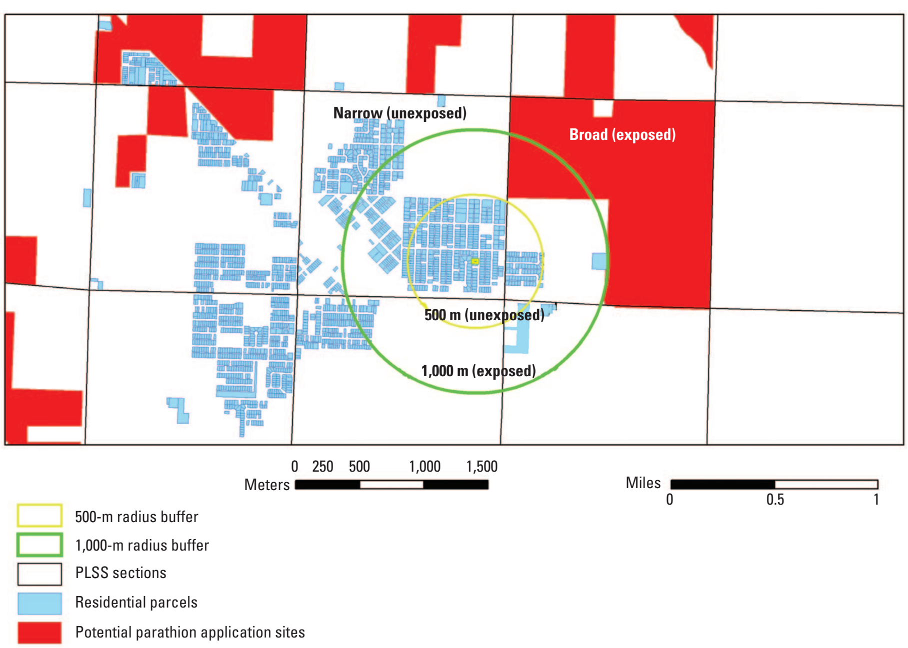

Figure 2. Example of exposure from residential proximity to parathion applications using the distance-based and narrow and broad zonal exposure definitions, 1988.

resulting from residential mobility and changes in exposure status.

Misclassification may also occur if pesticide exposure status is derived only from land-use information in the absence of pesticide-application data. Because parathion was predominantly used on orchard crops (65% of all treated acreage), specifically on almonds, apples, peaches, and nectarines, we evaluated whether the use of land-use information only (i.e., proximity to these orchard crops) would be a good indicator of exposure to this agent in the absence of PUR or any other pesticide-use data compared to the gold standard. With the help of the land-use survey, we identified the locations of the orchard crops on which the agent was predominantly applied, calculated proximity of residences to these orchard crops, and then compared the exposure status derived only from land-use surveys to that derived from the gold standard.

The use of annual aggregates of exposure may lead to misclassification if the exposure period of interest is less than 1 year. We disaggregated the 1988 PUR data into four 3-month seasons (i.e., winter: 1 January–31 March, spring: 1 April–30 June, summer: 1 July–30 September, and autumn: 1 October–31 December) and linked the seasonal PUR to the land-use maps. Seasonal exposure status was assessed using the same buffer-radius sizes (500 and 1,000 m) as the PUR plus land-use model. Using the seasonal exposure model now as the gold standard, we compared the exposure classifications based on annual aggregates using the PUR plus land-use model to those derived from the seasonal models.

Results

Table 1 summarizes the application patterns of the five selected pesticides in western Kern County based on 1988 PUR data and provides the crop acreage, residential proximity to crops, and the crop-specific distributions of pesticide applications. In terms of total treated acreage (from the PUR data, including multiple applications), paraquat was the most commonly applied agent (253,203 acres), followed by methomyl (147,584 acres), parathion (102,517 acres), endosulfan (18,765 acres), and maneb (11,494 acres). Parathion, however, was the most heavily applied pesticide (110,082 lb active ingredient), followed by methomyl (90,777 lb), paraquat (82,895 lb), endosulfan (22,243 lb), and maneb (14,223 lb).

Field crops were the most common crops by acreage in the region, followed by vineyards. Almonds, oranges, and pistachios represented the most common orchard crops by acreage. Among eligible residential parcels (i.e., residences within or adjacent to a section with any crops), more than 40% were within 500 m from any field crops, while only 8 and 9%, respectively, were within this distance from vineyards and almond orchards. We estimated the annual percentage of specific crop acres treated by a specific pesticide by dividing the total crop acres treated (PUR data) by the total acreage of the crop (land-use survey data). Paraquat was used on a variety of crops, including field crops (28% of total acres), grapes (11%), and various other orchard crops. Methomyl was applied on field crops (15%), grapes (29%), oranges (21%), and peaches and nectarines (64%). Parathion was predominantly applied on orchard crops, including almonds (155%), apples (167%), and peaches and nectarines (158%). (Percentages > 100% reflect multiple applications on acreage over the course of a year, whereas percentages < 100% may also reflect multiple applications.) Endosulfan and maneb were applied on grapes (12 and 6%, respectively) and, to a lesser extent, on field crops (approximately 1%).

