Journal of Exposure Analysis and Environmental Epidemiology (2001) 11, 56–65


www.nature.com/jea

On predicting multi-route and multimedia residential exposure to chlorpyrifos and diazinon

D.J. MOSCHANDREAS,[a] S. KARUCHIT,[a] Y. KIM,[a] H. ARI,[a] M.D. LEBOWITZ,[b] M.K. O'ROURKE,[b] S. GORDON[c] AND G. ROBERTSON[d]

[a]*Illinois Institute of Technology, Chicago, Illinois*
[b]*University of Arizona, Tucson, Arizona*
[c]*Battelle, Columbus, Ohio*
[d]*U.S. EPA, Las Vegas, Nevada*

This paper formulates regression models and examines their ability to associate exposures to chlorpyrifos and diazinon in residences with information obtained from questionnaires and environmental sampling of the National Human Exposure Assessment Survey Arizona (NHEXAS-AZ) database. A knowledge-based list of 29 potential exposure determinants was assembled from information obtained from six questionnaires administered in the course of the study. This list was used to select the independent variables of each model statistically and electronically. Depending on the data type of dependent and independent variables, four classes of regression models were developed to determine desired associations. Route-specific exposures were estimated using the indirect method of exposure estimation and measurements from the NHEXAS-AZ field study. The stepwise procedure was used to construct regression models. Significance level at $P=0.10$ was used for entry and retention of independent variables in a model. Twelve significant regression models were formulated to quantify associations among exposures and other variables in the NHEXAS-AZ database. Route-specific exposures to pesticides associate significantly with questionnaire-based variables such as preparation of pesticides, use of pesticide inside the house, and income level; and with concentration variables in three media: dermal wipe, sill wipe, and indoor air. Models formulated in this study may be used to estimate exposures to each of the pesticides. Yet, the use of these models must incorporate clear statements of the assumptions made in the formulation as well as the coefficient of determination and the confidence and prediction intervals of the dependent variable. *Journal of Exposure Analysis and Environmental Epidemiology* (2001) **11,** 56–65.

Keywords: *chlorpyrifos, diazinon, exposure determinants, NHEXAS study, regression model, residential exposure.*

Introduction

Exposure assessment scientists and scientists of other related fields such as industrial hygiene, environmental health, toxicology, environmental policy, and epidemiology use the term exposure to describe different ideas or processes (Zartarian et al., 1997). To avoid the resulting confusion, we define exposure to denote an event that occurs when there is contact at a boundary between a human receptor and a pollutant at a certain concentration for a certain interval of time (NAS, 1991). Moreover, in this paper, the boundary of contact refers to the receptor nose, mouth, skin, and eyes; thus, the boundary of contact is the visible exterior of the human receptor. Dose, a different but related concept, denotes the amount of a substance available for interaction with metabolic processes or biologically significant receptors *after crossing* the outer boundary. Dose is not addressed in this paper.

The National Human Exposure Assessment Survey (NHEXAS) investigates exposures to multiple pollutants in multiple media and multiple routes of exposure. NHEXAS employed a multi-stage probability-based experimental design. The study objectives were to formulate exposure distributions of subject contaminants, identify the upper 10^{th} percentile of such distributions, provide baseline information for future studies, evaluate and, if possible, rank exposure pathways, identify predictors of human exposure and compare estimated exposures with biomarker levels (Lebowitz et al., 1995). This paper focuses on multimedia, multi-route exposures to two pesticides, chlorpyrifos and diazinon, and employs the NHEXAS database generated in Arizona (NHEXAS-AZ).

Pesticides are natural or synthetic substances applied to control, destroy, repel and mitigate pests. Over a billion pounds of synthetic pesticides are used in the USA, 69% of the estimated billion is herbicides, 19% is insecticides, and 12% is fungicides (Kupchella and Hyland, 1992). Pesticides are applied indoors to protect residents from pests (Hodgson and Levi, 1996). A national survey

1. Address all correspondence to: Prof. Demetrios J. Moschandreas, PhD, Department of Chemical and Environmental Engineering, Illinois Institute of Technology, Perlstein Hall, Room 127, 10 West 33rd St., Chicago, IL 60616. Tel.: +1-312-567-3532. Fax: +1-312-567-8874. E-mail: djm@iit.edu

conducted by the U.S. EPA revealed that 90% of the U.S. households use pesticides; 83.7% of the users used them inside the house, 21.4% in the garden, and 28.7% on the lawn (U.S. EPA, 1979). Moreover, in a study of family pesticide use in the home, garden, orchard, and yard in 238 Missouri families, nearly all families used pesticides at least once a year and two-thirds used pesticides more than five times per year (Davis et al., 1992). Mixing and spraying pesticides, eating food or drinking water containing pesticide residues, breathing airborne pesticide particles and vapors, and coming in contact with surfaces on which pesticides have been deposited are activities that cause human exposure to pesticides mostly in the residential microenvironment. Consequently, exposure to pesticides in residences, the subject of this paper, appears to be the largest component of the exposure to pesticides in all nonoccupational microenvironments.

