Environmental Research 109 (2009) 193–199

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Examination of the relationships between environmental exposures to volatile organic compounds and biochemical liver tests: Application of canonical correlation analysis ☆

Jing Liu [a,b], Wanzer Drane [a], Xuefeng Liu [a], Tiejian Wu [a,c,*]

[a] Department of Biostatistics and Epidemiology, East Tennessee State University, Johnson City, TN 37614, USA
[b] Institute of Epidemiology and Biostatistics, Shandong University, Shandong Province, China
[c] Department of Family Medicine, East Tennessee State University, USA

ARTICLE INFO

Article history:
Received 23 July 2008
Received in revised form
28 October 2008
Accepted 10 November 2008
Available online 30 December 2008

Keywords:
Canonical correlation
Environmental exposure
Liver function
Volatile organic compound

ABSTRACT

This study was to explore the relationships between personal exposure to 10 volatile organic compounds (VOCs) and biochemical liver tests with the application of canonical correlation analysis. Data from a subsample of the 1999–2000 National Health and Nutrition Examination Survey were used. Serum albumin, total bilirubin (TB), alanine aminotransferase (ALT), aspartate aminotransferase (AST), lactate dehydrogenase (LDH), alkaline phosphatase (ALP), and γ-glutamyl transferase (GGT) served as the outcome variables. Personal exposures to benzene, chloroform, ethylbenzene, tetrachloroethene, toluene, trichloroethene, *o*-xylene, *m*-,*p*-xylene, 1,4-dichlorobenzene, and methyl *tert*-butyl ether (MTBE) were assessed through the use of passive exposure monitors worn by study participants. The first two canonical correlations were 0.3218 and 0.2575, suggesting a positive correlation mainly between the six VOCs (benzene, ethylbenzene, toluene, *o*-xylene, m-,p-xylene, and MTBE) and the three biochemical liver tests (albumin, ALP, and GGT) and a positive correlation mainly between the two VOCs (1,4-dichlorobenzene and tetrachloroethene) and the two biochemical liver tests (LDH and TB). Subsequent multiple linear regressions show that exposure to benzene, toluene, or MTBE was associated with serum albumin, while exposure to tetrachloroethene was associated with LDH and total bilirubin. In conclusion, exposure to certain VOCs as a group or individually may influence certain biochemical liver test results in the general population.

1. Introduction

Volatile organic compounds (VOCs) are contained in a wide variety of commercial, industrial, and residential products including fuel oils, gasoline, solvents, cleaners and degreasers, paints, inks, dyes, refrigerants, and pesticides. Because of their ubiquitous presence in the environment, accurate assessment of the risk to public health posed by VOCs requires both the quantification of these exposures on a population-wide basis and the evaluation of potential health effects associated with varying exposure levels.

The liver is the major site for processing chemicals and drugs which enter the blood stream. The liver helps by removing these chemicals from the blood stream and changing them into products that can be readily removed through the bile or urine. In this process, unstable toxic products are sometimes produced, which can attack and injure the liver. Many VOCs could cause toxic chemical injury to the liver through this type of mechanism, in addition to direct toxicity (Brautbar and Williams, 2002; Xiao and Levin, 2000).

Biochemical liver tests include tests that are routinely measured in all clinical laboratories (Cahill, 1999). Several tests including serum aspartate aminotransferase, γ-glutamyltransferase, and alkaline phosphatase can serve as sensitive indicators of liver injury (Giannini et al., 2005). A recent cohort study (Kim et al., 2004) found that there was a positive association between aminotransferase concentration, even within normal range, and mortality from liver disease, suggesting that moderately increased aminotransferase activity is significant in predicting liver disease.

In exploring the health effects of environmental exposures, observational epidemiologic studies often deal with data that include both a set of exposure variables and a set of outcome variables. Routine approaches such as multiple linear regressions to analyze such data are usually challenged as they are plagued by the potential issues including multicollinearity and multiple testing. Since canonical correlation analysis (CCA) assesses the

☆ *Funding*: National Institutes of Health (R03ES016368). *Human subjects*: Five hundred and sixty four human subjects involved. *Animal subjects*: Not applicable.
* Corresponding author at: Department of Biostatistics and Epidemiology, East Tennessee State University, Johnson City, TN 37614, USA. Fax: +1 423 439 6491.
E-mail address: wut@etsu.edu (T. Wu).

0013-9351/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.envres.2008.11.002

correlation of two canonical variates (latent variables), one representing a set of the exposure variables and the other a set of outcome variables, it is potentially a useful method to evaluate the health effects of environmental exposures. Although canonical correlation has often been applied to social sciences and bioinformatics (Stevens, 1986; Pugh and Hu, 1991; Steinfath et al., 2007), the method has been rarely used in environmental health assessment.

