Journal of Exposure Analysis and Environmental Epidemiology (2004) 14, 92–107

www.nature.com/jea

Variability of environmental exposures to volatile organic compounds

STEPHEN M. RAPPAPORT AND LAWRENCE L. KUPPER

School of Public Health, University of North Carolina, Chapel Hill, North Carolina, USA

Although studies of occupational exposure to volatile organic compounds (VOCs) often partition variability across groups, and between and within persons, those of environmental exposure to VOCs have not involved such partitioning. Using data from the Environmental Protection Agency's total exposure assessment methodology (TEAM) studies, we partitioned exposure variability across cities, and between and within persons for nine VOCs. The estimated variance components decreased in the order: within-person > between-person > across city. Despite their smaller magnitudes, estimates of between-person and across-city variance components were sufficiently large to provide reasonable contrast for informative epidemiology studies of most VOCs. Estimates of between-person variance components for environmental VOCs were similar to those published for occupational VOCs (groups defined by job and factory). However, estimates of within-person variance components were much greater for environmental VOCs, probably due to the greater diversity of locations (including the workplace) visited by the general public over time. For benzene and perchloroethylene, we used a simple model to calculate numbers of personal measurements required to relate the exposure level to health outcome statistically. About 10 times more personal measurements would be required to investigate perchloroethylene exposure as compared to benzene exposure; this disparity reflects the greater within-subject variability of perchloroethylene data compared to benzene data. We conclude that variability should be partitioned for environmental VOC exposures in much the same manner as for occupational exposures. There should be sufficient variability in the levels of most VOCs across cities and between subjects to provide reasonable contrast for informative epidemiology studies, as we illustrate for exposures to benzene. Yet, epidemiologists should be wary of investigating environmental VOCs without preliminary data with which to estimate the variance structure of exposure variables.
Journal of Exposure Analysis and Environmental Epidemiology (2004) **14,** 92–107. doi:10.1038/sj.jea.7500309

Keywords: *volatile organic compounds, environmental exposure, exposure variability*.

Introduction

Volatile organic compounds (VOCs) are generated from many human activities, notably smoking, chlorination of water, operating and refueling motor vehicles, and using solvents (reviewed by Wallace, 2001). Some VOCs are known or suspected carcinogens, including benzene, 1,3-butadiene, formaldehyde, naphthalene, *p*-dichlorobenzene, perchloroethylene, styrene, and trichloroethylene. Thus, it would be reasonable to relate environmental VOC exposures to morbidity and mortality or to relevant biomarkers of effect. Yet, the epidemiology of VOCs has hardly been explored in nonoccupational populations.

People are exposed to VOCs in their homes and workplaces, while moving about, and during the varied activities of everyday life. Sources of VOCs can vary across locations (cities or neighborhoods, say), between persons at a particular location (due to different individual sources and lifestyle factors), and within persons over time (due to differences in activities, mobility, and environmental factors). Thus, in order to assess VOC levels properly, one should repeatedly monitor personal exposures while carefully documenting subjects' locations, sources of VOCs, lifestyle factors, and activities. Then, random- or mixed-effects models can be used to identify the determinants of exposure and to partition the variability of air levels across locations, and between and within persons. The fixed effects and variance components estimated under these models can be invaluable in designing and interpreting epidemiology studies.

Random- and mixed-effects models have been widely applied to investigate workplace exposures to all classes of airborne contaminants (including VOCs) (see, for example, Kromhout et al., 1993, 1994; Symanski et al., 1996, 2001; Peretz et al., 1997, 2002; van Tongeren et al., 1997; Rappaport et al., 1999) as well as environmental exposures to NO_2 (Brunekreef et al., 1987; Spengler et al., 1994; Rijnders et al., 2001) and particulate matter (Janssen et al., 1997, 1998, 1999). However, we are not aware of any applications of these models to environmental VOC exposures. This is surprising because several environmental studies have collected the repeated personal measurements to VOCs that are needed to apply such models, including the classic total exposure assessment methodology (TEAM)

1. Address all correspondence to: Dr. Stephen M. Rappaport, School of Public Health, University of North Carolina, Chapel Hill, North Carolina, USA. Tel.: +1-919-966-5017. Fax: +1-919-966-0521.
E-mail: smr@unc.edu
Received 21 March 2003; accepted 28 July 2003

studies conducted in the US in the 1980s (reviewed by Wallace, 2001) and more recent investigations (e.g., Hoffmann et al., 2000; Kinney et al., 2002).

In the current investigation, we use data from early TEAM studies to estimate variance components of random effects describing the variation in environmental exposures to nine VOCs[2] across US cities, between subjects within these cities, and within subjects over time. We also investigate the contributions of seasonal effects and assay errors to exposure variability. We then compare estimated between- and within-person variance components from environmental VOC exposures with those observed in occupational studies. Finally, we use a simple exposure–health-effect model to estimate sample sizes required to study environmental exposures to benzene and perchloroethylene, two volatile carcinogens expected, on the basis of estimated variance components, to produce good and poor exposure information, respectively, for informative epidemiological studies.

Methods

Sources and Limitations of Data

Environmental exposure data were obtained from the TEAM studies conducted by the US Environmental Protection Agency between September 1981 and July 1987. The data were downloaded from an online database (THERdbASE (Total Human Exposure Risk database and Advanced Simulation Environment) at website: http://www.epa.gov/heasd/edrb/therd/therd-home.htm). They consisted of concentrations of 20–26 VOCs in personal and outdoor air of residents of Bayonne, NJ; Elizabeth, NJ; Greensboro, NC; Devils Lake, ND; and Los Angeles, CA. Subject-identification numbers were included along with dates and periods of collection (day and night). These data are described in detail by Wallace (1987) in a report and by Wallace et al. in several published papers (Wallace, 1987, 2001; Wallace et al., 1988, 1991). For our analyses of TEAM data, we selected nine VOCs (see footnote 2) that were designated by Wallace as either ''ubiquitous chemicals'' or ''compounds very often but not always found'' (Wallace, 1987). After comparing the downloaded data to published descriptions of the TEAM studies, we concluded that the online database included most TEAM measurements between September 1981 and May 1984, plus those from Los Angeles in 1987; the database is missing measurements from Denver, CO (1983), Pittsburgh & Antioch, CA (1984), Valdez, AK (1986), and Elizabeth, NJ and Baltimore, MD (1987). We excluded one set of NJ samples, collected in July–August 1982, because it had become contaminated prior to laboratory analysis (discussed in Wallace, 1987, p. 21).

Briefly, TEAM investigators randomly selected subjects from each city based on stratified probability sampling of households. Personal air samples were then collected from each subject for two consecutive 12-h periods (day and night), and outdoor air samples were collected at the residences of a subset of subjects. Repeated measurements of some subjects were obtained in Bayonne and Elizabeth, NJ (2 measurements/subject) and in Los Angeles, CA (2–3 measurements/subject); the time lag between measurements was 25 days to 3 years. Although TEAM investigators collected extensive information from each subject regarding location, occupation, smoking, daily activities, and lifestyle factors, these variables were not included in the online database.

Personal and outdoor measurements were extracted from the TEAM database for the nine VOCs to be investigated. A small number of observations with air concentrations listed as zero were deleted (depending upon the compound the following numbers were deleted: 0–13 personal day measurements, 0–11 personal night measurements, 0–4 outdoor day measurements, and 0–5 outdoor night measurements). Depending upon the VOC, analyses of personal samples were performed with 398–558 pairs of day and night measurements, obtained from 356 to 445 subjects (some subjects had 2 or 3 measurements), while analyses of outdoor samples were performed with 158–289 pairs of day and night measurements, involving 150–179 subjects. Appendix A summarizes the sample sizes available for each VOC and city. The number of subjects with repeated personal measurements ranged from 42 to 111 (median = 110), depending upon the compound.

For comparison with occupational exposure data, between- and within-person variance components of personal exposures, measured in a variety of workplaces over the full work shift, were obtained from a published analysis by Kromhout et al. (1993). These variance components had been estimated with a one-way random effects model, using data from industrial surveys of 12 VOCs[3] from 62 occupational groups defined by job and factory (10,746 personal measurements were available from 855 workers.).

Statistical Analyses

Sources of variability across cities, and between and within persons, were first investigated using 24-h personal exposures of each VOC. Since histograms of VOC levels were right skewed (i.e., long tails to the right) and were heteroscedastic, models were applied after natural logarithmic transformation of the air levels. Let X_{hij} represent the VOC concentration ($\mu g/m^3$) measured during the jth repetition for the ith person

2. The following VOCs were investigated: benzene, chloroform, ethylbenzene, methylchloroform, *p*-dichlorobenzene, perchloroethylene, styrene, trichloroethylene, and *o*-xylene.

3. The following VOCs were reported by Kromhout et al.: benzene, diphenyl, diphenylether, ethanal, formaldehyde, an unspecified organic vapor, perchloroethylene, styrene, toluene, total solvents, trichloroethane, and xylene.

in the hth city, and let Y_{hij} represent the natural logarithm of X_{hij}. The following model was used to define three random sources of variation:

$$Y_{hij} = \ln(X_{hij}) = \mu_y + \alpha_h + \beta_{hi} + \varepsilon_{hij} \quad (1)$$

(for $h = 1, \ldots, 5$ cities; $i = 1, \ldots, k_h$ subjects in the hth city; $j = 1, \ldots, n_{hi}$ measurements of a particular VOC from the ith subject in the hth city). In Model (1), μ_y represents the true unknown mean of logged VOC levels, α_h is the random effect of the hth city, β_{hi} is the random effect of the ith person in the hth city, and ε_{hij} is the random-error effect of the jth measurement from the ith person in the hth city. It is assumed that α_h, β_{hi}, and ε_{hij} are each normally distributed, and mutually independent, with means of zero and variances of σ^2_α, σ^2_{B}, and σ^2_{W}, respectively; we use the 'B' and 'W' subscripts to emphasize that the latter two variance components represent variation between- and within-persons, respectively. We designate the estimates of μ_y, σ^2_α, σ^2_{B}, and σ^2_{W} as $\hat{\mu}_y$, $\hat{\sigma}^2_\alpha$, $\hat{\sigma}^2_{\mathrm{B}}$, and $\hat{\sigma}^2_{\mathrm{W}}$, respectively. Assumptions of uniform variance and normality of random effects were evaluated with histograms and q–q plots of residuals from fitting Model (1).