Two of the most widely used pesticides are chlorpyrifos and diazinon. Chlorpyrifos is an organophosphorus compound used as broad-spectrum insecticide. It is placed in cracks and crevices and is used as a spot treatment for roaches, fleas, and termites; chlorpyrifos is also found in pet flea and tick collars and shampoos. Diazinon, a contact organophosphorus insecticide, is widely used outdoors to control turf and garden soil insects. It is also used in agriculture and as a drug in veterinary medicine. Thus, the major source of diazinon residues is in edible crops, meat, offal, and other animal products. Diazinon residues in vegetables, fruits, and animal products are very low; in fact, all environmental levels of diazinon are low (WHO, 1998).

In this paper, we formulate regression models and examine their ability to associate exposures to chlorpyrifos and diazinon in residences with a series of independent variables obtained from questionnaire-based information and environmental sampling using the NHEXAS-AZ database. We built models that may be used to predict residential exposures estimated using the indirect method.

Methods

Model formulation requires a database and a strategy for building and evaluating the model. In this section we introduce the NHEXAS-AZ study and our treatment of the database generated by the study, discuss the framework of estimating exposure to a pollutant using the indirect method of exposure estimation, and present the approach of model formulation and evaluation used in this study.

Experimental Design

Arizona was the location of all fieldwork. A four-stage probability sampling design was used for sample selection. The primary sampling units (PSU) of the study were all the tracts of every Arizona county, except one (La Paz county, which has very small population). In stage 1, 49 PSU were selected using probability proportional to size (PPS) technique. In stage 2, secondary sampling units (SSU) were chosen by selecting five blocks, the SSU, from each PSU using the PPS technique. From each SSU five households were randomly selected in stage 3 and from each household one primary respondent was randomly selected in stage 4.

Field Sampling

Questionnaires were distributed for response by the primary respondent or were completed by field technicians inspecting subject residences. A total of 955 households (out of 1225 contacted) responded to the descriptive questionnaire. Of these, 525 agreed to respond to the next questionnaire, baseline questionnaire. Subsequently, 403 of the 525 primary respondents agreed to respond to four additional and more intensive questionnaires: (1) time/activity questionnaire, (2) technician questionnaire describing specific data about housing and product usage, (3) follow-up questionnaire describing daily activity during the week of sampling, and (4) diet diary. Three hundred households were sampled to measure pesticide concentrations in residences, a few in all media and most in at least one medium. Media sampled included indoor air, food, sill wipes, house dust, and yard soil (Lebowitz et al., 1995).

Actively pumped air samples were evaluated and used for pesticides from fixed sites inside and outside homes. Integrated samples were collected over 3 days (12 and 24 cumulative hours for indoors and outdoors, respectively) at 4 l/min. Pesticides were collected on PUF and Teflon-coated glass fiber filters contained in a URG-2000 sampler with a 10 μm cutpoint. Floor dust was sampled with a portable vacuum and a special sampling head, developed by Battelle scientists, made from Teflon and stainless steel. Composite floor dust samples were collected by vacuuming 4 m^2 of floor area in the main room and a bedroom of each home. Dust was sieved in the laboratory using a 62 μm mesh screen, and a 1 g aliquot of the fine fraction was evaluated for pesticide content. Dust from window sills and dermal surfaces was obtained by moistening gauze pads and wiping the surface. Water was used as the moistening agent for sills and isopropanol was used to remove soil from both hands of the primary respondent. Soil samples were collected from the yard and along the foundation of each home. Subsamples were collected and composited in a prescribed manner to provide a single representative sample from the home. The yard and foundation soils were sieved using a 62 μm sieve, and 1 g of the fine clay/silt fraction was evaluated for pesticide content. The chemical evaluation summary: Each sample was spiked with a surrogate recovery standard (250 ng fenchlorphos) and extracted using acetone for each medium. Then 100 ng of trichloronate was added as an internal standard for

calibration. Samples underwent a solid phase extraction and cleanup. Pesticide content was analyzed using gas chromatography with electron capture detection for some samples (stage 2=GC/ECD) and gas chromatography mass/spectroscopy for other analyses (stage 3=GC/MS). Only yard soil samples analyzed by GC/ECD were used in this paper. Field and laboratory quality assurance spikes, duplicates and blanks were also evaluated. Details of the field and laboratory evaluations for pesticides by media are reported in the literature (Lebowitz et al., 1995; Moschandreas et al., 2001).

Censored Data

Censored values, concentration values below limit of detection (LOD), were treated using the robust method developed by Helsel (1990). The robust method assumes that all pesticide concentrations in each medium follow one distribution, the one that best fits the above detection limit values. The above LOD values were fit using Crystal Ball, commercial software that fits the data points to several distribution types. The chi-squared test, the Kolmogorov–Smirnov test, and the Anderson–Darling test were used to assess the goodness of fit. At least one of these tests must consider the fit acceptable. A modification of the robust method was used to incorporate into the method the adjustment for portion of the censored data. Based on extrapolation of the fitted distribution, the robust method creates "fill-in" values for samples that are below the detection limit. Such values were assigned once to a residence with below detection limit concentrations and were used for all subsequent analyses in the study.

The Indirect Method of Exposure Estimation

The dependent variable of the regression models constructed in this study is the exposure to pesticides calculated by the indirect method of exposure estimation. To estimate

Table 1. List of 29 questionnaire-based independent variables.