Using the PUR data only, we further examined the distribution of pesticide applications on field crops (Table 2). Based on the total acres treated as reported in the PUR, we observed that paraquat treatments almost exclusively occurred on cotton, whereas for endosulfan the predominant field crop applications were on alfalfa and lettuce. Maneb was applied on flowers, lettuce, and potatoes; methomyl on alfalfa, cotton, and lettuce; and parathion on lettuce and onions.

In Table 3, we list the mean exposure prevalences for 200 randomly selected Kern County residences in 1988 and each selected pesticide for our own PUR- and land-use-based model and the two different exposure definitions previously used by Bell et al. (2001). For our PUR- and land-use-based model with a 500-m radius around each residence, annual exposure prevalences ranged from 0.9% for maneb to 17% for methomyl. Under the narrow definition used by Bell et al. (2001), based only on PUR data and PLSS

Table 2. Percentage of all field crop acres treated by specific pesticide.[a]

Field crop	Endosulfan (7,754)[b]	Maneb (6,474)	Methomyl (109,023)	Paraquat (196,213)	Parathion (35,681)
Alfalfa	46.3	NR	39.2	3.4	2.4
Beans	1.5	NR	5.3	NR	4.2
Carrots	NR	7.6	NR	NR	7.4
Cotton	1.3	NR	12.3	94.4	1.1
Flowers	NR	26.0	1.4	NR	NR
Lettuce (head)	29.4	34.5	28.9	NR	36.3
Melons	2.7	NR	1.3	NR	NR
Onions	NR	3.6	NR	NR	43.6
Peppers (bell)	3.6	NR	NR	NR	NR
Potatoes	5.6	19.2	3.5	NR	NR
Squash	5.8	NR	NR	NR	NR
Sugarbeet	NR	NR	NR	NR	3.6
Tomatoes	NR	5.1	NR	NR	NR
Turnips	NR	2.3	NR	NR	NR
Watermelon	3.0	NR	2.4	NR	NR

NR, no reported pesticide applications on the field crop.
[a]Crop acres treated by pesticide (PUR) ÷ total field crop acres treated (PUR). [b]Values in parentheses are total number of field acres treated.

Table 1. Summary of pesticide applications and frequently treated crop types: residential proximity, application distribution, and rates, western Kern County, 1988.

Crop	Percent of residences within 500 m[a]	Total crop acres (land-use data)	Percent of crop acres treated (application rate)[b]				
			Paraquat	Parathion	Methomyl	Endosulfan	Maneb
Field crops	41.1	709,919	27.6 (0.27)	5.0 (0.43)	15.4 (0.55)	1.1 (0.86)	0.9 (1.30)
Grapes	7.7	91,944	11.2 (0.59)	11.4 (0.81)	29.0 (0.78)	12.0 (1.42)	5.5 (1.15)
Almonds	8.9	88,553	29.2 (0.53)	155.4 (1.26)	NR	NR	NR
Oranges	3.0	40,027	9.4 (0.39)	2.5 (1.68)	20.7 (0.81)	NR	NR
Pistachios	1.0	24,207	33.6 (0.44)	NR	NR	NR	NR
Apples	0.5	5,646	36.4 (0.77)	166.5 (0.93)	NR	NR	NR
Peaches and nectarines	0.6	4,422	87.9 (0.42)	157.5 (1.02)	63.6 (0.86)	NR	NR
Plums	0.9	3,361	25.9 (0.32)	NR	NR	NR	NR

NR, no reported pesticide applications on the crop.
[a]All residential parcels (1998) within or adjacent to PLSS sections containing agricultural-use polygons (1990; n = 77,619 out of 105,330 total). [b]Crop acres treated by pesticide (PUR) ÷ total crop acres (land use); total pounds applied (PUR) ÷ total crop-section acres (land use).

sections, prevalences ranged from 1 to 7%, but under the broad definition, the prevalence range expanded to between 7 and 49%. Comparing sensitivity and specificity using our gold standard, we observed that the narrow zonal definition had relatively low sensitivity (between 37 and 54%) that tended to rise with decreasing exposure prevalence, but almost perfect specificity (approximately 99%) for each pesticide. Perfect sensitivity but much lower specificity (between 62 and 94%) was observed under the broad definition, and specificity rose with decreasing exposure prevalence.