The National Health and Nutrition Examination Survey conducted during 1999 and 2000 (NHANES 1999–2000) included a sub-project which collected detailed information on personal exposures to 10 VOCs (CDCa). NHANES 1999–2000 also collected blood samples and performed a number of biochemical liver tests (CDCa). In this study, we applied CCA to the NHANES data and explored if there were associations between the environmental VOC exposures and the biochemical liver tests.

2. Materials and methods

2.1. Data source and study sample

Data from NHANES 1999–2000 were used for this study. Briefly, NHANES 1999–2000 used a stratified multistage probability design to obtain a representative sample of the civilian non-institutionalized US general population. A total of 12,160 persons were asked to take part in NHANES 1999–2000. Interview and physical examination data were collected on 9282 of the eligible participants, and the overall response rate was 76.3% (CDCb). Since the NHANES data sets are accessible to the public, the Institutional Review Board of East Tennessee State University granted the exemption of review and approved the study.

The VOC Project of personal exposures to air toxics was conducted among a representative subsample of the NHANES participants between the ages of 20 and 59 years (CDCc). The project was designed to characterize exposures to these air toxics and to determine predictors of exposure. Among 851 subjects eligible for the VOC Project during 1999 and 2000, about 75% had measurements of personal VOC exposure.

To limit potential confounding, we excluded the subjects who had liver conditions, heart disease, stroke, cancer, or diabetes. These conditions were confirmed if a subject reported in NHANES that a physician had ever told him/her having these conditions. Those tested serum positive to hepatitis C virus (HCV) in NHANES were also excluded from the study. Of the 851 subjects who participated in the VOC project, 115 who had one or more of the above conditions were excluded from the study. We then excluded 170 subjects, who did not have VOC assessments and the two subjects with very extreme values in ALT (1163 U/L) or AST (827 U/L), resulting in a sample size of 564. However, depending on the variables analyzed, the sample size varied slightly because some of the variables of interest had missing values. A comparison of the study sample with the NHANES 1999–2000 participants in several characteristics are shown in Table 1.

2.2. Study variables

2.2.1. Demographic and behavioral characteristics

Demographic information such as gender, age, ethnicity, family income, education attainment, and health risk behaviors such as cigarette smoking and alcohol consumption were collected in NHANES 1999–2000 through interviews or questionnaires.

2.2.2. Biochemical liver tests

In NHANES 1999–2000, venous blood specimens were collected, centrifuged, refrigerated at 4–8 °C, and then shipped weekly to a central laboratory where they were tested upon arrival (CDCd). Serum markers of liver function assessed in NHANES 1999–000 include: albumin, total bilirubin (TB), alanine aminotransferase (ALT), aspartate aminotransferase (AST), lactate dehydrogenase (LDH), alkaline phosphatase (ALP), and γ-glutamyl transferase (GGT). Laboratory procedures and quality control of the measurements are detailed by the NHANES Laboratory Procedure Manual (CDCd). In this study, the biochemical liver tests served as continuous outcome variables in the analyses.

2.2.3. Personal exposure to VOCs

Personal exposure to 10 VOCs including benzene, chloroform, ethylbenzene, tetrachloroethene, toluene, trichloroethene, *o*-xylene, *m*-,*p*-xylene, 1,4-dichlorobenzene, and methyl *tert*-butyl ether (MTBE) were assessed through the use of passive exposure monitors (or badges) worn by participants for a period of 48–72 h. The actual duration of exposure ranged from 43.3 to 76.0 h with a mean of 56.4 h. The number of cubic meters for each subject was derived from the duration the exposure monitor was used. The value for each analyte was expressed as the weight of the analyte per cubic meter. If a VOC test result was below the limit of detection, a value was imputed in NHANES 1999–2000 using the detection limit adjusted for the actual duration the badge was exposed divided by the square root of two. The procedures of sample collection and lab tests were detailed in the NHANES documentation (CDCc).

2.3. Statistical analysis

2.3.1. Descriptive analysis

Descriptive statistics including mean, standard deviation, and proportion were used to describe the characteristics of both the study sample and the NHANES participants. Descriptive statistics including mean, median, percentiles, minimum, and maximum values were used to show the distributions and numerical characteristics of the seven liver function test results and the exposure levels of the 10 VOCs.

2.3.2. Canonical correlation analysis (CCA)

CCA is an exploratory statistical method to assess correlations between two sets of variables (Stevens, 1986). One of the assumptions of CCA for the significance testing of canonical correlations is that the variables in both sets follow a multivariate normal distribution. Therefore, in this study, we transformed the VOC and liver function test variables to Blom normal scores from their ranks to assure that the multivariate normality is not violated. The transformation formula was $Z_i = \Phi^{-1}((r_i-3/8)/(n+1/4))$, where Φ^{-1} is the inverse cumulative normal (PROBIT) function, r_i is the rank of the ith observation, and n is the number of observations for the ranking variable (Blom, 1958; Tukey, 1962).

