ESTIMATED DISTRIBUTIONS OF PERSONAL EXPOSURE TO RESPIRABLE PARTICLES

RICHARD LETZ, P. BARRY RYAN, and JOHN D. SPENGLER

Department of Environmental Science and Physiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, U.S.A.

(Received 8 August, 1983)

Abstract. A method of estimating distributions of exposure to respirable particles is presented. Using pollutant monitoring data from outdoors and indoors, time-activity data and a time-weighted exposure model, means and variances for exposure distributions are generated. Variances are estimated using Gauss' law of error propagation. The model is calibrated using data from a personal monitoring study. Estimated distributions of exposure to respirable particles for children in six cities living in homes with and without smokers are presented. The implications of these estimates for air pollution epidemiology and needs for further research are discussed.

1. Introduction

The Harvard Air Pollution Health Study is a prospective epidemiologic study involving about 20 000 people in six communities. Respiratory symptoms and pulmonary function have been measured on these people for nine years (Ferris *et al.*, 1979). Air pollutant concentrations have been measured at central sites in these communities. Respiratory symptoms as well as pulmonary function changes have been related to exposure categories based on health questionnaire items such as the presence of smokers or gas cooking stoves in the home (Speizer *et al.*, 1980). Previous studies by our group were directed toward characterizing indoor pollutant concentrations associated with these simple dichotomous categories and attempted to identify better descriptors to estimate exposure (Spengler *et al.*, 1979).

Perhaps the best method of determining air pollution exposure for the people in our health study would be to do personal monitoring on everyone. Our group has conducted personal monitoring studies of respirable particle (Spengler *et al.*, 1981) and nitrogen dioxide (Quackenboss *et al.*, 1982) exposures of some participants. However, personal monitoring on all participants is not feasible for many reasons. The most important reason is cost; i.e., personal monitoring is very expensive. Also, personal dosimeters are not available for all the pollutants of interest and all have (size, weight, accuracy) limitations. (For a review, see Wallace and Ott, 1982.)

Experience with personal monitoring studies shows not only that personal monitoring is expensive, but also that personal exposures can be poorly correlated with central site ambient concentrations. Also, current ambient air quality or personal exposure monitoring may not reflect past exposures. Therefore, our group has focused on developing models of indoor pollutant concentrations (Ryan *et al.*, 1983). Exposure models allow estimation of pollutant exposure for groups of people and time periods for which personal monitoring has not been conducted. A simple approach for estimating distributions of exposure to respirable particles (RSP) is presented in this paper. Some evidence for validation of the model with data from a personal monitoring study of adults in Kingston-Harriman, Tennessee, is presented, and the model is applied to provide estimates of RSP exposure for children in our six communities.

2. Basic Approach

The general approach used in this paper is that of a time-weighted average concentration summed over microenvironments (Fugas, 1975; Duan, 1982), i.e.,

$$E = \sum_i E_i = \sum_i f_i C_i \tag{1}$$

where E is the mean exposure and E_i, f_i, and C_i are exposure, fractions of time and pollutant concentrations, respectively, in the i^{th} microenvironment. Pollutant concentrations are estimated by extrapolation from existing measurements or calculation from knowledge of source strengths, ventilation, removal and mixing volume. Our group has conducted indoor and outdoor monitoring in a large number of homes in the six cities (Spengler *et al.*, 1979) and has developed a framework for estimating indoor concentrations from outdoor concentrations (Ryan *et al.*, 1983; Sexton *et al.*, 1983):

$$C_i = p_i C_{OUT} + S_i \tag{2}$$

where in the i-th microenvironment, C_i is the concentration, p_i the 'effective' penetration, and S_i the 'effective' indoor source strength, and C_{OUT} is the pollutant concentration outdoors. The p_i and S_i are called 'effective' because they include factors for air exchange as well as pollutant deposition due to chemical and physical action (see Ryan *et al.*, 1983). Effective penetration is affected by various home characteristics such as infiltration rate and presence of active surfaces for deposition. Indoor effective source strengths can be affected by human activities such as smoking and hobbies. The utility of expressing the relationship in this way is that indoor pollutant concentrations are considered to be a function of outdoor concentrations and two parameters that can be estimated from indoor/outdoor monitoring. This approach allows extrapolation to locations other than those monitored, where outdoor pollutant concentrations may be different.