We used the prevalence, sensitivity, and specificity estimates derived from the comparison of the PUR- and land-use-based model (our gold standard) to the Bell et al. (2001) model to assess the degree of bias from nondifferential misclassification of three presumed true effect estimates (Table 4). Observed ORs tended to be the most attenuated when using the broad definition of exposure (due to low specificity), for larger true ORs, and for the least commonly applied pesticides. Under the narrow definition (with near-perfect specificity), observed ORs were less attenuated than those in the broad definition, but attenuation of the true ORs ranged between 34 and 56% for pesticides less commonly used (i.e., endosulfan and maneb).

We then estimated the impact of residential mobility on exposure classification using our PUR- and land-use-based model. For paraquat applications and residences within a 500-m radius, sensitivity rapidly declined with increasing mobility: from 91% at 10% mobility to 68% at 40% mobility. Specificity, however, was nearly perfect and declined only slightly with mobility from 99 to 97%. Similar trends were observed for the other pesticides using a 1,000-m radius around each residence. Note that the results observed in this simulation exercise were identical to estimates derived from formulas given by Khoury et al. (1988) in which sensitivity and specificity are treated as functions of the population mobility rate and the exposure prevalence.

The ORs observed after accounting for nondifferential misclassification as a result of residential mobility are shown in Table 5. Attenuation increased with an increase in both the mobility rate and true OR (e.g., from 96% with 10% mobility and true OR = 1.5 to 73% with 40% mobility and true OR = 3.0). In addition, we assessed the extent of misclassification occurring as a result of residential mobility when using the narrow or broad PUR-only zonal exposure models. Compared to the observed ORs under the assumption of no residential mobility in either the narrow or broad definitions, effect estimates were generally more attenuated with increasing mobility. In fact, for the broadly classified exposure setting a mobility rate of 40% diminished the true OR by 80%.

For parathion, we created another exposure definition based only on land-use survey data (Table 6). When exposure was defined as almonds, apples, or peaches and nectarines grown within a 500-m radius around a residence, we observed a sensitivity of 60% and a specificity of 94% compared to the PUR plus land-use model. When the radius was expanded to 1,000 m, sensitivity increased to 72% and specificity fell to 87%. At both distances, the true ORs were attenuated by approximately 57–59%.

We evaluated the validity of annual pesticide exposure status as a proxy for classifying seasonal exposure using our PUR plus land-use model with a 500-m radius. Because some of the parcel samples for the less commonly applied pesticides endosulfan and maneb had zero annually-exposed residences, we conducted these comparisons only for the more commonly applied agents: parathion, paraquat, and methomyl. The prevalence of exposure varied by season, with different trends of peak exposure prevalence for each of the pesticides (Figure 3). Exposure to parathion was most likely to occur in the winter (5.8% prevalence) and least likely in the spring (0.9%). Paraquat exposure was most likely to occur in the fall (6.1%) and least likely in the spring (0.8%), and methomyl exposure was most likely in the

Table 3. Simulated estimates (percentage) based on 1,000 replicates of 200 randomly sampled residential parcels.

	Annual exposure prevalence ± SD			Sensitivity and specificity ± SD of zonal PUR-only model vs. PUR/land-use model			
	PUR/land-use	PUR-only model		Narrow[b]		Broad[c]	
Pesticide	model 500 m[a]	Narrow[b]	Broad[c]	Sensitivity	Specificity	Sensitivity	Specificity
Methomyl	17.1 ± 2.6	7.0 ± 1.8	48.6 ± 3.5	36.9 ± 8.4	99.1 ± 0.7	100.0 ± 0	62.0 ± 3.7
Paraquat	10.8 ± 2.3	4.5 ± 1.5	36.2 ± 3.4	35.3 ± 10.6	99.3 ± 0.6	100.0 ± 0	71.5 ± 3.4
Parathion	8.4 ± 2.0	5.0 ± 1.5	27.1 ± 3.2	45.4 ± 12.9	98.7 ± 0.8	100.0 ± 0	79.6 ± 3.0
Endosulfan	5.3 ± 1.7	3.2 ± 1.3	24.5 ± 3.0	42.8 ± 16.0	99.0 ± 0.7	100.0 ± 0	79.7 ± 2.9
Maneb	0.9 ± 0.7	1.0 ± 0.7	6.9 ± 1.8	54.8 ± 38.9	99.4 ± 0.5	100.0 ± 0	93.9 ± 1.7

[a]Residential buffer radius. [b]Residence within a PLSS section with a reported pesticide application. [c]Residence within or adjacent to a PLSS section with a reported application.

Table 4. Matrix of simulated OR estimates (percent attenuation) based on a true OR and the prevalence, sensitivity, and specificity estimates presented in Table 3.