The fundamental principle behind CCA is the creation of a number of canonical variates, each consisting of a linear combination of one set of variables (X_i), which has the form:

$$U_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$$

and a linear combination of the other set of variables (Y_i), which has the form:

$$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \cdots + b_{iq}Y_q$$

The goal is to determine the coefficients, or canonical weights (a_{ij} and b_{ij}), that maximize the correlation between canonical variates U_i and V_i. The first canonical correlation, Corr(U_1, V_1), is the highest possible correlation between any linear combination of the variables in the exposure set and any linear combination of the variables in the outcome set. Further pairs of maximally correlated linear combinations are chosen in turn, and they are orthogonal to those already identified. The maximum number of canonical correlation is equal to the number of variables in the smaller set, which is seven in this study (the number of biochemical liver tests of interest). Significance test of a canonical correlation coefficient was performed using likelihood ratio test. Since this was an exploratory study of the health effects of VOC exposures, the significance level of the test was set at 0.10, instead of 0.05, to limit the chance of failing to detect an effect.

Structure correlation coefficients, also called canonical loadings, are used to interpret the importance of each original variable in the canonical variates. A structure correlation is the correlation of a canonical variate with the variable in its set. Variables that are highly correlated with a canonical variate should be considered more important when deriving a meaningful interpretation of the related canonical variate. This way of interpreting canonical variates is the same as the interpretation of factors in factor analysis (Shafto et al., 1997). As a rule of thumb, an absolute value of 0.3 or greater in canonical loading was used to select the variables that are thought to have a meaningful interpretation of the related canonical variate (Lambert and Durand, 1975; Thompson, 1984). We chose a cutoff value of 0.35 to select important loadings in this study.

2.3.3. Linear regression analysis

Restricted to the important VOC and liver function test variables derived from CCA, multiple linear regressions were used to examine the relationship between an individual VOC variable and an individual biochemical liver test, without and with adjustment for sex, age, BMI, ethnicity, educational level, family income, alcohol drinking, and cigarette smoking. These covariates are described in Table 1. Logarithm-transformations of ALP, GGT, LDH, and total bilirubin were performed to reduce the skewness of the variables. The level of significance was set at 0.10. All the analyses described above were performed using SAS statistical software (SAS Institute, 1989; Afifi et al., 2004).

3. Results

3.1. Characteristics of the study participants

The proportion of men in the sample was 44.33% and women 55.67%. Non-Hispanic whites accounted for 40.78%, Mexican

Table 1
Characteristics of the study subjects and the participants in NHANES 1999–2000.

Characteristic		Study sample		NHANES 1999–2000 (age 20–59)	
		n	Proportion (%) or mean (SD)	*n*	Proportion (%) or mean (SD)
Gender					
Male		250	44.33	1371	45.01
Female		314	55.67	1675	54.99
Age at exam in years		564	37.05 (10.92)	2834	38.42 (11.10)
BMI (kg/m^2)					
Male		249	27.54 (5.82)	1261	27.68 (5.61)
Female		314	28.63 (7.10)	1552	29.02 (7.27)
Ethnicity					
Mexican American		161	28.55	836	27.45
Non-Hispanic White		230	40.78	1264	41.50
Non-Hispanic Black		112	19.86	606	19.89
Other Hispanic		43	7.62	217	7.12
Other race–including multi-racial		18	3.19	123	4.04
Education					
Less than high school		173	30.67	983	32.28
High school diploma (including GED)		137	24.29	696	22.86
More than high school		254	45.04	1359	44.63
Refused		0	0	3	0.10
Don't know		0	0	4	0.13
Annual family income					
Over $20,000		373	66.49	1971	65.63
Under $20,000		161	28.70	884	29.44
Refused		15	2.67	81	2.7
Don't know		12	2.14	67	2.23
Alcohol drinking					
⩾12 drinks/ 1year	Yes	363	68.23	1847	69.91
	No	169	31.77	794	30.05
	Don't know	0	0	1	0.04
Smoking status					
Never		326	57.90	1683	55.36
Past		117	20.78	587	19.31
Current		120	21.31	770	25.33

American for 28.55%, and non-Hispanic blacks for 19.86%. About 68.23% of study participants reported drinking 12 or more times every year, and 21.31% were current smokers. Most characteristics of the study sample are comparable to the NHANES 1999–2000 participants (Table 1).

3.2. Biochemical liver test results

The median, 95th percentile, and percentage of abnormal values diagnosed according to the lab test standard for each of the seven lab tests are shown in Table 2. For example, the median and the 95th percentile were 45 and 50 g/L for albumin, 20 and 57 U/L for ALT, and 20 and 74 U/L for GGT. The percents of abnormal values for albumin, total bilirubin, ALT, AST, ALP, GGT, and LDH were 0.54%, 4.33%, 14.98%, 7.40%, 7.40%, 15.16%, and 0.90%, respectively.