The approach outlined so far should allow estimation of mean exposures, given estimates of the microenvironment fractional times (f_i) and pollutant concentrations (C_i). However, to calculate distributions one needs an estimate of the variance about the mean exposure. If not only the population mean but also the population variance of each parameter in the model is known, Gauss' law of error propagation (Bevington, 1969) can be used to approximate the variance about the estimated mean exposure:

$$\operatorname{var}(E) \approx \sum_i \left\{ \left(\frac{\partial E}{\partial C_i} \right)^2 \sigma_{C_i}^2 + \left(\frac{\partial E}{\partial f_i} \right)^2 \sigma_{f_i}^2 \right\}. \tag{3}$$

This approximation is good only if the uncertainties of the parameters are small relative to their means and the parameters are uncorrelated. In the current analysis the model parameters are assumed to be uncorrelated because no data were available to estimate the correlations. If such data were available, then the appropriate covariance terms could be added to Equation (3) in a straight-forward manner.

3. Model Validation

A distribution of exposures to respirable particles was available from a personal monitoring study of 88 non-smoking adults in Kingston-Harriman, Tennessee (Spengler *et al.*, 1983). To predict this exposure distribution, we will use a model with 5 microenvironments: outdoors (OUT), indoors at home while awake (HA), indoors at home while asleep (HS), other non-home indoor environments (OI), and vehicular travel (T). Two distinct microenvironments are associated with the home because the RSP concentrations are substantially different indoors when people are active and when they are not (NAS, 1981). People generate particles with their activities; e.g., cooking, cleaning, smoking. The expanded expression for estimating mean RSP exposure thus becomes:

$$E = f_{HA}C_{HA} + f_{HS}C_{HS} + f_{OI}C_{OI} + f_TC_T + (1 - f_{HA} - f_{HS} - f_{OI} - f_T)\,C_{OUT} \tag{4}$$

where E, f, and C are as in Equation (1). Note that because the fractional times must sum to unity, the fractional time spent in the last microenvironment is one minus the other fractional times. Thus, an n-microenvironment model has $2n$–1 free parameters.

Means and variances for fractional times, outdoor RSP concentrations and exposures were taken from Spengler *et al.* (1981). Personal exposures were monitored for three 24-hr periods using the Harvard/EPRI cyclone pump. The fractional times observed in the personal monitoring study are presented in Table I. The values agree well with those available from other sources (Koontz and Robinson, 1982; Chapin,

TABLE I

Time fractions used for 5-microenvironment adult's model

Micro-environment	Fractional time	Standard deviation
Home-awake (f_{HA})	0.38	0.25
Home-asleep (f_{HS})	0.30	0.05
Other-indoors (f_{OI})	0.15	0.16
Vehicular travel (f_T)	0.05	0.06
Outdoors ($1-f_{HA}f_{HS}f_{OI}f_T$)	0.12	–

1974). Data from our indoor/outdoor monitoring network (Spengler *et al.*, 1979) were used to estimate indoor pollutant concentrations. Annual average indoor RSP concentrations were regressed on outdoor RSP concentrations for 57 homes across 6 cities, yielding:

$$C_{IN} = 0.385\, C_{OUT} + 29.4(Smoking) + 13.8 \tag{5}$$

where smoking is a 0–1 dichotomous indicator variable and C_{OUT} and C_{IN} are in $\mu g\ m^{-3}$. The mean outdoor concentration was 19 $\mu g\ m^{-3}$ with a standard deviation of 11 $\mu g\ m^{-3}$, while the mean indoor concentration was 28 $\mu g\ m^{-3}$ with a standard deviation of 21 $\mu g\ m^{-3}$. The root mean square error was 16 $\mu g\ m^{-3}$ from the regression model, or about 60% of the mean.