	True OR = 1.5		True OR = 2		True OR = 3	
Pesticide	Narrow	Broad	Narrow	Broad	Narrow	Broad
Methomyl	1.37 (26)	1.18 (64)	1.70 (30)	1.35 (65)	2.27 (37)	1.70 (65)
Paraquat	1.38 (24)	1.15 (70)	1.73 (27)	1.30 (70)	2.36 (32)	1.60 (70)
Parathion	1.35 (30)	1.15 (70)	1.69 (31)	1.31 (69)	2.32 (34)	1.62 (69)
Endosulfan	1.33 (34)	1.11 (78)	1.66 (34)	1.22 (78)	2.27 (37)	1.44 (78)
Maneb	1.23 (54)	1.06 (88)	1.45 (55)	1.12 (88)	1.89 (56)	1.25 (88)

Table 5. Matrix of simulated OR estimates (percent attenuation) for paraquat based on Table 3[a] and rate of residential mobility.[b]

	True OR = 1.5			True OR = 2.0			True OR = 3.0		
Mobility rate	PUR/LU	Narrow	Broad	PUR/LU	Narrow	Broad	PUR/LU	Narrow	Broad
0%	1.50 (0)	1.38 (25)	1.15 (70)	2.00 (0)	1.73 (27)	1.30 (70)	3.00 (0)	2.36 (32)	1.60 (70)
10%	1.45 (11)	1.34 (33)	1.13 (74)	1.89 (11)	1.65 (35)	1.26 (74)	2.76 (12)	2.21 (40)	1.52 (74)
25%	1.38 (24)	1.30 (40)	1.12 (77)	1.75 (25)	1.57 (43)	1.23 (77)	2.47 (26)	2.07 (47)	1.45 (77)
40%	1.31 (37)	1.26 (48)	1.10 (80)	1.62 (38)	1.50 (50)	1.20 (80)	2.19 (40)	1.92 (54)	1.39 (80)

[a]Gold standard: PUR/land-use model with 0% residential mobility rate. [b]Mobility-adjusted sensitivity and specificity estimates with SDs for the PUR/LU model: 10% mobility: 90.8% (6.0%) sensitivity, 98.9% (0.8%) specificity; 25% mobility: 79.3% (8.9%) sensitivity, 97.6% (1.1%) specificity; 40% mobility: 67.8% (10.2%) sensitivity, 96.2% (1.4%) specificity.

Table 6. Parathion exposure prevalence, sensitivity, and specificity ± SD and simulated OR estimates and attenuation, land-use-only model vs. PUR/land-use model.[a]

Buffer radius	PUR/land-use model (%)	Land-use-only model (%)	Sensitivity (%)	Specificity (%)
500 m	8.4 ± 2.0	10.5 ± 2.1	60.1 ± 12.3	94.0 ± 1.8
1,000 m	15.8 ± 2.6	22.3 ± 2.9	72.2 ± 8.0	87.0 ± 2.6

[a]Based on proximity to crops typically treated with parathion (land-use data only): apples, almonds, peaches/nectarines, plums, pears, and apricots. For the true OR of 1.5 at both buffer radii, the observed OR and attenuation were 1.22 and 57%, respectively; for the true OR of 2.5, the observed OR and attenuation were 1.43 and 57%, respectively; and for the true OR of 3.0, the observed OR and attenuation were 1.82 and 59%, respectively.

summer (10.0%) and least likely in winter (2.4%). With seasonal exposure classification as the gold standard, annual exposure status had 100% sensitivity. Specificity, however, ranged from 85.3% for methomyl in the winter season to 97.4% for parathion in the winter and increased with seasonal exposure prevalence for each pesticide. The observed seasonal ORs followed the exposure prevalence trend, with the greatest attenuation occurring in the seasons with the smallest exposure prevalence and thus the lowest specificity (Figure 4). The greatest attenuation of the true OR was observed for paraquat in the spring (92%).

Discussion

In this simulation exercise, we found that the effect estimates were less attenuated if exposure status for a residence was defined using a smaller and thus more restricted geographic area yielding 100% specificity rather than relying on a larger area to increase sensitivity. This result is expected because high specificity leads to less attenuation of the effect estimate when the true prevalence of exposure is low in the population (Kelsey et al. 1996). Accordingly, Bell et al. (2001) reported stronger elevated ORs for different pesticide physicochemical groups and fetal deaths from congenital anomalies when using a narrower definition of exposure.

We evaluated the effect of changes in residence (i.e., residential mobility) on misclassification error and observed that, for relatively rare exposures to pesticides, when the residential mobility rate is low, the result is a high specificity with minimal attenuation of the true OR. Furthermore, nondifferential exposure misclassification increases with the residential mobility rate and prevalence of exposure in the population. For the purpose of this exercise, our model for residential mobility limited the geographic extent of address changes to that of all eligible parcels (i.e., near agricultural land in western Kern County). In reality, individuals may also choose to move into local metropolitan areas (e.g., Bakersfield) or urban or rural regions outside Kern County with no or unknown opportunity for environmental pesticide exposure. If residents tended to move to residential areas with a lower likelihood of exposure to agricultural pesticides, we would expect the exposure classification based on the incorrect rural addresses to have lower specificity.