Variable	Description	Type
adult	age 18 or older	indicator
agricult	surrounding area is agricultural area	indicator
aplypest	percent of days that the subject applied pesticides during the sampling week	categorical
carpeted	more than 50% of rooms are carpeted	indicator
child	age 13 or younger	indicator
contactp	contact pesticide in primary job	indicator
dustin	spent time dusting during the sampling week	indicator
farmer	primary business is farmer or involves pesticide use	indicator
firepl	use a fireplace during the sampling week	indicator
flea	use chemicals to control flea on pets	indicator
gardenin	spent time gardening during the sampling week	indicator
grasleav	percent of days that the subject had skin contact with grass or leaves from yard during the sampling week	categorical
hidust	overall dust level in home is rated as "high"	indicator
inside	pesticide was used inside home in the past 6 months	indicator
lawn	had regular lawn treatment in the past 6 months	indicator
lawnmow	used a lawnmower during the sampling week	indicator
lowinc	has annual income less than $20,000	indicator
male	is male	indicator
mobile	house type is mobile	indicator
nodust	overall dust level in home is rated as "no dust"	indicator
noschool	has no education	indicator
outside	pesticide was used outside home in the past 6 months	indicator
preppest	percent of days that the subject prepared pesticides during the sampling week	categorical
sit_rugs	average time that the subject spent laying down or sitting on carpet during the sampling week	continuous
soildirt	percent of days that the subject had skin contact with soil or dirt from yard during the sampling week	categorical
sweep	spent time sweeping during the sampling week	indicator
vaccu	spent time vacuuming during the sampling week	indicator
yd_grass	yard material is grass	indicator
yd_soil	yard material is soil	indicator

exposures, the indirect method combines pesticide concentration measurements obtained from chemical analysis of air, water, food, surface, or soil samples with information from questionnaires on subject demographics, food consumption, and time budgets. All exposure routes were estimated: inhalation, ingestion, and dermal exposure. Both the dietary and non-dietary ingestion pathways were estimated. Dietary ingestion exposure involves exposure *via* consumption of solid food, beverage, and water. Non-dietary ingestion exposure is due to unintentional ingestion of dislodgeable residue. Models for estimating exposure and relevant assumptions and factor values are presented in Appendix A.

Model Classification and Procedures

Regression models were formulated for each of the two pesticides in each of the routes and pathways: inhalation, dietary ingestion, non-dietary ingestion, and dermal absorption. From the six study questionnaires, we selected 29 questions as likely to bring about potential exposure determinants, which are measured characteristics that may affect exposure to pesticides. This knowledge-based list of 29 potential exposure determinants (see Table 1) was used as the set of 29 candidate regressors from which the independent variables of the regression models were to be selected electronically and statistically. These variables were not used for estimating exposures.

The dependent variable of all regression models is the natural logarithm ($\log_e$) of the estimated exposure. Because the exposure values estimated for each route appear to be lognormally distributed, the $\log_e$-transformed values were used to reduce the influence of few extreme values. Route-specific exposures used in this study (see Table 2) are determined using either *all* subject dwellings, those with above and below detection limit measured pesticide concentration, or those subject dwellings with *only* above detection limit measured pesticide concentration.

The model formulation scheme applied in this study consists of four classes that represent conditions of increasing complexity and data requirements, and combine different categories of dependent and independent variables.

1. Class 1: The dependent variable is the $\log_e$-transformed exposure to a pesticide estimated only for residences with measured pesticide concentrations above detection limit. Independent variables are selected from the list of 29 potential exposure determinants obtained exclusively from questionnaire information. This class is the easiest to use because it is based singularly on data already in the database.

2. Class 2: This class uses the same dependent variable as class 1. The list of independent variables is enlarged by the addition of pesticide concentrations in media different from the one(s) used for estimating the dependent variable (see Table 3). This is a slightly more complex class than class 1 because it requires selection and use of specific pesticide concentrations. For the inhalation route, additional potential independent variables include concentration in floor dust, dermal wipe, sill wipe, and yard soil. For dietary ingestion route, the additional independent variables used are the ones used in inhalation route plus concentration in indoor air. Dermal wipe and indoor air concentrations are additional independent variables for both dermal and non-dietary ingestion routes. Dermal wipe concentrations are used in all routes because they are not used in any route exposure estimation. However, food residues are not included as additional independent variable because, we assume, they are not likely to affect inhalation, dermal, or nondietary ingestion exposure. While by the residence yard soil concentrations were used in this class condition, ambient (by the residence) pesticide concentrations were not used in the model formulations because about 80% of these concentrations were below the detection limit. Out of 42 samples collected for each pesticide, only 4 and 9 ambient by the residence chlorpyrifos and diazinon concentrations were above their perspective detection level.

3. Class 3: The dependent variable is the $\log_e$-transformed exposure to a pesticide estimated in all residences, those with measured pesticide concentrations

Table 2. Mean values of route-specific exposure of subjects with above detection limit (ADL) concentrations only and all subjects.