3.3. Personal exposure to VOCs

Descriptive statistics of the personal exposure to each of the 10 VOCs are shown in Table 3. For example, the median and the 95th percentile were 2.8775 and 17.64 μg/m^3 for benzene, 1.1450 and 10.66 μg/m^3) for chloroform, and 0.6074 and 23.80 μg/m^3 for MTBE.

3.4. Canonical correlation between biochemical liver tests and individual exposures to VOCs

The first two canonical correlations were statistically significant ($F = 1.78$, $P < 0.0001$ and $F = 1.25$, $P = 0.10$, respectively), indicating that the two sets of variables were correlated. The first canonical correlation coefficient was 0.3248 and the second was 0.2587. The pooled sum of the squares of all the canonical correlation coefficients was 0.2370, which was contributed 46.00% and 27.97%, respectively, by the first and the second canonical correlations.

The canonical structures of the first two pairs of canonical variates were shown in Table 4. The first canonical variate of VOCs mainly represented the five aromatic hydrocarbons (benzene, ethylbenzene, toluene, *o*-xylene, and *m*,*p*-xylene) and MTBE and the first canonical variate of biochemical liver tests mainly represented albumin, ALP, and GGT. The second canonical variate of VOCs mainly represented 1,4-dichlorobenzene and

Table 2
Descriptive statistics of the biochemical liver tests.

Liver function test	*n*	Mean	Median	90th percentile	95th percentile	Abnormal[a]	
						n	%
Albumin (g/L)	554	44.39	45.00	49.00	50.00	3	0.54
Bilirubin,total (μmol/L)	554	9.21	8.60	13.70	17.10	24	4.33
ALT (U/L)	554	25.74	20.00	45.00	57.00	83	14.98
AST (U/L)	554	23.53	21.00	34.00	40.00	41	7.40
ALP (U/L)	554	81.35	78.00	112.00	129.00	41	7.40
GGT (U/L)	554	29.84	20.00	51.00	74.00	84	15.16
LDH (U/L)	554	147.22	144.00	182.00	195.00	5	0.90

[a] Albumin: <34 g/L; total bilirubin: >17.10 μmol/L; ALT: >41 U/L for males, > 31 U/L for females; AST: ⩾37 U/L for males, ⩾31 U/L for females; ALP: >117 U/L; GGT: >49 U/L for males, >32 U/L for females; LDH: >225 U/L for males, >214 U/L for females.

Table 3
Descriptive statistics of personal exposure to 10 VOCs.

VOCs (μg/m^3)	*n*	Detection limit[a]	Min.[b]	Mean	Median	90th percentile	95th percentile	Max.
Benzene	556	1.774	0.7294	5.33	2.88	11.32	17.64	119.47
Ethylbenzene	551	0.277	0.1381	7.71	2.33	11.55	23.08	837.13
Toluene	547	3.808	1.6551	34.59	16.07	54.70	92.30	1610.77
o-Xylene	555	0.434	0.1315	6.18	2.14	12.17	21.60	202.29
m,*p*-Xylene	555	0.480	0.2393	17.78	5.97	34.26	64.22	728.73
Chloroform	560	0.450	0.2075	2.69	1.15	6.17	10.66	41.13
Tetrachloroethene	552	0.257	0.1202	4.59	0.77	6.25	16.46	659.13
Trichloroethene	555	0.383	0.1222	4.08	0.32	1.20	7.38	327.28
1,4-Dichlorobenzene	555	0.882	0.3075	47.03	2.08	106.90	259.99	2235.57
MTBE	554	0.846	0.3647	6.09	0.61	13.83	23.80	181.68

[a] Detection limits were standardized for 48-h sample duration.
[b] Calculated as the detection limit adjusted for the duration of sample collection divided by the square root of two if below the detection limit.

Table 4
Canonical structures of the first two pairs of canonical variates.

First pair, $Rc_1 = 0.3248$ (Approx. $F = 1.78$, $P < 0.0001$)				Second pair, $Rc_2 = 0.2588$ (Approx. $F = 1.25$, $P = 0.10$)			
VOCs		Liver tests		VOCs		Liver tests	
Variable	Loading	Variable	Loading	Variable	Loading	Variable	Loading
Benzene	**0.7621**	**Albumin**	**0.7795**	**1,4-Dichlorobenzene**	**0.6866**	**LDH**	**0.6740**
***o*-Xylene**	**0.6776**	**ALP**	**0.3832**	**Tetrachloroethene**	**0.4596**	**TB**	**0.3991**
***m*,*p*-Xylene**	**0.6683**	**GGT**	**0.3827**	Trichloroethene	0.1775	ALP	0.3365
Ethylbenzene	**0.6553**	ALT	0.3242	*o*-Xylene	0.1747	ALT	−0.1718
MTBE	**0.5623**	AST	0.1734	Chloroform	−0.1285	Albumin	0.1710
Toluene	**0.4386**	LDH	−0.0817	Benzene	−0.1102	AST	0.1644
Chloroform	−0.2958	TB	−0.0442	*m*,*p*-Xylene	0.0899	GGT	0.0343
Tetrachloroethene	−0.0190			MTBE	0.0615		
1,4-Dichlorobenzene	−0.0101			Toluene	−0.0381		
Trichloroethene	0.0096			Ethylbenzene	0.0050		