Pesticide active ingredients are used on a wide variety of crops to treat pests and diseases. In the absence of detailed records of pesticide applications (as is the case outside of California), the utility of crop maps for identifying the locations of specific crops at a point in time is limited by the fact that nonorchard crops are often seasonally or annually rotated (e.g., cotton and tomatoes in California) and the 7–10 year gap between land-use surveys. Even for parathion, an agent predominantly applied on specific orchard crops, we observed 60% sensitivity, 94% specificity, and attenuation > 50% for the true OR when defining exposure simply as residence within 500 m of these specific orchards, compared to the PUR plus land-use model. Remotely-sensed satellite imagery can be used to produce crop maps with higher temporal resolution, but its current spatial resolution is limited by the fact that distinctions between crop types are not as clear as ground-verified land-use surveys.

Parathion was also applied on approximately 5% (35,496 acres) of all field crops (especially lettuce and onions) and 11% (10,298 acres) of all vineyard acreage. As a result, residences would be incorrectly classified as unexposed if they were located within the buffer-radius distance of field crops, vineyards, or other orchards where parathion may have been applied but were not included in the land-use-only exposure model or if the residences were located beyond the buffer-radius distance from almond, apple, peach, and nectarine orchards.

We aggregated PUR data to generate annual estimates of pesticide use by section and crop. These annual summaries, however, obscure seasonal application patterns. For example, we observed that parathion was predominantly applied in the winter months, and, correspondingly, the exposure prevalence was highest during this time of year. Using exposure estimates based on annual summaries of parathion use as a proxy for seasonal exposure, effect estimates were most strongly attenuated in the seasons when the agent was least applied. While such annual summaries may be useful for assessing long-term historical exposures, they will be insufficient for assessing short-term seasonal exposures for such outcomes as congenital malformations, acute pesticide-related illnesses, or asthma attacks and other respiratory symptoms. Therefore, by not taking into account seasonal trends in pesticide use, studies of such shorter-term outcomes using proximity to agricultural land-use as a proxy indicator for pesticide exposure and ignoring the timing of applications will reduce specificity for exposure and thus attenuate effect estimates.

There are several limitations to our PUR plus land-use exposures model that affect its validity as the gold standard in our simulation exercise. We used two existing data sets developed by independent state agencies for different purposes and with different frequencies of data collection. Individual PUR records represent individual pesticide applications with information on the date of the event, whereas

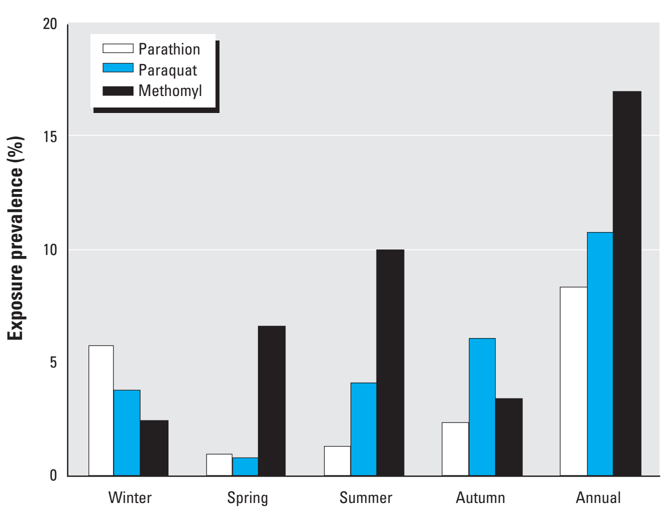

Figure 3. Seasonal and annual exposure prevalence by pesticide, random samples of 200 residential parcels, 1988.

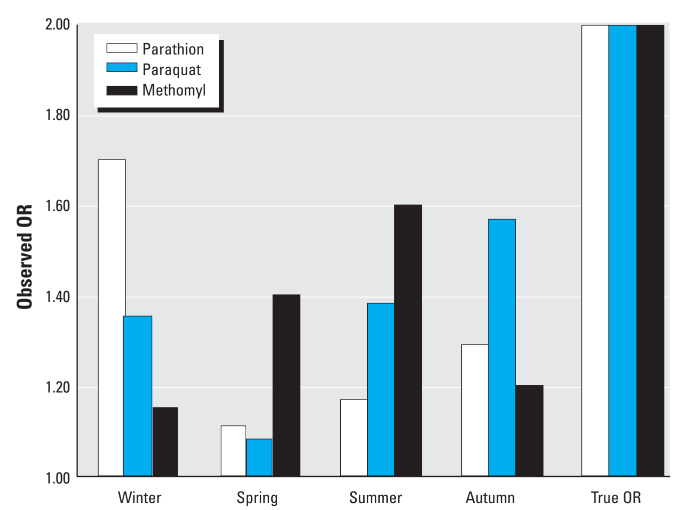

Figure 4. Observed seasonal ORs using annual exposure status as a proxy for seasonal exposure, true seasonal OR = 2.