This analysis indicated that indoor RSP concentrations would be about 40% of outdoor values when there are no indoor sources from smoking and people's activities. The smoking source of 29.4 $\mu g\ m^{-3}$ and the 13.8 $\mu g\ m^{-3}$ source due to people's activities have to be adjusted upward because these values represent the contributions of these sources to the 24-hr integrated value. Both sources are assumed 'off' during the HS fraction of time and 'on' during the HA fraction of time. The adjusted effective smoking source for the HA microenvironment thus becomes the observed value, 29.9 $\mu g\ m^{-3}$, divided by the fraction of time spent at home awake (f_{HA}), 0.38, or equal to 78.7 $\mu g\ m^{-3}$. The source due to people's activities was estimated to be 36.6 $\mu g\ m^{-3}$ ($= 13.9 \div 0.38$). Thus, for homes with non-smokers when the outdoor concentration was 18 $\mu g\ m^{-3}$ the indoor concentration was estimated as:

$$C_{HA} = 0.385(18) + 78.7(0) + 36.6 = 43.5. \tag{6}$$

The corresponding C_{HA} for homes with smokers would be 122.2 $\mu g\ m^{-3}$, and C_{HS} for both types of homes would be 6.9 $\mu g\ m^{-3}$. It was also assumed that C_{OI} is the same as C_{HA} in non-smoking homes, a conservative assumption, especially for groups with occupational exposure or passive smoke exposure away from home.

In addition to mean RSP concentrations for the microenvironments, the variances about these means are needed as input for the model. An attempt to use the standard errors on the p_i and S_i estimated from the outdoor/indoor regression model (Equation (5)) according to Gauss' law produced results that were inconsistent with observed values. It is likely that the estimates of the standard errors on the regression coefficients were poor because of the non-symmetric nature of the observed distributions and the violation of the assumption of homogeneity of variance. However, the violation of these assumptions should have little effect on the parameter estimates themselves. Results from indoor/outdoor monitoring (Spengler *et al.*, 1979, 1983) indicate that variances of indoor concentrations increase with increasing mean concentration. The root mean square error in the regression model was 16 $\mu g\ m^{-3}$ on an average indoor concentration of 28 $\mu g\ m^{-3}$, so 60% of the mean concentration for each microenvironment was used as the standard deviation for that C_i.

The observed RSP means and standard deviations for non-smokers living in homes with either non-smokers or smokers are presented in the first row of Table II, and the

TABLE II
Observed and predicted exposures to respirable particles in a personal monitoring study in Kingston-Harriman, Tenn.

	Non-smoke exposed		Smoke-exposed	
	mean ($\mu g\ m^{-3}$)	std. dev. ($\mu g\ m^{-3}$)	mean ($\mu g\ m^{-3}$)	std. dev. ($\mu g\ m^{-3}$)
Measured personal exposures	36	21	64	46
5-micro model estimates	33	18	60	42

corresponding predicted values are in the second row. The predicted values agree well with the observed values for both the means and the standard deviations for both groups. The predicted values are slightly lower than the observed values, but this may be attributable to occupational exposures which would not be well-characterized by our assumption that the mean and variance for the C_{OI} is the same as the C_{HA}.

4. Application to Children in Six Cities

The children's 5-microenvironment model is the same as that presented for the adults except that the adults' other-indoor (OI) microenvironment is changed to a school (S) microenvironment. The fractional times for the children's model are presented in Table III. These fractions were observed in an as-yet-unpublished personal monitoring study of children's exposure to NO_2 conducted in Watertown, Massachussets, during the fall of 1982. They agree closely with values observed in another study in Portage, Wisconsin, (Quackenboss *et al.*, 1981) and other values calculated from school hours and absence figures.

