Documents Responsive to RFP #6, 8, 9, 11

conoco

Interoffice Communication

To D. A. Kuhn

From W. D. Broddle

Date August 9, 1979

Subject Literature Search for Reproductive Effects of Benzene

I have reviewed all available experimental data on reproductive studies involving benzene and have found no additional information not contained in your memo to R. E. Lehmkuhl on July 13, 1979. I have also communicated with Dr. Cliff Conaway (Texaco) who is chairman of API's Toxicity Committee dealing with reproductive effects of benzene. Dr. Conaway stated that API's studies would not be completed until 1980, but that all previously reported reproductive effects on benzene, below 80 ppm, are equivocal.

Thus, I agree that 1 ppm TLV (5 ppm ceiling) should prevent unreasonable worker risk, but the potential toxicity of benzene merits close adherence to these atmospheric limits, especially when one desires a reasonable factor of safety below toxic experimental doses.

W. D. Broddle

jlk

CC:
R. C. Allred
T. R. Samsell

CONOCO

SAT
READ

Interoffice Communications

To R. E. Lehmkuhl

From D. A. Kuhn

Date July 13, 1979

Subject EFFECTS OF BENZENE HARMFUL TO FEMALES

As you requested I have reviewed several recent surveys[1-3] and the information in our files on benzene toxicity. There are no reports implicating benzene as especially toxic to women in human studies although there are several case reports describing the sudden onset of aplastic anemia when workers exposed to benzene become pregnant[4].

There are, however, several reports of reproductive effects from benzene inhalation in female experimental animals[5-8]. Analyzing these collectively, benzene is observed to have weak toxic effects on pregnant females and on fetuses at about 50 ppm in the experimental animals. In one study at 500 ppm, weak teratologic effects were also observed.

These findings should be put in perspective. Many toxicological effects in both humans and experimental animals, males and females, are found after exposure to benzene in the range 20 ppm and higher. Among them are acute affects such as vertigo, drowsiness, headache and nausea, as well as chromosomal breakage, leukopenia, pernicious and aplastic anemia and leukemia. There are also three articles dealing with effects on males.[9-11] While the information on males is sparse and inconclusive, limited data suggest a gonadal effect in male animals at 80 ppm.[10]

I conclude that there is only a small amount of information about the effects of benzene on a woman or her developing fetus. None indicate a special sensitivity. What we have tells us that the harmful effects occur within the same concentration range found for many of the toxic effects due to benzene exposure for both men and women. Consequently, our decision to set a Conoco maximum permissible exposure level at 1 ppm benzene, far below the levels where these effects have been noted, is the proper course to adequately protect men, women, and their potential offspring. Because these effects have been found at relatively low concentrations we should redouble our efforts to assure that the 1 ppm maximum permissible exposure level is being scrupulously maintained.

Dave

D. A. Kuhn

DAK:ajo

cc R A Darling, R T Ferrell, D L Norwood

REFERENCES

1. US EPA, Office of Research and Development, "Benzene Health Effects Assessment, Washington, D.C., (October 1977).

2. Department of Labor, Occupational Safety and Health Administration, "Occupational Exposure to Benzene, Fed. Reg. 42, 22516, (May 3, 1977).

3. Department of Labor Occupational Safety and Health Administration, "Occupational Exposure to Benzene, Fed. Reg. 43, 5918, (February 10, 1978)".

4. A. Hricko and M. Brunt, "Working for Your Life: A Womens Guide to Job Health Hazards", P. C-20, Labor Occupational Health Program and Public Citizens Health Research Group. (1976).

5. Hazleton Laboratories America, Inc., Vienna, Virginia, Project No. 145-552, Teratology Study in Rats, Final Report (September 25, 1975).

6. Dow Chemical U.S.A., Midland, Michigan, Report Code HET K-002661-(11), Embryotoxicity of Inhaled Benzene in Mice and Rabbits, Final Report (August 3, 1978).

7. Green, J. D., Leong, B. K., J. and Laskin, S., "Fetotoxicity of Inhaled Benzene in Rats," Toxicol. Appl. Pharmacol. (1977, In Press).

8. Litton Bionetics, Inc., Kensington, Maryland, Project No. 20698-3, "Teratology Study in Rats", Final Report (November, 1977).

9. Hett, J., "Benzol/Und Keimdrusen", Klinische Wochenschrift 17:1376 (1938).

10. Wolf, M. A., Rowe, V. K., McCollister, D. D., Hollingsworth, R. L., and Ogen, F., "Toxicology Studies of Certain Alkylated Benzenes and Benzene", Arch. Ind. Health 14:387-398 (1956).

11. Deichmann, W. B., MacDonald, W. E. and Bernal, E., "The Hemapoietic Tissue Toxicity of Benzene Vaport", Toxicol. Appl. Pharmacol. 5:201-224 (1962).

RECEIVED

NOV 0 9 1982

MEDICAL DIV.

All Plant Managers

R. E. Lehmkuhl

January 28, 1977

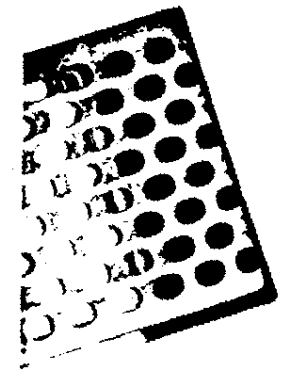
Standard for Employee Exposure to Benzene

Effective immediately, Conoco Chemicals standard for maximum permissible exposure to benzene will be an 8-hour time weighted average concentration of 1-ppm in air.

Feasible engineering controls will be implemented to reach this level in the plants. If engineering controls are not sufficient to reach the permissible exposure level they will be supplemented by work practice controls and respiratory protection. While engineering controls are being designed and constructed, work practice controls and respiratory protection sufficient to meet permissible level should be employed.

. Lehmkuhl

L. Lembke, M.D.
L. Whetstone, M.D.
Doyle, Legal
. Kuhn



DUPONT

HASKELL LABORATORY

Common Name: Benzene

Chemical Name: Benzene

CAS Registry No.: 71-43-2

Chemical Structure:

Physical and Chemical Properties:

Form:	Liquid
Molecular Weight:	78.12
Boiling Point:	80.1°C @ 760 mm Hg
Melting Point:	5.5°C
Density:	0.87865 (20/4°C)
Vapor Pressure:	100 mm Hg @ 26.1°C
	40 mm Hg @ 7.6°C
Conversion Factors:	1 mg/l = 312.5 ppm
	1 ppm = 3.2 mg/m^3

Recommended AEL:

10 ppm (8-hour TWA)

Chronic exposure to high concentrations of benzene is capable of causing an effect on the blood producing organs. Exposure to concentrations greater than 30 ppm has been reported to produce leukemia. Changes in the blood of exposed animals confirm this finding (1). These changes are the basis of the current 10 ppm TLV® (1) and OSHA PEL (2). OSHA has proposed a lowering of the PEL to 1 ppm. This proposed standard has been challenged in the courts as unreasonable (3). Recently, Exxon has reported to EPA that benzene may have an adverse effect on reproduction of laboratory animals (4). Exxon's conclusions were based primarily on a study performed by Litton-Bionetics and sponsored by the American Petroleum Institute. In this study pregnant rats were exposed to 0, 8.7 or 43.4 ppm of benzene on days 6-15 of gestation. Litton reported a significant effect on

Concern 1 ppm TWA (know why) OSHA 1969 10 TWA 10 ppm Ceiling 25 ppm 10 min peak 50 ppm

resorptions, resorptions/site and live fetuses/site at both levels (5). A review of this data by Du Pont biostatisticians indicates it to be equivocal.

Benzene is listed by Du Pont, ACGIH and OSHA as a carcinogen based in part on reports suggesting long-term exposure as a possible cause of leukemia in rubber industry workers (6). When the Du Pont classification was proposed it was recommended that benzene levels be maintained below the OSHA-recommended limit of 10 ppm. Efforts taken to comply with this recommendation should not be relaxed.

The Committee's recommendation was to maintain benzene levels as low as possible. However, we are of no unequivocal data which shows that an 8-hour TWA of 10 ppm is not adequate to protect worker health.

References

1. Documentation of TLV®s, 355-356 (1974-75).

2. Code of Federal Regulations, Title 29, Section 1910.1028.

3. Chemical Regulation Reporter, 1(53), 1955-1956 (1978); Ibid., 2(3), 73-74.

4. Section 8e (TSCA) Report, Exxon Corporation to EPA dated 9/28/78.

5. Unpublished Data, Litton-Bionetics Inc., Project No. 20698-3 (1977).

6. McMichael, A. J. et al., J. Occup. Med., 17, 234-239 (1975).

Richard C. Graham:md
December 12, 1979

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF BIOMEDICAL AND
ENVIRONMENTAL HEALTH SCIENCES

EARL WARREN HALL
BERKELEY, CALIFORNIA 94720

October 22, 1986

Dr. James Vail
American Petroleum Institute
1220 L Street, N.W.
Washington D.C. 20005

Dear Jim:

Enclosed is a manuscript just submitted to Applied Industrial Hygiene, the new ACGIH journal. It is a rewrite and condensation of the descriptive parts of the report and I don't believe it contains anything controversial. I'd appreciate any comments you might have in 6-8 weeks so I can incorporate them in whatever further rewrite we have to do.

The intent is to write a second paper dealing with compliance versus risk and I will send you a draft of that at an earlier stage in the process.

Hope all is well with you.

Sincerely,

Robert C. Spear, Ph.D.
Professor

RCS/p
Encl

SAL 000001376

Benzene exposure in the petroleum refining industry

ROBERT CLINTON SPEAR, STEVE SELVIN, JANE SCHULMAN and MARCIE FRANCIS

Northern California Occupational Health Center: School of Public Health: University of California: Berkeley CA 94720

Benzene exposure in the petroleum refining industry

Benzene exposure data, submitted by nine petroleum refining companies, were studied with the general objective of characterizing the distribution of exposures within work operations and job categories. The data were collected by company hygienists over the period of years between 1978 and 1984. All measurements were of personal exposures determined using charcoal tubes or organic vapor monitors. Of the 123 location and unit specific job groups studied most eight-hour TWA exposures were below 1 ppm. For some groups, however, the variability in exposure was such that 10% or more of exposures exceeded 1 ppm. For nineteen of the most highly exposed groups sufficient data was available to study the variability in exposure associated with individual workers versus the common work environment. Results indicate that there were some cases in which variability in exposure was mostly associated with the environment and others in which it was mostly associated with differences between workers, but it was most often approximately evenly split between the two sources. This analysis provides guidance in determining whether control strategies should be targetted at the work environment or at work requirements or practices of individual workers within the group. Some short-term exposure data were included for study. It was found that most fifteen-minute TWA exposures were less than 1 ppm , but again, there were groups with highly variable exposure with some measurements in excess of 5 ppm. These were often in jobs involving loading and unloading operations of barges or tanker trucks.

Introduction

In the latter part of 1984 we were asked by the American Petroleum Institute to conduct a study of benzene exposure data available from some of their member companies. These data had been collected by company hygienists over the years 1978 to 1984. The goals of the study were to:

1. Characterize the distribution of exposures within work operations and job categories.

2. Estimate the proportions of individuals exposed to benzene at or above levels ranging from 0.5 to 2.0 ppm.

3. Investigate the usefulness and suitability of specific statistical distributions as summary descriptors of exposure, e.g. normal, log-normal, etc.

4. Investigate intra-personal variation as a proportion of total variation in exposure levels within job categories.

5. Assess the implications of the results as they relate to air sampling strategies and the definition of compliance with OSHA standards.

These goals can be broadly classified into descriptive, analytic and policy-related aspects of the study. In this paper we deal first with a description of the benzene exposure levels in various parts of the industry during the time period in question. We then present an analytical framework which we found to be necessary to characterize exposures adequately and to deal with issues of exposure variability within and between workers. The policy implications of these analyses, in the context of item 5 above, will be treated in a subsequent paper.

The Data Base

The data base of personal exposure measurements which was available for analysis was comprised principally of eight-hour time weighted average (TWA) values, although some short-term exposures measurements were submitted. An early decision was made to focus on the eight-hour data to avoid the additional variability that would be introduced by different sampling durations.[1] All measurements were collected with either charcoal tubes or passive organic vapor monitors.

Data were submitted by nine petroleum refining companies. The measurements were originally collected for many different reasons by the company hygienists and, for the majority of the data, it was impractical to consider separately samples collected under routine conditions versus those collected to characterize exposures during unusual circumstances. In that regard, respiratory protection was used in some of these exposure situations although we did not collect data on the frequency of respirator use. In general, each measurement was classified by worker identifier, location, unit and job. The fact that more than one observation was collected on most workers presented the possibility of estimating the variability in exposure due to the common work environment versus that associated with a particular individual.

Early in the study considerable effort was given to working with the company hygienists in an attempt to define job categories that would at least approximate equal exposure groups. In some companies the specificity of job titles made it necessary

to pool several job categories into a single exposure group for analysis. In the end, all job groups were specific to a particular refinery, a unit within the refinery and an operation within the unit. For example, one particular job group was comprised of all benzene production operators (24 persons and 58 measurements) on the benzene and cyclohexane unit at location B.

Descriptive Results: Eight-hour TWA Measurements

Once we had come to an understanding of how to specify job groups, we asked the companies to submit eight-hour exposure data for all such groups for which there were 30 or more measurements collected from roughly 10 or more workers. There were 123 such groups. We have chosen to describe these data by giving the proportion of the measurements below the limits of detection and the proportions of measurements observed above 0.5,1.0 and 2.0 ppm, values of potential regulatory significance. Table 1 contains the descriptive exposure data for groups where at least 10% of measurements exceeded 0.5 ppm. That is, data are reported for the most highly exposed groups. The limits of detection vary by sampling medium and by analytical method. Since these factors differed from company to company and over time within companies, detection limits were reported separately for each data submission.

For 30 of the 123 groups all of the measurements were below 0.5 ppm. For many of these 123 groups the fraction below the limit of detection was also quite high. On the other end of the scale, roughly fifteen percent of these groups have significant

fractions of the measurements in excess of 1 ppm. The groups with relatively high exposures and with adequate numbers of measurements will be considered below in some detail. The high exposure groups are comprised of operators of production units within the refineries, bulk transfer or loading operations and laboratory quality control activities.

There was some interest in the exposure of maintenance workers and a special effort was made to obtain data on this employee group. The data submitted were eight-hour TWA values and are summarized in Table 2. However, the nature of maintenance work is such that these groups are not unit specific. Most, in fact, are only company specific so that there can be no pretence of uniformity of exposure. Nevertheless, it should be noted that only about 5% of the measurements exceed 1 ppm and a large percentage of the measurements are below detection limits.

Descriptive Results: Short-term Exposures

An effort was made, late in the study, to obtain data on short-term benzene exposures. In general, short-term refers to exposures measured over intervals of 5 to 30 minutes when a particular task or work condition lead the hygienist to expect high transient exposures. The format for data collection and analysis was, as with the eight-hour data, focussed on unit and site specific job groups. In this case we sought at least ten measurements per group, but no further constraints were imposed. The majority of data was submitted by one company, but thereare some data from a total of four companies.

Table 3 contains a summary of the data in the same format as used for the eight-hour TWA data. The exceedance fractions are based only on a count of the number of measurements above each of the limits where each measurement is treated as an independent value. In all cases the sampling medium was charcoal tubes. Because the sampling time was not uniform the values were normalized to a 15 minute standard. If the measurement duration was less than 15 minutes it was assumed that the task was complete and no further exposure would have occurred. Hence the TWA value was adjusted downward by the ratio of the sampling duration to the 15 minute standard. If the duration exceeded 15 minutes the TWA value was left unaltered under the assumption that exposure was uniform over time.

As can be seen from Table 3, most of the short-term exposures are very low. However, as with the eight hour data there are some groups that show relatively high exceedance fractions even at 5 ppm. Those groups showing exceedance fractions of 0.10 or above at 5 ppm are involved in tank gauging, loading and unloading operations, barge transfer operations and tanker truck loading.

For some groups in Table 3, groups 29 to 31 for example, the limits of detection were relatively high. This leads to high proportions below the detection limit, but also high proportions of the measurements in excess of 0.5 ppm. These entries are not in error and simply indicate short sample duration and correspondingly high detection limit situations.

Analytical Framework

The descriptive data leaves one with the impression that, to the extent that these data are representative of the industry, most operations are associated with exposures below 1 ppm most of the time. Some operations, however, have significant fractions of their exposures above 1 ppm. To reach any more specific conclusion, on the basis of these descriptive data alone, is difficult. Indeed, we cannot even respond to the second goal of the study which was to estimate the proportions of individuals that were exposed to various benzene levels since these data relate to the proportion of measurements rather than of individuals. What is needed is a framework to disentangle the individual worker's contribution to overall variability from those environmental factors that contribute to the variability in exposure common to the group.

Oldham and Roach were apparently the first to apply analysis of variance techniques to exposure data to investigate worker to worker versus day to day variability in exposure.(2) The concept is shown schematically in Figure 1. The idea is that each worker's exposure differs from day to day, and that, over time, the exposure is best described by a distribution of eight-hour TWA measurements which will generally be different for each worker. These are the distributions shown in the middle of the Figure. If these distributions are not the same for each worker, then one's view of the exposure of the group, based on a sample of exposure data of the sort presented above, obviously depends upon which workers were selected for measurement and how many

times each was measured, hence, the difficulty in interpreting the benzene data without directly addressing this issue.

We adopted an analysis of variance model based on the proposition that the distribution of individual exposures was log-normal and that the variance of each of these individual distributions is the same for each worker, but with differing mean levels of exposure. Because the individual variances are assumed identical, the only remaining element of the model to be specified concerns the distribution of mean values among workers. It is not unreasonable to postulate that this distribution is also log-normal. This is the distribution shown at the bottom of the Figure. That is, if one were to determine the mean value of each worker's exposure distribution and there were a large number of workers, the distribution of means would also be log-normal. (The means referred to here are the statistical expectations, estimated by the arithmetic means, or the first moments of the log-normal distributions, not the geometric means or the means of the logarithms.) Consequently, the exposure distribution obtained by randomly selecting workers, and randomly selecting days on which measurements were to be taken, is also log-normal.[3] This latter distribution is that shown at the top of the Figure and includes both the day to day and worker to worker components of variability in exposure. There is empirical evidence that these distributional assumptions are valid in many exposure situations, but we shall subsequently address the goodness of fit issue in the context of the benzene data.