Under Model (1), the jth (logged) VOC level $Y_{hij} = \ln(X_{hij})$ for the ith person in the hth city is normal with variance σ^2_{W}. Hence, ${}_{\mathrm{W}}R_{0.95} = e^{(3.92\sigma_{\mathrm{W}})}$, the ratio of the 97.5th percentile to the 2.5th percentile of the corresponding lognormal distribution of X_{hij}, is the so-called fold-range of variation of VOC levels within persons (e.g., see Rappaport, 1991). Our estimate of ${}_{\mathrm{W}}R_{0.95}$ is ${}_{\mathrm{W}}\hat{R}_{0.95} = e^{(3.92\hat{\sigma}_{\mathrm{W}})}$. Similarly, the underlying mean (logged) VOC level $\mu_{yhi} = (\mu_y + \alpha_h + \beta_{hi})$ for a randomly chosen person in the hth city is normal with variance σ^2_{B}, leading to the estimated between-person fold-range ${}_{\mathrm{B}}\hat{R}_{0.95} = e^{(3.92\hat{\sigma}_{\mathrm{B}})}$. And, the underlying mean (logged) VOC level for the hth city $\mu_{yh} = (\mu_y + \alpha_h)$ is normal with variance σ^2_α, so that ${}_{\alpha}\hat{R}_{0.95} = e^{(3.92\hat{\sigma}_\alpha)}$ is the estimated across-city fold-range. For example, values of ${}_{\alpha}\hat{R}_{0.95} = 6$, ${}_{\mathrm{B}}\hat{R}_{0.95} = 17$, and ${}_{\mathrm{W}}\hat{R}_{0.95} = 27$ for personal samples of benzene suggest, respectively, that about 95% of US cities have mean benzene exposures within a roughly six-fold range, that about 95% of persons in a given city have mean benzene concentrations within approximately a 17-fold range, and that about 95% of daily air concentrations for a given person lie within a roughly 27-fold range.

We investigated the variability of 24-h personal concentrations of VOCs separately in NJ and CA samples during different seasons of the year (NJ: winter, summer, and fall; CA: winter, spring, and summer) using the following mixed model:

$$Y_{ij} = \ln(X_{ij}) = \mu_y + \sum_{m=1}^{2} \delta_m C_m + \beta_i + \varepsilon_{ij} \quad (2)$$

where X_{ij} is the VOC concentration ($\mu g/m^3$) measured during the jth repetition for the ith person in either NJ or CA, μ_y represents the true unknown mean (logged) VOC level, and δ_m, the regression coefficient for the dummy variable C_m, represents the fixed effect of the mth season. In Model (2), β_i and ε_{ij} are normally distributed and mutually independent, with means of zero and variances σ^2_{B} and σ^2_{W}, respectively. To quantify the contribution of seasonal effects to σ^2_{W}, we compared the values of $\hat{\sigma}^2_{\mathrm{W}}$ obtained under Model (2) with those from the one-way random-effects model (also applied separately to NJ and CA data) of the form:

$$Y_{ij} = \ln(X_{ij}) = \mu_y + \beta_i + \varepsilon_{ij} \quad (3)$$

where β_i and ε_{ij} are assumed to be normally distributed and mutually independent, with means of zero and variances σ^2_{B} and σ^2_{W}, respectively. The variability associated with seasonal effects was estimated as $\hat{\sigma}^2_{\mathrm{S}} = \hat{\sigma}^2_{\mathrm{W}(3)} - \hat{\sigma}^2_{\mathrm{W}(2)}$, where $\hat{\sigma}^2_{\mathrm{W}(3)}$ and $\hat{\sigma}^2_{\mathrm{W}(2)}$ represent the estimates of σ^2_{W} obtained under Models (3) and (2), respectively. The proportional contribution of seasonable variability to the within-subject variance component was estimated as $\hat{\sigma}^2_{\mathrm{S}}/\hat{\sigma}^2_{\mathrm{W}(3)}$.

The contributions of assay error to σ^2_{W} were estimated from median coefficients of variation (CVs) reported from analyses of duplicate VOC samples in the NJ and CA studies (134 duplicates for NJ and 34 duplicates for CA) (Wallace, 1987; Wallace et al., 1988). The variance component associated with assay error was estimated as $\hat{\sigma}^2_{\mathrm{A}} = \ln(1 + CV^2)$ (Rappaport, 1991), and its proportional contribution to the within-subject variance component as $\hat{\sigma}^2_{\mathrm{A}}/\hat{\sigma}^2_{\mathrm{W}(3)}$.

Model (3) was also applied to 12-h day and night personal measurements from a given city to compare estimates of σ^2_{B} and σ^2_{W} from the TEAM database with those from 62 occupational groups reported by Kromhout et al. (1993), who employed the same model. Since there were several high correlations between different sets of VOC levels from the TEAM data, only one VOC was chosen from each group of chemicals having Spearman's correlations (r_s) > 0.7. This resulted in 18 pairs of estimates of σ^2_{B} and σ^2_{W} from day measurements and 21 pairs from night measurements. Cumulative distributions of the estimated fold ranges (${}_{\mathrm{B}}\hat{R}_{0.95}$ and ${}_{\mathrm{W}}\hat{R}_{0.95}$) for these environmental variance components were compared to those derived from 62 pairs ($\hat{\sigma}^2_{\mathrm{B}}$, $\hat{\sigma}^2_{\mathrm{W}}$) from occupational studies.

Finally, Model (1) was applied in balanced form for H cities, $k_h = k$ subjects/city, and $n_{hi} = n$ measurements/subject, in conjunction with a measurement error model reported by Tielemans et al. (1998), to predict sample sizes for epidemiological studies of selected VOCs. In our context, this model assumes that a continuous health outcome for the ith person in the hth city (R_{hi}) is related to exposure by

$$R_{hi} = \theta_0 + \theta_1 \mu_{yhi} + e_{hi} \quad (4)$$

where θ_0 is the background outcome, θ_1 is the true regression coefficient for exposure, μ_{yhi} is the true unobservable (logged) exposure of the ith person ($\mu_{yhi} = \mu_y + \alpha_h + \beta_{hi}$ under Model

(1)) in the hth city, and e_{hi} is the error term assumed to be normal with mean zero and variance σ_e^2. When $\theta_1 > 0$, we show in Appendix B that the following inequality can be used to estimate sample sizes (i.e., sets of values of H, k, and n, and hence total numbers of measurements (Hkn)) required so that the lower limit of a $100(1-\alpha)\%$ confidence interval for θ_1 exceeds zero with probability at least $(1-\beta)$:

$$\frac{(H-3)(k\sigma_\alpha^2+\sigma_B^2)}{(\sigma_e^2/\theta_1^2)(1+\sigma_W^2/n(k\sigma_\alpha^2+\sigma_B^2))+(\sigma_W^2/n)} \geq (z_{1-\alpha/2}+z_{1-\beta})^2 \quad (5)$$

In Eq. (5), $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard-normal distribution and $z_{1-\beta}$ is defined accordingly. Equation (5) was used to predict combinations of $H(>3)$, k, and $n(\geq 2)$ given values of θ_1^2/σ_e^2, σ_α^2, σ_B^2, σ_W^2, $z_{1-\alpha/2}$ and $z_{1-\beta}$. The quantity θ_1^2/σ_e^2, analogous to a squared 'signal-to-noise ratio', was arbitrarily set to values of 0.25, 1.0, and 4.0, a range that should contain most realistic scenarios. For example, in a relevant biomarker study, the (logged) level of benzene oxide–albumin adducts was found to be linearly related to (logged) occupational benzene exposure, with an estimated value of $\theta_1^2/\sigma_e^2 = 1.28$ (Rappaport et al., 2002), which is roughly in the middle of this range.

All statistical procedures employed SAS software (v8.2, SAS Institute, Cary, NC, USA). Estimates of parameters in Models (1)–(3) were obtained using either Proc NESTED or Proc MIXED, as indicated. The correlation between pairs of VOCs was investigated with Spearman's coefficients via Proc CORR. Ratios of medians of the distributions of personal air concentrations and outdoor air concentrations, and the corresponding interquartile ranges, were estimated via Proc UNIVARIATE, using data only from subjects having both types of measurements.

Results

Personal/Outdoor Ratios

Fig. 1 shows median ratios and interquartile ranges of ratios of personal- to outdoor-air concentrations for the subset of subjects having both types of measurements. All personal/outdoor ratios were greater than unity, with a typical VOC having a personal air concentration about twice the outdoor value (median ratio = 2.10, range: 1.79–3.85).

Air Concentrations Aggregated by City

The median 24-h air concentrations, as well as minima, maxima, and 25th and 75th percentiles, are given in Appendices C and D for the nine VOCs measured in five US cities based upon personal and outdoor measurements, respectively. Table 1 summarizes the highest and lowest median concentrations for these cities and gives the concentration ratio of the high city to the low city in each case. These high- /low-city concentration ratios were much smaller for personal measurements (median = 3.06, range: 1.74–3.77) than for outdoor measurements (median = 13.4, range: 5.06–987).