land-use surveys are conducted in the summer every 7–10 years. For the purpose of our simulation exercise, we used the 1990 Kern County land-use survey to approximate land use for 1988. Crop locations indicated in this survey were assumed to be the same as the years immediately before and after the year of the survey. For each of the five pesticides chosen, we found a high degree of exact (first tier) matching between what crops were reported for applications in the PUR and the land-use data documenting crop types within a section (Miller et al. 2002a). Mismatches between PUR and land-use survey crops in a section most likely resulted from land-use changes that occurred between the years the Kern County surveys were conducted (i.e., 1984 and 1990), including urban or suburban development expanding into agricultural areas as well as short-term, market-driven changes in crop production.

Our method for linking the PUR and land-use data collapses all field and vegetable crops into a single category to acknowledge our uncertainty regarding crop rotations that took place between surveys. As a result, erroneous matches between PUR and land-use polygons in a section could occur if the collapsing of field crops obscures true discordances between the PUR and land-use data. This may occur, for example, if the PUR reports a treatment on carrots in a section, but the land-use survey instead reports tomatoes and potatoes. Additionally, in sections of exact matches between PUR and land-use data, fields or orchards may enlarge, shrink, or change location within a section during the years between surveys and thus lead to invalid (first-tier) matches. As a consequence of these mismatches and erroneous PUR and land-use matches, we may have been overconfident in the accuracy of our gold standard and incorrectly classified residences as being exposed or unexposed. We would expect reduced accuracy in PLSS sections with a wide variety of field and vegetable crops, especially if the distribution of treated acres for a given pesticide varies considerably for these different crops.

We arbitrarily selected our buffer radius as 500 m, a distance that previous studies using historical data to assess pesticide exposures defined as an intermediate distance range for nonspecific pesticide applications (Ward et al. 2000). Furthermore, the perfectly circular shape of the buffer attempts to capture an individual's potential for exposure from pesticide drift. Only agents applied within the buffer are assumed to have the potential to drift to the residence at the center; those applied outside the buffer are assumed not to drift to the residence. The fate of pesticides in the environment, including drift potential, however, depends on several factors, including *a*) persistence (i.e., half-life) of the pesticide in the environment; *b*) solvents and adjuvants that may increase the adherence of the active ingredient to the soil or target crop or increase the volatilization of the active ingredient once released into the environment; *c*) application method (ground or air), equipment (e.g., aircraft, nozzle size), and droplet size; *d*) wind speed and direction at the time of application; and *e*) weather conditions, including temperature, sunlight, and precipitation (Menzie 1972).

A heavy rain after a ground or aerial application, for example, may eliminate the applied pesticide from the air, thus preventing exposure to the agent for residences near the application site. As a result, these residences would be misclassified as exposed. A strong wind may carry a pesticide far beyond 1,000 m (Zabik and Seiber 1993). Residences located downwind from the application site would be misclassified as unexposed if no other applications of the pesticide occurred within the residential buffer. As a result, ignoring the potential environmental fate of agents will lead to further nondifferential misclassification of exposure.

Our simulation exercise was based on a dichotomous exposure model that defines exposure as any specific pesticide use within the buffer around the residential parcel, including PUR-polygon matches where the application is slightly greater than zero. The exposed groups for each pesticide may include truly exposed residences within a certain distance of heavy applications as well as those with negligible opportunity for exposure that are near agricultural areas treated with minute amounts of pesticides. In this case, the exposure prevalences in the various models used in the simulation may be inflated. Although we intend to quantify potential pesticide exposure for subjects in our future studies, this approach was too computationally intensive for a simulation exercise involving more than 77,000 parcels. In our quantitative model, a map of pesticide application rates will be transformed from a vector (i.e., polygon) to a raster (i.e., grid cell or pixel) format (Huxhold 1991). Pixels will be assigned the value of the application rate at the pixel center, or zero for pixels in areas reporting no pesticide applications. Within each residential buffer, the application-rate values within the buffer zone will be averaged to estimate an annual intensity score for the respective residence. We will then be able to rank potential pesticide exposure for residences as a function of distance from and intensity of agricultural pesticide applications. Subsequently, we will estimate total exposure over several years.

The parcel data set we used to select residential locations represented all parcels in 1998 and thus included parcels that did not exist in 1988 and excluded those that existed in 1988 but were razed or converted for use before 1998. As a result, there may be some parcels included in the samples that did not exist in 1988 but were classified as either exposed or unexposed. One advantage of using parcel data, however, is that we avoided the mapping errors associated with geocoding procedures that interpolate addresses within a range of street numbers and potentially place residences in imprecise locations. This interpolation method, which is the default geocoding procedure commonly used in GIS software packages, may have limited utility in rural areas (Ratcliffe 2001). 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 separated from the street by a crop field, vineyard, or orchard (Ward et al. 2000). We created buffers around the geometric centroids instead of the parcel boundaries because we wanted to generate buffers of equal distance as well as area for each parcel. Although this decision may potentially lead to exposure misclassification for larger parcels and smaller buffer sizes (e.g., 100 m), we did not expect this to be a problem for the prespecified buffer sizes in this simulation.