In order to apply this model to the analysis of the benzene data it is necessary to give its analytical formulation and specify the parameters that will be estimated from the data. To this end let h(x) be the overall distribution of exposure levels (the distribution at the top of the Figure), $f(x|\mu)$ be the distribution of the exposure of an individual worker whose mean exposure is μ (the distributions in the middle of the figure illustrating several values of μ and $g(\mu)$ the distribution of means across the worker population (the distribution at the bottom of the figure). These three distributions are related by the expression:

$$h(x) = \int_0^\infty f(x|\mu)g(\mu)d\mu \qquad (1)$$

To explore the properties of h(x) it is necessary to specify the parameters associated with two basic log-normal distributions. Specifically,

$g(\mu)$ is log-normal: geometric mean = GMB
geometric standard deviation = GSDB

$f(x|\mu)$ is log-normal: geometric mean = GMW
geometric standard deviation = GSDW

Further, represent the means and variances of the normal distributions of $\log(\mu)$ as $\mu_B = \log(GMB)$, $\sigma_B = \log(GSDB)$, and $\log(x)$ as $\mu_W = \log(GMW)$ and $\sigma_W = \log(GSDW)$. Note that B refers to the distribution between individuals which is the distribution at the bottom of Figure 1. Similarly, W refers to within individual distributions and T to the total distribution i.e. the distributions in the middle and at the top of Figure 1, respectively.

Under these conditions (g and f both log-normal) h(x) is also a log-normal distribution with geometric mean equal to $\exp(\mu_B - \frac{1}{2}\sigma_W^2)$ and geometric standard deviation $\exp(\sqrt{\sigma_W^2 + \sigma_B^2})$.[3] The expected value of the exposure of a random individual selected from the distribution h(x) is $\exp(\mu_B + \frac{1}{2}\sigma_B^2)$ which is identical to the mean of the distribution g(μ). Table 4 gives an illustration of these three distributions for specific values of the log-normal parameters. It is the estimation and manipulation of these parameters that allows separate assessment of exposure due to individual differences (GSDB) versus those due to common environmental factors (GSDW).

Highly Sampled Groups

In order to estimate the parameters of the model it is clear that the exposure data have to include multiple measurements on each worker and multiple workers per group. In the benzene data we required, somewhat arbitrarily, that each group analyzed should have at least 45 eight-hour measurements on at least 15 individuals with an average of at least two measurements per individual and not more than 30% of the measurements below the limit of detection. These criteria provide sufficient numbers of observations to produce stable estimates of the log-normal parameters. The first nineteen groups in Table 1 met these criteria and are the subjects of the subsequent analysis.

Standard statistical procedures were used to estimate the model parameters.[3] These procedures provide consistent estimates while taking account of the fraction of the measurements below the limits of detection. Before presenting

these results, let us deal with the goodness of fit issue, that is, the degree to which the model adequately summarizes the data. We chose to focus on the right tail of the top, h(x), and bottom, g(μ), distributions of Figure 1, clearly the most critical part of any exposure distribution. In particular, we chose to contrast the proportion of measurements above 1 ppm predicted by the model with that observed and to contrast the predicted proportion of individual means above 1 ppm with those observed in the data set. That is, the means and variances of the model were estimated from the data and these estimates were used to predict the fraction of measurements or the fraction of individual means above 1 ppm. These estimates were then contrasted with the observed exceedances. These contrasts are presented in Tables 5 and 6. These values are calculated using the relationships in the previous section. For example, the values in Table 5 come from h(x) where:

$$p = P\{Z > [\log(1.0) - \log(GMT)]/\log(GSDT)\}$$

where Z has a standard normal distribution.

As can be seen from Tables 5 and 6 the observed and estimated exceedances are generally in close agreement. Where there are differences it is not clear whether they are due to inadequacies of the model or problems inherent in the data. That some data-related differences should exist is not surprising since, as noted earlier, these data were collected for many reasons and it is unlikely that considerations of representativeness and independence were high priority criteria.

We know that the benzene data are right-skewed and bounded below by zero, both characteristics of the log-normal distribution. These facts, together with the data presented in these tables are the principal evidence that can be offered on the distributional questions in any rigorous sense. The quantity of independent data was not sufficient to allow a more thorough evaluation of the distributional properties of the observed data. Nevertheless, we felt comfortable in adopting the log-normal model as a basis for subsequent analyses.

Table 7 contains the results of an analysis of variance for the nineteen groups. The mean value reported in Table 7 is the estimated arithmetic mean of both the total distribution and of the distribution of individual means. This value and the appropriate GSD allow one to calculate the geometric mean for either the total distribution, $h(x)$, or the distribution of means, $g(\mu)$.

The PERCENT column of Table 7 gives the percent of the total variability that is attributable to differences between workers as opposed to that associated with the common environment. [This value is based on the variance of the logs of the measurements rather than the GSDs or the second moments of the log-normal distributions.] If this number is near 100% it indicates that the group was heterogeneously exposed and that there were differences in the exposure of individual workers due to different work practices or to different tasks. A PERCENT value of 100 corresponds to a GSDW of 1.00 which, in turn, corresponds to zero variability associated with the environment. As can be

seen from the Table, there are groups where the variability is principally environmental, e.g. group 1, and others where variability is totally between workers. In general, however, a somewhat greater proportion of the total variability seems to be associated with differences between workers than with the common environment.

Three out of four cases in which the variability (GSDT) is almost totally associated with differences between workers are associated with groups from a single company. We suspected that this might be due, at least in part, to the sampling strategy used by this company. For example, workers might have been preferentially selected for monitoring because their exposure was highly variable. To explore this issue we calculated the correlation between the number of measurements per individual and the variance of those values for each group. Correlation coefficients are given in Table 7 under the column labelled r. As can be seen, there is a clear tendency for high values of r to be associated with high values of GSDT indicating that the total variability was inflated in some of these groups by repeated sampling of workers with the most variable exposures.

For some of these groups we were able to pursue this issue further and explore the possibility that these high variances were associated with non-routine exposure conditions. For groups 5, 6, 7, and 8, all associated with one company, it was possible to determine that measurements taken during turn-around and inspection conditions were over-represented in the sample and

contributed disproportionately to the variance estimates. Similarly, the most extreme value of GSDT, 9.62, comes from group 18 and results from the dominance of the data from only three individuals who have highly variable levels of exposure and were sampled repeatedly (r = .402). These results underscore the caution that must be exercised when attempting to form useful conclusions based on statistical analysis of data that were not originally collected with such analyses in mind.

Conclusions

From a descriptive point of view, we found that average eight-hour TWA exposures in most job groups that we analysed were below 1 ppm. However, in a number of cases, the variability was sufficiently high to result in 1 ppm exceedance fractions of 0.1 or above. This finding is true for both the distribution of means, $g(\mu)$, and the distribution of measurements, $h(x)$, in those cases where adequate data were available to partition the variability. In a subsequent paper we will explore the implications of this finding to health risk assessment and to compliance monitoring.

Although the data were not adequate to allow a conclusive investigation of the goodness-of-fit of the log-normal model, our limited exploration of the issue gave no indication that the assumption was misleading in this application. In fact, the conceptual clarification arising out of the analysis of variance model is a principal result of the benzene study. If the data have been collected so as to be representative of the exposure of the group, the log-normal model offers the hygienist an

opportunity to identify the most promising strategy for exposure reduction. If, for example, it is found that there is little difference in the exposure of the workers in the group, then environmental controls are indicated. Conversely, large differences in the exposures of the individual workers suggests that an analysis of job tasks or of work practices may be more useful.

REFERENCES

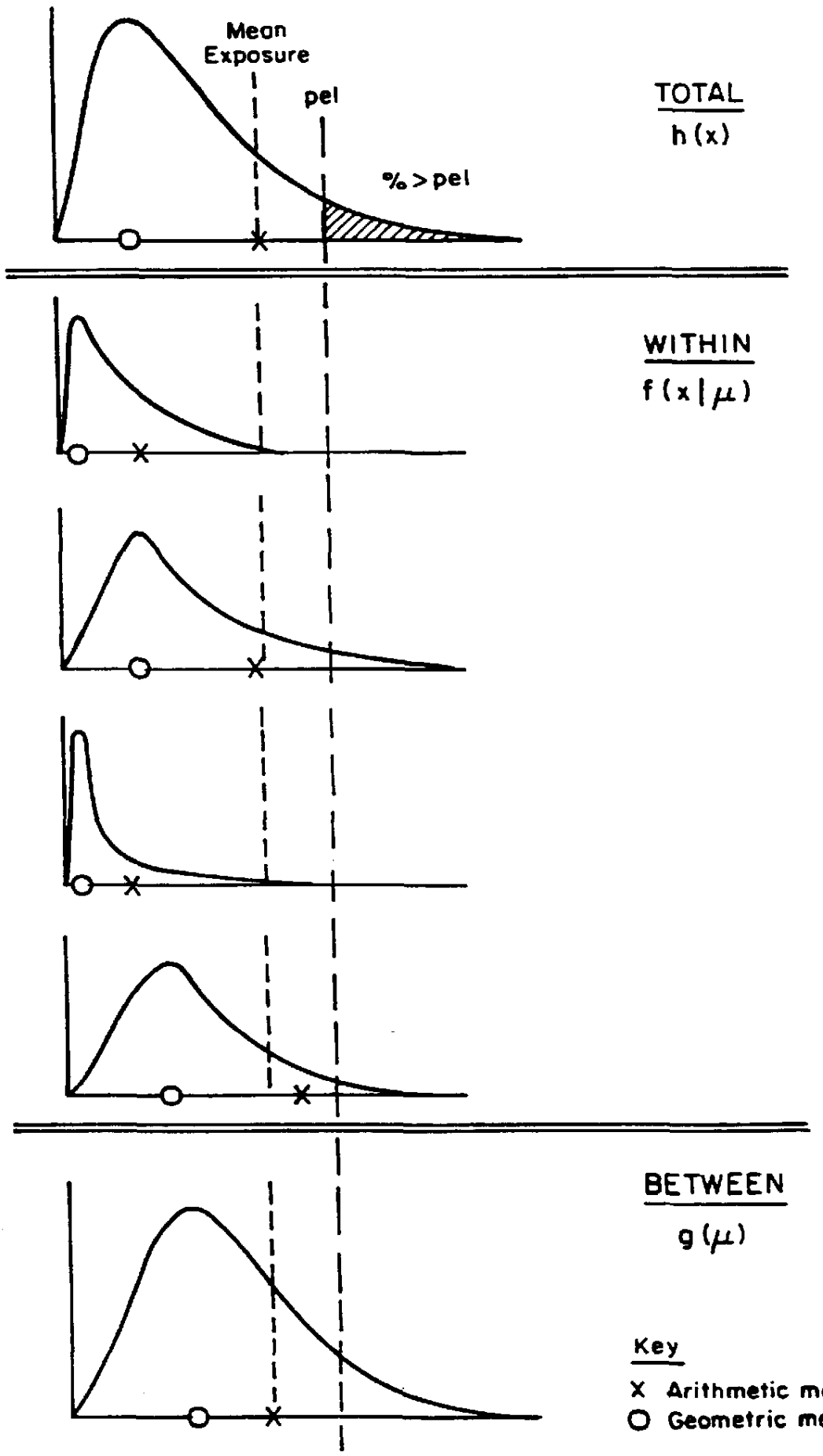
1. Spear, R.C., S.Selvin and M. Francis: The Influence of Averaging Time on the Distribution of Exposures. Am. Ind. Hyg. Assoc. J., 47:365 (1986)

2. Oldham, P.D. and S.A. Roach: A Sampling Procedure for Measuring Industrial Dust Exposure. Brit. J. Industr. Med. 9:112 (1952)

3. Aitchison, J. and J.A.C. Brown: The Lognormal Distribution. pp 110-111, Cambridge University Press, Cambridge (1976)

FIGURE CAPTION

Figure 1: Lognormal exposure model



SAL 00000139

TABLE I

Empiric estimates of exceeding certain limits (ppm) for groups of refinery workers exposed to benzene

GROUP	DESCRIPTION	MEDIUM[a]	SAMPLES	PERSONS	P(X>0.5)	P(X>1.0)	P(X>2.0)	% BLD[b]
1	Location A, Catalytic Cracker Refining Operators	OV	202	39	0.01	0.00	0.00	27.7
2	Location B, Ethylene Unit Benzene Production Operators	OV	170	44	0.09	0.01	0.01	8.2
3	Location B, Waste Water Treatment Refining Operators	OV	83	22	0.37	0.13	0.06	15.7
4	Location B, Benz. & Cyclohex. Unit Benzene Production Operators	OV	58	24	0.41	0.14	0.03	0.0
5	Location E, CRU Platformer Head Operator, Controlman	CT	55	16	0.13	0.09	0.05	3.6
6	Location E, CRU Platformer Misc. Operators	CT	66	22	0.20	0.17	0.14	7.6
7	Location E, Petrochemical Plant Head Operator, Controlman	CT	80	20	0.28	0.20	0.10	2.5
8	Location E, Petrochemical Plant Misc. Operators	CT	161	38	0.35	0.25	0.17	3.1
9	Location F, Benzene Extraction Unit, Operator	OV	124	23	0.26	0.12	0.03	10.5
10	Location F, Aromatics-East ACU Operator	OV	97	25	0.20	0.11	0.07	18.6
11	Location F, Aromatics Gauger/Operator	OV	113	33	0.16	0.04	0.00	15.9
12	Location F, Aromatics Column Operator	OV	98	25	0.24	0.07	0.01	14.3
13	Location F, Aromatics-West ACU Operator	OV	79	19	0.11	0.06	0.01	25.3
14	Location F, Aromatics, Foreman	OV	89	17	0.24	0.11	0.03	14.6
15	Location F, Cumene-Phenol/ Acetone, Operator	OV	177	25	0.24	0.12	0.08	24.3
16	Location A, Cracking Department Still-Platformer, Refining Oper.	OV	90	18	0.10	0.02	0.00	2.2
	[a] Petrochemical Dept., Benz. Prod. Oper.	OV	59		0.49	0.19	0.05	0.0

SAL 0(

18	Location S, Benzene Transfer/Movement Operators	CT	48	24	0.	0.25	0.19	6.2
19	Location T, Lab Unit, Technical	CT	55	18	0.23	0.14	0.02	9.1
20	Location G, Main Deck, Transportation-Marine	CT	49	n/a[c]	0.61	0.55	0.41	0.0
21	Location B, Bulk Oil Pumphouse Refining Operators	OV	33	14	0.36	0.12	0.00	15.2
22	Location B, Bulk Oil Pump Station Refining Operators	OV	45	16	0.18	0.13	0.07	26.7
23	Location B, Laboratory Staff	OV	37	34	0.35	0.08	0.03	0.0
24	Location E, Petrochemical Unit	CT	96	56	0.11	0.08	0.05	14.6
25	Location M, Lab QC Unit All Staff	CT	38	32	0.37	0.21	0.13	2.6
26	Location F, Dispatching Dock Dockman	OV	45	23	0.36	0.22	0.13	26.7
27	Location F, Phenol/Acetone Foreman	OV	116	19	0.14	0.08	0.04	49.1
28	Location F, Oil Recovery-Environmental, Operator	OV	35	22	0.11	0.06	0.03	42.9
29	Location Q, Platformer, Hydro-desulfurizer, Stillman, Refining	CT	56	10	0.14	0.00	0.00	73.2
30	Location Q, Unit D Stillman, Refining	CT	42	17	0.14	0.05	0.00	64.3
31	Location Q, Control Room Laborer, Yardman, Roustabout	CT	36	8	0.11	0.00	0.00	0.0
32	Location R, Compound Plant Yardman	CT	36	8	0.11	0.00	0.00	0.0
33	Location U, Maintenance	CT	37	34	0.14	0.03	0.03	16.2
34	Location U, Lab, Quality Control	CT	52	47	0.19	0.10	0.08	19.2
35	Location W, Receipt, Storage, and Movement	CT	55	36	0.49	0.25	0.20	10.9
36	Location X, Lab Unit, Technical	CT	35	17	0.11	0.03	0.00	11.4
37	Location Y, Catalytic Cracking Unit, Outside Journeyman	CT	34	22	0.12	0.09	0.09	11.8
38	Location Y, Aromatics Recovery Unit, Outside Journeyman	CT	34	18	0.50	0.26	0.09	0.0
39	Location AA, Benz. Unit Operators	CT	58	39	0.31	0.16	0.03	6.9
40	Location AA, Ethylene Unit Feedstock Operators	CT	30	24	0.10	0.10	0.00	13.3
41	Location BB, Lab Tech	CT	46	32	0.11	0.00	0.00	41.3

42	Location CC, Lab Tech	CT	48	12	0.60	0.50	0.27	20.8
43	Location DD, Lab Tech	CT	106	63	0.21	0.07	0.03	8.5
44	Location II, Lab Tech	CT	44	20	0.11	0.02	0.00	52.3

a Medium, OV = Passive Organic Vapor Monitor, CT = Charcoal Tube
b BLD = Below the limit of detection
c n/a = not applicable since there are no ID numbers

TABLE II

Empiric estimates of exceeding certain limits (ppm) for groups of maintenance workers

GROUP	DESCRIPTION	MEDIUM[a]	SAMPLES	PERSONS	P(X>1.0)	P(X>2.0)	% BLD[b]
M1	All Locations Pipefitter/Welder, Pipefitter/Plumber,Welder	CT	133	34	0.00	0.00	11.3
M2	Location O Contract Workers, Mostly Maintenance	OV	138	n/a[c]	0.16	0.06	22.5
M3	Location O Maintenance, Misc.	OV	56	34	0.05	0.00	58.9
M4	Location P Contract Workers, Misc. Maintenance	OV	70	n/a[c]	0.04	0.01	78.6
M5	Location F Pipefitter-Phenol/Acetone-Maintenance	OV	104	47	0.03	0.02	56.7
M6	All Locations Maintenance unit	CT	58	49	0.09	0.07	51.7
M7	Location B Instrument Shop, Machine Shop, Paint Shop	OV	53	16	0.00	0.00	45.3
M8	All Locations Maintenance Operations	OV	56	40	0.09	0.09	41.1
M9	All Locations Mechanic-General Plant	CT	118	36	0.02	0.00	29.7
M10	All Locations Maintenance-Leadman	CT	52	20	0.04	0.04	59.6

a Medium, OV = Passive Organic Vapor Monitor, CT = Charcoal Tube
b BLD = Below the Limit of Detection
c n/a = not applicable since there are no unique ID #'s in these groups