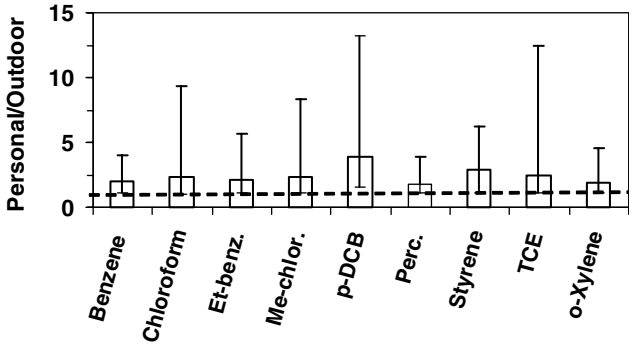

Figure 1. Median ratios of personal to outdoor-air concentrations for nine VOCs among subjects having both types of measurements (error bars represent interquartile ranges). The dashed line represents a ratio of one. Abbreviations: Et-benz. = ethylbenzene, Me-chlor. = methylchloroform, p-DCB = p-dichlorobenzene, Perc. = perchloroethylene, TCE = trichloroethylene. (See Appendix A for sample sizes.).

Correlations Between VOC Levels for Personal Measurements

Correlations of personal measurements were computed for pairs of the nine VOCs. Spearman's correlation coefficients (r_s) are given in Appendix E for 24-h data by city for all pairs of VOCs. The correlations tended to be small, with $r_s < 0.5$ in most cases; the only consistent exceptions were for ethylbenzene and o-xylene, where $r_s \geq 0.9$ for all cities, except Devils Lake.

Exposure Variability Partitioned Across Cities and Between and Within Persons

Table 2 summarizes variance components, and the corresponding fold ranges and mean values, estimated from 24-h personal measurements under Model (1) (Proc NESTED). The results show the dominance of the estimated within-person variance component in all cases (mean contribution = 74.1% of the total variance), with the remaining variability partitioned between persons (mean contribution = 15.7% of the total) and across cities (mean = 10.2% of the total).

Contributions of Seasonal Effects and Assay Errors

Seasonal effects were estimated under Model (2) for samples from NJ and CA and were highly significant for all nine VOCs ($P < 0.0001$). The results are summarized in Table 3 as the air levels of the nine VOCs predicted during the measured seasons. For all chemicals, Los Angeles' air levels were greater in the winter than in the spring or summer. Seasonal differences in NJ samples, although less pronounced than those in CA, followed a similar pattern, with fall and winter

Table 1. Ranges of median 24-h air concentrations ($\mu g/m^3$) of nine VOCs measured in five US cities.

	Personal concentrations			Outdoor concentrations		
Chemical	Low city	High city	High/low	Low city	High city	High/low
Benzene	7.30	22.4	3.06	4.45	14.4	3.23
Chloroform	1.14	4.30	3.77	0.055	0.829	15.2
Ethylbenzene	3.10	10.2	3.29	0.379	3.64	9.60
Methylchloroform	14.5	48.1	3.32	0.055	54.3	987
p-Dichlorobenzene	2.55	4.43	1.74	0.079	1.06	13.4
Perchloroethylene	3.10	8.85	2.85	0.858	4.34	5.06
Styrene	0.880	2.50	2.84	0.105	0.763	7.27
Trichloroethylene	1.12	3.49	3.12	0.083	1.78	21.5
o-Xylene	3.55	9.71	2.74	0.058	3.70	63.8

Table 2. Estimated variance components and mean values for nine VOCs from 24-h samples collected in five US cities.

Chemical	$\hat{\sigma}^2_{\alpha}$ ($_{\alpha}\hat{R}_{0.95}$)	$\hat{\sigma}^2_{B}$ ($_{B}\hat{R}_{0.95}$)	$\hat{\sigma}^2_{W}$ ($_{W}\hat{R}_{0.95}$)	$\hat{\mu}_y$
Benzene	0.209 (6.00)	0.527 (17.2)	0.707 (27.0)	2.61
Chloroform	0.567 (19.1)	0.000 (1.00)	1.26 (81.9)	0.761
Ethylbenzene	0.178 (5.22)	0.203 (5.85)	0.927 (43.6)	2.04
Methylchloroform	0.072 (2.85)	0.893 (40.6)	1.17 (68.8)	3.22
p-Dichlorobenzene	0.000 (1.00)	1.06 (56.5)	2.00 (256)	1.73
Perchloroethylene	0.063 (2.66)	0.067 (2.76)	1.37 (98.1)	2.10
Styrene	0.100 (3.45)	0.070 (2.81)	1.08 (58.9)	0.736
Trichloroethylene	0.292 (8.32)	0.064 (2.69)	2.20 (337)	0.806
o-Xylene	0.068 (2.77)	0.000 (1.00)	1.07 (57.6)	2.02

Parameters estimated using logged air concentrations; see Appendix A for sample sizes. $\hat{\sigma}^2_{\alpha}$, $\hat{\sigma}^2_{B}$ and $\hat{\sigma}^2_{W}$ are the variance components estimated under Model (1) (across cities, between persons and within persons, respectively); $_{\alpha}\hat{R}_{0.95}$, $_{B}\hat{R}_{0.95}$ and $_{W}\hat{R}_{0.95}$ are the corresponding estimated fold ranges; $\hat{\mu}_y$ represents the estimated mean (logged) air concentration ($\mu g/m^3$).

Table 3. Personal 24-h air concentrations ($\mu g/m^3$) of nine VOCs predicted under Model (2) in Los Angeles, CA and Bayonne and Elizabeth, NJ.

Location	Chemical	Winter	Spring	Summer	Fall
CA	Benzene	14.5	7.10	8.80	NM
CA	Chloroform	0.74	0.27	0.62	NM
CA	Ethylbenzene	8.29	5.24	4.51	NM
CA	Methylchloroform	27.8	18.6	11.5	NM
CA	*p*-Dichlorobenzene	3.17	NM	1.31	NM
CA	Perchloroethylene	8.35	3.51	2.98	NM
CA	Styrene	2.45	1.05	1.37	NM
CA	Trichloroethylene	1.55	0.79	0.52	NM
CA	*o*-Xylene	10.5	3.78	6.09	NM
NJ	Benzene	16.6	NM	8.49	14.9
NJ	Chloroform	2.05	NM	1.50	3.36
NJ	Ethylbenzene	9.20	NM	5.13	9.80
NJ	Methylchloroform	18.2	NM	14.3	22.2
NJ	*p*-Dichlorobenzene	7.04	NM	2.75	4.67
NJ	Perchloroethylene	11.0	NM	8.77	9.74
NJ	Styrene	2.04	NM	0.99	2.44
NJ	Trichloroethylene	0.68	NM	0.90	3.34
NJ	*o*-Xylene	0.86	NM	1.42	7.77

Parameters estimated using logged air concentrations; see Appendix A for sample sizes. NM — not measured.

exposures being greater than those in the summer for all compounds. As shown in Table 4, the estimated variances (from Proc MIXED) representing seasonal effects ($\hat{\sigma}^2_S$) contributed 3.4–21.0% of $\hat{\sigma}^2_{W(3)}$ in CA samples (median = 12.4%) and 0–6.8% of $\hat{\sigma}^2_{W(3)}$ in NJ samples (median = 0%). Following a similar analysis of $\hat{\sigma}^2_B$ (data not shown), the effect of season was more pronounced in NJ samples, contributing 0–20.1% of the Model (3) estimate

(median = 5.6%) than in CA samples, where season only reduced the Model (3) estimate of $\hat{\sigma}^2_B$ for trichloroethylene (7.5% reduction).

Table 4 also summarizes the median CVs, reported by TEAM investigators for assays of the nine VOCs. These CVs correspond to contributions of assay error ($\hat{\sigma}^2_A$) ranging from 1.3 to 5.2% of $\hat{\sigma}^2_{W(3)}$ in NJ (median = 2.3%) and from 1.0 to 7.8% of $\hat{\sigma}^2_{W(3)}$ in CA samples (median = 2.0%).

Comparisons Between Environmental and Occupational Exposures

We compared between- and within-person variance components, estimated separately for each of three cities (Bayonne, Elizabeth, and Los Angeles) under Model (3) using 12-h personal day and night measurements (Proc MIXED), with those compiled from occupational studies by Kromhout et al. (1993) for workers categorized by job and factory. These comparisons are illustrated in Figure 2, which shows cumulative distributions of the fold ranges corresponding to $\hat{\sigma}^2_B$ (top) and $\hat{\sigma}^2_W$ (bottom). Regarding the *between-person cumulative distributions*, fold ranges were similar for environmental and occupational exposures (median: 3.3- to 4.0-fold range between persons), although day exposures were noticeably more variable than night and occupational exposures in half the data sets. For the *within-person cumulative distributions*, fold ranges were much greater for environmental exposures (median: day = 136-fold range within persons, night = 80-fold range) than for occupational exposures (median = 15-fold range), and day exposures were more variable than night exposures in 80% of the cases.

Sample Size Calculations

We used Eq. (5) to calculate sample sizes needed to detect a positive coefficient ($\theta_1 > 0$) in Eq. (4), with exposure to benzene and perchloroethylene, using $\alpha = 0.05$ and ($1-\beta$) = 0.80. For both exposures, values of σ^2_α, σ^2_B, and σ^2_W were set equal to the estimates under Model (1) (Table 2); and, the number n of measurements/subject was between 2 and 4, since it would generally be impractical to collect larger numbers of repeat measurements/subject. Various combinations of H, k, and n were evaluated to arrive at the minimum total number of measurements (Hkn) for values of $\theta^2_1/\sigma^2_e = 0.25$, 1.0, and 4.0. As shown in Table 5, the number of personal measurements required for benzene would lie between 36 and 208, while those for perchloroethylene would lie between 312 and 2912. The disparity in sample sizes basically reflects the greater within-subject variability of perchloroethylene data compared to benzene data (Table 2).

Discussion

The first TEAM studies represented a landmark in the assessment of environmental exposures to chemical substances (Wallace et al., 1988, 1991; Wallace, 1987, 2001). These investigations ushered in personal monitoring to quantify environmental exposures accurately (personal sampling had

Table 4. Contributions of seasonal effects and assay errors to the estimated within-person variance component ($\hat{\sigma}^2_W$).