TABLE III
Time fractions assumed for 5-microenvironment children's model

	Annual	School year	Summer
f_{HA}	0.40	0.40	0.40
(s.d.)	(0.20)	(0.20)	(0.20)
f_{HS}	0.30	0.30	0.30
(s.d.)	(0.10)	(0.10)	(0.10)
f_S	0.12	0.16	0.01
(s.d.)	(0.02)	(0.03)	(0.0003)
f_T	0.03	0.03	0.03
(s.d.)	(0.03)	(0.03)	(0.03)
f_O	0.18	0.14	0.29

Indoor concentrations and their variances were calculated as for the adult model. Annual averages and (spatial) variances were taken from data available from previous indoor/outdoor monitoring in each city (Spengler *et al.*, 1979). The means and standard deviations (spatial distribution) of outdoor RSP concentrations for each city as well as predicted mean exposures and their standard deviations are presented in Table IV. Exposure estimates are presented separately for children not living with smokers and children living with smokers. Note that although the mean outdoor concentrations of RSP vary by a factor of about four, estimated mean exposures for children living in non-smoking homes vary by less than a factor of two. For example, the mean exposure of children in smoking homes where outdoor RSP concentrations are lowest (Portage) is higher than the mean exposure of children in non-smoking homes where outdoor concentrations are on average four times higher (Steubenville).

Distributions of exposures observed in our personal monitoring studies can be fit well with a gamma distribution. Gamma distributions have variances proportional to their means, cannot assume negative values, are generally skewed to the right, and approach the Gaussian distribution as the variance becomes small relative to the mean. Further, gamma distributions are additive in the sense that the sum of two gamma-distributed

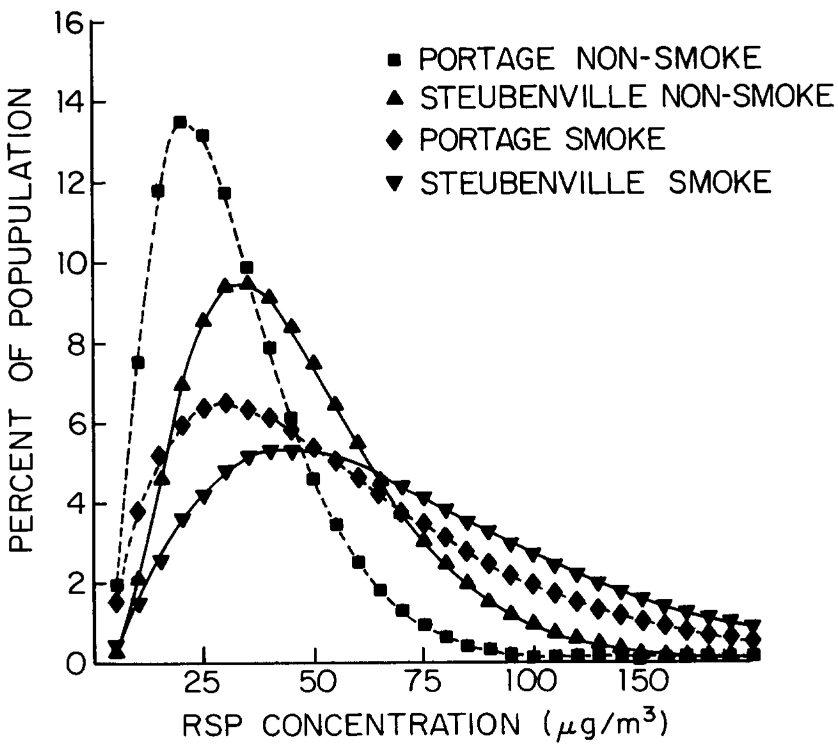

Fig. 1. Estimated distributions of RSP exposure for children living in homes with and without smokers in Portage, Wisconsin, and Steubenville, Ohio.