TABLE

Probability of exceeding selected limits (ppm) for short term samples

GROUP	DESCRIPTION	MEDIUM	SAMPLES	PERSONS	0.5	1.0	2.0	5.0	10.0	% BLD[a]
1	Company 1, Location 2, RS&M[b] Loading-Routine	CT	12	2	0.17	0.08	0.00	0.00	0.00	17
2	Company 1, Location 11, RS&M Unloading-Routine	CT	19	3	0.11	0.11	0.05	0.05	0.00	32
3	Company 1, Location 11, RS&M Manually Gauging Tanks-Routine	CT	11	3	0.27	0.18	0.00	0.00	0.00	9
4	Company 1, Location 12, RS&M Loading-Routine	CT	16	4	0.38	0.13	0.00	0.00	0.00	13
5	Company 1, Location 12, RS&M Unloading-Routine	CT	20	3	0.15	0.00	0.00	0.00	0.00	10
6	Company 1, Location 13, RS&M Loading-Routine	CT	15	3	0.87	0.47	0.27	0.00	0.00	7
7	Company 1, Location 13, RS&M Unloading-Routine	CT	10	3	0.20	0.10	0.10	0.10	0.00	10
8	Company 1, Location 14, RS&M Loading-Routine	CT	12	3	0.50	0.17	0.00	0.00	0.00	0
9	Company 1, Location 14, RS&M Unloading-Routine	CT	14	3	0.29	0.14	0.07	0.00	0.00	14
10	Company 1, Location 15, RS&M Unloading-Routine	CT	24	3	0.33	0.17	0.00	0.00	0.00	13
11	Company 1, Location 16, RS&M Receipt, Storage and Movement Loading-Routine	CT	13	2	0.15	0.00	0.00	0.00	0.00	0
12	Company 1, Location 17, RS&M Loading-Routine	CT	10	3	0.20	0.10	0.00	0.00	0.00	10
13	Company 1, Location 17, RS&M Unloading-Routine	CT	13	2	0.15	0.00	0.00	0.00	0.00	23
14	Company 1, Location 18, RS&M Unloading-Routine	CT	19	3	0.21	0.11	0.00	0.00	0.00	0
15	Company 1, Location 18, RS&M Manually Gauging Tanks-Routine	CT	14	4	0.64	0.64	0.50	0.21	0.00	7
16	Company 1, Location 22, RS&M Unloading-Routine	CT	11	6	0.18	0.00	0.00	0.00	0.00	36
17	Company 1, Location 24, RS&M Loading-Routine	CT	10	3	0.90	0.90	0.70	0.20	0.10	10

.8	Company 1, Location 25, RS&M Loading-Routine	CT	44	11	.30	0.16	0.05	0.00	0.00	25
19	Company 1, Location 26, Petrochemical Plant Area Visual Inspection of Automatic Gauges-Routine	CT	27	4	0.26	0.07	0.00	0.00	0.00	11
20	Company 1, Location 27 Laboratory, Quality Control Testing-Routine	CT	15	9	0.33	0.00	0.00	0.00	0.00	33
21	Company 1, Location 28, RS&M Manually Gauging Tanks-Routine	CT	10	5	0.30	0.10	0.00	0.00	0.00	0
22	Company 3, Location 1 Barge Transfer Operations	CT	18	9	0.94	0.94	0.89	0.72	0.56	0
23	Company 6, Location 1 Tanker Truck Loading-Bottom Without Recovery	CT	18	10	0.22	0.17	0.17	0.11	0.00	67
24	Company 6, Location 2 Tanker Truck Loading-Bottom With Recovery	CT	12	9	0.25	0.08	0.00	0.00	0.00	67
25	Company 6, Location 4 Tanker Truck Loading-Top Without Recovery	CT	10	10	0.40	0.20	0.20	0.10	0.00	40
26	Company 2, Location 6 Crude Processing Unit Outside Journeyman-Routine	CT	14	7	0.14	0.07	0.07	0.07	0.07	0
27	Company 2, Location 7 ARO Recovery Unit Outside Journeyman-Routine	CT	26	9	0.65	0.54	0.35	0.04	0.04	0
28	Company 2, Location 7 Lab Unit Technician-Routine	CT	16	7	0.19	0.06	0.06	0.00	0.00	37
29	Company 2 Marketing Ops-RS&M Admin/Sup-Routine	CT	18	6	0.44	0.33	0.17	0.00	0.00	61
30	Company 2 Marketing Ops-RS&M Terminal Operators-Routine	CT	68	22	0.62	0.26	0.04	0.01	0.01	79
31	Company 2 Marketing Ops-RS&M Transport Driver-Routine	CT	98	46	0.78	0.42	0.11	0.02	0.00	58

a MEDIUM, CT = Charcoal Tubes, OV = Passive Organic Vapor Monitors
b BLD = Below the Limit of Detection
c RS&M = Receipt, Storage and Movement

TABLE IV

Example of the relationship between distributional parameters in the lognormal analysis of variance model

	GEOMETRIC		NORMAL		
DISTRIBUTION	MEAN	VARIANCE	MEAN	VARIANCE	EXPECTATION
BETWEEN=g(u)	2.000	3.000	0.693	1.099	3.657
WITHIN=f(x|u)	2.500	2.000	0.916	0.693	3.178
TOTAL=h(x)	1.573	3.666	0.453	1.299	3.657

TABLE V

Empiric and parametric estimates for the proportion of samples exceeding 1 ppm

GROUP	P(X>1.0) (EMPIRIC)	P(X>1.0) (PARAMETRIC)
1	0.005	0.007
2	0.012	0.042
3	0.133	0.192
4	0.138	0.154
5	0.091	0.046
6	0.167	0.122
7	0.200	0.184
8	0.255	0.260
9	0.121	0.102
10	0.113	0.108
11	0.035	0.023
12	0.071	0.066
13	0.063	0.045
14	0.112	0.090
15	0.124	0.123
16	0.022	0.012
17	0.186	0.193
18	0.250	0.274
19	0.145	0.112

TABLE VI

Empiric and parametric estimates for the proportion of means exceeding 1 ppm

GROUP	P(μ>1.0) (EMPIRIC)	P(μ>1.0) (PARAMETRIC)
1	0.000	0.000
2	0.023	0.023
3	0.182	0.263
4	0.125	0.172
5	0.000	0.039
6	0.136	0.127
7	0.200	0.192
8	0.342	0.315
9	0.087	0.099
10	0.120	0.106
11	0.000	0.015
12	0.000	0.012
13	0.053	0.059
14	0.059	0.095
15	0.160	0.144
16	0.000	0.001
17	0.200	0.198
18	0.333	0.380
19	0.056	0.098

TABLE VII

Parameters of the lognormal model for benzene exposure and results from the one-way analysis applied to 19 specific job groups

GROUP	GSDB g(μ)	GSDW f(x|μ)	GSDT h(x)	PERCENT	μ_B [a]	r
1	1.87	2.83	3.36	26.6	0.09	0.041
2	2.26	2.52	3.43	43.7	0.28	0.106
3	3.36	3.42	5.62	49.2	0.96	0.065
4	1.91	1.54	2.18	68.9	0.67	0.101
5	3.62	1.98	4.30	77.9	0.24	0.275
6	6.52	1.00	6.52	100.0	0.68	0.327
7	5.52	1.00	5.52	100.0	0.23	0.316
8	6.31	1.26	6.41	98.5	2.24	0.311
9	2.28	2.01	2.94	58.0	0.49	0.030
10	2.67	2.44	3.76	54.7	0.48	0.182
11	2.11	1.69	2.49	67.0	0.26	0.051
12	2.02	1.97	2.66	51.5	0.26	0.088
13	3.11	1.66	3.46	83.5	0.32	0.145
14	2.41	2.08	3.15	59.0	0.46	0.103
15	4.15	2.22	5.12	76.0	0.61	0.130
16	1.56	1.73	2.02	40.1	0.26	0.031
17	1.75	1.54	2.03	62.3	0.73	0.016
18	9.29	1.49	9.62	96.9	6.07	0.402
19	3.01	2.72	4.43	54.9	0.44	0.281

a This column is included so the entire parametric model may be reconstructed from the given data.

Benzene Risk Assessment: An Overview

C. J. DiPerna
Mobil Oil Corporation
Chairman API Benzene Issues Group

Benzene has become the representative chemical for initiating or increasing controls for a wide variety of health and environmental regulatory issues, being:

- The subject of federal air and occupational regulations for over ten years, often providing the state-of-the-art for setting important judicial and regulatory precedents on acceptable risks or margins of safety.

- The pollutant of choice to set groundwater contamination allowances which in turn drive clean up standards in CERCLA and RCRA.

- The primary pollutant for examining further air toxic control needs and as a model for Title III CERCLA emissions reporting in 1988.

- The subject of a major new rulemaking under Section 112 of the Clean Air Act in 1988.

- The primary subject for debate about the need to remove aromatics from gasoline in California where regulations are expected in the third quarter of 1988. California Proposition 65 also requires establishment of a safe level of exposure by the fall of 1988.

Industry now is being regulated by the ultraconservative risk assessments that result, for example, in a one in a million risk for benzene in ambient air ranging from 0.006 ppb (California) to 0.04 ppb (EPA).

In order to influence these regulations and provide credible risk assessment alternatives, API has undertaken major efforts to gain a fundamental understanding of the health effects of benzene and to provide information into the regulatory and legislative process. These programs include:

- Risk Assessment - The 1985 EPA linear-based risk assessment has been reworked to incorporate updated human exposure information and refinements in risk assessment methodology. Meanwhile, efforts are underway to develop and gain acceptance of non-linear, biologically based risk assessment that makes maximum use of the available animal and human mechanism of action and leukemia data on benzene. This work also will show the uncertainties of the ultra-conservative risk assessments being used today in EPA and California rulemaking.

- Risk Perspective - Projects are underway to quantify the contribution and relative importance of various sources of benzene to humans. Sources of benzene include ambient air, indoor air, smoking, water and food. A specific objective is to assess the relative impact that small reductions in ambient air levels have on the total body burden of benzene.

- Threshold Level - Longer term studies are being supported on benzene mechanism of action, pharmacokinetics and other toxicological research to help define a carcinogenic threshold level for benzene exposure.

API and WOGA are working closely together to provide the results of these research activities into EPA, California and other state agencies as they consider various regulations on benzene in air, water and wastes.

2538P

Mobil Oil Corporation

ENVIRONMENTAL HEALTH
AND TOXICOLOGY LABORATORY
P.O. BOX 1025
PRINCETON, NEW JERSEY
TELEPHONE [illegible]

M.A. MEHLMAN
DIRECTOR
TOXICOLOGY DIVISION

April 5, 1988

Dr. Robert Drew
API
1220 L St., NW
Washington, DC 20005

Dear Bob:

Enclosed for your information is a copy of "Occupational Exposure to Benzene at Automotive Service Stations" by R.W. Hartle.

Best regards,

Sincerely,

M. A. Mehlman

MAM/kt
Enc.
0200c

Occupational Exposure to Benzene at Automotive Service Stations

by

Richard W. Hartle

A thesis submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Public Health

University of Washington

1980

Approved by _____

(Chairperson of Supervisory Committee)

Program Authorized to
Offer Degree _____

Date _____

Master's Thesis

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this thesis is allowable only for scholarly purposes. It is understood, however, than any copying or publication of this thesis for commercial purposes, or for financial gain, shall not be allowed without my written permission.

Signature_____

Date_____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Professors Morgan and Horstman for their assistance and understanding in the preparation of this manuscript. In addition, special thanks are due to Mr. Ronald J. Young, whose guidance was helpful during the early phase of this undertaking.

CHAPTER I

Introduction

The permanent final federal standard for benzene [1] requires that worker exposures be limited so as not to exceed one part per million (ppm) averaged over the duration of the work shift. In light of the benzene content of gasoline, environmental surveys were conducted to evaluate the occupational benzene exposures of service station attendants. Personal and general area environmental air samples were collected under a wide range of climatic conditions and at a variety of types and sizes of service stations. Certain variables with potential influence on exposure variations were measured to facilitate overall evaluations while controlling for these effects. Service stations equipped with vacuum assist-type vapor recovery systems were also surveyed to determine the relative reduction in exposures.

CHAPTER II

Background

As with numerous chemical and physical agents once considered relatively innocous, gasoline vapor is now regarded as potentially hazardous, primarily due to its benzene content. Benzene has received notoriety as a hazardous substance since the early part of this century, but only recently have the federal regulatory agencies concluded that the cumulative medical and epidemiological evidence is sufficient to cite benzene as a carcinogen. This, from an enforcement standpoint, incriminates even minute exposures as potentially detrimental to health.

On May 3, 1977, the U.S. Department of Labor, specifically the Occupational Safety and Health Administration (OSHA) issued an Emergency Temporary Standard (ETS) [2] for occupational exposure to benzene. The ETS provided for a permissible employee exposure limit of one ppm benzene on an 8-hr. time-weighted basis, and a 5 ppm ceiling limit. Although a judicial restraining order was immediately issued, the ETS set the framework for a permanent final standard involving the same exposure limits plus, among other aspects, requirements for measurement of employee exposure, personal protective equipment, employee training, work practices, medical surveillance, signs, labeling, and recordkeeping. As indicated in the Proposed Permanent Standard, OSHA exempted automotive service stations and certain other aspects of transportation, storage, distribution and sale of gasoline and other motor fuels from the permanent final standard. OSHA explained its intention to assess the regulatory action to be taken to protect workers involved in these activities after conclusion of the deliberations of a joint EPA-NIOSH-OSHA Task Force and to deal with this area separately, subsequent to the promulgation of the permanent final standard. Included in the issues which the Task Force was attempting to resolve, were occupational exposure levels related to the sale of gasoline.

Prior to 1976, the general time when the 10 ppm benzene TLV was initially challenged, [3,4,5] little literature was available concerning

benzene in gasoline and related exposures, and that which was available was mostly related to exposures at various types of bulk loading facilities. The only pertinent study was reported by Parkinson in 1971,[6] and dealt with an investigation of working conditions at "typical" filling stations and at selected oil company bulk loading installations in England. Of nine filling stations visited, the benzene content of the gasoline ranged from 2.8 to 5.8% v/v. Results of breathing zone environmental air samples indicated benzene exposures ranging from 0.3 to 3.2 ppm. The average of the mean responses was 1.2 ppm. The samples were collected under a wide range of climatic conditions and during a period when sales were expected to be above average (throughput during tests ranged from 28 to 850 gallons of fuel; average = 390 gals.). The author concluded that only "trivial" exposure to benzene occurred when gasoline with normal benzene content of up to 5% v/v was dispensed, with respect to the then current TLV and ceiling standards of 10 and 25 ppm, respectively.

Subsequent to 1976, additional studies were undertaken to document employee exposures to benzene as a result of distribution and sale of gasoline. In 1979, McDermott et al.[7] reported an environmental survey involving 84 charcoal tube personal samples collected at seven Shell service stations on six or more days of sampling at each station. The authors concluded that the time-weighted exposure determinations were well below one ppm, except for one exposure of 2.08 ppm, and that eight 15 minute peak determinations resulted in exposures of 1.21 ppm or less. They also concluded that no correlations could be established based on the following variables; benzene content, Reid Vapor Pressure, storage tank temperature, and location of auto fill opening. Additionally, the authors stated that other variables which were not measured, such as work practices and micro-climatic conditions, may be implicated in exposure variations. In an evaluation of the effectiveness of vapor recovery, the authors concluded that exposures may not be dramatically reduced.

In a brief study reported by Runion,[8] 20 charcoal tube samples involving nine employees collected hydrocarbons over two-hour sampling periods. Total two-hour integrated benzene levels showed concentrations

ranging from 0 to 1.7 ppm. Only two of these 20 samples produced results above 0.6 ppm, and the average concentration for all samples taken was 0.32 ppm.

Berlin et al.[9] reported a European study of chromosome aberrations among workers exposed to benzene via gasoline. Tank truck drivers, tanker ship crews, and service station attendants were involved in the survey. Of ten service station attendants monitored for benzene exposure, results indicated concentrations ranging from 0.19 to 1.47 ppm; averaging 0.545 ppm.

In a brief preliminary study conducted by Battelle Laboratories for the Environmental Protection Agency,[10] three short term environmental samples (less than three minutes) were collected while individuals filled their tanks with gasoline. Exposures ranged from 0.043 to 0.647 ppm, averaging 0.254 ppm.

In light of the small amount of available data, and the exposure levels indicated by the data which does exist (with reference to the permanent final standard), a scientific endeavor was undertaken involving numerous environmental surveys at automotive service stations. The purpose of this document is to present a composite, overall evaluation based on the data collected during these surveys, with major interest in documenting occupational exposure levels and as a secondary objective, to determine those factors which contribute to the extent of exposure. Also presented is a limited evaluation of the effectiveness of vapor recovery.

CHAPTER III
Methods

To surmount potential bias of a particular geographical region the environmental survey proceeded from Ohio to Florida, again to Ohio, to Nevada, and finally to California. In all, the survey encompassed 35 days of environmental sampling.

Environmental air samples were collected in the breathing zone of the service station attendants (personal samples) and in the general work area (area samples). The majority of the personal samples were collected for approximately 150 minutes, or three per 8-hr. shift. Area samples were collected for approximately 240 minutes, or two per shift. MSA low-flow sampling pumps calibrated at 600 cc/min. were connected to SKC 150 mg activated cocoanut shell charcoal tubes contained in MSA tube holders. The tube holders were clipped to the attendant's lapel in a vertical position for personal sample collection and to the top of the fuel pump in a vertical position for area samples. Sample times and flow rates were determined by the NIOSH analytical laboratories to account for the low benzene concentrations relative to the other constituents of gasoline.

Sample analysis was performed via the NIOSH P&CAM Method #127 [11] by the NIOSH Analytical Laboratories in Cincinnati and by the Utah Biomedical Test Laboratories in Salt Lake City. Variations of the method (ie., column type, operational temperature, use of an internal standard) were used depending on the analyst and data requirements.

A number of variables with potential influence on the extent of benzene exposure received by the attendants were monitored during sample collection. These were percent nozzle time, dispensed gallons of gasoline per hour, liquid volume percent benzene, temperature, windspeed, and relative humidity (barometric pressure was also monitored, for use in microgram to ppm conversions, but was not considered as having any effect on exposure variations). Each of these parameters was noted and recorded for each environmental sample.

Most variables were selected essentially from an intuitive process based on experience and a knowledge of the interactions between exposures and environmental factors. However, percent nozzle time was

selected subsequent to a brief pilot study which indicated that work practices varied considerably between workers and between stations, and that a measure of this variability seemed appropriate. Percent nozzle time represents the portion of the environmental sampling period that the attendant remained in the vicinity of the nozzle while gasoline was being dispensed. "Vicinity" was defined as within reach of the nozzle.

Temperature, barometric pressure, and relative humidity values were obtained by averaging a number of readings taken periodically (15-20 min. intervals) throughout the sampling period. Windspeed values were obtained using a Weather-Measures Corporation anemometer equipped with a strip-chart recorder. The number of gallons and type of fuel dispensed (regular, premium, or unleaded) was recorded after each sale. Bulk samples were obtained and analyzed for liquid volume percent (LV %) benzene. Weighted averages of benzene content were calculated for subsequent use in data analysis.

CHAPTER IV

Results and Discussion

The basic objective of this research was to estimate service station attendant's exposure to benzene. Accordingly, initial data analysis involved descriptive statistics which summarize benzene exposure concentrations obtained at the various sampling locations. Included are mean responses, ranges, and standard deviations plus information on the distributions of the reported values. Secondly, regression analysis was utilized to examine the relationships between exposure values and those factors contributing to variations in exposures. Finally, data collected at service stations equipped with vapor recovery systems were evaluated as to the effectiveness in controlling exposures. The Statistical Package for the Social Sciences (SPSS) was used as the primary means of data analysis.