Location	Chemical	$\hat{\sigma}^2_{W(3)}$	$\hat{\sigma}^2_{W(2)}$	Est. seasonal component $\hat{\sigma}^2_S$	$\hat{\sigma}^2_S/\hat{\sigma}^2_{W(3)}$ (%)	Assay CV	Est. assay component $\hat{\sigma}^2_A$	$\hat{\sigma}^2_A/\hat{\sigma}^2_{W(3)}$ (%)
NJ	Benzene	2.32	2.30	0.022	1.0	0.36	0.122	5.2
NJ	Chloroform	1.85	1.77	0.085	4.6	0.20	0.039	2.1
NJ	Ethylbenzene	1.08	1.09	0.000	0.0	0.20	0.039	3.6
NJ	Methylchloroform	3.09	3.16	0.000	0.0	0.27	0.070	2.3
NJ	*p*-Dichlorobenzene	2.66	2.48	0.180	6.8	0.23	0.052	1.9
NJ	Perchloroethylene	1.97	1.97	0.000	0.0	0.21	0.043	2.2
NJ	Styrene	0.978	1.00	0.000	0.0	0.18	0.032	3.3
NJ	Trichloroethylene	1.45	1.48	0.000	0.0	0.14	0.019	1.3
NJ	*o*-Xylene	1.10	1.07	0.025	2.3	0.19	0.035	3.2
CA	Benzene	0.669	0.541	0.129	19.2	0.130	0.017	2.5
CA	Chloroform	2.09	1.87	0.223	10.7	0.286	0.079	3.8
CA	Ethylbenzene	0.954	0.883	0.071	7.4	0.122	0.015	1.5
CA	Methylchloroform	1.62	1.50	0.126	7.7	0.137	0.019	1.1
CA	*p*-Dichlorobenzene	1.81	1.43	0.380	21.0	0.277	0.074	4.1
CA	Perchloroethylene	1.55	1.31	0.238	15.4	0.126	0.016	1.0
CA	Styrene	1.25	1.10	0.156	12.4	0.320	0.097	7.8
CA	Trichloroethylene	3.55	3.43	0.122	3.4	0.274	0.072	2.0
CA	*o*-Xylene	1.61	1.39	0.218	13.6	0.124	0.015	1.0

$\hat{\sigma}^2_{W(3)}$ is the variance component estimated without a fixed effect for season (Model (3)). $\hat{\sigma}^2_{W(2)}$ is the variance component estimated with a fixed effect for season (Model (2)). $\hat{\sigma}^2_S$ is the estimated variance component representing seasonal effects [$\hat{\sigma}^2_{W(3)} - \hat{\sigma}^2_{W(2)}$]. CV is the coefficient of variation estimated from duplicate measurements of the chemical ($N = 134$ for NJ and $N = 34$ for LA). $\hat{\sigma}^2_A$ is the estimated variance component representing assay error.

been used in occupational settings since the 1960s (Sherwood and Greenhalgh, 1960)), and they employed stratified random sampling to allow valid inferences to be drawn about the levels of exposure in the US population. The striking conclusion of the TEAM studies was that personal VOC measurements were poorly correlated with and, indeed, much larger in magnitude than outdoor air measurements (as shown in Figure 1 and Table 1).

Given the success of TEAM investigators in collecting personal measurements from members of the general public, it is surprising that so few subsequent studies of environmental VOCs have relied upon personal sampling. Rather, most investigators gathered 'microenvironmental' samples at fixed indoor and outdoor locations, in the apparent hope that such data would be highly correlated with personal VOC levels. However, in light of the extreme variability in VOC levels observed across indoor and outdoor environments, we expect only modest correlations to exist between personal and microenvironmental samples, save for members of the public who rarely leave their residences. For example, using data from an extensive study of VOC levels in many microenvironments[4] by Kim et al. (2001), the ratio of high mean air concentration/low mean air concentration across microenvironments ranged from 22 to 234 with a median value of 68 (for 13 VOCs).

People are dynamic, moving from home to vehicle to workplace, visiting shops, bars, restaurants, and a host of other locations where sources and levels of VOCs vary tremendously. This notion is supported by our finding that daytime exposures to VOCs were much more variable than night time exposures across TEAM subjects (Figure 2). Given such diversity of locations and sources of exposure, the comprehensive investigation of microenvironments for a given VOC appears to be an open-ended and probably futile exercise. Thus, we encourage investigators to forsake microenvironmental measurements of VOCs in favor of personal measurements, a sentiment recently voiced by Kromhout and van Tongeren (2003). Since passive monitoring devices allow subjects to measure their own VOC exposures with a minimum of professional involvement (Tielemans et al., 1999; Egeghy et al., 2000, 2002; Liljelind et al., 2000, 2001), this goal is realistic.

Sources of VOCs

The TEAM studies identified a few associations between particular sources and VOC levels in Los Angeles (gleaned from questionnaires and location/activity diaries), notably

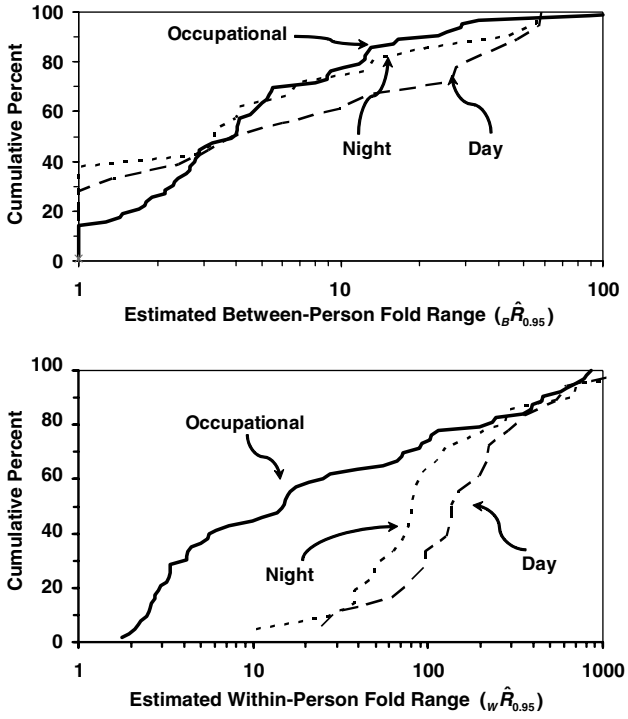

Figure 2. Cumulative distributions of fold ranges of personal VOC exposures between and within persons estimated under Model (3). Top: Estimated fold range $({}_B\hat{R}_{0.95} = e^{(3.92\hat{\sigma}_B)})$ containing 95% of mean exposures received by the population. Bottom: Estimated fold range $({}_W\hat{R}_{0.95} = e^{(3.92\hat{\sigma}_W)})$ containing 95% of the 24-h air concentrations experienced by a given person. "Day" and "Night" refer to 12-h environmental exposures of 9 VOCs from measurements reported by the TEAM study during daytime and nighttime, respectively. "Occupational" refers to shift-long personal measurements to 12 VOCs reported by Kromhout et al. (1993).

those between smoking and aromatic hydrocarbons (benzene, ethylbenzene, styrene, and the xylenes) (Wallace et al., 1988). Based upon the studies reviewed in Wallace (2001) plus more recent studies (Egeghy et al., 2000; Hoffmann et al., 2000; Kinney et al., 2002), we suspect the following sources of the nine VOCs we investigated: benzene — active and passive smoking and gasoline (e.g., attached garages, automobile refueling, automobile exhaust); chloroform — washing clothes and bathing/showering; *p*-dichlorobenzene — moth crystals and deodorizers; ethylbenzene, styrene and *o*-xylene — driving and solvent use; methylchloroform and perchloroethylene — dry cleaned clothing & solvent use; and trichloroethylene — solvent use. (Note that solvent use can involve both occupational sources and residential use of commercial products.) The above-identified sources of VOC exposure further emphasize the relatively minor contributions of outdoor air to personal air levels.

Ranges of Exposure Across Cities

The median personal exposure levels for VOCs varied between three- and five-fold across the cities investigated (Table 1). Upon closer inspection of these median (city)

4. The VOCs routinely measured above detection limits were 1,3-butadiene, benzene, toluene, ethylbenzene, *p*-, *m*-, and *o*-xylene, 1,3,5-trimethylbenzene, styrene, *p*-isopropyltoluene, 1,2,4-trimethylbenzene, *p*-dichlorobenzene, and naphthalene. Sampled locations included homes, offices, restaurants, pubs, department stores, cinemas, perfume shops, libraries, laboratories, train stations, trafficked roads, cars, trains, and buses.

Table 5. Sample sizes required to detect $\theta_1>0$ for Model (4) with exposure to benzene or perchloroethylene and data grouped by city.

Chemical	σ_α^2	σ_B^2	σ_W^2	θ_1^2/σ_e^2	Cities (H)	Subjects per city (k)	Meas. per subj. (n)	Total meas. (Hkn)
Benzene	0.209	0.527	0.707	0.25	52	2	2	208
Benzene	0.209	0.527	0.707	1.0	18	2	2	72
Benzene	0.209	0.527	0.707	4.0	9	2	2	36
Perchloroethylene	0.063	0.067	1.369	0.25	56	13	4	2912
Perchloroethylene	0.063	0.067	1.369	0.25	52	14	4	2912
Perchloroethylene	0.063	0.067	1.369	1.0	40	7	3	840
Perchloroethylene	0.063	0.067	1.369	1.0	35	8	3	840
Perchloroethylene	0.063	0.067	1.369	4.0	78	2	2	312

Sample size calculations based on Eq. (5) with $\alpha=0.05$ and $(1-\beta)=0.80$. σ_α^2, σ_B^2, σ_W^2, representing across-city, between-person, and within-person variance components, respectively, were set equal to their estimated values in Table 2; the ratio θ_1^2/σ_e^2 was arbitrarily set to values of 0.25, 1.0, and 4.0.

exposure levels, no pattern could be discerned with regard to geographic differences or degrees of urbanization. For example, despite its small size and rural location, Devils Lake, ND registered the highest exposure level for methylchloroform, while Los Angeles, CA (the largest city) did not register the highest exposure level to any VOC.