Pesticide	Route	Unit	ADL subjects only		All subjects	
			Sample Size	Mean	Sample size	Mean
Chlorpyrifos	Inhalation	ng/h/m^3	77	1387.09	119	905.28
	Dietary Ingestion	ng/kg/day	34	54.98	145	16.07
	Dermal Absorption	μg/day	18	102.31	54	34.22
	Non-dietary Ingestion	ng/kg/day	18	11.43	54	3.84
Diazinon	Inhalation	ng/h/m^3	75	6663.69	119	4204.33
	Dietary Ingestion	ng/kg/day	13	75.85	145	8.45
	Dermal Absorption	μg/day	24	3.84	54	2.52
	Non-dietary ingestion	ng/kg/day	24	5.92	54	4.80

Table 3. Measured concentrations used as independent variables.

Pesticide	Variable	Description	Unit	Used as independent variables for:			
				Inhalation	Dietary Ingestion	Dermal Absorption	Non-dietary Ingestion
Chlorpyrifos	xffpinc	chlorpyrifos concentration in indoor air	ng/m^3		✓	✓	✓
	xfdp2c	chlorpyrifos concentration in floor dust	$\mu g/m^2$	✓	✓		
	xdwp2c	chlorpyrifos concentration in dermal wipe	$\mu g/m^2$	✓	✓	✓	✓
	xswpc	chlorpyrifos concentration in sill wipe	$\mu g/m^2$	✓	✓		
	xyspc	chlorpyrifos concentration in yard soil	$\mu g/g$	✓	✓		
Diazinon	xffpind	diazinon concentration in indoor air	ng/m^3		✓	✓	✓
	xfdp2d	diazinon concentration in floor dust	$\mu g/m^2$	✓	✓		
	xdwp2d	diazinon concentration in dermal wipe	$\mu g/m^2$	✓	✓	✓	✓
	xswpd	diazinon concentration in sill wipe	$\mu g/m^2$	✓	✓		
	xyspd	diazinon concentration in yard soil	$\mu g/g$	✓	✓		

above and below detection limit. Independent variables of the regression model are selected from the list of class 1. The use of the robust method increases the complexity of requirements for model construction.

4. Class 4: This class uses the same dependent variable as class 3. The independent variables are selected from the same list of potential independent variables as class 2. This class uses both dependent and independent variables that require data handling.

Given a set of potential independent variables, the stepwise procedure was used to construct regression models by selecting only those independent variables that contribute significantly to the model. The criterion for significant contribution is the calculation of significant F-statistics at $P=0.10$ for entry and retention in the model. In addition, model adequacy was evaluated using the square of the multiple correlation coefficient, R^2, value of each model and the root mean square error (RMSE). The square of the multiple correlation coefficient denotes the percentage of variance of the dependent variable explained by the model. The RMSE reflects model uncertainty in units of the dependent variable. An uncertainty indicator for each model is estimated by the ratio of the model RMSE value over the corresponding $\log_e$-transformed mean value of the exposure. This ratio value provides an indication of the model uncertainty relative to the mean value of the dependent variable. Sample size was used as an arbitrary criterion for model selection; only models with sample size equal to or larger than 15 were considered as acceptable models. Residual analysis was performed on each selected model to further evaluate whether it is appropriate for the corresponding data set. In addition to standardized residuals, predicted/residual and predicted/observed scatter plots were also constructed to examine if they display any particular pattern. All regression procedures were performed with SPSS, a commercially available statistical package. Unweighted analyses were performed for this paper though the use of sampling weights would lead to more accurate estimates of exposure. We believe that uncertainties associated with models used and assumptions made do not justify the use of sampling weights and specialized software.

Results and discussion

Models formulated in this study are route-specific linear models that associate exposure to a pesticide with one or more independent variables. The list of potential exposure determinants was generated from questionnaire information and from sampling data. The assumption is that exposure is related linearly with potential exposure determinants. The model construction was based on the scheme of four classes discussed in the methods section. Route-specific regression exposure models are presented and discussed in this section.

Inhalation Models

Table 4 depicts inhalation exposure models with statistically significant association between in-residence inhalation exposure to each pesticide and a set of independent variables. Exposure to chlorpyrifos for classes 2 and 3 did not lead to a significant model. A significant multiple regression model, $P\cong 0.0$, with four independent variables, predictors, was formulated with the class 1 approach. This model explains 49% of the variance of residential inhalation exposures to chlorpyrifos. The RMSE is 0.97, which is a rather small value in comparison with 7.23, the $\log_e$-transformed mean value of the exposure to chlorpyrifos estimated using the inhalation deterministic model. The model uncertainty estimator is 13%. The class 4 model is significant, $P\cong 0.0$; it has only one independent variable: chlorpyrifos concentration in dermal wipe, and has an R^2 value of 0.25. Its uncertainty estimate is rather small, 3%. Inhalation exposure to diazinon for classes 2 and 3 did not lead to significant models. The in-residence inhalation

Table 4. Summary of significant exposure models.