Descriptive Statistics

A total of 141 breathing zone samples were collected at 28 service stations at various geographical locations. During the initial survey, samples were collected for the duration of the attendant's shift. During subsequent surveys, personal samples were collected for no more than 150 minutes (or usually three per shift) to account for 1) the possibility of "break-through" and/or 2) in the event of equipment malfunction or subsequent loss of a particular sample, data would still be available for a portion of the affected shift. Appropriate consecutive samples were combined and labeled as "TWA" (Time-weighted Average).

Table 1 illustrates the catagorizations or subgroups of samples according to site and time of year of sample collection. Also presented is the number of samples collected at each site.

Descriptive statistics by location and time of year are presented in Table 2. Table 3 presents composite descriptive statistics for selected variables.

Included in Table 2 and 3 is a calculated variable, Estimate of Source Exposure (ESE), which is an estimate of the exposure concentrations received while the attendant was actually pumping fuel (disre-

garding the cumulative effect of area exposures). As depicted in Figure 1, this variable was mathematically derived (derivation of weighted average calculation) by multiplying the total sample time by the exposure concentration, subtracting the product of total sample time and the area concentrations, and dividing by the product of percent nozzle time and total sample time. Simply stated, if personal exposure concentration, general area exposure concentration, total personal sample time, and time near source (% Nozzle Time) are all known, an estimate of source exposure can be obtained through the above calculation.

This value is at best a conservative estimate, due to an unrecorded parameter, involving that portion of the sample period in which there was zero exposure (making change, restroom visits, etc.) and the effect of surging area concentrations while fuel was being dispensed. Another closely related calculated variable is ESEA, which is ESE plus the general area concentration. This value roughly indicated results which would have been obtained had samples been collected while the attendant actually pumped fuel.

Also presented in the tables are personalized area concentrations (P-Area). General area benzene concentrations were obtained at all locations except the Nevada site. At all other locations, samples were collected at one to four stationary positions in the general area, usually on top of the fuel pumps. Sample periods usually ranged from three to four hours. Since personal sample collection periods and area sample collection periods did not coincide, "personalized" area benzene concentration values were calculated for the periods during which personal samples were collected.

Improvements in analytical methods were made during the course of the surveys which decreased (improved) the analytical limit of detection. However, the Ohio-January data was collected prior to any substantial improvements, resulting in 34 of 36 area sample results being reported as below the 10 microgram limit of detection (the two detectable results were reported at 0.020 and 0.022 ppm). Therefore, this precluded evaluation of the ESE or ESEA variables for this data set.

Figure 1

Mathematical Derivation of Estimate of Source Exposure

$$\frac{(\text{ESE})((\%\ \text{Nozzle Time}/100)(\text{Sample Time}))+(\text{Area Concentration})(\text{Sample Time})}{\text{Sample Time}} = \text{Exposure Concentration}$$

Algebraic Manipulation:

$$\frac{(\text{Semple Time})(\text{Exposure Concentration})-(\text{Sample Time})(\text{Area Concentration})}{(\%\ \text{Nozzle Time}/100)(\text{Sample Time})} = \text{Estimated Source Exposure}$$

Data analysis was complicated due to a "suspect" analytical report, involving the Florida-February and Ohio-February, March data sets, which were submitted to the UBTL analytical laboratories. Reporting errors were initially suspected due to a dramatic and sustained increase in the microgram benzene/tube values appearing approximately mid-way through the 183 analytical results. UBTL records were incomplete and gave no indication as to the cause of the problem. Attempts to explain the sustained increase revealed that, on the average, those values following the increase are 4.01 times greater than the values prior to the increase. This suggests an error involving the gas chromatograph attenuation control, which moderates the amplitude of the electrical signal in a series of 2X, 4X, 8X, 16X, etc., so that the displayed graphics remain within the confines of the output paper or, conversely, to enhance the output thus producing a readable and measurable graph.

Although it is questionable that any of the suspect data should be used in an overall evaluation, attempts were made at appropriate modification through the use of regression models, since the suspect data comprises such a large percentage of the entire data set.

Regression Analysis

The focus of the initial regression analysis was the evaluation of the overall dependence of exposure concentrations on the independent variables. Subsequently, attention was given to individual independent variables while controlling for the other variables.

A nonstructured (standard) program was developed for the multiple regression of time-weighted personal exposure values upon the independent variables.[12] Suspect and vapor recovery data were excluded from the initial run. As mentioned earlier, individual consecutive samples were combined (usually three per shift) on a time-weighted basis. However, a number of time-weighted samples were not composed of three individual samples as a result of lost or mis-labeled charcoal tubes or malfunctioning equipment, yielding total time-weighted sample times

ranging from 75 to 490 minutes. SPSS provides a method of weighting when cases involved in the analysis exhibit characteristics which require them to be considered more or less heavily. For the purposes of multiple regression, it was desirable to assign greater importance to sample results obtained over a period of, for instance, 490 minutes than to those obtained for only 75 minutes. Therefore, to comply with SPSS methodology, each case's total sample time was divided by the mean of all total sample times, and the resultant figure, or "weight factor", was used to assign relative weights.

Personal Environmental Samples

The SPSS subprogram Regression, in the non-structured mode, initially selects the variable with the highest zero-order correlation, and subsequently selects single variables for inclusion in the model which have the highest zero-order correlating in a stepwise fashion. A summary of the program output of the initial analysis involving nonsuspect, nonvapor recovery data is depicted in Figure 2. The first variable to enter the equation was percent nozzle time. It was reported as highly significant, with an R-square value of 0.58. Temperature, relative humidity, windspeed, liquid volume percent benzene, and gallons per hour were then sequentially entered, with a total R-square value of 0.91. While the overall equation was shown to be significant at each step, percent nozzle time and gallons per hour were shown not to be significant in the final equation ($p = .05$). However, significance of a particular variable in the final equation is determined not only by its correlation with the dependent variable, but also with the degree of correlation and inter-relationships with the other independent variables. In this instance (although there is no theoretical support for the occurrence) percent nozzle time is highly correlated with relative humidity, and to a lesser extent, with temperature. Also, in reference to an overall evaluation, less concern is given to a particular variable's contribution, but rather the cumulative effect of all variables; which in this case accounts for approximately 91% of the variation in exposure concentrations of the sample data.

FIGURE 2

Print-out Summary: Regression
Analysis of Nonsuspect Data

Dependent Variable Exposure Concentration

Variable Entered on Step Number 1 % Nozzle Time

Multiple R.	0.762	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.580	Regression	1.	0.093	0.093	35.932
Adjusted R. Square	0.564	Residual	26.	0.067	0.003	
Standard Error	0.051					

Variables in the Equation

Variable	B	Beta	Std. Error B.	F
%Nozzle Time	0.00957	0.762	0.002	35.932
(Constant)	0.00713			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
L.V.%	0.271	0.392	0.878	4.541
Temp.	0.280	0.427	0.981	5.584
WA	-0.157	-0.240	0.984	1.524
RH	0.324	0.375	0.563	4.101
GPH	0.232	0.347	0.937	3.426

Variable Entered on Step Number 2 Temp.

Multiple R.	0.810	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.657	Regression	2.	0.105	0.053	23.925
Adjusted R. Square	0.629	Residual	25.	0.054	0.002	
Standard Error	0.047					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
%Nozzle Time	0.0091	0.723	0.001	37.333
Temp.	0.0007	0.280	0.000	5.584
(Constant)	-0.0334			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
L.V.%	0.141	0.175	0.528	0.754
WS	-0.174	-0.294	0.980	2.273
RH	0.631	0.716	0.442	25.287
GPH	0.217	0.359	0.934	3.543

Figure 2, cont.

Variable Entered on Step Number 3 RH

Multiple R.	0.913	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.833	Regression	3.	0.133	0.044	39.874
Adjusted R. Square	0.812	Residual	24.	0.027	0.001	
Standard Error	0.033					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
ΣNozzle Time	0.00345	0.274	0.002	5.006
Temp.	0.00131	0.502	0.000	27.799
RH	0.00258	0.631	0.001	25.287
(Constant)	-0.15220			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
L.V.%	0.103	0.184	0.526	0.802
WS	-0.203	-0.491	0.976	7.314
GPH	0.125	0.287	0.886	2.066

Variable Entered on Step Number 4 WS

Multiple R.	0.934	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.873	Regression	4.	0.140	0.035	39.601
Adjusted R. Square	0.851	Residual	23.	0.020	0.001	
Standard Error	0.030					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
ΣNozzle Time	0.00292	0.232	0.001	4.441
Temp.	0.00136	0.521	0.000	37.609
RH	0.00266	0.651	0.000	33.868
WS	-0.00110	-0.203	0.004	7.314
(Constant)	-0.13354			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
L.V.%	0.274	0.503	0.428	7.469
GPH	0.044	0.107	0.735	0.254

Figure 2, cont.

Variable Entered on Step Number 5 L.V.%

Multiple R.	0.952	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.905	Regression	5.	0.145	0.029	42.085
Adjusted R. Square	0.883	Residual	22.	0.015	0.001	
Standard Error	0.263					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
%Nozzle Time	0.00195	0.155	0.001	2.337
Temp.	0.00094	0.358	0.000	14.056
RH	0.00262	0.641	0.000	41.868
WS	-0.01566	-0.290	0.004	15.531
L.V.%	0.06850	0.274	0.025	7.469
(Constant)	-0.15325			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
GPH	0.108	0.289	0.683	1.912

Variable Entered on Step Number 6 GPH

Multiple R	0.956	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.913	Regression	6.	0.146	0.024	36.844
Adjusted R. Square	0.888	Residual	21.	0.104	0.001	
Standard Error	0.257					

Variables in the Equation

Variable	0.00192	Beta	Std. Error B	F
%Nozzle Time	0.00082	0.152	0.001	2.351
Temp	0.00245	0.314	0.000	10.048
RH	-0.01402	0.599	0.000	34.849
WS	0.07785	-0.260	0.004	11.874
L.V.%	0.00025	0.312	0.025	9.341
GPH	-0.16049	0.108	0.000	1.912
(Constant)				

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F

Figure 2, cont.

Dependent Variable Exposure Concentration

Summary Table

Variable	Multiple R	R Square	RSQ Change	Simple R	B	Beta
%Nozzle Time	0.762	0.580	0.580	0.762	0.00192	0.152
Temp	0.810	0.657	0.077	0.380	0.00082	0.314
RH	0.913	0.833	0.176	0.687	0.00245	0.599
WS	0.934	0.873	0.040	-0.252	-0.01402	-0.260
L.V.%	0.952	0.905	0.032	0.505	0.07785	0.312
GPH	0.956	0.913	0.008	0.409	0.00025.	0.108
(Constant					-0.16049	

Determination of Relative Reporting Errors of the "Suspect" Data

Reflecting back to the suspect data sets and the question concerning the distributions of the ppm exposure values relative to actual, it is now possible to predict ppm exposure values, given the model shown in Figure 2 and a set of independent variables. The mean responses of the initial (lower) portion of the suspect data set were fit to the equation, resulting in a predicted ppm value of 0.148. The average "reported" value was only 0.032 ppm. The same procedure was conducted for the second (higher) portion of the suspect data, resulting in an expected ppm value of 0.103, which approximated the average reported value of 0.107 ppm. When compared to the average ppm value of the nonsuspect nonvapor recovery data, it appears that the initial portion of the suspect data is .25X actual (assuming that G.C. attenuation was the cause of the error) and that the second portion, appearing after the sustained increase, is not in error. Exposure values appearing in Table 2 for the Florida-February data were calculated using "modified" data; ie., lower portion of suspect data X4.

Regression Analysis Including the "Modified" Suspect Data Set

Following appropriate modifications of the suspect data (lower portion X4) a new regression model was developed which included all non-vapor recovery data ($n=51$). Again, the nonstructured (standard) mode was used, and again percent nozzle time was the first variable entered. However, this run produced a comparatively lower R-square value for this variable (.24) as was the case for the overall regression (.68). Also, variables were entered in a different order; liquid volume percent second, then sequentially windspeed, relative humidity, temperature, and finally gallons per hour. All variables except gallons per hour were shown to be signficant in the final equation (Figure 3).

While a "structured" regression model could have been developed which would have facilitated inclusion of independent variables in a step-wise fashion based on a preconceived notion of each variable's relative importance to the overall relationship, the nonstructured mode, which enters a variable at each step depending on its relative absolute

FIGURE 3

Printout Summary - Regression Analysis
Including "Modified" Suspect Data

Dependent Variable: Exposure Concentration

Variable Entered on Step Number 1: % Nozzle Time

Multiple R.	0.487
R. Square	0.237
Adjusted R. Square	0.221
Standard Error	0.062

Analysis of Variance	DF	Sum of Squares	Mean Square	F
Regression	1.	0.058	0.058	15.250
Residual	49.	0.188	0.004	

Variables in the Equation

Variable	B	Beta	Std. Error B	F
%Nozzle Time	0.00620	0.487	0.002	15.250
(Constant)	0.06264			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
L.V.%	0.371	0.413	0.944	9.867
Temp	0.307	0.350	0.993	6.707
WS	-0.341	-0.388	0.987	8.479
RH	0.256	0.237	0.656	2.866
CPH	0.034	0.034	0.794	0.057

Variable Entered on Step Number 2: L.V.%

Multiple R	0.606
R. Square	0.367
Adjusted R. Square	0.341
Standard Error	0.057

Analysis of Variance	DF	Sum of Squares	Mean Square	F
Regression	2.	0.090	0.045	13.938
Residual	48.	0.156	0.003	

Variables in the Equation

Variable	B	Beta	Std. Error B	F
%Nozzle Time	0.00731	0.574	0.002	23.666
L.V.%	0.05435	0.371	0.017	9.867
(Constant)	-0.02138			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
Temp	0.228	0.275	0.925	3.850
WS	-0.480	-0.576	0.909	23.317
RH	0.254	0.259	0.656	3.378
CPH	0.188	0.198	0.709	1.916

Figure 3, cont.

Variable Entered on Step Number 3 WS

Multiple R	0.760	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.577	Regression	3.	0.142	0.047	21.384
Adjusted R. Square	0.550	Residual	47.	0.104	0.002	
Standard Error	0.047					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
%Nozzle Time	0.00702	0.552	0.001	31.947
L.V.%	0.07455	0.509	0.015	25.046
WS	-0.02551	-0.480	0.005	23.317
(Constant)	0.00664			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
Temp	0.183	0.270	0.916	3.609
RH	0.240	0.298	0.655	4.495
CPH	0.087	0.110	0.683	0.562

Variable Entered on Step Number 4 RH

Multiple R	0.784	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.615	Regression	4.	0.151	0.038	18.355
Adjusted R. Square	0.581	Residual	46.	0.095	0.002	
Standard Error	0.045					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
%Nozzle Time	0.00524	0.412	0.001	12.770
L.V.%	0.07421	0.507	0.014	26.666
WS	-0.02523	-0.475	0.005	24.480
RH	0.00108	0.240	0.001	4.495
(Constant)	-0.03436			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
Temp	0.271	0.398	0.830	8.471
CPH	0.068	0.090	0.678	0.370

Figure 3, cont.

Variable Entered on Step Number 5 Temp

Multiple R	0.822	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.676	Regression	5.	0.167	0.033	18.763
Adjusted R. Square	0.640	Residual	45.	0.080	0.002	
Standard Error	0.042					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
INozzle Time	0.00402	0.316	0.001	7.991
L.V.I	0.06228	0.425	0.014	19.941
WS	-0.02371	-0.446	0.005	24.822
RH	0.00152	0.338	0.001	9.426
Temp	0.00085	0.271	0.000	8.471
(Constant)	-0.08631			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
GPH	0.034	0.050	0.670	0.108

Variable Entered on Step Number 6 GPH

Multiple R	0.823	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.677	Regression	6.	0.167	0.028	15.344
Adjusted R. Square	0.632	Residual	44.	0.080	0.002	
Standard Error	0.043					

Variables in the Eduation

Variable	B	Beta	Std. Error B	F
INozzle Time	0.00389	0.306	0.001	6.850
L.V.I	0.06371	0.435	0.015	18.660
WS	-0.02342	-0.441	0.005	23.004
RH	0.00150	0.334	0.001	8.904
Temp	0.00084	0.268	0.000	0.108
GPH	0.00008	0.034	0.000	
(Constant)	-0.08919			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F

Figure 3, cont.

Dependent Variable Exposure Concentration

Summary Table

Variable	Multiple R	R. Square	RSQ Change	Simple R	B	Beta
INozzle Time	0.488	0.237	0.237	0.487	0.00389	0.306
L.V.X	0.606	0.367	0.130	0.236	0.06371	0.435
WS	0.760	0.577	0.210	-0.392	-0.02342	-0.441
RH	0.784	0.615	0.038	0.454	0.00150	0.344
Temp	0.822	0.676	0.061	0.344	0.00084	0.268
CPH	0.823	0.677	0.001	0.248	0.00008	0.034
(Constant)					-0.08919	

partial correlation value, entered variables in descending importance in a manner similar to that expected. For example, percent nozzle time, the first variable to enter the equation, is a measure of the time the attendant remained in the vicinity of the nozzle, or near the source of exposure. It seems reasonable to assume that this value should account for the majority of variation in exposures. The second variable entered, liquid volume percent, is a measure of the amount of benzene present in the gasoline; and again it is reasonable that this should account for a large portion of exposure variation. Next to enter were the three climatic variables. Temperature and windspeed are related to the volatility of gasoline and "ventilation effects". The relationship between humidity and exposures is somewhat more difficult to understand or define, but it is interesting to note that Parkinson, in his study of benzene exposures to service station attendants in England, also found a positive correlation between relative humidity and exposures. The last variable to enter the model, gallons per hour, was found to have a non-significant coefficient in the final equation. This is partially due to its correlation with percent nozzle time (more fuel pumped per hour leads to higher percent nozzle time). Thus, while on an individual basis GPH is positively correlated with exposures, in the overall regression model this correlation was included within that of percent nozzle time in the initial entry. It is important to note that this may occur for all independent variables entered into a particular regression model, and a particular variable's contribution to the overall regression can only be interpreted in light of the other variables in the equation.

In light of the small R-square increase with the addition of GPH to the regression model, an F-Test was conducted to determine if the increase, and thus GPH, significantly added to the overall R-square value. Results showed that a test of the hypothesis H_0: GPH adds no significant increase, could not be rejected. Through this method, the regression equation was simplified by deleting the GPH variable with no significant decrease in the overall R-square value. An identical procedure was applied to the regression equation involving the nonsuspect data sets, and the same result was obtained.

The role of individual variables was analyzed with particular interest in those factors which can be regulated and potentially controlled (versus climatic variables); namely percent nozzle time and liquid volume percent. The method utilized in analyzing and depicting the effects of individual independent variables on variations in exposure concentrations is identical to the method used in construction of regression lines and confidence intervals for regression models containing only one independent variable. However, through use of the multiple regression model, predictions for a selected independent variable's effect on exposures can be made while controlling for the remaining variables and thus account for independent variable interactions. Figures 4 through 7 present graphic illustrations of this method, individually depicting the effect of percent nozzle time and liquid volume percent on exposure concentrations: initially utilizing the regression model derived from the nonsuspect data, and subsequently including the "modified" suspect data. Constants used in the regression equations are mean variable responses of the respective data sets. Regression lines depicted in the figures are for the actual range of data. Variance values used in the confidence interval formula for both dependent and independent variables were derived from unweighted data. While the use of unweighted variance results in confidence intervals which are not actual, only indicative, any deviation from actual is conservative (ie., further removed from the mean).