Rc_1: the first canonical correlation; Rc_2: the second canonical correlation. VOCs and liver tests with loading >0.35 are given in bold.

tetrachloroethene, and the second canonical variate of biochemical liver tests mainly represented LDH and total bilirubin. These results helped to narrow down the relationship between the VOC exposure and liver function to fewer numbers of VOCs and liver function tests. This implies that exposure to a cluster of certain VOCs might be associated with certain biochemical liver tests as a group. That is, personal exposure to the aromatic hydrocarbons and MTBE as a group might affect the serum levels of albumin, ALP, and GGT, while personal exposure to 1,4-dichlorobenzene and tetrachloroethene might affect serum concentrations of LDH and total bilirubin.

3.5. Relationship between individual VOCs and biochemical liver tests

The results of individual relationships from the linear regressions for the important exposure and outcome variables selected from the above canonical analysis are shown in Tables 5 and 6. After adjustment for several covariates, personal exposure to benzene, toluene, and MTBE were significantly associated with serum albumin under the significance level of 0.10, while tetrachloroethene was significantly associated with LDH and total bilirubin.

Table 5
Multiple linear regressions for the outcome variables derived from first pair of canonical variates.

VOCs (100 μg/m^3)	Albumin(g/L)				Natural logarithm of alkaline phosphate (U/L)				Natural logarithm of γ-glutamyl-transpeptidase (U/L)			
	Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]	
	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P
Benzene	6.279 (2.006)	0.002	4.722 (2.372)	0.046	0.124 (0.171)	0.468	−0.207 (0.231)	0.370	0.387 (0.367)	0.293	-0.034 (0.462)	0.941
Ethylbenzene	0.789 (0.414)	0.058	0.869 (0.869)	0.317	0.039 (0.035)	0.264	−0.052 (0.084)	0.536	0.113 (0.076)	0.137	0.032 (0.170)	0.851
Toluene	0.366 (0.140)	0.009	0.195 (0.116)	0.093	0.004 (0.012)	0.762	−0.001 (0.011)	0.928	−0.013 (0.026)	0.626	−0.027 (0.023)	0.240
o-Xylene	3.600 (0.996)	0.001	0.793 (0.999)	0.427	0.071 (0.086)	0.409	−0.059 (0.097)	0.543	0.363 (0.184)	0.048	0.029 (0.196)	0.882
m,p-Xylene	1.165 (0.331)	0.001	0.307 (0.305)	0.314	0.017 (0.028)	0.556	−0.022 (0.030)	0.463	0.130 (0.061)	0.033	0.021 (0.060)	0.726
MTBE	1.752 (0.958)	0.068	2.095 (0.911)	0.021	0.036 (0.081)	0.936	−0.026 (0.089)	0.770	−0.042 (0.173)	0.220	−0.099 (0.178)	0.578

[a] Adjusted by gender, age, BMI, ethnicity, education, annual family income, drinking alcohol, and smoking. SE, standard error.

Table 6
Multiple linear regressions for the outcome variables derived from the second pair of canonical variates.

VOCs (100 μg/m^3)	Natural logarithm of lactate dehydrogenase (U/L)				Natural logarithm of total bilirubin (μmol/L)			
	Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]	
	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P
Tetrachloroethene	0.070 (0.026)	0.007	0.062 (0.025)	0.012	0.136 (0.064)	0.035	0.137 (0. 059)	0.020
1,4-Dichlorobenzene	0.002 (0.004)	0.613	−0.002 (0.004)	0.651	−0.007 (0.010)	0.523	−0.001 (0.010)	0.887

[a] Adjusted by gender, age, BMI, ethnicity, education, annual CPS family income, drinking alcohol and smoking. SE, standard error.

4. Discussion

In this study, we found that there was a significant correlation between the set of VOCs and the set of biochemical liver tests. That is, individual exposure to the five aromatic hydrocarbons and MTBE tended to be associated with serum levels of albumin, ALP, and GGT, while individual exposure to 1,4-dichlorobenzene and tetrachloroethene tended to be associated with serum concentrations of LDH and total bilirubin. The ability of CCA to analyze the relationship between the two sets of variables has several advantages. First, as it used the information from all the variables in both the exposure and outcome variable sets and maximized the estimation of the relationship between the two sets, CCA could be more efficient in assessing the effect of the VOCs on biochemical liver tests than the methods used previously such as multiple linear regressions. As a result, CCA could be more suitable for detecting a subtle effect of VOCs at relatively low exposure levels. Second, CCA started with simultaneously including both the exposure and outcome variables in the analysis; therefore, the issue of multiple testing was largely limited. Third, the latent variable approach employed in CCA also helped to avoid multicollinearity, an issue which might be troublesome in some analyses such as multiple linear regressions.