Route	Pesticide	Class	No. of samples	R^2	RMSE	P	Independent variable	Coefficient
Inhalation	Chlorpyrifos	1	38	0.490	0.968	0.000	(constant)	3.918
							preppest	17.129
							contactp	1.496
							inside	0.765
							sit_rugs	0.392
		4	33	0.248	1.787	0.000	(constant)	4.577
							xdwp2c	0.002
	Diazinon	1	38	0.325	1.428	0.001	(constant)	3.539
							contactp	2.220
							preppest	18.268
		4	53	0.433	1.807	0.000	(constant)	3.818
							xswpd	0.039
Dietary ingestion	Chlorpyrifos	1	31	0.309	1.097	0.006	(constant)	2.518
							adult	1.323
							vacuu	0.983
	Diazinon	4	40	0.400	1.410	0.000	(constant)	−0.417
							xswpd	0.028
Dermal absorption	Chlorpyrifos	3 and 4	50	0.461	1.937	0.000	(constant)	−5.934
							preppest	29.587
							male	2.038
							lawnmow	3.028
	Diazinon	4	30	0.551	1.633	0.000	(constant)	−3.683
							xffpind	0.001
							firepl	4.490
							lowinc	1.719
Non-dietary ingestion	Chlorpyrifos	3 and 4	54	0.169	2.239	0.002	(constant)	−3.828
							preppest	26.832
	Diazinon	4	30	0.424	2.013	0.001	(constant)	−3.778
							xffpind	0.001
							firepl	5.583
							preppest	0.546

exposure to diazinon model for class 1 conditions is a multiple regression model with two independent variables, pesticide contact in primary job and percent of days that subjects prepared pesticides during the sampling week. This model explains 33% of the variation of the dependent variable. The RMSE is relatively small, 1.43, resulting to a model uncertainty estimator of 16%. Class 4 conditions lead to a simple linear regression model with sill wipe concentration as the independent variable, it explains 43% of the variance of the dependent variable, and it has 22% uncertainty estimate.

Dietary Ingestion Models

Dietary ingestion exposure to chlorpyrifos associates significantly with vacuuming activity during sampling week and adult subjects only for class 1 conditions, $P=0.006$ and $R^2=0.31$ (see Table 4). The uncertainty estimator of this model is 27%. For class 4 conditions ingestion exposure to diazinon associates significantly with sill wipe concentration only, $P\cong 0.0$, $R^2=0.40$. The uncertainty estimator of 66% indicates considerable model uncertainty. All other classes of model formulation led to regression models that did not satisfy our criteria of acceptable associations.

Dermal Exposure Models

For class 3 conditions, dermal exposure to chlorpyrifos associates significantly, $P\cong 0.0$, $R^2=0.46$, with percent of days a subject spent preparing pesticide during the sampling week, gender, and regular lawn treatment in the past 6 months (see Table 4). The model uncertainty estimate is 55%. Exactly the same multiple linear regression model is formulated for class conditions 3 and 4, that is the addition of concentrations in media other than the ones used to

estimate the dependent variable did not affect the association. Conditions 1 and 2 did not lead to a significant model for dermal exposure to chlorpyrifos. Dermal exposure to diazinon associates significantly with indoor air concentration, use of fireplace during the sampling week and low income, $P \cong 0.0$, and $R^2 = 0.55$, for class 4 conditions. The uncertainty estimate of the model is large, 177%.

Non-dietary Ingestion Exposure Models

One non-dietary exposure to chlorpyrifos model was constructed for both class conditions 3 and 4. This model is a simple linear model that associates the exposure with portion of time spent in preparing for pesticide use during the sampling week; the addition of relevant concentration levels did not change the model (see Table 4). The R^2 value of these models is 0.17, a very low value, yet the model is significant, $P = 0.002$. The uncertainty estimate is considerably large, 166%. A significant model for non-dietary exposure to diazinon was obtained from class 4 conditions only. The multiple linear regression model, $P \cong 0.0$, $R^2 = 0.42$, has three independent variables, indoor air concentration, use of fireplace during the sampling week and percent of days spent preparing pesticides during the sampling week. The uncertainty estimate is equal to 128%. All other class conditions led to models that did not satisfy study acceptance criteria.

Residual Analysis

Residual analysis was used to assess the appropriateness of each of the models discussed above. This analysis identified one outlier, a residual value that is detached from the rest of the residuals, for inhalation exposure to diazinon, one for dermal exposure to diazinon, and one for non-dietary ingestion exposure to diazinon. No other outliers were identified for the other exposure route models. In an effort to determine if our approach should be modified, we removed the identified outliers; the impact of removing the detached value is model dependent. Removal of the outlier in the inhalation exposure to diazinon did not change the independent variables included in the model. The model under class 1 conditions increased the R^2 value but the one under class 4 conditions decreased the R^2 value. In the case of dermal exposure to diazinon, removal of the outlier resulted to a significant model similar to the one constructed before removal of the outlier. One independent variable was excluded and the R^2 value decreased. Removal of the outlier in the nondietary ingestion exposure to diazinon changed the model from multiple linear regression to simple linear regression with much reduced R^2 value. Review of the QA/QC program of the NHEXAS-AZ and inspection of the values suggests that all values are both valid and reasonable, though the identified outlier values are large. Thus, we conclude that all values should be used in the linear regression model formulation process to reflect the range of measured concentrations and estimated exposures. This approach reflects the emphasis of environmental concerns that focus on large values and not on low and moderate exposure levels.