Regression Analysis of Estimated Source Exposures

To examine possible interactions between the set of independent variables and estimated source exposures, a regression analysis similar to the one just described was conducted utilizing Estimate of Source Exposure (ESE) as the dependent variable. Due to the large variability of source exposure times (time in vicinity of nozzle) responses were weighted in a similar manner as previously described. Percent nozzle time was, of course, excluded from the independent variable list due to its prior use in the ESE calculations. Because area values were also essential in the calculations, analysis was restricted to those data

Figure 4

EXPOSURE CONCENTRATION
VS.
PERCENT NOZZLE TIME
(nonsuspect data)

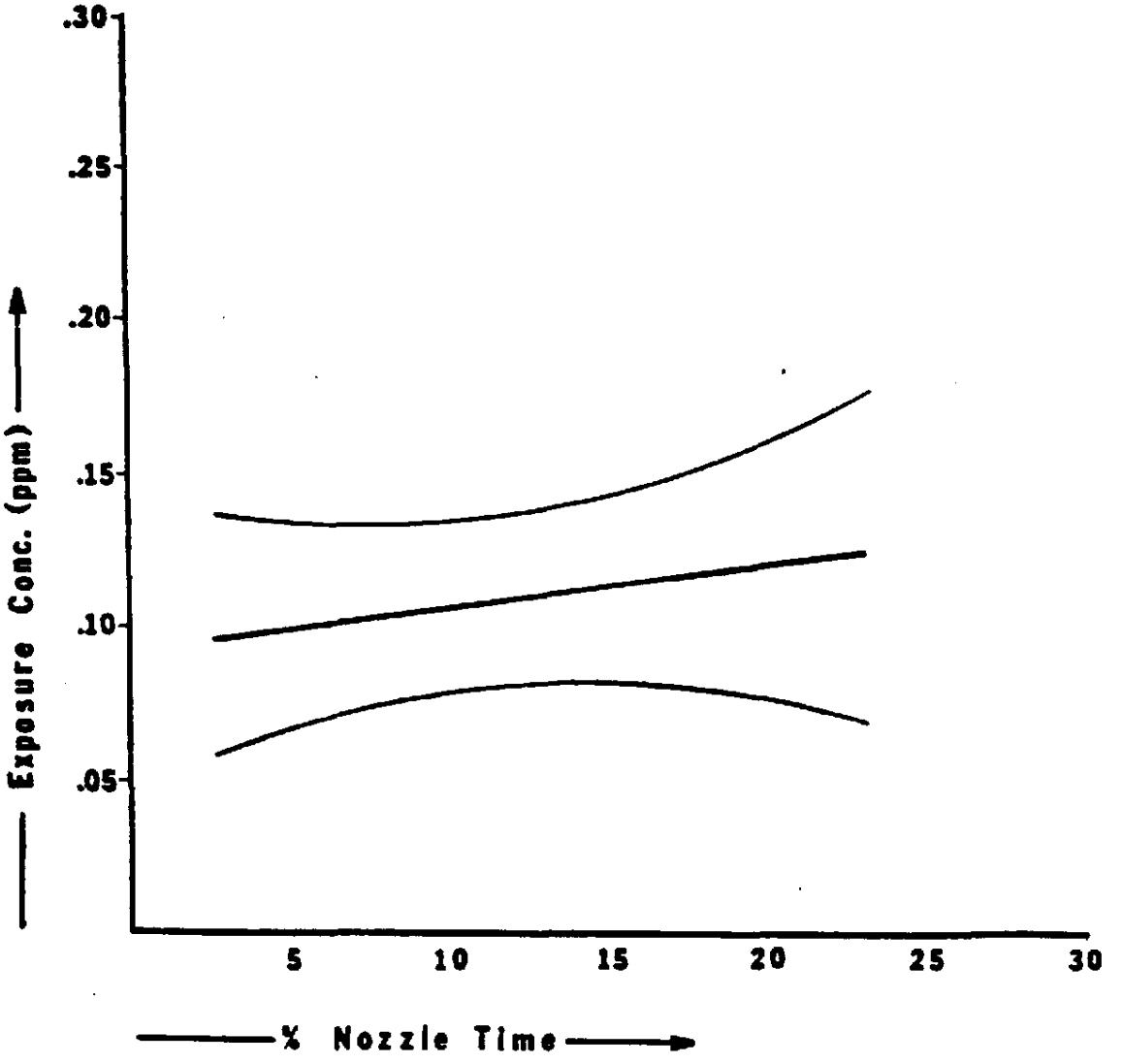


Figure 5

EXPOSURE CONCENTRATION
VS.
L.V.% BENZENE
(nonsuspect data)

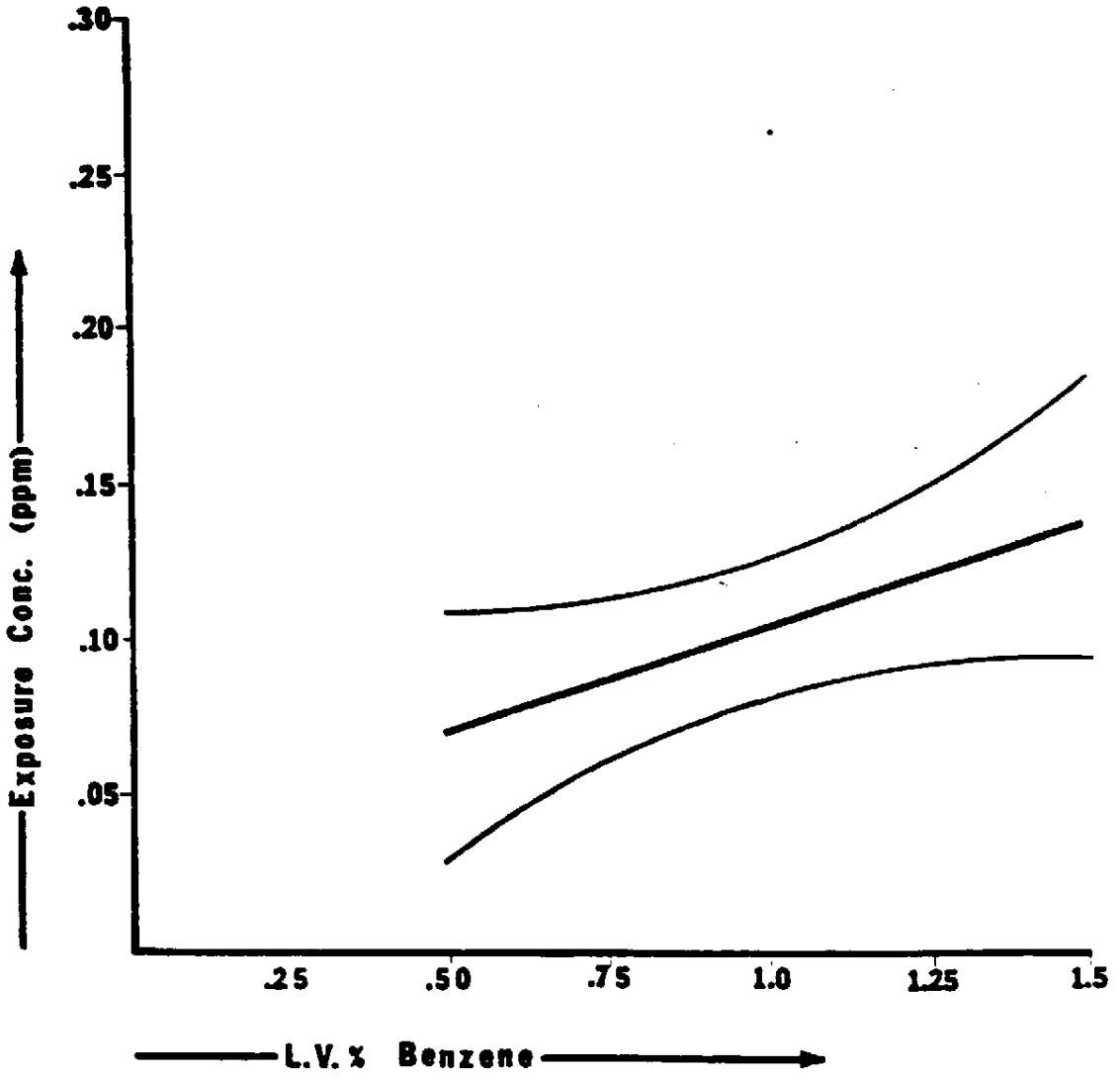


Figure 6

EXPOSURE CONCENTRATION
VS.
PERCENT NOZZLE TIME
(all data)

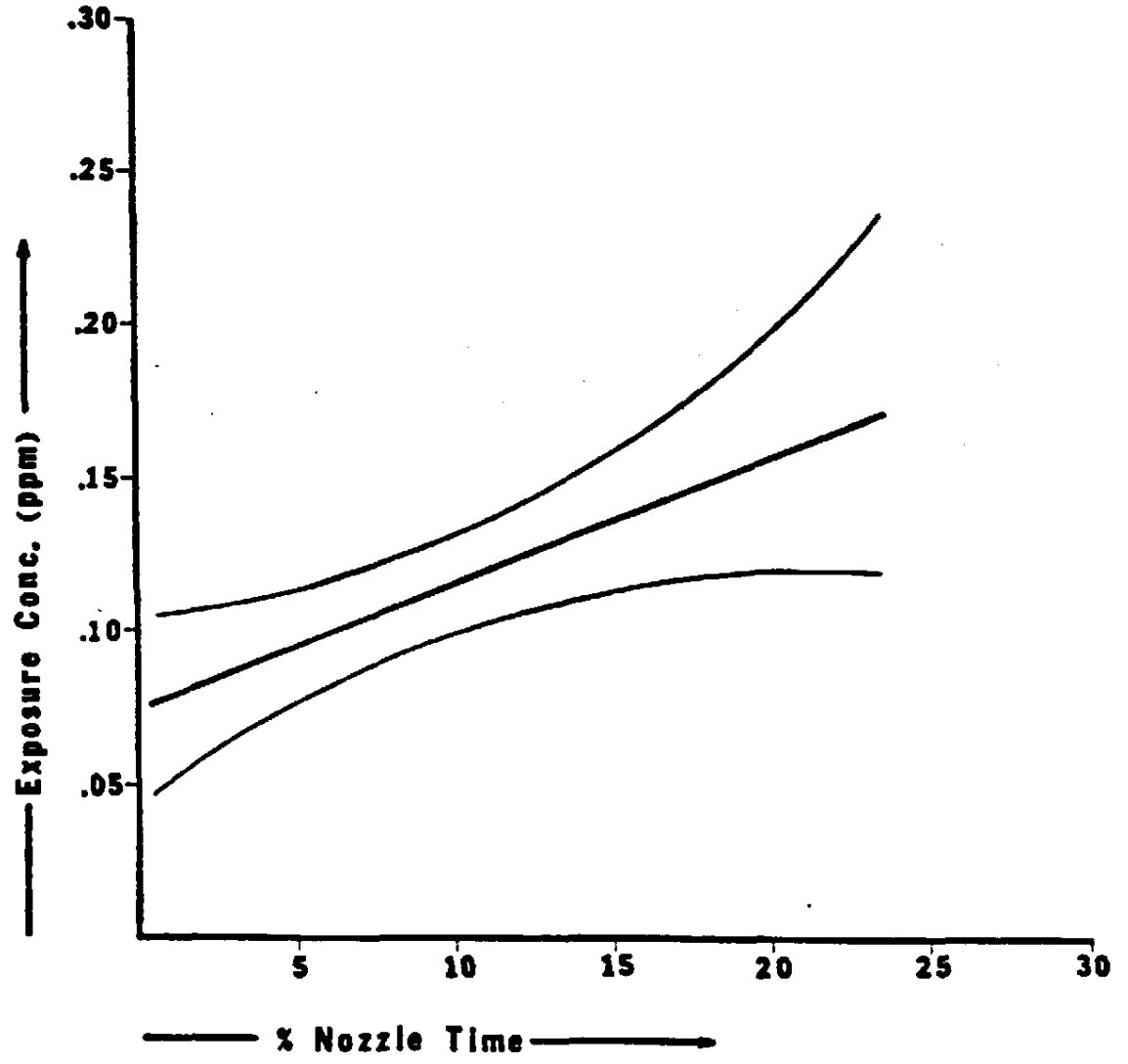
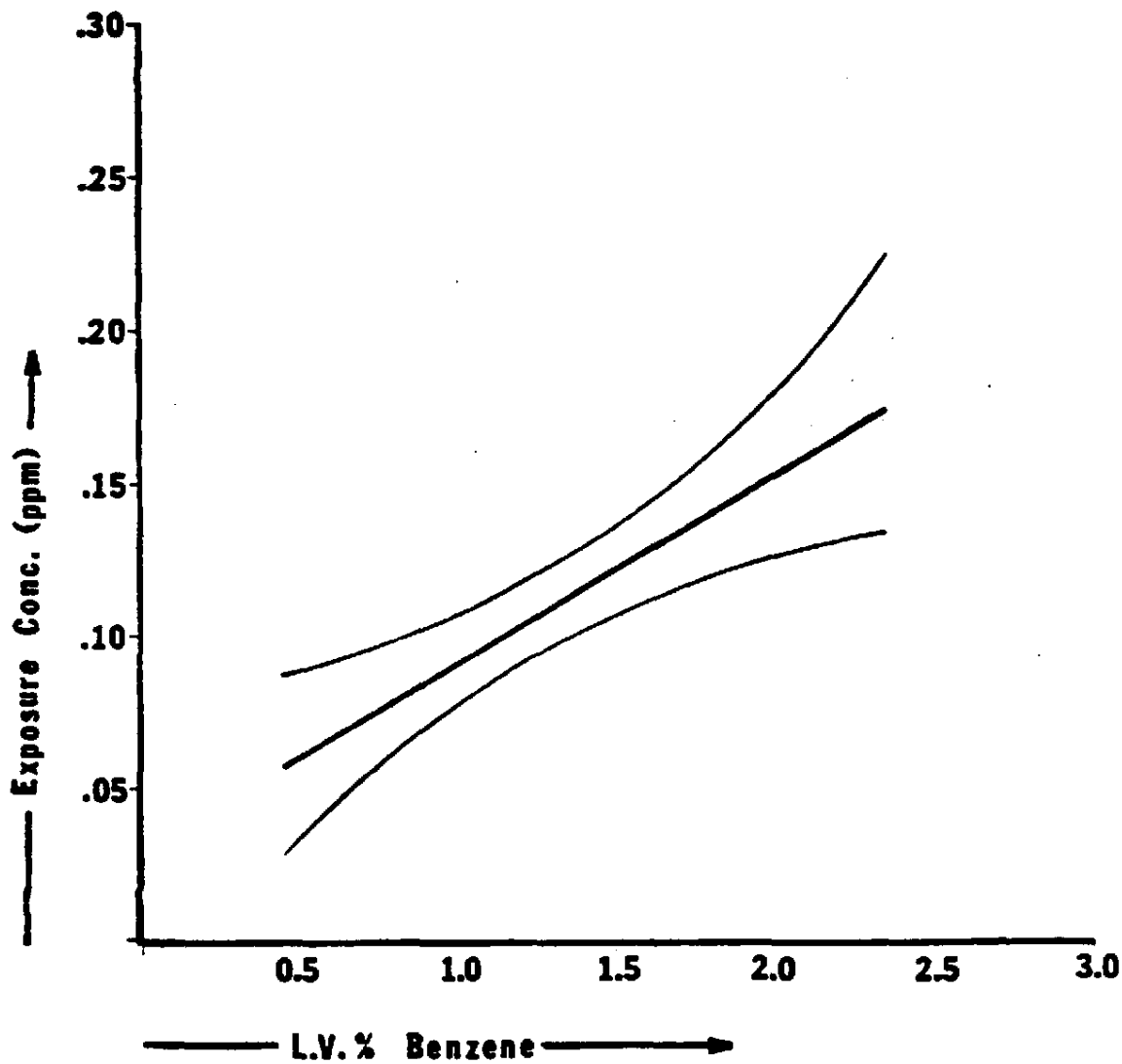


Figure 7

EXPOSURE CONCENTRATION
VS.
L.V.% BENZENE
(all data)



sets with available area data. The field of independent variables included liquid volume percent, temperature, windspeed, and relative humidity. The "modified" suspect data sets were included in the analysis, resulting in a total sample size of 32 cases. Windspeed, temperature, and liquid volume percent were all shown to add significantly to the regression equation (Figure 8). With regard to possible control measures, major interest was given to liquid volume percent. Figure 9 presents a graphic illustration of the effect of liquid volume percent on source exposures while controlling for temperature and windspeed. Again, variances used in the confidence interval formula are from unweighted data.

Vapor Recovery

A total of four days of environmental sampling was conducted at two service stations equipped with vacuum assist-type vapor recovery systems. A service station in the same vicinity without vapor recovery was also surveyed to aid in the initial evaluation. Sampling methodology at the vapor recovery stations was essentially the same as that for all other service stations, except flow rates were increased to 1000 cc/min. and sample periods were increased to approximately 240 minutes to account for the expected decrease in concentrations.

Configuration of the vapor recovery systems was essentially the same at both stations: the fuel nozzle was enshrouded with a contractable hose which created a partial seal upon insertion of the nozzle into an auto fuel tank. The contractable hose was connected to an in-line air pump, and with operation of the fuel pump, the air pump was activated and vapors were exhausted to the roof of the service station and burned.

Slight mechanical difficulties were encountered with one of the vapor recovery systems and, as described by the system's technician, a reduction of approximately 70% of the system's potential was realized during sampling. At both stations, however, a "hissing" noise was created as surrounding air was drawn through gaps around the retractable

Figure 8

Print-out Summary: Regression Analysis

of Estimated Source Exposure

Dependent Variable Estimate of Source Exposure

Variable Entered on Step Number 1 WS

Multiple R	0.354	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.126	Regression	1.	2.247	2.247	4.738
Adjusted R. Square	0.099	Residual	33.	15.651	0.474	
Standard Error	0.689					

Variables in the Equation

Variable	B	Beta	Std. Error B.	F
WS	-0.19563	-0.35433	0.08988	4.738
(Constant)	1.43713			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
GPH	-0.575	-0.477	0.602	9.440
L.V.I	0.591	0.500	0.626	10.656
Temp	0.092	0.090	0.817	0.261
RH	-0.177	-0.184	0.947	1.122

Variable Entered on Step Number 2 L.V.I

Multiple R.	0.587	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.344	Regression	2.	6.157	3.078	8.390
Adjusted R. Square	0.303	Residual	32.	11.741	0.367	
Standard Error	0.606					

Variables in the Equation

Variable	B	Beta	Std. Error B.	F
WS	-0.39515	-0.71569	0.09992	15.638
L.V.I	1.16325	0.59078	0.35635	10.656
(Constant)	0.12704			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
GPH	-0.354	-0.287	0.430	2.778
Temp	0.498	0.474	0.595	8.992
RH	0.089	0.092	0.704	0.267

Figure 8, cont.