The relationships between the clusters of VOCs and the biochemical liver tests have important implications. Although the biochemical tests may indicate different biological responses to liver injury, the combination of different tests would capture more information in certain liver damage/response than one individual biochemical liver test. Furthermore, the cause of liver damage may lie in a number of factors rather than in isolated exposure (Dossing and Skinjob, 1985; Chen et al., 1997). In most settings, exposures occur to mixtures of solvents. In addition, because many VOCs might have a common mechanism of action, there can be additive effects. For instance, the VOCs we studied may share a common toxicity pathway that includes metabolism by cytochrome P450 2E1 (Brautbar and Williams, 2002; Dennison et al., 2004). Therefore, even when exposure to each agent individually is at a level that would not be expected to cause adverse effects, the combination of several VOCs may lead to inhibition of detoxifying enzymes or induction and increased activity of enzyme systems (Alessio, 1996). As a result, the combination effect could be significant particularly when exposure levels of individual VOCs are low, as shown in our study. Therefore, further investigation of possible joint effects of these select VOCs on liver function is warranted.

In addition to providing an assessment of the association between two sets of variables, the application of CCA helped narrowing down fewer exposure and outcome variables that might contribute to the relationship between the VOC exposure and liver function, based on the variable loadings to the canonical variates. CCA could thus serve as a "data mining tool" to identify the key exposure and outcome variables, so that the assessment of the relationships between an individual exposure and an outcome can be further preceded. In this study, following the canonical analysis, the linear regressions indicated that the level of exposure to benzene, toluene, or MTBE was associated with

serum albumin, and the level of exposure to tetrachloroethene was associated with serum LDH, suggesting that more specific effects on the biochemical liver tests due to the exposure to the individual VOCs (benzene, toluene, MTBE, or tetrachloroethene) might exist.

All VOCs we studied are listed by the US Environmental Protection Agency as hazardous air pollutants (US Environmental Protection Agency, 2008b). The liver is regarded to be one of the target organs for the critical effects of many VOCs (Caldwell et al., 1998; US Environmental Protection Agency, 2008a; California Environmental Protection Agency, 2002). For example, VOCs including benzene, ethylbenzene, MTBE, and tetrachloroethene have been found to cause hepatotoxicity in animal studies (Kim et al., 1990; Elliott and Strunin, 1993).

Biochemical liver tests have been used in human studies to assess the effects of occupational exposure to VOCs (Kaukiainen et al., 2004; Rees et al., 1993; Michailova et al., 1998; Chen et al., 1991). For example, a study of 29 solvent-exposed workers and 19 controls found that serum alanine aminotransferase and aspartate aminotransferase correlated positively with cumulative solvent exposure in the past 5 years, while total bilirubin correlated with current exposure (Kaukiainen et al., 2004). Increased γ-glutamyl transferase activity was found to be associated with severity of exposure to the mixture of solvents (Chen et al., 1991). However, to date there was no observational study that examined the effects of background VOC exposure on the liver in non-occupational settings according to our literature review.

Personal exposures to VOCs were assessed in NHANES 1999–2000 and employed in this analysis. It has been demonstrated repeatedly that personal exposures typically exceed outdoor air concentrations and that levels of human exposure to VOCs depend on people's locations, especially indoors (Payne-Sturges et al., 2004). Therefore, personal monitoring to assess VOC exposure has advantages to fixed-site ambient monitoring. However, personal exposure to VOCs often involves a mix of exposure to relatively constant levels and exposure to highly intermittent sources. Therefore, if the actual exposure duration was not able to characterize the exposure, uncertainty could be introduced into risk estimates. The reliability of time-averaging measures of exposure is an issue of concern (Lorenzana et al., 2005). In this study, the actual duration of exposure ranged from 43.3 to 76.0 h, which might capture an individual's routine activities and well represent the exposure to relatively constant sources. A recent study of the NHANES data shows significant associations between the levels of the VOCs in blood and air except for ethylbenzene and toluene in smoking subjects (Lin et al., 2008). Accurate exposure assessment is critical to a credible and scientifically sound assessment of risk (National Research Council, 1983; Sexton et al., 1995).