Variance Components Related to Cities, Subjects, and Errors

While TEAM investigators explored associations between personal exposures to particular VOCs and some sources, they did not partition exposure variability into components that might be used to optimize exposure assessment strategies. Assuming that the five cities in our study represent a random sample of all possible US cities, we applied Model (1) to estimate components of variation associated with cities, persons (within cities), and random variation (within persons). The results, summarized in Table 2, are intriguing because the estimated within-person variance component represented, on average, about three-fourths of the total observed variability for these compounds. This indicates that fluctuations in air levels experienced by a typical person over time were greater than the combined variation arising from cities and individuals. Nonetheless, collective differences in subject-specific factors were sufficient to offer reasonable contrast for epidemiological investigations, except for perchloroethylene, styrene, and *o*-xylene. This is illustrated in Table 2 by the magnitude of the estimated across-city and between-person fold ranges, at least one of which was large for six of the nine VOCs.

Contributions of Seasonal Effects and Assay Errors

Given the large contribution of the estimated within-person variance component to the overall variation in personal exposure to VOCs, we investigated two effects that should contribute to σ_W^2, namely, the season of the year and random assay errors. Seasonal effects were readily apparent in the application of Model (2) to personal measurements of all chemicals in Los Angeles and New Jersey (Table 3). Kinney et al. (2002) noted that VOC exposures tended to be higher in the winter than in the summer in New York City, possibly because of reduced air-exchange rates in homes and work places during the colder months. This explanation is not entirely consistent with our results, because the contribution of seasonal variation to $\hat{\sigma}_W^2$ was greater in Los Angeles (median = 12.4%), where the winters are mild, than in New Jersey (median = 0%), where the winters are cold. In any case, seasonal effects contributed only marginally to the within-subject variance component in our data. Likewise, assay errors were small, contributing only 1.0–7.8% of $\hat{\sigma}_W^2$. The combined contributions of seasonal effects and assay errors to $\hat{\sigma}_W^2$ ranged between 1.3 and 31.7% with a median value of only 7.7%. Thus, we conclude that seasonal effects and assay errors contributed little to within-subject variability in the TEAM studies.

Comparisons Between Environmental and Occupational Exposures

Since random- and mixed-effects models have been widely used to estimate between- and within-person variance components from occupational exposures to VOCs (defined by job and factory), we wished to compare the cumulative distributions of $\hat{\sigma}_B^2$ and $\hat{\sigma}_W^2$ (reported in occupational groups by Kromhout et al. (1993)) with those observed here from environmental groups (defined by city). As shown at the top of Figure 2, the cumulative distributions representing between-person variability were quite similar for occupational and environmental exposures to VOCs. We conclude from this that the variation in exposure to VOCs between residents in a given city is comparable to that observed between workers in a given factory and job. The cumulative distributions suggest that VOC exposure typically varies about three- to four-fold between persons in a given city, but varies more than 20-fold in about a quarter of the cases. These systematic differences in exposure to a particular VOC across city residents can be identified and used for epidemiological purposes in much the same manner that tasks or work locations are exploited to assign groups in occupational epidemiology (e.g., see Burstyn et al., 2002,

2003a, b). The key is to develop databases of personal measurements that include information about covariates during the period of monitoring. Although this notion was pioneered by the TEAM studies for environmental VOCs, most recent applications have involved occupational exposures (e.g., see Kromhout et al., 1994; Kromhout and Heederik, 1995). Nonetheless, Hoffmann et al. (2000) reported models that explained 39% of the variation in environmental exposure levels for benzene and 60% of the variation in exposure levels for ethylbenzene and the xylenes.

The major difference between occupational and environmental VOC exposures relates to the extreme variability of environmental VOCs within persons (bottom of Figure 2). Indeed, a typical environmental study would record exposure varying within subjects by about 100-fold from one survey to the next (day = 136-fold, night = 80-fold), while a typical occupational study would find only about 15-fold variation. One reason for the greater within-subject variation of TEAM subjects is probably the fact that some of these persons were exposed both in and out of the workplace (Wallace, 1991). However, occupational exposures would have occurred primarily during daytime; and since both day and night exposures reflected much greater within-person variation than occupational exposures, the contributions of workplace sources to our analyses were probably minor. However, regardless of its origins, within-person variability can lead to nonignorable bias effects (from exposure measurement errors) in studies of exposure-related health effects (e.g., see Navidi et al., 1994; Navidi and Lurmann, 1995; Spiegelman et al., 1997; Armstrong, 1998; Zeger et al., 2000); thus, larger sample sizes would be needed to assess accurately exposure–response relationships in environmental studies as compared to occupational studies. This issue will be considered next.

Sample Size Requirements

To illustrate one epidemiological application of our work, we predicted sample sizes needed to detect a true positive association between a continuous health outcome and a VOC exposure, after grouping by city. The enabling relationship (Eq. (5)) predicts total sample sizes (*Hkn*) involving *H* cities, *k* subjects per city, and *n* personal measurements per subject. As representative VOCs, we chose benzene and perchloroethylene, two carcinogenic substances that represent extremes of likely contrast for epidemiological studies, based upon the estimated variance components from the TEAM data. In particular, the proportions of total variation explained by $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_B^2$ were much larger for benzene than for perchloroethylene (Table 2). Since large across-city and between-person variance components help to reduce biasing measurement error effects under Model (4), we anticipated that values of *Hkn* predicted under Eq. (5) would be much smaller for benzene than for perchloroethylene. This was, in fact, the case, with values of *Hkn* for benzene being about 1/10th those for perchloroethylene at a given value of θ_1^2/σ_e^2 (Table 5). From the estimated variance components given in Table 2, it appears that the sample size requirements for benzene and perchloroethylene should bracket, in general, the likely range required for VOCs.

Sample sizes predicted for benzene were surprisingly modest, with a predicted range of 36–208 personal measurements for values of θ_1^2/σ_e^2 between 0.25 and 4. Such numbers should be achievable in most investigations. The corresponding predictions for perchloroethylene were much larger (312–2912 personal measurements) and could well be beyond the reach of many investigators unless θ_1^2/σ_e^2 were large.

The relationship shown in Eq. (5) assumes a grouped analysis with city as the grouping variable. Other types of grouping could be preferred in some cases. For example, since environmental exposure to benzene is heavily influenced by cigarette smoking, one could envision grouping by smoking status and by categories of exposure to environmental tobacco smoke. It would also be possible to perform similar calculations assuming an individual-based study where the continuous health outcome would be regressed directly upon the individual exposure level, using a sample of subjects randomly selected from the population. Tielemans et al. (1998) provide relationships that can be used, along the lines shown in Appendix B, to predict sample sizes for individual-based studies.

Limitations of our Study

Since our analyses employed secondary data, they are not without limitations. In particular, we only had access to about half of the TEAM data described in published reports (cited in the Introduction). Also, our random effects models assume that days, subjects, and cities represent random samples from the US population and, since the TEAM studies employed stratified probability sampling, we cannot be sure to what extent randomness assumptions were satisfied, particularly across the cities represented in the public database. The public database also did not include information about subjects' occupations, activities, smoking, and mobility that might have explained some sources of variability. Finally, the number of subjects with repeated measurements was relatively small (typically about 100 for a given VOC) and repeated measurements were not available from two of the five cities represented (Devils Lake and Greensboro).

In conclusion, we used data from early TEAM studies to show that random- and fixed-effects models can be applied to environmental VOC exposures in much the same manner as they are to occupational VOC exposures. The key to such applications is the collection of repeated measurements from representative persons grouped, as desired, by location (e.g., city) or lifestyle factors. We found sufficient variability in levels of most VOCs across cities and between subjects to provide reasonable contrast for epidemiology studies. Yet, the estimated within-subject variance component was large for environmental VOC data, indeed, much larger than for

occupational VOC data. Since within-subject variability leads to biasing measurement error effects when assessing exposure–response relationships, we caution epidemiologists to be wary of investigating the effects of environmental VOCs without having preliminary data with which to partition exposure variability into appropriate variance components. Using estimated across-city, between-person, and within-person variance components from the TEAM data, we calculated the numbers of personal measurements needed to investigate exposure-response relationships for environmental exposures to benzene and perchloroethylene. Under simple but realistic assumptions about the true regression model (relating exposure level to health outcome), the calculations suggest that sample sizes needed for valid and precise epidemiology studies would be readily achievable for benzene, but not for perchloroethylene.

Acknowledgments

This work was supported by the National Institute for Environmental Health Sciences through Grants P42ES05948 and P30ES10126. The authors appreciate the assistance of Dr. Peter Egeghy in downloading data from EPA's online database.

References

Armstrong B.G. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med* 1998: 55(10): 651–656.

Brunekreef B., Noy D., and Clausing P. Variability of exposure measurements in environmental epidemiology. *Am J Epidemiol* 1987: 125(5): 892–898.

Burstyn I., Boffetta P., Burr G.A., Cenni A., Knecht U., Sciarra G., et al. Validity of empirical models of exposure in asphalt paving. *Occup Environ Med* 2002: 59(9): 620–624.

Burstyn I., Boffetta P., Kauppinen T., Heikkila P., Svane O., Partanen T., et al. Performance of different exposure assessment approaches in a study of bitumen fume exposure and lung cancer mortality. *Am J Ind Med* 2003a: 43(1): 40–48.

Burstyn I., Boffetta P., Kauppinen T., Heikkila P., Svane O., Partanen T., et al. Estimating exposures in the asphalt industry for an international epidemiological cohort study of cancer risk. *Am J Ind Med* 2003b: 43(1): 3–17.

Egeghy P.P., Nylander-French L., Gwin K.K., Hertz-Picciotto I., Rappaport S.M. Self-collected breath sampling for monitoring low-level benzene exposures among automobile mechanics. *Ann Occup Hyg* 2002: 46(5): 489–500.

