ESTIMATING SERUM POLYCHLORINATED BIPHENYL LEVELS IN HIGHLY EXPOSED WORKERS: AN EMPIRICAL MODEL

Philip R. Taylor

Cancer Prevention Studies Branch, Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, Maryland

Andrew A. Reilly, Jeanette M. Stelma, Charles E. Lawrence

Statistical and Computer Science Laboratory, Wadsworth Center for Labs and Research, New York State Department of Health, Albany, New York

A regression model estimating high-homolog polychlorinated biphenyl (PCB) serum concentration on the basis of job exposure categorizations was developed. The model assumes first-order kinetics with a half-life determined empirically and uses variables that incorporate both intensity and duration of exposure over a 30-yr period. In order to compare the efficiency of these regression-based exposure estimates relative to often-used epidemiological parameters, models with dichotomized, ordinal, and continuous exposure surrogates were also investigated. Among the alternative exposure categorizations the most straightforward measure, ever versus never direct, was a particularly poor predictor of serum PCB level (r^2 = .01). Nearly all of the candidate exposure measures we tried predicted serum levels poorly. The best of these after the fact was with total months employed in direct-exposure jobs (r^2 = .43). None of the logical deductive models approached the predictability of the empirical model developed here (r^2 = .69).

INTRODUCTION

An idealized exposure estimate is based on actual measurements of a putative exposure over some biologically relevant time. Studies of health outcomes in occupational cohorts that lack actual exposure measures

The authors gratefully acknowledge the help of Marie Ross and Dr. Richard Lawton, General Electric Corporation; Dr. Ho-Ling Hwang, New York State Department of Health; Dr. Brenda Edwards, National Cancer Institute; Drs. George Hutchison and Marcello Pagano, Harvard School of Public Health; Dr. Edward Baker, National Institute for Occupational Safety and Health; and Emily Brooks and Lisa Licitra, Information Management Services.

This research was supported in part by a contract (210-81-5102) with the National Institute for Occupational Safety and Health and a grant from the New York State PCB Settlement Committee. During part of this investigation, Dr. Taylor was supported by a National Cancer Institute National Research Service Award (5T32-Ca-09001) from the Harvard School of Public Health.

Requests for reprints should be sent to Philip R. Taylor, NCI/NIH, 9000 Rockville Pike, Executive Plaza North, Room 211, Bethesda, MD 20892.

Journal of Toxicology and Environmental Health, 34:413–422, 1991
Copyright © 1991 by Hemisphere Publishing Corporation

typically consider the entire cohort to be exposed or use a simple dichotomy such as "exposed" versus "unexposed" within the cohort.

In the present article we describe a two-stage approach to estimate exposure to PCBs in an occupational cohort. First, a sample of the cohort was selected for which both individual exposure measures and surrogate exposure variables were available. These data were used to develop a regression model for individual exposure based on surrogate exposure measures, which was applied in the second stage to estimate exposures for the whole cohort solely on the basis of the surrogate variables.

MATERIALS AND METHODS

Population

Between 1946 and 1977 two facilities of the same company located in upstate New York manufactured capacitors using PCBs with Aroclors 1254, 1242, and 1016 as their primary dielectric fluid. In 1976 a study was initiated to examine the possible human health effects of PCB exposure. This was done by conducting an examination of all highly exposed workers, those most likely to exhibit adverse health effects. One hundred ninety-four individuals were evaluated, including all current employees whose jobs required direct PCB contact in zones of high PCB air concentration, were in the immediate periphery of the high-exposure zone, or had high but intermittent exposure. The health evaluation has been described in detail elsewhere and included serum PCB measurement (Lawton et al., 1985a). A similar evaluation was conducted in 1979 and 1983 on participants still available for reexamination. Because of concern over a laboratory error in the 1976 data, only the data from 1979 were used in this analysis. Complete job history records through 1976 and serum Aroclor 1254 measurements from 1979 were available on 157 employees. These records form the data base for the first stage. The half-life of Aroclor 1254 was empirically estimated as 3.32 yr by comparing serum levels from 1979 with additional follow-up values obtained in 1983 on 150 participants. Calculations assumed first-order kinetics (Morrow, 1977). Once the half-life was determined, it was incorporated into all of the regression models described here in order to account for excretion of PCBs.