This study is the first to report a relationship between exposure to certain VOCs and the level of serum albumin in the general population. It suggests liver synthesis of albumin is increased in response to the VOC exposure. Albumin plays a role in VOC toxicokinetics by forming protein adducts to VOCs (Heinrich-Ramm et al., 2000; Yeowell-O'Connell, et al., 1998). However, certain biochemical tests are not highly specific for liver function which could limit the implications of our findings. Therefore, the relationships between exposure to VOCs and biochemical liver tests found in this study need be interpreted with caution. For example, albumin is produced only by liver cells, its concentration in the serum, however, is influenced by other factors such as dehydration, burns or renal disease (Cahill, 1999). The liver is usually responsible for the detoxification of bilirubin and its excretion into bile. However, bilirubin is not only increased in liver disease but other conditions that cause an increased breakdown of red blood cells (Cahill, 1999).

In summary, the study supports that personal exposures to certain VOCs as a group or individually may have an effect on liver function in the general population. The finding adds new insights on the potential health effects of VOC exposure in the general population and may thus have important implications for future research. The study also indicates that canonical correlation could be a useful method for environmental health risk assessment.

Acknowledgments

The authors thank Dr. Matthew P. Longnecker, Dr. Phillip R. Scheuerman, and Dr. Kenneth Z. Silver for their helpful comments on an earlier version of the manuscript. The study was supported by a grant from the National Institutes of Health (R03ES016368).

References

Afifi, A., Clark, V., May, S. (Eds.), 2004. Computer-Aided Multivariate Analysis, forth ed. Chapman & Hall/CRC, Boca Raton, FL, pp. 233–248.
Alessio, L., 1996. Multiple exposure to solvents in the workplace. Int. Arch. Occup. Environ. Health 69 (1), 1–4.
Blom, G. (Ed.), 1958. Statistical Estimates and Transformed Beta-Variables. Wiley, New York.
Brautbar, N., Williams, J., 2002. Industrial solvents and liver toxicity: risk assessment, risk factors and mechanisms. Int. J. Hyg. Environ. Health 205 (6), 479–491.
Cahill, M. (Ed.), 1999. Handbook of Diagnostic Tests, second ed. Springhouse Corporation, Springhouse, PA.
California Environmental Protection Agency, 2002. Toxicity Criteria Database. California Environmental Protection Agency, Sacramento, CA. Available: 〈http://www.oehha.ca.gov/〉 (accessed April 23, 2008).
Caldwell, J., Woodruff, T.J., Morello-Frosch, R., Axelrad, D.A., 1998. Application of health information to hazardous air pollutants modeled in EPA's cumulative exposure project. Toxicol. Ind. Health 14 (3), 429–454.
CDCa National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey 1999–2000. US Department of Health and Human Services, Center for Disease Control and Prevention, Hyattsville, MD, 1999–2000. Available: 〈http://www.cdc.gov/nchs/data/nhanes/survey_content_99_08.pdf〉 (accessed April 23, 2008).
CDCb National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey 1999–2000. US Department of Health and Human Services, Center for Disease Control and Prevention, Hyattsville, MD, 1999–2000. Available: 〈http://www.cdc.gov/nchs/data/nhanes/guidelines1.pdf〉 (accessed April 23, 2008).
CDCc National Center for Health Statistics. (NCHS). National Health and Nutrition Examination Survey 1999–2000. US Department of Health and Human Services, Center for Disease Control and Prevention. Hyattsville, MD, 1999–2000. Available: 〈http://www.cdc.gov/nchs/data/nhanes/frequency/lab21_doc.pdf〉 (accessed April 23, 2008).
CDCd National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey 1999–2000. US Department of Health and Human Services, Center for Disease Control and Prevention, Hyattsville, MD, 1999–2000. Available: 〈http://www.cdc.gov/nchs/data/nhanes/frequency/lab18_met_biochemistry_profile.pdf〉 (accessed April 23, 2008).
Chen, J.D., Wang, J.D., Jang, J.P., Chen, Y.Y., 1991. Exposure to mixture of organic solvents among paint workers and biochemical alterations of liver function. Br. J. Ind. Med. 48 (10), 696–701.
Chen, J.D., Wang, J.D., Tsai, S.Y., Chao, W.I., 1997. Effects of occupational and non-occupational factors on liver function tests in workers exposed to solvent mixtures. Arch. Environ. Health 52 (4), 270–274.
Dennison, J.E., Bigelow, P.L., Andersen, M.E., 2004. Occupational exposure limits in the context of solvent mixtures, consumption of ethanol, and target tissue dose. Toxicol. Ind. Health 20, 165–175.
Dossing, M., Skinjob, P., 1985. Occupational liver injury. Int. Arch. Occup. Environ. Health 56 (1), 1–21.
Elliott, R.H., Strunin, L., 1993. Hepatotoxicity of volatile anaesthetics. Br. J. Anaesth. 70 (3), 339–348.
Giannini, E.G., Testa, R., Savarino, V., 2005. Liver enzyme alteration: a guide for clinicians. Can. Med. Assoc. J. 172 (3), 367–379.
Heinrich-Ramm, R., Jakubowski, M., Heinzow, B., Christensen, J.M., Olsen, E., Hertel, O., 2000. Biological monitoring for exposure to volatile organic compounds. Pure Appl. Chem. 72, 385–436.
Kaukiainen, A., Vehmas, T., Rantala, K., Nurminen, M., Martikainen, R., Taskinen, H., 2004. Results of common laboratory tests in solvent-exposed workers. Int. Arch. Occup. Environ. Health 77 (1), 39–46.
Kim, H.C., Nam, C.M., Jee, S.H., Suh, I., 2004. Normal serum aminotransferase concentration and risk of mortality from liver diseases: prospective cohort study/commentary: liver function tests: defining what's normal. Br. Med. J. 328 (7446), 983–985.