Egeghy P.P., Tornero-Velez R., and Rappaport SM. Environmental and biological monitoring of benzene during self-service automobile refueling. *Environ Health Perspect* 2000: 108(12): 1195–1202.

Hoffmann K., Krause C., Seifert B., and Ullrich D. The German Environmental Survey 1990/92 (GerES II): sources of personal exposure to volatile organic compounds. *J Expos Anal Environ Epidemiol* 2000: 10(2): 115–125.

Janssen N.A., Hoek G., Harssema H., and Brunekreef B. Childhood exposure to PM10: relation between personal, classroom, and outdoor concentrations. *Occup Environ Med* 1997: 54(12): 888–894.

Janssen N.A., Hoek G., Harssema H., and Brunekreef B. Personal sampling of airborne particles: method performance and data quality. *J Expos Anal Environ Epidemiol* 1998: 8(1): 37–49.

Janssen N.A., Hoek G., Harssema H., and Brunekreef B. Personal exposure to fine particles in children correlates closely with ambient fine particles. *Arch Environ Health* 1999: 54(2): 95–101.

Kim Y.M., Harrad S., and Harrison R.M. Concentrations and sources of VOCs in urban domestic and public microenvironments. *Environ Sci Technol* 2001: 35(6): 997–1004.

Kinney P.L., Chillrud S.N., Ramstrom S., Ross J., and Spengler J.D. Exposures to multiple air toxics in New York City. *Environ Health Perspect* 2002: 110 (Suppl. 4): 539–546.

Kromhout H., and Heederik D. Occupational epidemiology in the rubber industry: implications of exposure variability. *Am J Ind Med* 1995: 27(2): 171–185.

Kromhout H., Swuste P., and Boleij J.S. Empirical modelling of chemical exposure in the rubber-manufacturing industry. *Ann Occup Hyg* 1994: 38(1): 3–22.

Kromhout H., Symanski E., and Rappaport S.M. A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. *Ann Occup Hyg* 1993: 37(3): 253–270.

Kromhout H., and van Tongeren M. How important is personal exposure assessment in the epidemiology of air pollutants? *Occup Environ Med* 2003: 60(2): 143–144.

Liljelind I.E., Rappaport S.M., Levin J.O., Stromback A.E., Sunesson A.L., and Jarvholm B.G. Comparison of self-assessment and expert assessment of occupational exposure to chemicals. *Scand J Work Environ Health* 2001: 27(5): 311–317.

Liljelind I.E., Stromback A.E., Jarvholm B., Levin J.O., Strangert B.L., and Sunesson A.-L.K. Self-assessment of exposure: a pilot study of assessment of exposure to benzene in tank truck drivers. *Appl Occup Environ Hyg* 2000: 15: 195–202.

Navidi W., and Lurmann F. Measurement error in air pollution exposure assessment. *J Expos Anal Environ Epidemiol* 1995: 5(2): 111–124.

Navidi W., Thomas D., Stram D., and Peters J. Design and analysis of multilevel analytic studies with applications to a study of air pollution. *Environ Health Perspect* 1994: 102 (Suppl. 8): 25–32.

Peretz C., Goldberg P., Kahan E., Grady S., and Goren A. The variability of exposure over time: a prospective longitudinal study. *Ann Occup Hyg* 1997: 41(4): 485–500.

Peretz C., Goren A., Smid T., and Kromhout H. Application of mixed-effects models for exposure assessment. *Ann Occup Hyg* 2002: 46(1): 69–77.

Rappaport SM. Assessment of long-term exposures to toxic substances in air. *Ann Occup Hyg* 1991: 35(1): 61–121.

Rappaport S.M., Waidyanatha S., Qu Q., Shore R., Jin X., Cohen B., et al. Albumin adducts of benzene oxide and 1,4-benzoquinone as measures of human benzene metabolism. *Cancer Res* 2002: 62(5): 1330–1337.

Rappaport S.M., Weaver M., Taylor D., Kupper L., and Susi P. Application of mixed models to assess exposures monitored by construction workers during hot processes. *Ann Occup Hyg* 1999: 43(7): 457–469.

Rijnders E., Janssen N.A., van Vliet P.H., and Brunekreef B. Personal and outdoor nitrogen dioxide concentrations in relation to degree of urbanization and traffic density. *Environ Health Perspect* 2001: 109 (Suppl. 3): 411–417.

Sherwood R.J., and Greenhalgh D.M.S. A personal air sampler. *Ann Occup Hyg* 1960: 2: 127–132.

Spengler J., Schwab M., Ryan P.B., Colome S., Wilson A.L., Billick I., et al. Personal exposure to nitrogen dioxide in the Los Angeles Basin. *J Air Waste Manage Assoc* 1994: 44(1): 39–47.

Spiegelman D., McDermott A., and Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* 1997: 65 (Suppl. 4): 1179S–1186S.

Symanski E., Chan W., and Chang C.C. Mixed-effects models for the evaluation of long-term trends in exposure levels with an example from the nickel industry. *Ann Occup Hyg* 2001: 45(1): 71–81.

Symanski E., Kupper L.L., Kromhout H., and Rappaport S.M. An investigation of systematic changes in occupational exposure. *Am Ind Hyg Assoc J* 1996: 57(8): 724–735.

Tielemans E., Heederik D., Burdorf A., Vermeulen R., Veulemans H., Kromhout H., et al. Assessment of occupational exposures in a general population: comparison of different methods. *Occup Environ Med* 1999: 56(3): 145–151.

Tielemans E., Kupper L.L., Kromhout H., Heederik D., and Houba R. Individual-based and group-based occupational exposure assessment: some equations to evaluate different strategies. *Ann Occup Hyg* 1998: 42(2): 115–119.

van Tongeren M., Gardiner K., Calvert I., Kromhout H., and Harrington J.M. Efficiency of different grouping schemes for dust exposure in the European carbon black respiratory morbidity study. *Occup Environ Med* 1997: 54(10): 714–719.

Wallace L., Nelson W., Ziegenfus R., Pellizzari E., Michael L., Whitmore R., et al. The Los Angeles TEAM Study: personal exposures, indoor-outdoor air concentrations, and breath concentrations of 25 volatile organic compounds. *J Expos Anal Environ Epidemiol* 1991: 1(2): 157–192.

Wallace L.A. *The Total Exposure Assessment Methodology (TEAM) Study: Summary and Analysis: Volume I*. Office of Research and Development, US Environmental Protection Agency, Washington, DC, 1987.

Wallace L.A. Personal exposure to 25 volatile organic compounds. EPA's 1987 team study in Los Angeles, California. *Toxicol Ind Health* 1991: 7(5–6): 203–208.

Wallace L.A. Human exposure to volatile organic pollutants: implications for indoor air studies. *Annu Rev Energy Environ* 2001: 26: 269–301.

Wallace L.A., Pellizzari E.D., Hartwell T.D., Whitmore R.W., Zelon H., Perritt R., et al. The California TEAM study: breath concentrations and personal exposures to 26 volatile compounds in air and drinking water of 188 residents of Los Angeles, Antioch, and Pittsburgh, CA. *Atmos Environ* 1988: 22(10): 2141–2163.

Zeger S.L., Thomas D., Dominici F., Samet J.M., Schwartz J., Dockery D., et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000: 108(5): 419–426.

Appendix A.

Samples of data obtained from the TEAM database, providing personal and outdoor air concentrations, day and night, for nine VOCs are summarized in Table A1.

Table A1. Samples of data obtained from the TEAM database, providing personal and outdoor air concentrations, day and night, for nine VOCs.

VOC	City	No. pers. meas.	No. pers. subjects	No. outdoor meas.	No. outdoor subjects	Meas. per subj.	Time frame
Benzene	Bayonne	137	120	38	36	1–2	9/81–2/83
Benzene	Elizabeth	191	167	60	58	1–2	9/81–2/83
Benzene	Greensboro	24	24	6	6	1	5/82
Benzene	Los Angeles	171	110	167	70	1–3	2/84–7/87
Benzene	*All*	*523*	*421*	*271*	*170*	*1–3*	*9/81–7/87*
Chloroform	Bayonne	138	121	38	36	1–2	9/81–2/83
Chloroform	Elizabeth	191	167	60	58	1–2	9/81–2/83
Chloroform	Greensboro	24	24	6	6	1	5/82
Chloroform	Devils Lake	23	23	5	5	1	10/82
Chloroform	Los Angeles	177	108	177	74	1–3	2/84–7/87
Chloroform	*All*	*553*	*443*	*286*	*179*	*1–3*	*9/81–7/87*
Ethylbenzene	Bayonne	138	121	36	34	1–2	9/81–2/83
Ethylbenzene	Elizabeth	190	166	60	58	1–2	9/81–2/83
Ethylbenzene	Greensboro	24	24	6	6	1	5/82
Ethylbenzene	Devils Lake	23	23	5	5	1	10/82
Ethylbenzene	Los Angeles	180	111	179	74	1–3	2/84–7/87
Ethylbenzene	*All*	*555*	*445*	*286*	*177*	*1–3*	*9/81–7/87*
Methylchloroform	Bayonne	139	122	38	36	1–2	9/81–2/83
Methylchloroform	Elizabeth	191	167	60	58	1–2	9/81–2/83
Methylchloroform	Greensboro	24	24	6	6	1	5/82
Methylchloroform	Devils Lake	23	23	5	5	1	10/82
Methylchloroform	Los Angeles	181	111	180	74	1–3	2/84–7/87
Methylchloroform	*All*	*558*	*447*	*289*	*179*	*1–3*	*9/81–7/87*

Table A1. (*continued*)