Variable Entered on Step Number 3 Temp

Multiple R	0.701	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.491	Regression	3.	8.797	2.932	9.988
Adjusted R. Square	0.442	Residual	31.	9.101	0.294	
Standard Error	0.542					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
WS	-0.378	-0.684	0.090	17.793
L.V.X	1.747	0.888	0.374	21.878
Temp	0.029	0.498	0.001	8.992
(Constant)	-2.793			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
GPH	-0.211	-0.187	0.398	1.080
RH	-0.237	-0.233	0.488	1.716

Variable Entered on Step Number 4 RH

Multiple R	0.720	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.519	Regression	4.	9.289	2.322	8.093
Adjusted R. Square	0.455	Residual	30.	8.609	0.287	
Standard Error	0.536					

Variables in the Equation

Variable	B	Beta	Std. Error B	F
WS	-0.353	-0.640	0.091	15.187
L.V.X	1.624	0.825	0.381	18.157
Temp	0.0378	0.641	0.012	10.565
RH	-0.0133	-0.237	0.010	1.716
(Constant)	-2.4365			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F
GPH	-0.162	-0.143	0.379	0.609

Figure 8 cont.

Variable Entered on Step Number 5 CPH

Multiple R	0.727	Analysis of Variance	DF	Sum of Squares	Mean Square	F
R. Square	0.529	Regression	5.	9.466	1.893	6.511
Adjusted R. Square	0.448	Residual	29.	8.432	0.291	
Standard Error	0.539					

Variables in the Equation

Variable	B	Beta	Std. Error B.	F
WS	-0.385	-0.698	0.100	14.813
L.V.I	1.430	0.726	0.457	9.775
Temp	0.034	0.585	0.012	7.665
RH	-0.012	-0.206	0.010	1.215
CPH	-0.003	-0.161	0.004	0.609
(Constant)	-1.805			

Variables Not in the Equation

Variable	Beta In	Partial	Tolerance	F

Figure 8 cont.

Multiple Regression

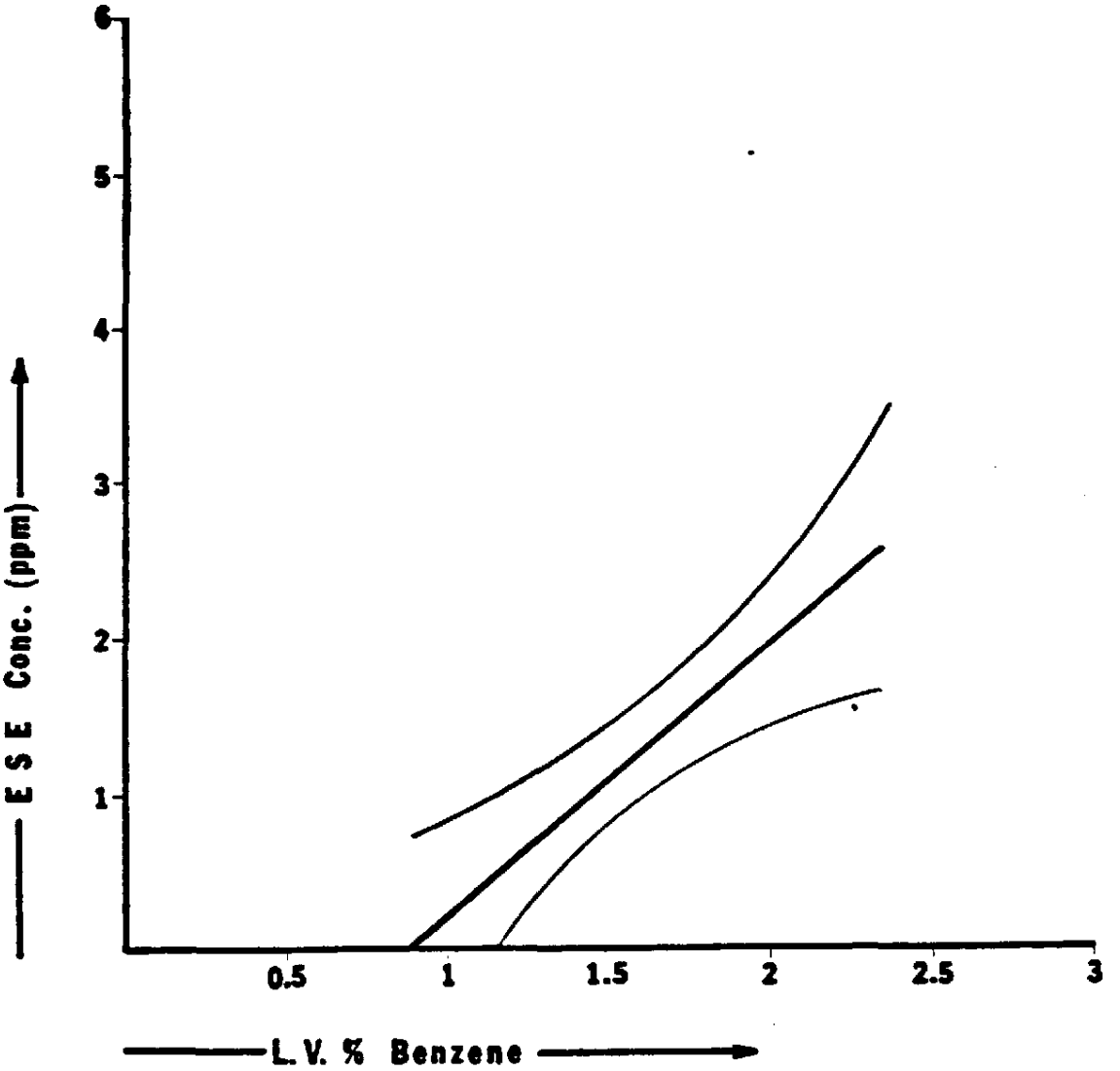
Dependent Variable Estimate of Source Exposure

Summary Table

Variable	Multiple R	R. Square	RSQ Change	Simple R	B	Beta
WB	0.354	0.126	0.126	-0.354	-0.385	-0.698
L.V.%	0.587	0.344	0.218	0.153	1.430	0.726
Temp	0.701	0.491	0.148	0.227	0.034	0.584
RH	0.720	0.519	0.028	-0.086	-0.012	-0.206
CPH	0.727	0.529	0.001	-0.123	-0.003	-0.161
(Constant)					-1.805	

Figure 9

ESTIMATED SOURCE EXPOSURE CONCENTRATIONS VS. L.V.% BENZENE



hose, indicating a high level of performance. Also, no odor of gasoline could be detected when very near (one inch) the nozzle at either station.

A preconceived aspect of the survey at this location was to estimate the "worst case" exposures at service stations equipped with vapor recovery. To facilitate this, automatic flow clips (devices which allow automatic flow and shut-off) were removed for one of the two days of sampling at each location, increasing percent nozzle time and thus theoretically increasing exposures.

While manipulating the percent nozzle time variable attempted to address the question of worst-case exposures, a comparative evaluation of vapor recovery would appear limited to only those stations involved in the manipulation. This rough comparison indicates an approximate four fold reduction in exposures. However, through use of the regression models previously developed, a more comprehensive comparison can be made involving all data sets previously discussed, while controlling for percent nozzle time. Toward this end, all vapor recovery data were combined and the mean variable responses were inserted first into the regression equation involving the nonsuspect data, and secondly into the model representing all the data.

The first procedure resulted in an expected ppm exposure value of 0.205 as compared to the actual average value of 0.046 ppm. The second fit, utilizing the regression equation obtained from the nonsuspect plus modified data sets resulted in an expected value of 0.213 ppm. As a general evaluation based on this data, it appears that vapor recovery reduces time-weighted exposures four to five fold.

A reduction of vapor recovery source exposures was also investigated via the regression model developed for the created ESE variable. The average liquid volume percent, temperature, and windspeed values of the vapor recovery data were fit to the equation (presented in Figure 8) resulting in an expected source exposure of 1.412 ppm, as compared to the actual average response of 0.116 ppm; a 12 fold reduction.

General Area Exposures

Area environmental air samples were collected at all survey locations except Nevada. Table 4 presents average area benzene concentrations for each of these sites which had available area data (Ohio-January area results reported below the analytical limit of detection). Although a number of variables suspected to affect personal exposures were monitored and averaged for each personal sample, only temperature and barometric pressure values were recorded and averaged for each area sample, for use in subsequent microgram to ppm conversions. However, as previously described, "personalized" (P-Area) values were calculated for each personal sample with available area data (Table 2). While the primary purpose of this adjustment was to facilitate ESE and ESEA variable generation, it also resulted in area values being roughly matched with all of the variables quantified for the personal samples, thus allowing a regression analysis. All data sets with available area data were used, including the modified suspect data. The non-structured regression mode was ulitized and liquid volume percent, temperature, windspeed, and gallons per hour were submitted as independant variables. Results of the analysis indicate that only temperature and windspeed significantly effect variations in P-Area values and surprisingly, temperature is negatively correlated.

CHAPTER V

Summary and Conclusions

As stated earlier, the primary goal of this research effort was to document and evaluate the occupational benzene exposures of service station attendants. As indicated in the attached tables, time-weighted benzene exposures range below one ppm, and average approximately 0.1 ppm.

Although these average exposure levels are below the final permanent benzene standard for non-gasoline work-place environments, regulatory decisions for controlling benzene exposures involving the distribution and sale of gasoline are forthcoming. Also to be considered are the peak exposures indicated by the ESE and ESEA values. If standards to reduce the average and/or peak exposures below the current levels are promulgated or if the oil companies initiate self-imposed guidelines for limiting exposures, control measures must be adopted.

This leads to the secondary objective of this research: those factors which contribute to variations in exposures. Through regression analysis techniques, it was shown that relative time near the source of exposure, benzene content of the gasoline, windspeed, temperature, and relative humidity all contribute to variations in time-weighted exposures, and that source exposures vary with benzene content, temperature, and windspeed. As evidenced in Figures 4 through 7 and Figure 9, emphasis was placed on those variables subject to modification. In a generalized fashion, inspection of these figures conveys the potential for reduced exposures resulting from modification of work practices or lowering the benzene content of gasoline.

The final consideration of this research was toward evaluating the effectiveness of vapor recovery. The data indicate a four to five fold reduction of time-weighted personal exposures and an approximate twelve-fold reduction of estimated source exposures from samples collected at stations equipped with vapor recovery systems. Although based on a limited number of samples, general area benzene concentra-

tions were not reduced at the two stations equipped with vapor recovery as compared to the area concentrations measured at the station in the same vicinity without vapor recovery (Table 4). This may be related to an unrecorded parameter; occurrences and extent of spills. This parameter would also account for the personal exposure levels found at the vapor recovery stations. Although these levels were substantially reduced, exposures still occurred even in the presence of apparently highly efficient vapor recovery systems.

If warranted, attempts at reducing occupational exposure to benzene at service stations should first be directed toward modification of work practices, including a reduction of the time in the vicinity of the nozzle, and concern with the occurrence of spills. Even in the presence of vapor recovery, work practices appear to offset exposures. In comparison to the estimated costs of alternate forms of exposure reduction (ie., initial cost of benzene removal from gasoline @ 5.3 billion) modifications of work practices appear quite attractive.

TABLE 1

Catagories By Site and Time of Collection

	# Visited	# Samples	# TWA'S
Ohio-July	3	5*	5
Ohio-January	5	22	12
Florida-February	11	47	19
Ohio-February, March	6	20	9
Nevada-June	1	17	7
California-August (no vapor recovery)	1	14	4
California-August (vapor recovery)	2	16	8
TOTALS	28	141	64

*6 to 8 Hr. samples

TABLE 2

DESCRIPTIVE STATISTICS
Ohio-July
(N=5)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA EXP. (ppm)	0.206	0.28	0.161	0.237	-1.15	2.56
LIQUID VOL. % BENZENE	1.02	.13	0.93	1.24	2.11	4.59
PERCENT NOZZLE TIME	15.1	5.08	8.3	20.5	-.31	-1.56
GALLONS PER HOUR	90.0	29.1	64.4	137.7	1.42	1.83
SAMPLE TIME (MIN)	431.8	25.9	408.0	470.0	-.56	.85
TEMP (^{o}F)	80.4	2.2	78.0	82.0	-.61	-3.33
WIND SPEED (MPH)	0.3	0.1	0.3	0.5	2.24	5.00
REL. HMDTY	69.6	1.67	68.0	72.0	.51	-.61
P-AREA (ppm)	0.044	.013	0.030	0.055	-.59	-3.30
ESE (ppm)	1.185	.567	0.885	2.194	2.19	4.82
ESEA (ppm)	1.229	.574	0.924	2.249	2.18	4.78

TABLE 2
(cont)

DESCRIPTIVE STATISTICS
Ohio- January
(N=5)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA EXP. (ppm)	0.064	.032	0.028	0.124	.81	-.29
LIQUID VOL.% BENZENE	0.68	.19	0.49	1.03	.69	-.87
PERCENT NOZZLE TIME	9.5	4.6	4.2	21.7	1.84	4.42
GALLONS PER HOUR	63.7	47.0	25.7	180.3	1.56	2.40
SAMPLE TIME (MIN)	261.8	143.2	75.0	440.0	.13	-1.87
TEMP (^{o}F)	22.4	9.0	11.0	33.0	-.06	-1.87
WIND SPEED (MPH)	2.0	1.63	0.0	5.3	.83	.03
REL. HMDITY	61.2	11.9	45.0	76.0	-.13	-1.79
P-AREA (ppm)	----	----	----	----	----	----
ESE (ppm)	----	----	----	----	----	----
ESEA (ppm)	----	----	----	----	----	----

TABLE 2
(cont)

DESCRIPTIVE STATISTICS
Florida-February*
(N=19)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA EXP. (ppm)	0.131	.064	0.058	0.321	1.36	3.08
LIQUID VOL.% BENZENE	1.80	.30	1.32	2.29	-.05	-.95
PERCENT NOZZLE TIME	5.0	2.8	0.4	9.5	-.13	-1.06
GALLONS PER HOUR	20.8	9.3	3.2	46.2	.74	2.34
SAMPLE TIME (min)	349.4	92.2	147.0	463.0	-.51	-.80
TEMP (^{o}F)	63.0	6.6	53	73	-.13	-1.35
WIND SPEED (mph)	2.2	1.3	0.3	4.3	.18	-1.23
REL. HMDITY	49.1	12.6	26.8	76.1	.16	-.08
P-AREA (ppm)	0.052	.025	0.018	0.096	.26	-.81
ESE (ppm)	2.101	1.513	0.427	6.533	1.47	2.79
ESEA (ppm)	2.153	1.517	0.523	6.593	1.47	2.78

*Lower portion of suspect data X4

TABLE 2
(cont)

DESCRIPTIVE STATISTICS
Ohio-February, March
(N=9)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA EXP. (ppm)	0.097	.030	0.067	0.163	1.51	2.30
LIQUID VOL.% BENZENE	1.64	.18	1.43	1.93	.53	-.99
PERCENT NOZZLE TIME	7.6	2.7	4.3	12.3	.56	-.31
GALLONS PER HOUR	49.5	35.2	18.0	129.0	1.63	2.80
SAMPLE TIME (min)	323.0	149.0	54.0	508.0	-.70	-.30
TEMP (°F)	50.9	8.7	38.0	60.0	-.45	-1.67
WIND SPEED (mph)	1.9	1.2	0.4	3.5	.31	-1.52
REL. HMDITY	43.0	8.8	34.0	56.6	.60	-1.38
P-AREA (ppm)	0.050	.008	0.040	0.063	.39	-1.05
ESE (ppm)	0.700	0.499	0.123	1.516	0.89	-0.14
ESEA (ppm)	0.751	0.494	0.177	1.560	0.89	-0.16

TABLE 2
(cont)

DESCRIPTIVE STATISTICS
Nevada-June
(N=7)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA EXP. (ppm)	0.042	.014	0.027	0.068	1.30	2.14
LIQUID VOL.% BENZENE	1.07	.05	1.03	1.17	.2.02	4.30
PERCENT NOZZLE TIME	5.3	2.8	2.5	9.0	.32	-2.30
GALLONS PER HOUR	42.8	29.9	13.4	99.3	1.14	1.39
SAMPLE TIME (min)	307.0	153.5	120.0	490.0	.09	-2.08
TEMP (^{o}F)	92.4	4.1	88.0	99.0	.41	-.91
WIND SPEED (mph)	2.2	1.5	0.8	4.1	.41	-2.52
REL. HMDITY	26.2	5.9	18.0	33.2	-.10	-2.00
P-AREA	----	----	----	----	----	----
ESE	----	----	----	----	----	----
ESEA	----	----	----	----	----	----

TABLE 2
(cont)

DESCRIPTIVE STATISTICS
California-August (No Vapor Recovery)
(N=4)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA EXP. (ppm)	0.193	.024	0.161	0.220	-.54	1.66
LIQUID VOL. % BENZENE	1.47	.005	1.46	1.47	-.03	-.548
PERCENT NOZZLE TIME	19.1	4.1	14.2	23.3	-.32	-2.3
GALLONS PER HOUR	31.0	9.9	21.7	41.4	.10	-5.19
SAMPLE TIME (min)	442.3	15.0	426.0	456.0	-.17	-4.86
TEMP (^{o}F)	75.6	0.9	75.0	76.0	.00	-5.99
WIND SPEED (mph)	2.9	0.3	2.7	3.1	.00	-5.99
REL. HMDITY	72.0	2.2	70.1	74.1	.02	-5.85
P-AREA (ppm)	0.020	0.001	0.019	0.021	.00	-6.00
ESE (ppm)	0.928	0.183	0.746	1.151	.43	02.41
ESEA (ppm)	0.948	0.183	0.766	1.170	.0431	-2.385

TABLE 2
(cont)

DESCRIPTIVE STATISTICS
California-August (Vapor Recovery)
(N=8)

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA Exp. (ppm)	0.046	.014	0.028	0.071	.37	-.08
LIQUID VOL. % BENZENE	1.56	.09	1.46	1.70	1.01	-.46
PERCENT NOZZLE TIME	21.8	7.6	12.2	34.1	.31	-.81
GALLONS PER HOUR	47.0	8.2	33.0	56.0	-.47	-.87
SAMPLE TIME (min)	449.9	22.7	403.0	474.0	-1.27	2.00
TEMP (^{o}F)	77.6	2.2	75.0	81.0	.18	-1.56
WIND SPEED (mph)	2.1	.6	1.4	2.6	-.32	-1.61
REL. HMDITY	64.5	3.2	59.4	67.5	-1.10	-.31
P-AREA (ppm)	0.020	.003	0.018	0.024	1.33	-.17
ESE (ppm)	0.116	.053	0.041	0.199	.22	-.89
ESEA (ppm)	0.136	.51	0.065	0.218	.28	-.89

TABLE 3

COMPOSITE DESCRIPTIVE STATISTICS

	MEAN	S.D.	MIN.	MAX.	SKEW.	KURT.
TWA* EXP. (ppm)	0.111	0.066	0.027	0.321	.83	.32
LIQUID* VOL. % BENZENE	1.31	0.50	0.49	2.29	.05	-.90
PERCENT● NOZZLE TIME	8.3	5.4	0.4	23.4	1.17	1.04
GALLONS* PER HOUR	44.3	36.0	3.17	180.3	1.76	3.26
SAMPLE* TIME (min)	335.1	125.7	54.0	508.0	-.68	-.81
TEMP* (^{o}F)	58.5	23.6	11.0	99.0	-.44	-.58
WIND* SPEED (mph)	2.0	1.4	0.0	5.3	.40	-.94
REL.* HMDITY	51.3	16.5	18.0	76.1	-.15	-1.08
P-AREA** (ppm)	0.047	0.021	0.018	0.096	.47	0.06
ESE** (ppm)	1.510	1.280	0.123	6.533	2.08	5.59
ESEA** (ppm)	1.557	1.285	0.177	6.593	2.08	5.56

● auspect data X4 (N=56)
** from non-vapor recovery data with available area data (N=37)

TABLE 4

AVERAGE AREA BENZENE CONCENTRATIONS

Site and Time of Collection	#Samples	Ave. Concentration (ppm)
Ohio-July	12	0.049
Ohio-January	36	———
Florida-February	64	0.060
Ohio-February, March	33	0.053
Nevada-June	0	———
California-August (no V.R.)	8	0.020
California-August (vapor-recovery)	8	0.020

Weighted Average of Area Concentrations: 0.052 ppm

REFERENCES

1. Code of Federal Regulations, Title 29: Part 1910. Occupational Exposure to Benzene, February 10, 1978.

2. Code of Federal Regulation, Title 29: Part 1910. Emergency Temporary Standard; Occupational Exposure to Benzene, Notice of Hearings, May 3, 1977.