Comprehensive assessments of pesticide exposure examine multiple sources, including voluntary occupational and residential use and involuntary environmental exposure to specific agents. California's PUR and land-use survey databases offer a unique opportunity to model historical environmental exposure. Despite the limitations of these data, including uncertainty of the locations of specific field crops and the long periods of time between land-use surveys, our PUR-plus land-use model represents a more comprehensive approach over previously developed methods for assessing historical pesticide exposures from residential proximity. Detailed reporting data are a critical and necessary (yet not sufficient) component of residential level exposure assessment in epidemiologic studies examining potential health effects from agricultural drift. Our simulation results indicate that in the absence of such information, including the likely application sites, substantial nondifferential exposure misclassification may occur, thus leading to bias and attenuation of true effect estimates. Due to the complex nature of pesticide environmental fate and exposure, models based on these data should incorporate additional information, including solvents and adjuvants mixed with the active ingredient, weather, wind speed and direction, and the application method. In addition, the model should appropriately cover the hypothesized critical exposure period and be used in conjunction with detailed residential histories.

Appendix

The exposure prevalence among cases (P_1) from the gold standard exposure model (i.e., PUR plus land use), based on the control exposure prevalence (P_0) and true odds ratios (OR) (Greenland and Rothman 1998) is

$$P_1 = \frac{1}{\{1 + [(1 - P_0)/\mathrm{OR}(P_0)]\}}.$$

The case exposure prevalence (P_{obs1}) from the comparison exposure models (e.g., broad definition), based on P_1 and the sensitivity (SE) and specificity (SP) of the comparison model (Goldberg 1975) is

$$P_{\mathrm{obs1}} = [P_1(\mathrm{SE})] + [(1 - P_1)(1 - \mathrm{SP})].$$

The observed odds ratio ($\mathrm{OR}_{\mathrm{obs}}$) using the comparison exposure model, based on P_{obs1} and the control exposure prevalence (P_{obs0}) (Greenland and Rothman 1998) is

$$\mathrm{OR}_{\mathrm{obs}} = \frac{P_{\mathrm{obs1}}/(1 - P_{\mathrm{obs1}})}{P_{\mathrm{obs0}}/(1 - P_{\mathrm{obs0}})}.$$

The attenuation (Att%) of the true OR, based on $\mathrm{OR}_{\mathrm{obs}}$ (Thompson and Walter 1988) is

$$\mathrm{Att\%} = 1 - \frac{\mathrm{OR}_{\mathrm{obs}} - 1}{\mathrm{OR} - 1}.$$

References

Ames RG, Howd RA, Doherty L. 1993. Community exposure to a paraquat drift. Arch Environ Health 48:47–52.

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

Byass JB, Lake JR. 1977. Spray drift from a tractor-powered field sprayer. Pestic Sci 8:117–126.

CDPR. 1999. Pesticide Use Report Form. Sacramento, CA:California Department of Pesticide Regulation. Available: http://www.cdpr.ca.gov/docs/enfcmpli/prenffrm/enf025.pdf [accessed 1 October 2002].

———. 2000a. Summary of Pesticide Use Report Data: 2000. Sacramento, CA:California Department of Pesticide Regulation.

———. 2000b. Pesticide Use Reporting: An Overview of California's Unique Full Reporting System. Sacramento, CA:California Department of Pesticide Regulation. Available: http://www.cdpr.ca.gov/docs/pur/purovrvw/ovr52000.pdf [accessed 1 October 2002].

CDWR. 2002. Land Use Surveys. Sacramento, CA:California Department of Water Resources. Available: http://www.waterplan.water.ca.gov/landwateruse/landuse/lusurveys.htm [accessed 1 October 2002].

Chester G, Ward RJ. 1984. Occupational exposure and drift hazard during aerial application of paraquat to cotton. Arch Environ Contam Toxicol 13:551–563.

Clary T, Ritz B. 2003. Pancreatic cancer mortality and pesticide use in California. Am J Ind Med 43:306–313.

Currier WW, MacCollom GB, Baumann GL. 1982. Drift residues of air-applied carbaryl in an orchard environment. J Econ Entomol 75:1062–1068.

Ecobichon DJ, Joy RM. 1994. Pesticides and Neurological Diseases. 2nd ed. Boca Raton, FL:CRC Press.

Engel LS, Checkoway H, Keifer MC, Seixas NS, Longstreth WT Jr, Scott KC, et al. 2001. Parkinsonism and occupational exposure to pesticides. Occup Environ Med 58:582–589.