VOC	City	No. pers. meas.	No. pers. subjects	No. outdoor meas.	No. outdoor subjects	Meas. per subj.	Time frame
p-Dichlorobenzene	Bayonne	139	122	35	34	1–2	9/81–2/83
p-Dichlorobenzene	Elizabeth	189	165	60	58	1–2	9/81–2/83
p-Dichlorobenzene	Greensboro	24	24	6	6	1	5/82
p-Dichlorobenzene	Devils Lake	23	23	5	5	1	10/82
p-Dichlorobenzene	Los Angeles	23	22	126	50	1–3	2/84–7/87
p-Dichlorobenzene	*All*	*398*	*356*	*232*	*153*	*1–3*	*9/81–7/87*
Perchloroethylene	Bayonne	139	122	38	36	1–2	9/81–2/83
Perchloroethylene	Elizabeth	191	167	60	58	1–2	9/81–2/83
Perchloroethylene	Greensboro	24	24	6	6	1	5/82
Perchloroethylene	Devils Lake	23	23	5	5	1	10/82
Perchloroethylene	Los Angeles	179	111	162	72	1–3	2/84–7/87
Perchloroethylene	*All*	*556*	*447*	*271*	*177*	*1–3*	*9/81–7/87*
Styrene	Bayonne	139	122	38	36	1–2	9/81–2/83
Styrene	Elizabeth	190	166	60	58	1–2	9/81–2/83
Styrene	Greensboro	24	24	6	6	1	5/82
Styrene	Los Angeles	179	111	177	74	1–3	2/84–7/87
Styrene	*All*	*532*	*423*	*281*	*174*	*1–3*	*9/81–7/87*
Trichloroethylene	Bayonne	139	122	37	36	1–2	9/81–2/83
Trichloroethylene	Elizabeth	191	167	60	58	1–2	9/81–2/83
Trichloroethylene	Greensboro	24	24	6	6	1	5/82
Trichloroethylene	Devils Lake	23	23	5	5	1	10/82
Trichloroethylene	Los Angeles	176	108	174	74	1–3	2/84–7/87
Trichloroethylene	*All*	*553*	*444*	*282*	*179*	*1–3*	*9/81–7/87*
o-Xylene	Bayonne	139	122	38	36	1–2	9/81–2/83
o-Xylene	Elizabeth	190	166	60	58	1–2	9/81–2/83
o-Xylene	Greensboro	24	24	6	6	1	5/82
o-Xylene	Devils Lake	23	23	5	5	1	10/82
o-Xylene	Los Angeles	180	111	179	74	1–3	2/84–7/87
o-Xylene	*All*	*553*	*444*	*282*	*179*	*1–3*	*9/81–7/87*

Appendix B. Derivation of a relationship to predict sample sizes

Assume exposure Model (1) with $k_h = k$ subjects in the hth city ($h = 1, 2, \ldots, H > 3$) and $n_{hi} = n$ measurements/subject), and assume that each person is assigned his/her city's observed mean (log-scale) exposure

$$\bar{Y}_h = \frac{1}{k}\sum_{i=1}^{k} \bar{Y}_{hi}$$

(since μ_{yhi} is unobservable in Model (4)), where

$$\bar{Y}_{hi} = \frac{1}{n}\sum_{j=1}^{n} Y_{hij}$$

is the observed (log-scale) mean exposure for the ith person in the hth city. Then the unweighted least-squares estimator of β_1 in Model (4), using a group-based analysis, is

$$\hat{\theta}_1^* = \frac{\sum_{h=1}^{H}(\bar{Y}_h - \bar{Y})\bar{R}_h}{\sum_{h=1}^{H}(\bar{Y}_h - \bar{Y})^2}$$

where

$$\bar{R}_h = \frac{1}{k}\sum_{i=1}^{k} R_{hi}$$

and

$$\bar{Y} = \frac{1}{Hk}\sum_{h=1}^{H}\sum_{i=1}^{k} \bar{Y}_{hi}$$

Using relationships given by Tielemans et al. (1998), we can show that the expected value of $\hat{\theta}_1^*$ is $E(\hat{\theta}_1^*) = c\beta_1$, where

$$c = \frac{\sigma_\alpha^2 + \sigma_B^2/k}{\sigma_\alpha^2 + (\sigma_B^2/k) + (\sigma_W^2/kn)}$$

and where σ_α^2, σ_B^2 , and σ_W^2 are defined under Model (1). Also the standard error of $\hat{\theta}_1^*$ has the form:

$$SE(\hat{\theta}_1^*) = \left[\frac{\frac{\sigma_e^2}{k}\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k} + \frac{\sigma_W^2}{kn}\right) + \beta_1^2\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k}\right)\frac{\sigma_W^2}{kn}}{(H-3)\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k} + \frac{\sigma_W^2}{kn}\right)^2}\right]^{1/2}$$

Using the above relationships for $E(\hat{\theta}_1^*)$ and $SE(\hat{\theta}_1^*)$ and assuming the variance components σ_e^2, σ_α^2, σ_B^2, and σ_W^2 are known (a standard assumption for such sample size derivations), a $100(1-\alpha)\%$ large-sample approximate confidence interval for θ_1 is

$$\frac{\hat{\theta}_1^*}{c} \pm z_{1-\alpha/2}\frac{SE(\hat{\theta}_1^*)}{c}$$

Assuming $\theta_1 > 0$ (indicating an adverse effect of exposure on the health outcome), we seek choices for $H(>3)$, k, and n (≥ 2) satisfying the following inequality:

$$\Pr\left\{\frac{\hat{\theta}_1^*}{c} - z_{1-\alpha/2}\frac{SE(\hat{\theta}_1^*)}{c} > 0 | \theta_1 > 0\right\} \geq (1-\beta)$$

where $(1-\beta)$ is the smallest probability we would accept for the event that the lower limit of the confidence interval exceeds zero. For large samples, we equivalently have

$$\Pr\left\{Z > z_{1-\alpha/2} - \frac{c\theta_1}{SE(\hat{\theta}_1^*)} > 0 | \theta_1 > 0\right\} \geq (1-\beta)$$

where Z is approximately a standard normal deviate. Hence, we require

$$z_{1-\alpha/2} - \frac{c\theta_1}{SE(\hat{\theta}_1^*)} \leq z_{1-\beta}$$

leading to

$$\frac{c^2\theta_1^2}{Var(\hat{\theta}_1^*)} \geq (z_{1-\alpha/2} + z_{1-\beta})^2$$

After substituting for c and $SE(\hat{\beta}_1^*)$, we obtain

$$\frac{\left(\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k}\right)\Big/\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k} + \frac{\sigma_W^2}{kn}\right)\right)^2\theta_1^2}{\left[\frac{\frac{\sigma_e^2}{k}\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k} + \frac{\sigma_W^2}{kn}\right) + \theta_1^2\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k}\right)\frac{\sigma_W^2}{kn}}{(H-3)\left(\sigma_\alpha^2 + \frac{\sigma_B^2}{k} + \frac{\sigma_W^2}{kn}\right)^2}\right]} \geq \left(z_{1-\alpha/2} + z_{1-\beta}\right)^2$$

After some manipulation, this leads to Eq. (5) in the text.

Appendix C.

Statistics for VOC concentrations ($\mu g/m^3$) from 24-h personal measurements in five US cities (1981–1987) are summarized in Table C1.

Table C1 Summary statistics for VOC concentrations ($\mu g/m^3$) from 24-h personal measurements in five US cities (1981–1987).

VOC	City	No. meas.	Median	Min.	Max.	Q1[a]	Q3[a]
Benzene	Bayonne	137	14.5	0.118	225	9.25	24.5
Benzene	Elizabeth	191	22.4	0.025	172	11.0	42.5
Benzene	Greensboro	24	7.31	0.073	30.5	0.934	15.0
Benzene	Los Angeles	171	13.0	1.16	51.5	7.87	19.0
Chloroform	Bayonne	138	2.63	0.275	26.0	1.50	5.09
Chloroform	Elizabeth	191	4.30	0.085	41.0	2.22	9.68
Chloroform	Greensboro	24	1.81	0.153	6.5	0.835	3.03
Chloroform	Devils Lake	23	0.440	0.128	25.1	0.150	1.20
Chloroform	Los Angeles	177	1.14	0.027	46.7	0.640	2.10
Ethylbenzene	Bayonne	138	6.30	0.884	212	3.75	10.8
Ethylbenzene	Elizabeth	190	10.2	0.456	752	5.90	18.9
Ethylbenzene	Greensboro	24	3.45	0.890	16.6	1.87	6.98
Ethylbenzene	Devils Lake	23	3.10	0.120	132	0.959	6.45
Ethylbenzene	Los Angeles	180	7.44	0.079	73.4	4.28	12.0
Methylchloroform	Bayonne	139	14.5	0.575	2260	7.44	29.5
Methylchloroform	Elizabeth	191	22.8	1.02	1.65E04	10.8	69.0
Methylchloroform	Greensboro	24	48.1	0.415	155	12.1	68.1
Methylchloroform	Devils Lake	23	43.0	0.245	610	8.15	182.5
Methylchloroform	Los Angeles	181	26.5	1.75	2020	13.1	55.5
p-Dichlorobenzene	Bayonne	139	4.43	0.140	1145	1.65	20.4
p-Dichlorobenzene	Elizabeth	189	3.80	0.266	1008	1.76	13.8
p-Dichlorobenzene	Greensboro	24	2.94	0.719	68.5	1.79	9.85

Table C1 (*continued*)

VOC	City	No. meas.	Median	Min.	Max.	$Q1^a$	$Q3^a$
p-Dichlorobenzene	Devils Lake	23	2.55	0.362	275	1.15	8.00
p-Dichlorobenzene	Los Angeles	23	3.79	0.587	216	1.37	48.9
Perchloroethylene	Bayonne	139	7.80	0.988	183	4.40	19.5
Perchloroethylene	Elizabeth	191	8.85	0.390	6010	4.30	20.5
Perchloroethylene	Greensboro	24	3.10	0.795	33.5	2.03	7.55
Perchloroethylene	Devils Lake	23	5.15	0.344	802	2.24	13.8
Perchloroethylene	Los Angeles	179	7.10	0.026	288	3.65	12.1
Styrene	Bayonne	139	1.70	0.075	24.0	0.980	2.97
Styrene	Elizabeth	190	2.50	0.284	3251	1.54	4.15
Styrene	Greensboro	24	0.880	0.192	12.6	0.51	1.63
Styrene	Los Angeles	179	2.40	0.023	114	1.46	4.20
Trichloroethylene	Bayonne	139	2.22	0.260	203	1.24	3.78
Trichloroethylene	Elizabeth	191	3.49	0.113	562	1.50	8.50
Trichloroethylene	Greensboro	24	1.12	0.263	38.7	0.608	2.61
Trichloroethylene	Devils Lake	23	0.493	0.153	25.6	0.175	1.86
Trichloroethylene	Los Angeles	176	1.61	0.022	175	0.383	6.69
o-Xylene	Bayonne	139	5.70	0.610	202	3.71	9.94
o-Xylene	Elizabeth	190	7.93	0.625	417	5.07	14.0
o-Xylene	Greensboro	24	5.18	1.06	16.3	3.21	8.98
o-Xylene	Devils Lake	23	3.55	0.088	249	2.15	8.40
o-Xylene	Los Angeles	180	9.71	0.026	63.4	5.65	14.8

[a]$Q1$ = 25th percentile concentration, $Q3$ = 75th percentile concentration.