Variable Definition

Although additional data were gathered on participants, variables used in the development of the prediction model were limited to items available for the study subjects at the second stage: age, sex, job codes, job description, and dates during which the job was held. The job description in combination with manufacturing process information and industrial hygiene data was used to categorize all jobs into two broad exposure groups and a total of four specific categories. Direct-exposure

jobs were defined as those in which direct contact with PCBs occurred during the manufacturing process. These jobs were further characterized into subcategories as follows:

Low: air contact only
Medium: air contact plus occasional dermal contact
High: air contact plus frequent dermal contact

All other jobs within the plants, including office and manufacturing areas where PCBs were not directly used, were termed indirect-exposure jobs.

Auxiliary information concerning exposure was available but not explicitly recorded in the company's personnel records. The facility started operation and began using PCBs in 1946. In 1954 use of the highly chlorinated Aroclor 1254 was phased out and replaced by Aroclor 1242; in 1965 major engineering changes occurred, including closed-process filling of capacitors and improvements in ventilation, which presumably resulted in substantial reduction in exposure to PCBs; and in 1971 Aroclor 1242 use was stopped in favor of Aroclor 1016. To account for the potential influence of these changes, separate variables corresponding to the periods 1946–1954, 1955–1965, 1966–1971, and 1972–1976 were incorporated into the model to indicate the era of exposure:

Era 1: 1946–1954 (25–33 yr before sampling)
Era 2: 1955–1965 (14–24 yr before sampling)
Era 3: 1966–1971 (8–13 yr before sampling)
Era 4: 1972–1976 (3–7 yr before sampling)

Sixteen variables corresponding to the four exposure levels by four time intervals were created. Each of the 16 variables was defined as the number of months served in the corresponding job category within the relevant era. People who changed job exposure categories within an era but were continuously employed thereby contribute a number of exposure months equal to the era's length. Their contribution to a specific exposure category is the number of months they worked at jobs of that category. The resulting regression coefficients therefore have units proportional to PCB concentration times (era-specific, exposure-category-specific months)$^{-1}$. Details of serum PCB determinations have been previously reported (Lawton et al., 1985b).

Regression Model

Classical multiple-regression techniques based on least squares were employed with serum high-homolog PCB as the dependent variable and era-specific exposure-category months as the independent variables. To satisfy the analytic assumptions of homoscedasticity and normality of the error, weighting was performed using weights equal to the inverse of the variance within deciles of predicted values. The amount of variation ex-

plained by the regressions for various models was compared using the r^2 statistic.

Potential elimination of PCBs from the body was evaluated by assuming first-order kinetics and incorporating this into the regression model (Appendix). Statistical analyses were conducted using the Statistical Analysis System (SAS) (SAS Institute, Inc., 1982).

Application to Reproductive Outcome Study (Second Stage)

The regression models were applied in a study of the birth weights and gestational ages of the infants of female employees. Results of the second stage are reported elsewhere (Taylor et al., 1989). In this report we describe only the application of the exposure model developed here to this second-stage study cohort. From among the 3,018 total women employees, groups with direct and indirect exposure were selected to evaluate estimated PCB exposure at the time of reproductive outcome. The direct-exposure group included a random sample of 200 women who had ever held a direct exposure job, while the indirect-exposure group included a random sample of 205 women who never held a job in the direct-exposure areas. Clerical and management jobs were excluded from both groups.

Alternative Models

A series of alternative univariate regression models using dichotomous, ordinal, and continuous variables to estimate serum PCB levels were also run for comparison. Examples of predictor variables examined include ever having a direct-exposure job, highest level ever exposed (from none to high), and duration of employment.

RESULTS

Population

The 147 individuals upon whom results are reported included 114 males and 33 females, ranging in age from 25 to 78 yr old (mean = 44) in 1979, who had a median serum high-homolog PCB level of 53 ppb.

Air Concentrations

Results of air sample monitoring performed in both direct- and indirect-exposure job areas during the industrial hygiene surveys showed that, in 1977, air concentrations of PCBs in the indirect-exposure job areas were an order of magnitude below those in the direct-exposure job areas (27 vs. 310 $\mu g/m^3$, respectively). The indirect-exposure jobs, in turn, had much higher PCB concentrations than did areas surrounding the plants, where values averaging 6.2 $\mu g/m^3$ were recorded prior to dis-

continuation of PCB use. These concentrations all exceed previously reported urban ambient air averages of 0.1 $\mu g/m^3$ (Kutz and Yang, 1975).