Kim, H.J., Odend'hal, S., Bruckner, J.V., 1990. Effect of oral dosing vehicles on the acute hepatotoxicity of carbon tetrachloride in rats. Toxicol. Appl. Pharmacol. 102 (1), 34–49.
Lambert, Z.V., Durand, R.M., 1975. Some precautions in using canonical analysis. J. Mark. Res. 7, 468–475.
Lin, Y.S., Egeghy, P.P., Rappaport, S.M., 2008. Relationships between levels of volatile organic compounds in air and blood from the general population. J. Expo. Sci. Environ. Epidemiol. 18, 421–429.
Lorenzana, R.M., Troast, R., Klotzbach, J.M., Follansbee, M.H., Diamond, G.L., 2005. Issues related to time averaging of exposure in modeling risks associated with intermittent exposures to lead. Risk Anal. 25 (1), 169–178.
Michailova, A., Kuneva, T., Popov, T., 1998. A comparative assessment of liver function in workers in the petroleum industry. Int. Arch. Occup. Environ. Health 71, s46–s49.
National Research Councils (Ed.), 1983. Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, DC, pp. 17–49.
Payne-Sturges, D.C., Burke, T.A., Breysse, P., Diener-West, M., Buckley, T.J., 2004. Personal exposure meets risk assessment: a comparison of measured and modeled exposures and risks in an urban community. Environ. Health Perspect. 112 (5), 589–598.
Pugh, R.C., Hu, Y., 1991. Use and interpretation of canonical correlation analyses in Journal of Educational Research articles: 1978–1989. J. Edu. Res. 84 (3), 147–152.
Rees, D., Soderlund, N., Cronje, R., Song, E., Kielkowski, D., Myers, J., 1993. Solvent exposure, alcohol consumption and liver injury in workers manufacturing paint. Scand. J. Work Environ. Health 19 (4), 236–244.
SAS Institute (Eds.), 1989. SAS Procedures Guide, Version 6. third ed., SAS Publishing.
Sexton, K., Callahan, M.A., Bryan, E.F., Saint, C.G., Wood, W.P., 1995. Informed decisions about protecting and promoting public health: rationale for a national exposure assessment survey. J. Expo. Anal. Environ. Epidemiol. 5 (3), 233–256.
Shafto, M.G., Degani, A., Kirlik, A., 1997. Canonical correlation analysis of data on human–automation interaction. In: Human Factors and Ergonomics Society Annual Meeting Proceedings, Aerospace Systems, pp. 62–65.
Steinfath, M., Repsilber, D., Scholz, M., Walther, D., Selbig, J., 2007. Integrated data analysis for genome-wide research. EXS 97, 309–329.
Stevens, J. (Ed.), 1986. Applied Multivariate Statistics for the Social Sciences. Lawrence Erlbaum Associates Inc., New Jersey, pp. 373–397.
Thompson, B., 1984. Canonical correlation analysis: uses and interpretation. In: Quantitative Applications in the Social Sciences, vol. 47, no. 47, Sage Publications, Thousand Oaks, CA.
Tukey, J.W., 1962. The future of data analysis. Ann. Math. Stat. 33 (1), 1–67.
US Environmental Protection Agency, 2008a. Integrated Risk Information System (IRIS). US Environmental Protection Agency, 2000, Washington, DC. Available: 〈http://www.epa.gov/iris/index.html〉 (accessed April 23, 2008).
US Environmental Protection Agency, 2008b. The original list of hazardous air pollutants. Technology Transfer Network Air Toxics Web site. Available: 〈http://www.epa.gov/ttn/atw/188polls.html〉 (accessed October 24, 2008).
Xiao, J.Q., Levin, S.M., 2000. The diagnosis and management of solvent-related disorders. Am. J. Ind. Med. 37 (1), 44–61.
Yeowell-O'Connell, K., Rothman, N., Smith, M.T., Hyyes, R.B., Li, G., Waidyanatha, S., 1998. Hemoglobin and albumin adducts of benzene oxide among workers exposed to high levels of benzene. Carcinogenesis 19, 1565–1571.