Appendix D.

Statistics for VOC concentrations ($\mu g/m^3$) from 24-h outdoor measurements in five US cities (1981–1987) are summarized in Table D1.

Table D1. Summary statistics for VOC concentrations ($\mu g/m^3$) from 24-h outdoor measurements in five US cities (1981–1987).

VOC	City	*n*	Median	Min.	Max.	$Q1^a$	$Q3^a$
Benzene	Bayonne	38	8.64	0.130	26.0	4.50	12.3
Benzene	Elizabeth	60	5.53	0.038	34.5	2.58	11.3
Benzene	Greensboro	6	14.4	0.011	63.5	0.435	21.5
Benzene	Los Angeles	167	4.45	0.750	29.0	2.40	9.20
Chloroform	Bayonne	38	0.829	0.048	6.15	0.375	1.50
Chloroform	Elizabeth	60	0.697	0.038	12.2	0.146	1.97
Chloroform	Greensboro	6	0.336	0.036	0.664	0.115	0.573
Chloroform	Devils Lake	5	0.055	0.048	0.828	0.050	0.058
Chloroform	Los Angeles	177	0.223	0.024	13.1	0.075	0.593
Ethylbenzene	Bayonne	36	3.64	0.260	11.0	1.82	5.05
Ethylbenzene	Elizabeth	60	3.48	0.190	13.5	1.91	4.95
Ethylbenzene	Greensboro	6	0.865	0.273	1.70	0.415	1.66
Ethylbenzene	Devils Lake	5	0.379	0.024	1.80	0.029	0.609
Ethylbenzene	Los Angeles	179	2.05	0.250	19.0	1.17	4.35
Methylchloroform	Bayonne	38	5.61	0.260	30.8	2.70	7.70
Methylchloroform	Elizabeth	60	3.72	0.066	255	1.32	6.73
Methylchloroform	Greensboro	6	54.3	13.0	256	43.0	158
Methylchloroform	Devils Lake	5	0.055	0.048	4.78	0.050	0.058
Methylchloroform	Los Angeles	180	6.33	0.990	112	3.26	13.7

Table D1. (*continued*)

VOC	City	*n*	Median	Min.	Max.	$Q1^a$	$Q3^a$
p-Dichlorobenzene	Bayonne	35	1.06	0.185	6.00	0.550	1.55
p-Dichlorobenzene	Elizabeth	60	0.616	0.110	12.2	0.363	1.05
p-Dichlorobenzene	Greensboro	6	0.516	0.149	1.90	0.335	0.703
p-Dichlorobenzene	Devils Lake	5	0.079	0.055	2.25	0.075	0.083
p-Dichlorobenzene	Los Angeles	126	0.500	0.050	10.7	0.313	1.40
Perchloroethylene	Bayonne	38	4.34	0.637	30.3	2.55	10.4
Perchloroethylene	Elizabeth	60	2.88	0.098	48.0	1.28	6.88
Perchloroethylene	Greensboro	6	0.858	0.378	1.44	0.603	1.23
Perchloroethylene	Devils Lake	5	1.00	0.228	3.28	0.666	1.09
Perchloroethylene	Los Angeles	162	2.02	0.313	28.5	0.999	4.92
Styrene	Bayonne	38	0.763	0.175	2.11	0.581	1.03
Styrene	Elizabeth	60	0.652	0.055	3.65	0.316	0.920
Styrene	Greensboro	6	0.105	0.064	0.199	0.071	0.157
Styrene	Los Angeles	177	0.575	0.040	9.50	0.250	1.78
Trichloroethylene	Bayonne	37	1.78	0.070	7.20	0.830	3.84
Trichloroethylene	Elizabeth	60	1.21	0.068	9.75	0.591	2.10
Trichloroethylene	Greensboro	6	0.562	0.139	1.49	0.145	0.875
Trichloroethylene	Devils Lake	5	0.083	0.073	1.23	0.075	0.085
Trichloroethylene	Los Angeles	174	0.140	0.023	2.45	0.060	0.265
o-Xylene	Bayonne	38	3.70	0.813	9.80	2.25	5.20
o-Xylene	Elizabeth	60	3.54	0.075	12.0	2.23	4.58
o-Xylene	Greensboro	6	1.39	0.378	2.15	0.863	1.91
o-Xylene	Devils Lake	5	0.058	0.050	0.966	0.055	0.483
o-Xylene	Los Angeles	179	3.43	0.313	27.1	1.95	6.78

[a]$Q1 = 25$th percentile concentration, $Q3 = 75$th percentile concentration.

Appendix E.

Spearman's correlation coefficients for 24-h personal measurements of nine VOCs in five US cities are summarized in Table E1. See Appendix A for nominal sample sizes.

Table E1. Spearman's correlation coefficients for 24-h personal measurements of nine VOCs in five US cities.

City	Chemical	Benzene	Chloroform	Et-Benz.	Me-Chlor.	p-DCB	Perc.	Styrene	TCE
Bayonne	Chloroform	0.651							
Bayonne	Et-Benz.	0.690	0.408						
Bayonne	Me-Chlor.	0.449	0.431	0.420					
Bayonne	*p*-DCB	0.213	0.263	0.260	0.315				
Bayonne	Perc.	0.592	0.459	0.519	0.499	0.249			
Bayonne	Styrene	0.509	0.369	0.658	0.349	0.348	0.494		
Bayonne	TCE	0.489	0.426	0.465	0.627	0.227	0.584	0.415	
Bayonne	*o*-Xylene	0.653	0.368	0.945	0.359	0.265	0.537	0.630	0.404
Elizabeth	Chloroform	0.426							
Elizabeth	Et-Benz.	0.513	0.318						
Elizabeth	Me-Chlor.	0.309	0.321	0.347					
Elizabeth	*p*-DCB	0.076	0.034	0.179	0.117				
Elizabeth	Perc.	0.379	0.264	0.401	0.456	0.116			
Elizabeth	Styrene	0.430	0.349	0.627	0.257	0.174	0.281		
Elizabeth	TCE	0.430	0.277	0.372	0.491	0.098	0.551	0.299	
Elizabeth	*o*-Xylene	0.460	0.167	0.920	0.322	0.226	0.394	0.579	0.379

Table E1. (*continued*)

City	Chemical	Benzene	Chloroform	Et-Benz.	Me-Chlor.	p-DCB	Perc.	Styrene	TCE
Greensboro	Chloroform	0.089							
Greensboro	Et-Benz.	0.198	−0.249						
Greensboro	Me-Chlor.	0.161	−0.080	0.270					
Greensboro	*p*-DCB	0.003	0.484	−0.198	−0.154				
Greensboro	Perc.	0.031	0.081	0.026	0.303	−0.129			
Greensboro	Styrene	0.393	0.083	0.340	0.052	0.039	0.040		
Greensboro	TCE	0.030	0.442	0.021	0.369	0.174	0.132	0.090	
Greensboro	*o*-Xylene	0.306	−0.252	0.955	0.162	−0.154	−0.008	0.409	−0.015
Devils Lake	Et-Benz.	.	0.250						
Devils Lake	Me-Chlor.	.	0.358	0.228					
Devils Lake	p-DCB	.	−0.075	0.373	0.525				
Devils Lake	Perc.	.	0.454	0.311	0.642	0.596			
Devils Lake	TCE	.	0.743	0.088	0.632	0.267	0.745	.	
Devils Lake	*o*-Xylene	.	0.297	0.576	0.504	0.608	0.561	.	0.370
Los Angeles	Chloroform	0.292							
Los Angeles	Et-Benz.	0.824	0.186						
Los Angeles	Me-Chlor.	0.516	0.283	0.537					
Los Angeles	*p*-DCB	0.401	−0.057	0.307	−0.025				
Los Angeles	Perc.	0.573	0.290	0.520	0.679	−0.217			
Los Angeles	Styrene	0.615	0.329	0.666	0.549	0.272	0.585		
Los Angeles	TCE	0.274	0.223	0.354	0.577	0.033	0.498	0.417	
Los Angeles	*o*-Xylene	0.839	0.154	0.928	0.491	0.345	0.532	0.646	0.305
All	Chloroform	0.489							
All	Et-Benz.	0.663	0.336						
All	Me-Chlor.	0.382	0.229	0.394					
All	*p*-DCB	0.127	0.135	0.201	0.174				
All	Perc.	0.489	0.343	0.455	0.495	0.166			
All	Styrene	0.509	0.287	0.653	0.388	0.242	0.432		
All	TCE	0.393	0.374	0.401	0.496	0.153	0.519	0.362	
All	*o*-Xylene	0.577	0.174	0.897	0.391	0.243	0.435	0.627	0.332

Et-benz. — Ethylbenzene, Me-chlor. — methylchloroform; *p*-DCB — *p*-dichlorobenzene; Perc. — perchloroethylene, TCE — trichloroethylene.