Discussion

The number of days that a subject spent preparing pesticide during the week of sampling is the determinant of exposure used most often; it associates with the inhalation, dermal (only to chlorpyrifos), and non-dietary routes of exposure. Probable explanations for inclusion of statistically selected variables can be advanced to justify each significant model, but such explanation must be provided with care. For instance, pesticide preparation, pesticide usage inside home, and chlorpyrifos concentration in dermal wipes are three independent variables that contribute significantly to the inhalation exposure. One may speculate that application of the pesticide has resulted in airborne chlorpyrifos and contaminated hands. Yet, further investigation is needed before such relationships can be concluded. Contact with pesticide in primary job is a determinant only for the inhalation route for both pesticides. It was thought that it would be a determinant for other exposure routes such as dermal and non-dietary ingestion, this was not the case. We speculate that this was caused by the fact that a rather large segment of subjects did not identify a primary job. This does not indicate a high unemployment level, rather it demonstrates the large number of retirees in Arizona and a population that works part time at several industries including tourism, agriculture, heavy industry and others but not at one primary job.

The addition of pesticide concentrations in media other than the one(s) used to calculate the route-specific exposure led to significant models for all routes of exposure to diazinon. The same process of adding pesticide concentrations did not contribute significantly in any of the route-specific models of exposure to chlorpyrifos, except in the inhalation route. Although it is difficult to explain this difference, we suspect that it may be associated with patterns and purpose of use of each pesticide.

Conclusions

Significant regression models were constructed using the NHEXAS-AZ database and the stepwise procedure for selecting independent variables. The dependent variable was a route-specific function of exposure to each of the two subject pesticides. The exposure to each pesticide was calculated using the indirect method that uses pesticide concentration in the appropriate media and information obtained from questionnaires. Exposures were calculated

using either only residences with concentrations above the detection limit of the analytical method, or all residences including those with below detection limit concentrations. In the later case, use of the Helsel's robust method randomly assigned low concentration values to residences with below detection limit concentrations. Independent variables were selected by two methods: The first was a list of 29 potential exposure determinants assembled by the authors strictly from subject responses to study questionnaires. The second method enlarged the list of potential exposure determinants by adding pesticide concentrations from media other than the one(s) used for calculating the dependent variable. Models formulated in this study are constrained by uncertainties associated with models used to estimate the route-specific exposures, assumptions made, and by the magnitude of the resulting associations, which though significant were not always strong.

The combination of two pesticides, four exposure routes, and four class conditions lead to the possibility of formulating as many as 32 significant models associating the dependent variable with its determinants. Only 12 significant regression models were formulated using the NHEXAS-AZ database. Of the 12 models constructed, only three were formulated using only residences with above detection limit pesticide concentrations. Use of the robust method helped in the formulation of nine additional significant models. We are cognizant of concerns that relate the robust method with models that lead to biased results, yet if the analyst is careful in the use of such models they are helpful investigative tools.

This study identifies several exposure determinants that are obtained from questionnaire-based information. Among them, determinants of exposure to chlorpyrifos include age, preparation of pesticides, and use of pesticide inside the house. In case of diazinon, determinants include preparation of pesticides and income level. Concentration levels of a few media are determinants in some regression models constructed in this study. Only dermal wipe concentration is identified for inhalation exposure to chlorpyrifos, whereas either sill wipe or indoor air concentrations are identified in each route as exposure determinants for diazinon.

From a general perspective, the selection procedure was statistical. While the original list of potential determinants was knowledge-based, selection decisions on whether to keep or eliminate an independent variable from the model were made on the test statistic of the estimated coefficient, which depends on the data and the specified significance levels. Thus, the final regression model selections were based on electronically, statistically, and data-driven selection procedures and, only later we could reason out all associations illustrated by the models. The strength of the association of these models, R^2, varies from 12% to 55%; it may be increased by expanding the models with the use of nonlinear terms including interactions, and by increasing the member of independent variables. The 12 significant models quantified the predictive power of the variables in the NHEXAS-AZ database; yet, these models must be used with care because they are constrained by the assumptions made, and the associations though significant were not always strong. The uncertainty indicator associated with these models varies form route to route and it varies from relatively small to large, equal to or sometimes larger than the mean of the dependent variable.

Models formulated in this study result in an association that is not always clear. In a few cases, the direction of the association is the opposite of what was expected, that is an increase in the value of an independent variable expected to result in an increase of the dependent variable resulted in a decrease. We attribute this to the lack of exact between sampling and questionnaire responses, to the time-integrated sampling techniques employed, and to the plethora of categorical independent variables. It may be that models formulated in this study do not represent a physically justifiable relationship, rather they constitute empirical equations that associate variables and explain portions of variability of the dependent variable as a linear function of the independent variables. Measurements of route-specific exposures are recommended using personal monitors when such sampling methods are available. Notwithstanding the limitations of empirical regression models, models formulated in this study may be used to estimate exposures to each of the pesticides. Use of the models formulated must incorporate clear statements of the strengths and limitations of each model, including the coefficient of determination and confidence and prediction intervals of the dependent variable.