Regression Model

The weighted least-squares regression model with 16 exposure-month variables identified 10 influential outlier values (standardized residual ≤ -2.5 or ≥ 2.5). After review of the original data these records were removed. The new regression resulted in an improvement in the r^2 from .56 to .71. Description of the participants by era-specific exposure category months variables is shown in Table 1. The "full" model with 16 independent variables was subsequently simplified by dropping the four statistically insignificant variables from era 4 and the single additional variable with a biologically implausible negative coefficient, era 2 low, which was also insignificant. The final "reduced" model with 11 variables is shown in Table 2 and is described in greater detail in the Appendix. The r^2 for this reduced model was .69 based on 147 observations.

TABLE 1. Description of the Era-Specific Months Variables Used in the Regression Models to Predict High-Homolog Serum PCB Level (n = 147)

Variable	Number of subjects with any time at given level	Range[a] (mo)	Mean (mo)
Era 1			
Indirect	30	4–95	32
Low	8	1–84	1
Medium	13	1–14	6
High	6	3–17	8
Era 2			
Indirect	73	1–132	96
Low	24	1–132	9
Medium	28	1–60	12
High	14	1–122	30
Era 3			
Indirect	91	1–72	43
Low	27	1–70	5
Medium	59	1–69	22
High	24	1–70	42
Era 4			
Indirect	84	1–60	37
Low	35	1–60	18
Medium	32	1–60	37
High	84	2–60	28

Note. See Materials and Methods for definition of variables.
[a]Among subjects with employment during the era specified.

TABLE 2. Final, Reduced, Weighted Regression Model for Predicting High-Homolog Serum PCB Levels (n = 147)

Variable	Beta	SE[a]	t Value	p Value
Intercept	24.71	2.63	9.37	0.0001
Era 1				
Indirect	75.22	55.34	1.35	0.1763
Low	652.11	220.65	2.95	0.0037
Medium	882.68	521.89	1.69	0.0931
High	77.50	694.56	0.11	0.9113
Era 2				
Indirect	6.58	3.44	1.91	0.0576
Medium	39.36	20.49	1.92	0.0569
High	95.84	15.94	6.01	0.0001
Era 3				
Indirect	1.69	0.80	2.09	0.0382
Low	0.69	1.96	0.35	0.7256
Medium	7.28	1.48	4.91	0.0001
High	16.01	2.78	5.76	0.0001
		r^2 = .69		

Note. See Materials and Methods for definition of variables.
[a]Standard error.

Application to Reproductive Outcome Study (Second Stage)

Using the model derived in stage one from the 147 subjects in the cohort (Table 2), estimated values ranged from 25 to 14,595 in the 200 women in the direct-exposure group, with a median of 52 ppb, while the range in the 205 women in the indirect-exposure group was from 25 to 2,240 with a median of 53 ppb. An example illustrating application of the final model to the work history of a single individual is shown in Figure 1.

Alternative Models

The predictability of the seven alternative models examined varied widely. The least predictive was the simple ever versus never direct-exposure model (r^2 = .01), while the most predictive used a single continuous variable, duration of employment in direct-exposure jobs (r^2 = .43).

DISCUSSION

We considered ever-direct versus only-indirect exposure as the simplest and most logical of the alternative exposure categorizations examined. However, it was a particularly poor predictor of serum PCB level (r^2 = .01). After the fact, the best we could do in predicting serum PCB

using a priori deductive models was with total months employed in direct-exposure jobs only (r^2 = .43). None of the logical deductive models approached the predictability of the empirical model developed here (r^2 = .69).

The coefficients observed in the final model generally show that time spent in higher exposure categories resulted in higher predicted serum PCB levels. They also indicate a substantial difference in exposure between eras, with increasing exposure for each subsequent era progressing back in time, and very high exposure during the most distant era. Individual coefficients should be interpreted with caution, however, in particular those from distant eras where data are sparse. Nevertheless, the indication of higher exposure in past eras is consistent with historical data from the plants denoting process and/or engineering changes that resulted in a reduction in exposure at the points in time used to demarcate eras for our analysis.