3. Letter to William J. Usery, Secretary of Labor, U.S. Dept. of Labor, from Peter Bommarito, President of United Rubber Workers, concerning; Emergency Temporary Standard on Benzene, April 23, 1976.

4. Health Effects of Benzene: A Review. National Academy of Sciences, 1976.

5. Memorandum from Director, NIOSH, to Assistant Secretary of Labor, OSHA, concerning: Recommendation for an Emergency Temporary Standard for Benzene, October, 1976.

6. Parkinson, G.S., Benzene in Motor Gasoline - An Investigation into Possible Health Hazards in and Around Filling Stations and in Normal Transport Operations. Ann Occup Hyg 14:145-153, 1971.

7. McDermott, H.J., Service Station Attendants' Exposure to Benzene and Gasoline Vapors. AIHAJ. 40: No. 4, 315-321, 1979.

8. Runion, H.E., Benzene in Gasoline II. Presented at the International Workshop on Benzene, Universite Rene Descarte, November 9, 1976.

9. Berlin, M., Fredya, K., Gage, J.C., Lagesson, V., Reitalu, J., and Tunek, A., Exposure to Benzene During the Handling of Motor Fuels. Report from the University of Lund, Lund, Sweden, 1976.

10. Letter to Richard Johnson, Office of Air Quality Planning and Standards, EPA, from C. W. Tounly, Concerning: Results of Self-Service Exposure Samples, May, 1977.

11. Organic Solvents in Air. NIOSH Manual of Analytical Methods, DHEW(NIOSH) Publication No. 77-157-A, 1977.

12. Statistical Package for the Social Sciences. 2nd Edition, pp. 320-360, 1975. McGraw-Hill, Inc., Publishers.

QUANTITATIVE RE-EVALUATION OF THE HUMAN LEUKEMIA RISK ASSOCIATED WITH INHALATION EXPOSURE TO BENZENE

5/24/88 FYI

PHASE 1 REPORT

TO: BFG,TACS,PROP. 65 STEERING COM.

FROM: MIKE WANG

Supported by:

American Petroleum Institute
1220 L Street, N.W.
Suite 900
Washington, D.C. 20005

Prepared by:

Clement Associates, Inc.
9300 Lee Highway
Fairfax, VA 22031

Principle Author: Todd W. Thorslund, Ph.D.

Clement Associates, Incorporatec

April 18, 1988

Health and Environmental Science

A Subsidiary of ICF Technology Inc

QUANTITATIVE RE-EVALUATION OF THE HUMAN LEUKEMIA RISK ASSOCIATED WITH INHALATION EXPOSURE TO BENZENE

PHASE 1 REPORT

Supported by:

American Petroleum Institute
1220 L Street, N.W.
Suite 900
Washington, D.C. 20005

Prepared by:

Clement Associates, Inc.
9300 Lee Highway
Fairfax, VA 22031

Principle Author: Todd W. Thorslund, Ph.D.

April 18, 1988

PREFACE

This document represents the first of a two phased effort to re-evaluate the carcinogenic potency of benzene. Since the last unit risk estimate of cancer potency completed by CAG in 1985, there have been significant advances made in the development of risk assessment methodology, and important new information has become available concerning the effects of benzene on humans and experimental animals.

This phase I report adopts the same theoretical structure used in the EPA (1985) risk models to obtain an upper bound cancer risk estimate but refines the new model by the use of new data and methodological developments. Included in the new estimates are an update of the original epidemiological study by Infante et al. on pliofilm workers, new epidemiological studies conducted in China, and the use of more precise statistical estimation procedures made possible by the availability of exposure and vital status information on individual cohort members. It is important to note that the model developed in this study does not depart substantially from the basic structure of the original CAG model. It incorporates the same one-hit one stage model that was the basis for EPA's (1985) risk assessments. The new estimates reflect the agency's preference for using human data when available, and are based upon parameter estimation techniques which make the most efficient use of the available information.

An observation that demonstrates the consistency of one of the more important underlying assumptions of the model, is that for both the atomic bomb survivors in Japan and workers exposed to benzene in China, the peak leukemia risk for benzene occurs at 5 to 10 years after

initiating the final biological transformation. Thus, in the worst case, the risk of devaloping leukemia decreases steadily for those people who have had no benzene exposure for more than 10 years. This observation is being confirmed in the pliofilm workers cohort where we see no new cases of leukemia in the additional years of follow-up (1979-1981) included in this report.

The changes made in the model have a pronounced effect upon the unit risk estimate. The upper bound unit risk estimate obtained in this report is 2.7×10^{-3} which is almost an order of magnitude lower than the value of 2.6×10^{-2} that was adopted by EPA in 1985. Another advantage of the new model is that it provides logical way of testing the linear dose-response hypothesis. It was demonstrated that the dose-response relationship exhibited more curvature than could be explained by a simple linear relationship. This demonstration of a curvature, and the probable underestimation of exposure both strongly suggest that the upper bound unit risk estimate is very conservative.

The difference in the cancer potency estimates obtained with the current model compared with the previous approach is primarily due to the type of information used to estimate the model parameters. The present approach uses the vital status and exposure estimates at each point in time for each individual in the cohort as well as information on latency period from other studies. The previous approach grouped individuals into exposure ranges and used the average exposure, observed and expected number of leukemias, and person years in the exposure group in the analysis. The data used in the current study provides much more precise information for the estimation of the

parameters in the dose-response model.

Phase II of this program will involve the use of additional biological and exposure data and a more refined biologically motivated model that takes into consideration benzene's potential mechanisms of action. Since most of benzene's hypothesized mechanisms of action suggest non-linear dose effects, we feel confident the predictions of leukemia risk will be lower than the present upper bound risk estimates.

TABLE OF CONTENTS

I. INTRODUCTION

A complete analysis of the carcinogenic potency of benzene was performed by the U.S. Environmental Protection Agency almost ten years ago. Since this analysis, considerable new and relevant information on benzene has become available and major advances in risk assessment methodology have been achieved. However, a comprehensive reassessment of the carcinogenic potency of benzene has not as yet been undertaken.

In 1985, EPA performed an interim quantitative cancer risk assessment of benzene (EPA 1985). This interim assessment was done in response to a petition and, thus, had severe time constraints that forced the use of a secondary source (Crump and Allen 1984) for deriving the cancer potency estimates which did not permit optimal use of all the available information.

The purpose of this report, which shall be denoted as Phase I or first submittal, is to supply EPA with a new state-of-the-art upper bound estimate of benzene's cancer potency and to inform EPA of the more biologically relevant risk model that is under development which will be referred to as Phase II or second submittal which is scheduled for completion in mid June. The advice and suggestions of agency personnel are sought herefor how best to incorporate scientific and regulatory considerations into the new biologically relevant risk model to be developed in Phase II. In the Phase I report, the approaches that were taken to modify and complete EPA's interim cancer potency estimates are described and the approach to be taken in the development of a new more biologically relevant model in Phase II outlined.

The upper bound cancer dose-response model for benzene that is developed in the Phase I report includes the following extensions and improvements of that used in EPA's 1985 report:

(1) Inclusion of additional years of follow-up (1979-1981) from the Rinsky et al. (1981) cohort;

(2) Restricting the analysis to the most relevant data set and functional form of the potential models;

(3) Use of updated U.S. vital statistics to estimate rate of benzene-induced leukemia in the presence of competing mortality;

(4) Corrections and additions of job classification codes that were found on the Rinsky et al. (1981) computer tape to reconcile them with original work records;

(5) Use of actual data from several leukemogens in addition to benzene to define a probability distribution for the time from malignant transformation of the first cancer cell to death due to leukemia; and

(6) Use of more efficient techniques for cancer potency parameter estimation, which is made possible by the availability of information on exposure and vital status for every individual in the Rinsky et al. (1981) cohort.

The new best estimate biologically relevant model that is under development in Phase II will incorporate diverse biomedical information into a biologically based quantitative dose-response model. For example, information from observations of the leukemogenic effects of radiation (atomic bombs, therapeutic), chemotherapeutic agents, and experimental benzene studies will be used to estimate the biological latency period. In addition, the shape of the age-incidence curve for granulocytic leukemia for the general U.S. population

will be used to estimate other model parameters. The approach taken is to derive risk estimates for benzene that are based upon various forms of a two-stage model with clonal expansion of preneoplastic or 1st stage cells (Thorslund et al. 1987).

The two-stage biologically based model can be used to estimate risk from either human or animal data, which is an improvement over the present inconsistent EPA approach of using different models for the two types of data. The new approach has the further advantage of proving a means to incorporate theories of biological mechanisms into the dose response model. For example, we can incorporate benzene's clastogenic effects as well as its effect on cell proliferation into the model.

The following sections of this report provide a more detailed discussion of the approaches that are used here as well as those that are under development.

II. ANALYSIS OF EPA'S 1985 MATHEMATICAL MODELING APPROACHES USED FOR BENZENE

A. General Mathematical Framework

The models that were used in the EPA's cancer risk assessments for benzene (EPA 1985) are based on the implicit assumption that one molecule of benzene can cause the single necessary event that transforms a normal stem or progenitor cell in the bone marrow into a neoplastic cell (i.e., one-stage, one-hit models). The latency period or time between this cellular transformation and death due to leukemia is a variable that is affected by a number of factors such as immunological competence and the stage of hematopoietic maturation of the transformed cell. Thus the time-dependent dose-response model for benzene and leukemia is dependent upon two factors:

1. the dose-response relationship between benzene exposure and the cellular transformation; and
2. the probability distribution of the length of survival (i.e., time between the cellular transformation and death due to leukemia).

This joint biological process is mathematically modeled in a general form in the next section. The manner in which CAG's previous cancer risk models for benzene-induced leukemia fit into this general framework is then described.

1. Development of General Model

In this section, the general EPA cancer dose-response model for benzene is derived mathematically and biologically-based rationales for the model components are proposed.

The relative instantaneous probability that exposure to benzene at time v resulted in death due to leukemia at time t can be expressed as

$$h(v,t) = G[x(v)]w(t-v) \qquad \text{(II-1}$$

where G[x(v)] is the instantaneous dose-dependent probability that exposure to a specified level of benzene at time v, x(v), will induce the cellular transformation (e.g., chromosomal abnormality) that will lead to leukemia some time in the future; t-v is the length of time between the cellular transformation and death due to leukemia, which is considered to be a random variable with a definable probability distribution w(t-v); and w(t-v) is the probability density or weighting function that defines the relative probabilities of death due to leukemia at time t given that the cellular transformation occurred at time v (i.e., the probability density function of t-v).

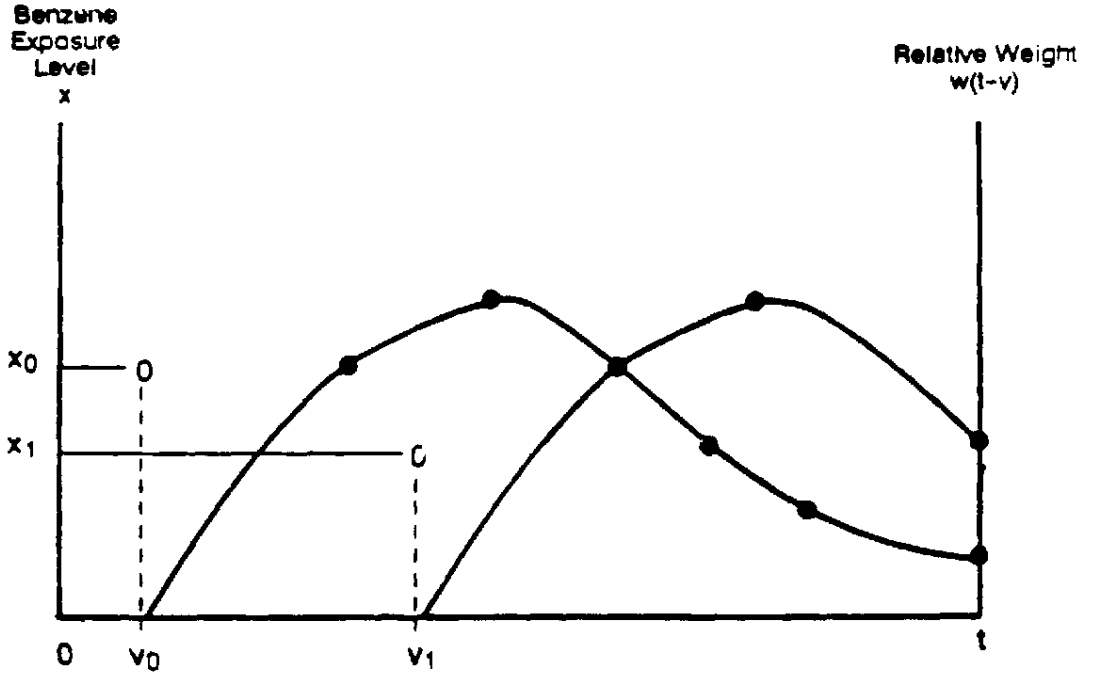
The relative probability that a series of exposures to benzene prior to time t will result in death due to leukemia at time t is simply the sum of the effect of the exposures at all individual points in time. Thus, if an individual were subjected to a series of r + 1 exposures ($x_0, x_1, \ldots, x_r$) occurring at times ($v_0, v_1, \ldots, v_r$), the instantaneous probability of death due to leukemia at time t is the hazard function or age-specific cancer rate:

$$h(t) = \sum_{j=0}^{r} h(v_j,t) = \sum_{j=0}^{r} G(x_j)w(t-v_j) \qquad \text{(II-1)}$$

The hazard function is illustrated in Figure II-1 for the simple case where the exposure level was x_0 at time v_0 and x_1 at time v_1, the weighting function has the simple unimodal form $w(t-v) = K^2(t-v)\exp-K(t-v)$, and the cellular transformation function is assumed to be $G(x) = \beta x^2$, which would be the case if the transformation requires two simultaneous one-hit events. An example of a transformation that requires two simultaneous events is reciprocal

FIGURE II-1

Illustration of Hazard Function for Leukemia at Time t Following Two Exposures to Benzene at Points In Time v_0 and v_1



$$h(t) = \beta x_0^2 w(t-v_0) + \beta x_1^2 w(t-v_1)$$

$$G[x(v)] = \beta x(v)^2$$

$$w(t-v) = K^2(t-v)\exp -K(t-v)$$

$$h(t) = \beta K^2[x_0^2(t-v_0)\exp -K(t-v_0) + x_1^2(t-v_1)\exp -K(t-v_1)]$$

translocation of chromosomes, which requires the breakage of two different chromosomes during a single phase of the cell cycle. For example, 95% of adults with chronic myelogenous leukemia exhibit the presence of a reciprocal translocation involving chromosomes 9 and 22 or a variant translocation of chromosome 22 (the Philadelphia chromosome) (Sandberg 1980, Mitelman and Levan 1981).