Frost KR, Ware GW. 1970. Pesticide drift from aerial and ground applications. Agric Eng 51:460–467.

Goldberg JD. 1975. The effects of misclassification on the bias in the difference between two proportions and the relative odds in the fourfold table. J Am Statist Assoc 123:736–751.

Greenland S, Rothman KJ. 1998. Measures of effect and association. In: Modern Epidemiology (Rothman KJ, Greenland S, eds). Philadelphia:Lipincott-Raven Publishers, 253–279.

Gunier RB, Harnly ME, Reynolds P, Hertz A, Von Behren J. 2001. Agricultural pesticide use in California: pesticide prioritization, use densities, and population distributions for a childhood cancer study. Environ Health Perspect 109:1071–1078.

Huxhold WE. 1991. An Introduction of Urban Geographic Information Systems. Oxford, UK:Oxford University Press.

Kelsey JL, Whittemore AS, Evans AS, Thompson WD. 1996. Methods in Observational Epidemiology. 2nd ed. New York:Oxford University Press.

Kern County Assessor. 2002. About the Assessor's Office. Bakersfield, CA:Kern County Government Services. Available: http://assessor.co.kern.ca.us/assessor/about.cfm [accessed 1 October 2002].

Khoury MJ, Stewart W, Weinstein A, Panny S, Lindsay P, Eisenberg M. 1988. Residential mobility during pregnancy: implications for environmental teratogenesis. J Clin Epidemiol 41:15–20.

Loewenherz C, Fenske RA, Simcox NJ, Bellamy G, Kalman D. 1997. Biological monitoring of organophosphorus pesticide exposure among children of agricultural workers in central Washington State. Environ Health Perspect 105:1344–1353.

MacCollom GB, Currier WW, Baumann GL. 1986. Drift comparisons between aerial and ground orchard application. J Econ Entomol 79:459–464.

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

Miller RS, Nuckols JR, Gunier RB, Hertz A, Reynolds P. 2002a. Assessing the spatial accuracy of the California Pesticide Use Reporting database for use in exposure assessment studies [Abstract]. In: Linking Exposures and Health: Innovations and Interactions. Proceedings of the 12th Conference of the ISEA and the 14th Conference of the ISEE, 11–15 August 2002, Vancouver, BC. Boston:JSI Research and Training Institute, 39.16.

Miller RS, Nuckols JR, Gunier RB, Hertz A, Ward MH, Reynolds P. 2002b. Linking the California Pesticide Use Reporting database with spatial land use data for exposure assessment [Abstract]. In: Linking Exposures and Health: Innovations and Interactions. Proceedings of the 12th Conference of the ISEA and the 14th Conference of the ISEE, 11–15 August 2002, Vancouver, BC. Boston:JSI Research and Training Institute, 39.15.

Mills PK. 1998. Correlation analysis of pesticide use data and cancer incidence rates in California counties. Arch Environ Health 53:410–413.

Mitchell JP, Lanini WT, Miyao EM, Brostrom PN, Herrero EV, Jackson J, et al. 2001. Growing processing tomatoes with less tillage in California. In: Acta Hort 542 (Hartz TK, ed). VII International Symposium on the Processing Tomato, 10–13 June 2000, Sacramento, CA. Leuven, Belgium:ISHS, 347–354.

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

Reynolds P, Von Behren J, Gunier RB, Goldberg DE, Hertz A, Harnly ME. 2002. Childhood cancer and agricultural pesticide use: an ecologic study in California. Environ Health Perspect 110:319–324.

Ritz B, Yu F. 2000. Parkinson's disease mortality and pesticide exposure in California 1984-1994. Int J Epidemiol 29:323–329.

Shaw GM, Velie EM, Katz EA, Morland KB, Schaffer DM, Nelson V. 1999. Maternal occupational and hobby chemical exposures as risk factors for neural tube defects. Epidemiology 10:124–129.

Thompson WD, Walter SD. 1988. A reappraisal of the kappa coefficient. J Clin Epidemiol 41:949–958.

USDA. 1997. Census of Agriculture, 1997. Washington, DC:U.S. Department of Agriculture, National Agricultural Statistics Service.

Ward MH, Nuckols JR, Weigel SJ, Maxwell SK, Cantor KP, Miller RS. 2000. Identifying populations potentially exposed to agricultural pesticides using remote sensing and a geographic information system. Environ Health Perspect 108:5–12.

Woods N, Craig IP, Dorr G, Young B. 2001. Spray drift of pesticides arising from aerial application in cotton. J Environ Qual 30:697–701.

Zabik JM, Seiber JN. 1993. Atmospheric transport of organophosphate pesticides from California's Central Valley to the Sierra Nevada Mountains. J Environ Qual 22:80–90.

Zahm SH, Ward MH, Blair A. 1997. Pesticides and cancer. Occup Med 12:269–289.