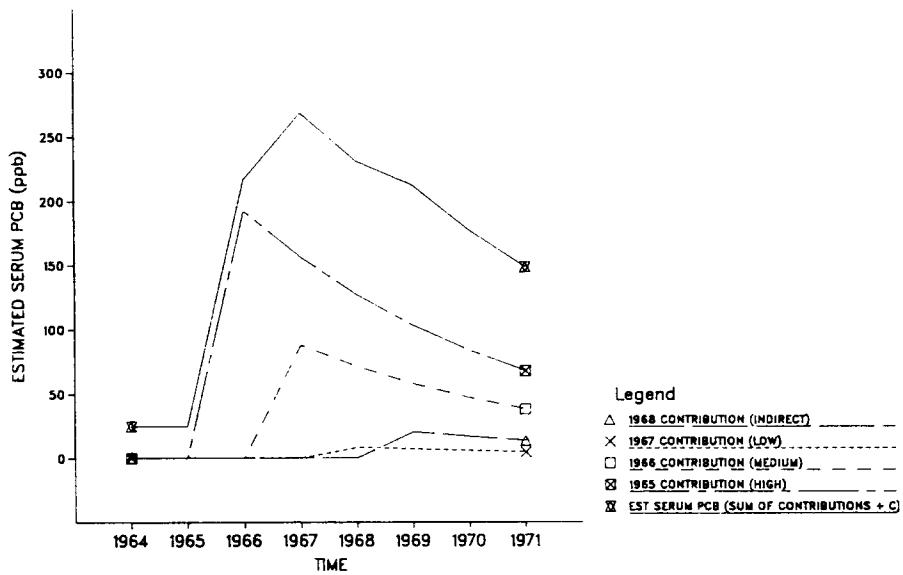

FIGURE 1. Background serum high-homolog PCB level is assumed to be equal to the intercept (24.71 ppb, Table 2). From era 3 coefficients shown in Table 2, each 1 mo of employment at a high-exposure job results in an increase above background of 16.01 ppb, so that at the end of a full 12 mo, serum PCB level due to high exposure in the year passed would be 12 × 16.01 ppb = 192.12 ppb. Cumulative exposure at the end of the year would equal background (24.71 ppb) plus the increment due to high exposure during the year (192.12 ppb) or 216.83 ppb. Similarly, a year at medium exposure would add 12 × 7.28 ppb = 87.36 ppb, a year at low exposure 12 × 0.69 ppb = 8.28, a year at indirect exposure 12 × 1.69 ppb = 20.28 ppb, and a year unemployed nothing. Excretion is assumed to follow first-order kinetics with a half-life estimated at 3.32 yr. This means that by April 1969 (3.32 yr later), half of the 192.12 ppb increment in serum PCB level due to the high exposure in 1965 will have been excreted. Estimation of total serum high-homolog PCB level at any point in time, therefore, takes into account level and duration of exposure as well as excretion.

The half-life of Aroclor 1254 estimated from serum measurements and used in the subsequent model development described here (3.32 yr) is in reasonably close agreement with the half-life of 4.8 yr estimated by Phillips et al. (1989) for Aroclor 1254. While changes in the half-life used in the model resulted in alteration of the relative magnitude of the coefficients between eras, with longer half-lives reducing the differences in coefficients between eras, the overall predictiveness of the model as measured by r^2 was little affected by changes in the half-life (data not shown).

The model described here was developed based on serum levels measured in 1979, 2–3 yr after exposure last occurred. Model development in this setting was possible because appropriate data were available to estimate the half-life of Aroclor 1254 (i.e., exposure measurements from two points in time without interval exposure), the half-life was long relative to the time since last exposure, and the exposure was high.

The model developed here is flexible and capable of estimating exposure in different ways to accommodate different postulated mechanisms of disease. Using information on estimated half-life from external data, exposure at a given point in time can be estimated as was done for the pregnancy study described. Similarly, by calculating exposure at multiple points in time, peak exposure can be estimated. This might be useful if the exposure under study was considered to act only above a certain threshold. And finally, by assuming a half-life of infinity, cumulative exposure up to any point in time can be estimated. Cumulative exposure would typically be the exposure of most interest in studies of cancer etiology.

In conclusion, we developed an empirical model that allows estimation of serum PCB levels in a variety of circumstances. Application of this empirical approach appears to offer substantial improvement to the alternative deductive approach in accurately estimating and categorizing exposure for epidemiologic studies.