TABLE IV

Outdoor concentrations of respirable particles and predicted exposures for children in six cities using a 5-micronenvironment model

City/Group	Outdoor		Exposure		% > 75
	Mean[a]	S.D.[a]	Mean[a]	S.D.[a]	μg m^{-3} [b]
Non-smoke explosed					
Portage	11	1	26	15	1
Topeka	12	2	26	15	1
Watertown	19	2	30	17	2
Kingston	17	2	29	16	2
St. Louis	20	3	30	17	2
Steubenville	41	10	40	22	7
Smoke-exposed					
Portage	11	1	56	41	25
Topeka	12	2	56	41	26
Watertown	19	2	60	42	28
Kingston	17	2	59	42	28
St. Louis	20	3	60	42	29
Steubenville	41	10	70	46	38

[a] in μg m^{-3}.

[b] Assuming a gamma distribution with the predicted mean and standard deviation.

variables also has a gamma distribution (Hays, 1973). Gamma distributions of predicted exposures of RSP for children living in homes with smokers and non-smokers in the cities with the highest and lowest outdoor concentrations are presented in Figure 1. The estimated percentage of children in the six cities having annual average RSP exposures greater than 75 μg m^{-3} (the NAAQS for *TSP*), assuming gamma distributions with the predicted means and standard deviations, are presented in the last column of Table IV. A sizeable percentage of the children living in homes with smokers is seen to have estimated exposures above the NAAQS for *TSP*, even in areas with low outdoor RSP concentrations.

5. Discussion

A simple approach to estimating distributions of RSP exposure for children in the Harvard Air Pollution Health Study is presented in this paper. Obviously, much work remains to be done. Model validation needs to be done with results from personal monitoring studies. Better estimates of the model parameters and especially their variances are needed. Some improvements can be made with analysis of existing data, but additional data on both pollutant concentrations and fractional times in other microenvironments are needed. Chemical and elemental analysis of indoor and personal RSP samples could be used as a means of quantifying source contributions. Also, there is a need to reconsider microenvironment definitions to minimize within-microenvironment variances rather than using arbitrary *a priori* schemes (Duan, 1982). Improvements in estimation of exposure variances may require attention to co-variances between model parameters. Finally, the implications of assumptions about the form of distributions of pollutant concentrations and personal exposures needs to be explored.

By choice, the approach presented in this paper rests on some simplifying assumptions. This analysis assumes that respirable particles from different sources are equivalent. In particular, particles from tobacco smoke are assumed to be equivalent to particles from outdoor sources. The chemical and elemental composition clearly differs between particles from outdoor sources and particles from indoor sources (see, e.g., Colome *et al.*, 1982). Ongoing research should provide information for determining whether to segregate these sources on the basis of their toxicity and how to do so. A second set of complications was avoided in the current analysis by restricting it to estimating annual average exposures. Outdoor concentrations and acidity, the percentage of time that people spend outdoors and the penetration of pollutants to indoor microenvironments all change during the course of the year. Not only do they change, but they co-vary, i.e., all are highest during summer in most places. These factors may prove to be important in the long run, and, conceptually, they can be incorporated into the exposure model. However, each additional factor added to the model has a multiplicative effect on the number of parameters that must be estimated. The approach presented here is both the simplest one that accounts for the fundamental components and the most complicated one that is reasonably supported by existing data.

The large variance of exposure within exposure categories and the large overlap of observed, as well as predicted, exposure distributions, has implications for the power of epidemiologic investigations (Shy *et al.*, 1978). It suggests that very large numbers of subjects are necessary to detect health effects differences between people grouped into exposure categories, even when the mean difference in pollutant concentrations between exposure categories appears large. Alternatively, it suggests that we might achieve greater efficiency by quantifying exposures well with personal monitoring on a much smaller number of subjects. Finally, this work points out the potential importance of indoor sources (e.g., percent of homes with smokers, amount smoked, etc.) as confounders in studies of the effects of outdoor air pollution across communities.

Acknowledgments

We thank Drs John S. Evans, James H. Ware, and Benjamin G. Ferris, Jr., for their critical reading of the manuscript and helpful suggestions. This work was suppported under general support provided to the Harvard Air Pollution Health Study through NIEHS grant ES-01108 Electric Power Research Institute contract RP 1001–1, and EPA grant EPA 68–02–3466.