Appendix A. Exposure Models and Assumptions

Exposure Models

Exposure models used were adapted from several U.S. EPA models (U.S. EPA, 1990, 1992, 1993, 1997) to focus on exposure estimation rather than dose estimation. The equation used for estimating inhalation exposure to a pesticide for each subject is:

$$E_{\mathrm{i}} = \sum_{\mathrm{j}} \left(C_{\mathrm{i,j}} \times t_{\mathrm{i,j}}\right) \qquad \text{(A1)}$$

where E_{i} is the inhalation exposure to one pesticide for each subject i, $\mathrm{ng/h/m^3}$; $C_{\mathrm{i,j}}$ is the concentration of the pesticide in the air, associated with the subject i in microenvironment j, $\mathrm{ng/m^3}$; $t_{\mathrm{i,j}}$ is the average time spent per day by the subject i in the microenvironment j, h.

The microenvironments of concern for this paper are indoor and outdoor at home. However, the majority of the outdoor air concentration values in the NHEXAS Arizona were below detection limit. Therefore, the exposure to the

chemicals in the outdoor air is assumed to be negligible and only the indoor microenvironment was included in the exposure estimation.

The equation used for the direct estimation of dietary ingestion exposure to a pesticide for each subject is:

$$E_{\mathrm{T,i}} = \frac{\sum_{\mathrm{F}} \left(C_{\mathrm{F,i}} \times W_{\mathrm{F,i}}\right)}{BW_{\mathrm{i}}} \times 10^{6} \quad \text{(A2)}$$

where: $E_{\mathrm{T,i}}$ is the total dietary ingestion exposure to the chemical residue, chlorpyrifos or diazinon, from the composited food items consumed by subject i during the day of measurement, ng/kg/day; $C_{\mathrm{F,i}}$ is the concentration of the chemical residue, chlorpyrifos or diazinon, in the composited food items F consumed by subject i during the day of measurement, mg/kg; $W_{\mathrm{F,i}}$ is the weight of composited food items F consumed by subject i during the day of measurement, kg/day; $\mathrm{BW_i}$ is the body weight of subject i, kg; F is the type of composited food items. There are three types: solid food, beverage, and water.

Only up to 2% of chlorpyrifos or diazinon concentrations in beverages and water are detectable. Therefore, it was concluded that the population exposures to the pesticides from consumption of beverages and water are negligible. Only solid foods were included in the equation.

The equation used for estimating dermal exposure to a pesticide for each subject is (U.S. EPA, 1993):

$$E = \left[\sum_{s=1}^{S} \left(C_{\mathrm{Ds}} \times A_{\mathrm{ps}} \times T_{\mathrm{ps}} \times 1 - \mathrm{DO_{ps}}\right)\right] + \left[\sum_{s=1}^{S} \left(C_{\mathrm{Ss}} \times \left(S_{\mathrm{ps}} \times SA_{\mathrm{ps}} - SO_{\mathrm{ps}}\right) \times M\right)\right] \quad \text{(A3)}$$

where E is the total dermal exposure to pesticide (chlorpyrifos or diazinon), μg/day; s is the type of surfaces contacted per day; S is the total number of surfaces contacted per day; C_{Ds} is the concentration of dislodgeable surface residue, $\mu g/m^2$; A_{ps} is the surface area contacted by subject p, m^2/day; T_{ps} is the transfer from surface by subject p, proportion; $\mathrm{DO_{ps}}$ is the dislodgeable residue transferred to oral route by subject p *via* hands, food, and objects, proportion; C_{Ss} is the concentration of pesticides in soil, μg/g; S_{ps} is the soil covering on skin from surface s for subject p, g/m^2 day; $\mathrm{SA_{ps}}$ is the surface area of skin exposed to surface s for subject p, m^2; $\mathrm{SO_{ps}}$ is the soil or dust from surface s transferred to oral route by subject p, g/day; **M** is the matrix effect of soil, proportion.

The first and second terms of the right-hand side of the model estimate dermal exposure *via* dislodgeable and soil/dust surfaces, respectively. Assumptions made and values used for the variables in the model are discussed in the section following the non-dietary ingestion exposure model.

The equation used for estimating non-dietary ingestion exposure to a pesticide for each subject is (U.S. EPA, 1993):

$$E = \left[\sum_{s=1}^{S} \left(C_{\mathrm{Ds}} \times A_{\mathrm{ps}} \times T_{\mathrm{ps}} \times \mathrm{DO_{ps}}\right)\right] \times \frac{P}{BW} \times 10^{3} \quad \text{(A4)}$$

where E is the total non-dietary ingestion exposure to pesticide (chlorpyrifos or diazinon), ng/kg/day; s is the type of surfaces contacted per day; S is the total number of surfaces contacted per day; C_{Ds} is the concentration of dislodgeable surface residue, $\mu g/m^2$; A_{ps} is the surface area contacted by subject p, m^2/day; T_{ps} is the transfer from surface by subject p, proportion; $\mathrm{DO_{ps}}$ is the dislodgeable residue transferred to oral route by subject p *via* hands, food, and objects, proportion; P is the dislodgeable residue transferred to oral route that is actually ingested, proportion; BW is the body weight, kg.

Assumptions and Values in Dermal and Nondietary Ingestion Exposure Models

Type of Surface Contacted, s There are two types of dislodgeable surfaces in the NHEXAS study. One is the window sill surface and the other is the floor surface. The former is assumed to be representative of all surfaces other than the floor in the house, such as furniture surfaces, and is called "non–floor surface." The soil/dust surface is the measured concentration of yard soil.