APPENDIX

If we assume that PCBs introduced into a compartment are subject to a single removal process and that the fractional removal rate from that compartment is constant, then first-order kinetics may apply. The general formula for first-order kinetics is

$$Y(T, t) = Y(t)e^{-k(T-t)} \tag{1}$$

where $Y(T, t)$ is the concentration at the time of interest T due to exposure at time t, $Y(t)$ is the concentration at time t due to exposure at time t, and k is an elimination constant such that $e^{-k(T-t)}$ indicates elimination proportional to concentration at any given time. Cumulative concentra-

tion in year $(t - 1)$ to t is $Y_0(t) = \int Y(t, x)\, dx$. Cumulative exposure at T is

$$CY(T) = \int Y(T, t)\, dt = \sum Y_0(t) e^{-k(T-t)} \tag{2}$$

Let

$$Y_0(t) = B_0 + \sum A_{i,t} M_{i,t} \tag{3}$$

where B_0 is the background serum PCB level, $A_{i,t}$ is the serum concentration resulting from an exposure of 1 month at level i in time $(t - 1)$ to t years, and $M_{i,t}$ is the number of months exposure at level i in time $(t - 1)$ to t years. Concentration is therefore dependent on both level and time. A total of four different levels of exposure and four different time periods were identified in which PCB exposure was believed to have differed, and coefficients must be estimated for each level and era. A description of these levels and eras as used in the cohort under study is detailed in the Materials and Methods section.

Let

$$A_{i,t} = B_i(t) \tag{4}$$

where $B_i(t)$ is the era exposure factor for $(t - 1)$ to t for the ith level of individual exposure. Beta can be expressed as a function of time and level of exposure as

$$B_i(t) = \begin{cases} B_{i,\text{era}\,1} & \text{for} \quad 25 \le t < 33 \\ B_{i,\text{era}\,2} & \text{for} \quad 14 \le t < 25 \\ B_{i,\text{era}\,3} & \text{for} \quad 8 \le t < 14 \\ B_{i,\text{era}\,4} & \text{for} \quad 3 \le t < 8 \end{cases} \tag{5}$$

where i = indirect, low, medium, and high exposure levels. All values of t are years of employment before the time at which serum PCB level is being estimated.

Substituting $B_i(t)$ from Eq. (4) for $A_{i,t}$ in Eq. (3):

$$Y_0(t) = B_0 + \sum B_i(t) M_{i,t} \tag{6}$$

Substituting $B_0 + \Sigma B_i(t) M_{i,t}$ from Eq. (6) for $Y_0(t)$ in Eq. (2) gives us the model:

$$CY(T) = B_0 + \sum_t e^{-k(T-t)} \sum_i B_i(t) M_{i,t} \tag{7}$$

If $M_{i,t} = 0$ for $t = 0$–7 (there was no PCB exposure in yr 0–2, and the era 4 variables for yr 3–7 had no significant impact), then

$$CY(T) = B_0 + \sum_{t=8}^{33} e^{-k(T-t)} \sum_i B_i(t) M_{i,t} \tag{8}$$

REFERENCES

Kutz, F. W., and Yang, H. S. C. 1975. A note on polychlorinated biphenyls in air. *Proc. National Conf. Polychlorinated Biphenyls*, p. 182. Environmental Protection Agency, Washington, D.C.

Lawton, R. W., Ross, M. R., Feingold, J., and Brown, J. F., Jr. 1985a. Effects of PCB exposure on biochemical and hematological findings in capacitor workers. *Environ. Health Perspect.* 60:165–184.

Lawton, R. W., Brown, J. F., Jr., Ross, M. R., and Feingold, J. 1985b. Comparability and precision of serum PCB measurements. *Arch. Environ. Health* 40:29–37.

Morrow, P. E. 1977. Clearance kinetics of inhaled particles. In *Respiratory Defense Mechanisms*, eds. J. D. Brain, D. F. Proctor, and L. M. Reid, p. 493. New York: Marcel Dekker.

Phillips, D. L., Smith, A. B., Burse, V. W., Steele, G. K., Needham, L. L., and Hannon, W. H. 1989. Half-life of polychlorinated biphenyls in occupationally exposed workers. *Arch. Environ. Health* 44:351–354.

SAS Institute, Inc., 1982. *SAS User's Guide*. Cary, N.C.: SAS Institute, Inc.

Taylor, P. R., Stelma, J. M., and Lawrence, C. E. 1989. The relation of polychlorinated biphenyls to birthweight and gestational age in the offspring of occupationally exposed mothers. *Am. J. Epidemiol.* 129:395–406.

Received April 30, 1991
Accepted July 13, 1991