References

Bevington, P. R.: 1969, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.
Chapin, F. S.: 1974, *Human Activity Patterns in the City*, Wiley-Interscience, New York.

Colome, S. D., Spengler, J. D., and McCarthy, S.: 1982, 'Comparisons of Elements and Inorganic Compounds Inside and Outside Residences', *Environ. Int.* **8**, 197–212.

Duan, N.: 1982, 'Models for Human Exposure to Air Pollution', *Environ. Int.* **8**, 305–309.

Ferris, B. G., Jr., Speizer, F. E., Spengler, J. D., Dockery, D. W., Bishop, Y. M. M., Wolfson, M., and Humble, C.: 1979, 'Effects of Sulfur Oxides and Respirable Particles on Human Health: Methodology and Demography of Populations in Study', *Am. Rev. Resp. Dis.* **120**, 767–779.

Fugas, M.: 1975, 'Assessment of Total Exposure to an Air Pollutant', Proceedings of the International Conference on Environmental Sensing and Assessment, IEEE # 75-CH 1004-1 ICESA, Las Vegas, Nevada.

Hays, W. L.: 1973, *Statistics for the Social Sciences*, Holt, Rinehart and Winston, New York.

Koontz, M. D. and Robinson, J. P.: 1982, 'Population Activity Patterns – St. Louis Study', *Environ. Mon. Assess.* **2**, 197–212.

NAS: 1981, *Indoor Pollutants*, National Academy Press, Washington, D.C.

Quackenboss, J. J., Kanarek, M. S., Spengler, J. D., and Letz, R.: 1982, 'Personal Monitoring for Nitrogen Dioxide Exposure: Methodological Considerations for a Community Study', *Environ. Int.* **8**, 249–258.

Ryan, P. B., Spengler, J. D., and Letz, R.: 1983, 'The Effects of Kerosene Heaters on Indoor Pollutant Concentrations: A Monitoring and Modeling Study', *Atmos. Environ.* **17**, 1339–1345.

Sexton, K., Letz, R., and Spengler, J. D.: 1983, 'Human Exposures to Nitrogen Dioxide: An Indoor/Outdoor Modeling Approach', *Environ. Res.* **32**, 151–166.

Shy, C. M., Kleinbaum, D. G., and Morgenstern, H.: 1978, 'The Effect of Missclassification of Exposure Status in Epidemiological Studies of Air Pollution Health Effects', *Proc. N. Y. Acad. Med.* **54**, 1155–1165.

Speizer, F. E., Ferris, B. G., Jr., Bishop, Y. M. M., and Spengler, J. D.: 1980, 'Respiratory Disease Rates and Pulmonary Function in Children Associated with NO_2 Exposure', *Am. Rev. Resp. Disease* **121**, 3–10.

Spengler, J. D., Ferris, B. J., Jr., Dockery, D. W., and Speizer, F. E.: 1979, 'Sulfur Dioxide and Nitrogen Dioxide Inside and Outside Homes and the Implications for Health Effects Research', *Environ. Sci. Tech.* **13**, 1276–1280.

Spengler, J. D., Treitman, R. D., Tosteson, T. D., and Mage, D. T.: 1981, 'Personal Exposures to Respirable Particulates: A Tale of Two Cities – Kingston and Harriman, Tennessee', paper presented at the *International Symposium on Indoor Air Pollution, Health and Energy Conservation*, Amherst, Mass.

Spengler, J. D., Duffy, C. P., Letz, R., Tibbits, T. W., and Ferris, B. G.: 1983, 'Nitrogen Dioxide Inside and Outside 137 Homes and Implications for Ambient Air Quality Standards and Health Effects Research', *Environ. Sci. Technol.* **17**, 164–168.

Wallace, L. A. and Ott, W. R.: 1982, 'Personal Monitors: A State-of-the-art Survey', *J. Air Poll. Control Assoc.* **32**, 601–610.