Surface Area Contacted, A_{ps} Eighty percent of the total floor surface area is assumed to be available for contact. The floor surface area contacted by adults and children is assumed to be 5% and 20% of the available floor surface area, respectively. The available non–floor surface area is assumed to be 20% of the available floor surface area. The non–floor surface area contacted by adults and children is then assumed to be equal to 5% and 20% of the available non–floor surface area, respectively.

Transfer from Surface, T_{ps} The transfer from the two dislodgeable surfaces is assumed to be 10%.

Dislodgeable Residue Transferred to Oral Route, DO_{ps} The dislodgeable residue transferred to oral route is calculated by (U.S. EPA, 1993):

$$\mathrm{DO_{ps}} = \mathrm{SA_{ph}}/\mathrm{SA_{ps}} \quad \text{(A5)}$$

where $\mathrm{SA_{ph}}$ is surface area of skin on the hands that is used to transfer material to mouth or food, m^2. This is assumed to be 25% of the hands area (Hawley, 1985). $\mathrm{SA_{ps}}$ is the surface area of skin exposed to surface s, m^2.

Soil Covering on Skin, S_{ps} The values used for soil covering on skin are 8 and 11.1 g/m^2 for children and adults, respectively (U.S. EPA, 1997).

Surface Area of Skin Exposed to Surface, SA_{ps} The total surface area of the skin is calculated from the following equation (U.S. EPA, 1997):

$$SA = 0.0239H^{0.417}W^{0.517} \quad (A6)$$

where H is height of subject, cm, and W is body weight of subject, kg. The surface area of skin exposed to surface for a subject is assumed to be the area of hands, feet, legs and arms of the body for children, and hands and feet for adults. The percentage of total body surface area by part is used to get the surface area of skin exposed for children and adults (U.S. EPA, 1997).

Soil Transferred to Oral Route, SO_{ps} The information on the soil from surface transferred to oral route is not available from the literature, so it is assumed to be zero.

Matrix Effect of Soil, **M** The matrix effect of soil is assumed to be 0.15 (U.S. EPA, 1993).

Dislodgeable residue transferred to oral route that is actually ingested, P This proportion represents the fraction of dislodgeable residue transferred from dermal to oral route that is actually ingested by the subject. It is assumed to be equal to 0.1 and 0.25 for adults and children, respectively.

Other Considerations

In accordance with the definition of exposure used in this paper, exposures at the receptor boundaries are considered in this paper, but dose is not. Therefore penetration factors such as inhalation rate and absorption rate are not included in the above models.

Acknowledgments

Preparation of this document has been funded wholly or in part by the United State Environmental Protection Agency under Contract 68-D-99-003 to the University of Arizona. It has been subjected to Agency review and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Davis J.R., Brownson R.C., and Garcia R. Family pesticide use in the home, garden, orchard, and yard. *Arch Environ Contam Toxicol* 1992: 22: 260–266.

Hawley J.K. Assessment of health risk from exposure to contaminated soil. *Risk Anal* 1985: 5: 289–302.

Helsel D.R. Less than obvious: Statistical treatment of data below the detection limit. *Environ Sci Technol* 1990: 24 (12).

Hodgson E., and Levi P.E. Pesticides: an important but underused model for the environmental health sciences. *Environ Health Perspect* 1996: 104 (Suppl 1).

Kupchella C.E., and Hyland M.C. Environmental Science: Living Within the System of Nature. 3rd ed. Prentice-Hall, 1992.

Lebowitz M.D., O'Rourke M.K., Gordon S.M., Moschandreas D., Buckley T., and Nishioka M. Population-based exposure measurements in Arizona: a phase I field study in support of the National Human Exposure Assessment Survey. *J Exposure Anal Environ Epidemiol* 1995: 5 (3): 297–324.

Moschandreas D.J., Kim Y., Karuchit S., Ari H., Lebowitz M.D., O'Rourke M.K., and Gordon S. Occurrence and distributions of residential exposure to chlorpyrifos and diazinon. 2001. Accepted for publication by Journal of Atmospheric Environment.

National Academy of Sciences (NAS). Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities. Washington, DC, 1991.

U.S. Environmental Protection Agency (U.S. EPA). National household pesticide usage survey, 1976–1977. Office of Pesticides and Toxic Substances, U.S. Government Printing Office, Washington, DC, 1979, EPA 540/9-800-002.

U.S. Environmental Protection Agency (U.S. EPA). Exposure Factors Handbook. 1990.

U.S. Environmental Protection Agency (U.S. EPA). Guidelines for Exposure Assessment. *Fed Reg* 1992: 57 (104): 22888–22938.

U.S. Environmental Protection Agency (U.S. EPA). Exposure Factors Handbook. 1997.

U.S. Environmental Protection Agency (U.S. EPA). Research Solicitation: Human Exposure Assessment. Office of Research and Development, 1993.

World Health Organization (WHO). Diazinon. Geneva, 1998.

Zartarian V.G., Ott W.R., and Duan N. A quantitative definition of exposure and related concepts. *J Exposure Anal Environ Epidemiol* 1997: 7 (4): 411–437.