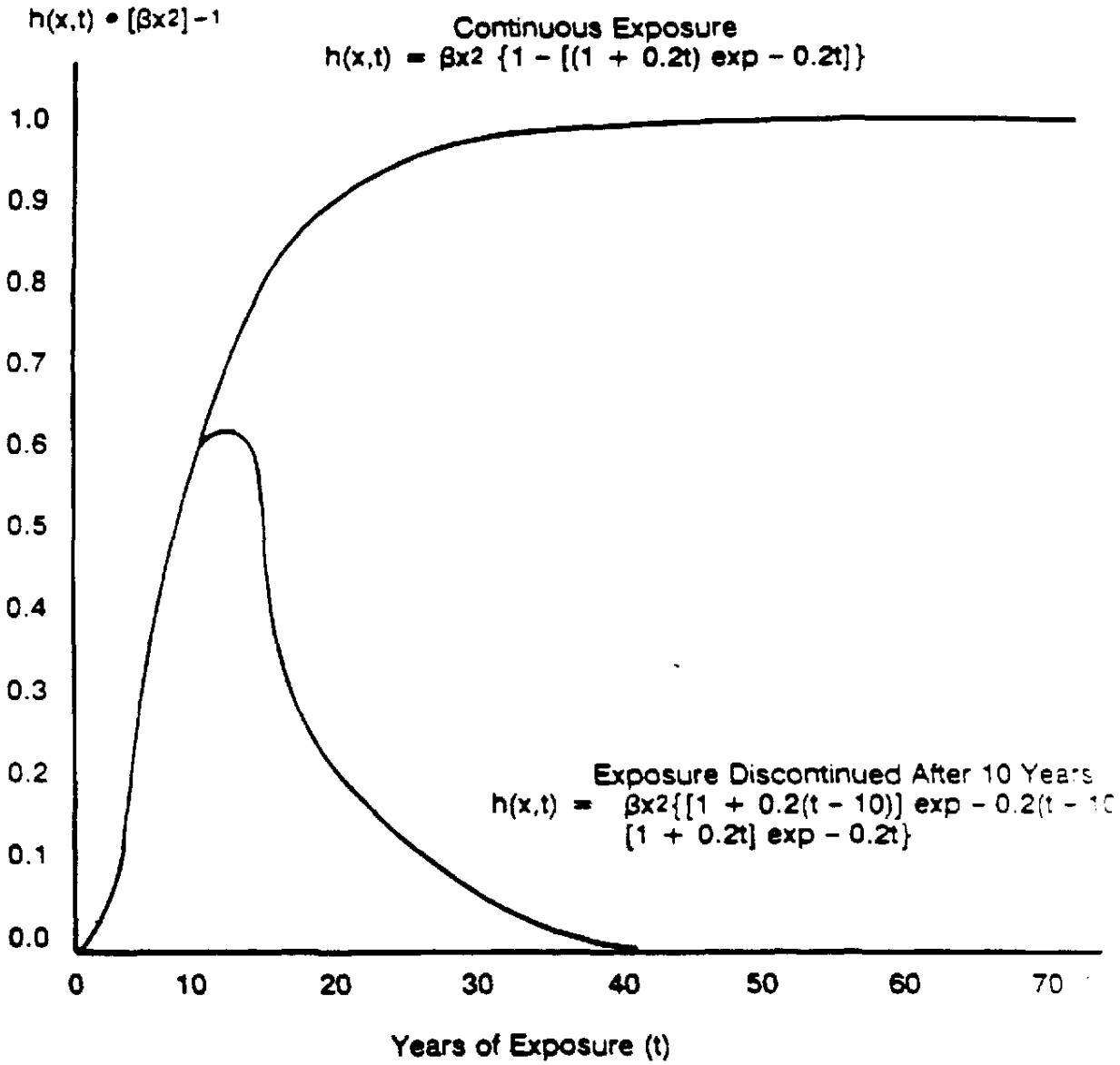
When exposure is continuous over time, the hazard function has the same functional form as equation (II-2) except that the summation sign is replaced by integrals, yielding the relationship:

$$h(t) = \int_0^t h[x(v),v] = \int_0^t G[x(v)]w(t-v)dv, \qquad \text{(II-3}$$

which may be viewed as weighting the functional effect of G[x(v)] by w(t-v) and adding the weighted effects from all exposures prior to t to obtain the total effect. The age-specific leukemia death rates due to continuous lifetime or continuous ten-year exposures, assuming the previously specified forms of G • and w(•), are derived from equation (II-3) and shown in Figure II-2, where a modal time (t-v) is assumed to be 5 years. Under this model, the age-specific leukemia death rate reaches a plateau under conditions of continuous exposure but rapidly declines upon cessation of exposure. Using a weighting function for latency to obtain a composite hazard function is discussed in more general terms by Whittemore and Keller (1978). Selecting functional forms for G(•) and w(•) should be guided to the greatest extent possible by what is known about the biological processes involved in benzene-induced leukemogenesis. In the next section, the choices that were made in EPA's benzene risk assessment (EPA 1985) are reviewed.

FIGURE II-2

Relationship between Age-Specific Leukemia Rate and Duration of Benzene Exposure [a]



[a] Assuming K = 0.2 (i.e., modal latency is 5 years)

2. Functional Forms for G(•) and w(•) Used by EPA

A total of six risk model forms were used by EPA to estimate benzene-induced leukemia risk based on all combinations of two forms of G(•) and three forms of w(•) (EPA 1985). The models were derived by transforming the basic models for an occupational exposure scenario developed by Crump and Allen (1984) for OSHA and subsequently reported in Crump et al. (1987) into models appropriate for environmental exposure. The dose-effect relationship for the cellular event used by EPA may be represented by either of the following (EPA 1985):

$$G[x(v)] = \begin{cases} \beta x(v) & \text{one-hit absolute risk model} \\ \beta x(v)\alpha(t) & \text{one-hit relative risk model} \end{cases}$$

where $\alpha(t)$ is assumed to be the background age-dependent leukemia rate at age t. The three functions specified for $w(t-v)$ by EPA (EPA 1985) may be expressed as step functions of the forms:

$$\begin{cases} 0 & t-v < 5 \\ 1 & 5 \le t-v \end{cases} \quad \text{cumulative dose}$$

$$w(t-v) = \begin{cases} 0 & t-v < 2.5 \\ 1 & 2.5 \le t-v < 7.5 \\ 1/3 & 7.5 \le t-v < 12.5 \\ 1/6 & 12.5 \le t-v \end{cases} \quad \text{weighted cumulative dose}$$

$$\begin{cases} 0 & t-v < 2.5 \\ 1 & 2.5 \le t-v \le 12.5 \\ 0 & 12.5 < t-v \end{cases} \quad \text{window dose}$$

The cumulative dose hypothesis is equivalent to a two-stage model (i.e., k=2 in multistage) in which the first stage is assumed to be exposure-dependent with a five year lag-time. The weighted cumulative dose is consistent in some vague unspecified sense with the latency pattern of leukemia in Japanese atomic bomb survivors and the window dose has no theoretical or intuitive rationale. Combining these three weighting functions with the two dose-response models for the final cellular transformation yields a total of six potential models. The biological assumptions supporting these models are evaluated in the next section.

B. Analysis of Underlying Assumptions

1. Linearity

A critical assumption made by all the dose-response models for benzene is that the probability of the cellular transformation is linearly related to dose. The simple one-hit (linearity) assumption for benzene is based on the compound assumption that the reactive benzene metabolite level at the site of action (i.e., bone marrow) is proportional to the level of benzene exposure in the air and that only one molecule of the benzene metabolite is required for the critical reaction that elicits the cellular transformation. No direct evidence for these assumptions has been presented for benzene so that their use in risk assessment models is strongly based on the rationale of health-conservatism. For the upper bound model developed in Phase I the assumption of linearity is retained; however, in the Phase II report, the linearity assumption will be critically evaluated by:

- obtaining estimates of the relationship between the level of benzene in the air and the dose of benzene's principle reactive metabolites predicted to be delivered to the bone marrow using pharmacokinetic models assuming that a parallel funded project is available by mid June;

- defining the dose-response relationships between benzene and events that are likely to be responsible for the cellular transformation such as chromosomal aberrations; and

- making a comparison of the log-likelihoods (i.e., goodness of fit) of linear and non-linear models using the epidemiological data on benzene and leukemia.

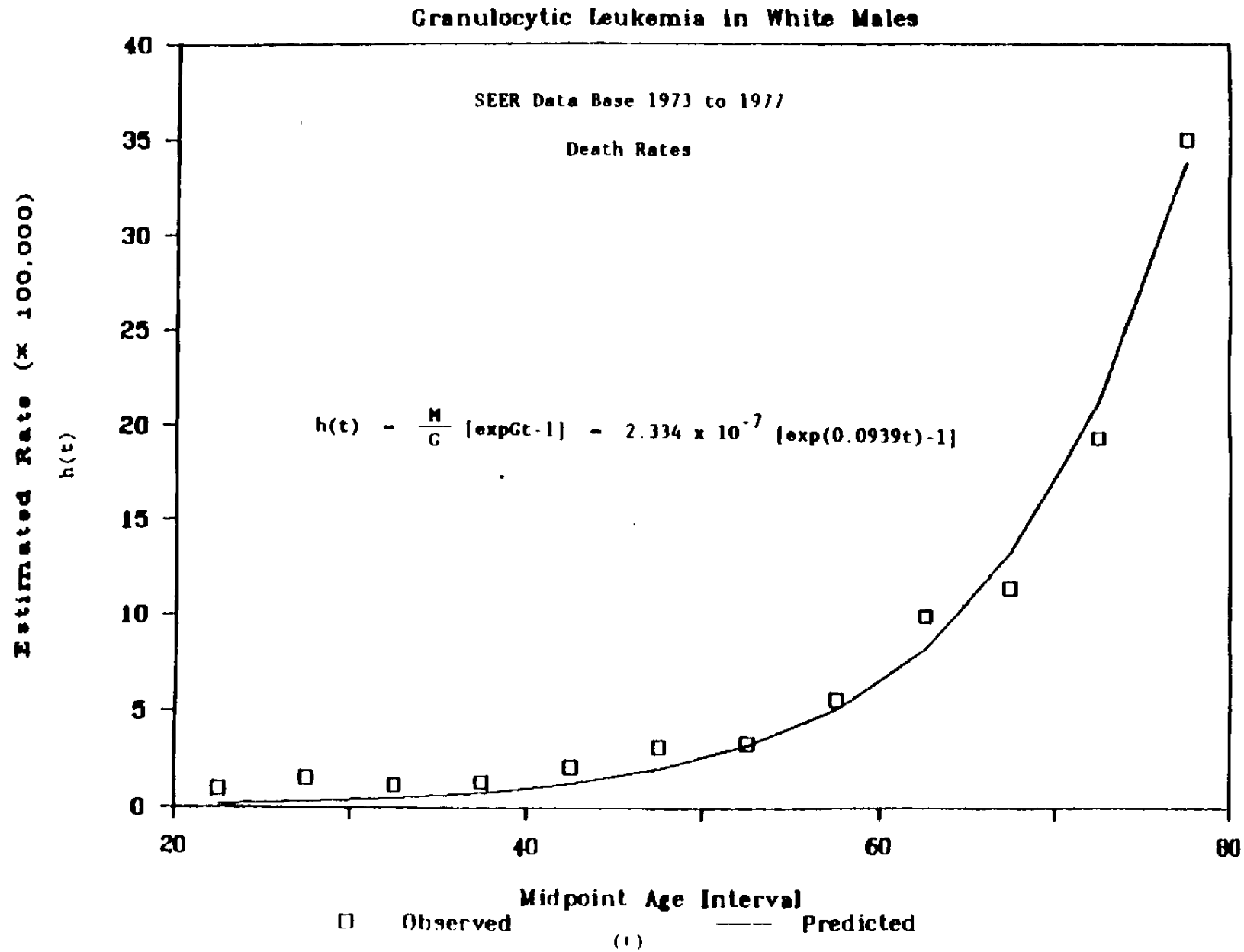
If these factors suggest non-linearity, more biologically relevant models that are under development may give more plausible estimates of leukemia risk at low doses.

2. Relative or Absolute Risk

The assumption of relative risk (i.e., a unit of benzene exposure changes the background leukemia rate by a constant function at all ages) is often made without critical analysis; however, for benzene and other agent-induced leukemias, the evidence is inconsistent with the relative risk assumption. The U.S. background death rate for granulocytic leukemia, the type most closely associated with radiation and benzene exposure, increases as an exponential function of age. This relationship is shown in Figure II-3 using the leukemia death rates in the SEER data base (excluding Puerto Rico) from 1973 to 1977. It is well recognized that the shape if not the magnitude of such curves are consistent between human populations. In contrast, leukemia death rates have been shown to reach a peak and then decline over time upon cessation of exposure to:

- benzene in animals and humans (Rinsky et al. 1981, Yin et al. 1987, Cronkite et al. 1985)

FIGURE II-3



- atomic bomb radiation (Radford 1983)
- therapeutic radiation (Cadman et al. 1977, Foucar et al. 1979, Boice et al. 1987)
- cancer chemotherapy (Cadman et al. 1977, Foucar et al. 1979)

This trend was shown previously in Figure II-2 in a schematic manner.

The ratio of the background and agent-induced leukemia death rates (i.e., rates in Figure II-2 divided by rates in Figure II-3 for the same year) are obviously not consistent over time. This time dependent relative risk adjusted to make the maximum value equal to ten is shown in Figure II-3a for both a ten-year and a lifetime exposure. As a result, a relative risk model in which the constant relative risk assumption is made should not be used for benzene or other leukemogens for the purpose of quantitative cancer risk estimation. In all subsequent upper bound model developments in the Phase I report, the absolute risk form will be employed, which is the standard form used by EPA in risk models based on animal bioassay data.

3. Form of the Weighting Functions

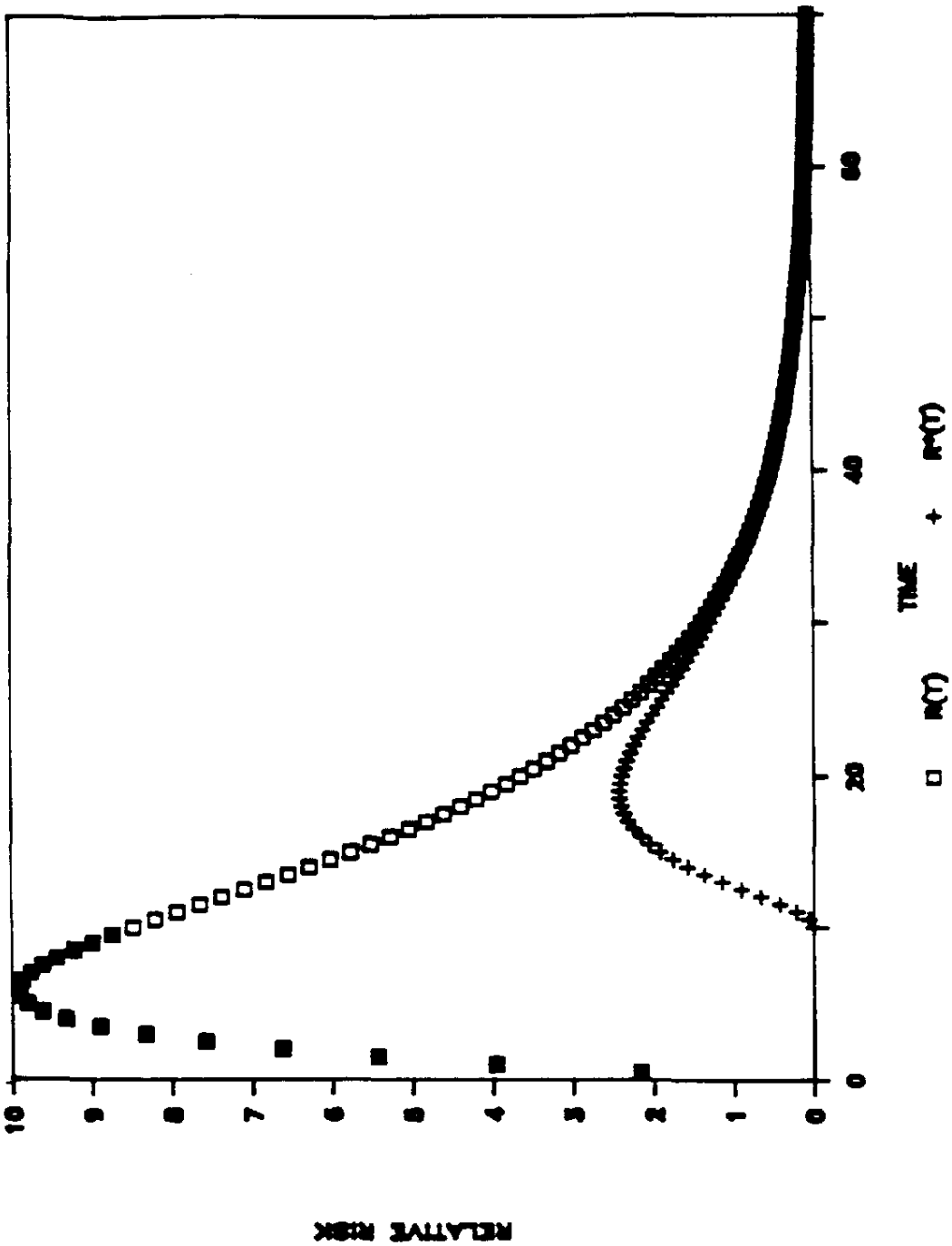
In this section, the weighting functions used by EPA to assess benzene risk are evaluated for their biological consistency with the observed latency period for leukemia.

Window Dose

As is suggested by Crump et al. (1987), the window dose assumption has utility for testing hypotheses about the duration of the exposure effect; however, since the window dose assumption appears to have little biological relevance and gives the lowest risk estimates of the three weighting functions (i.e., is least conservative), it will not be explored further here.

FIGURE II-3a

RELATIVE SCALED LEUKEMIA RISK FOR A BENZENE-EXPOSED POPULATION OVER TIME



Cumulative Dose

Within the context of weighting functions, cumulative dose is the measure of exposure that results from the assumption that all future age-specific cancer rates are affected equally by past exposures of the same magnitude, regardless of when they occurred. Cumulative dose is an attractive concept as a measure of exposure because of its simplicity and intuitive appeal. The concept of cumulative dose and the total effect of that dose over the observation period is illustrated in Figure II-4. Cumulative dose is also consistent with the assumption that the underlying dose-response model is multistage restricted to two stages, and that only the first stage is affected by the agent. To derive this two-stage form, the result of Crump and Howe (1984) giving the age-specific cancer rate for exposure to a constant level x of an agent from age s to f in which only the first stage is affected is used.

In this case, the age-specific leukemia death rate may be expressed:

$$h(t) = \alpha t^{k-1} + \beta x[(t-s)^{k-1} - (t-f)^{k-1}] \qquad \text{(II-4}$$

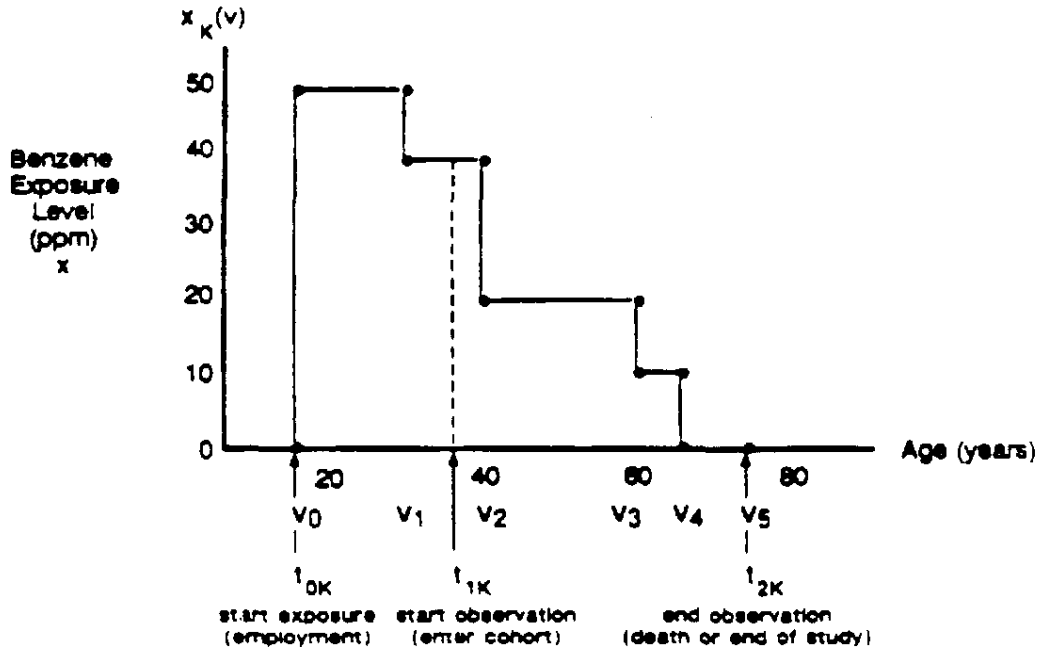
where k is the number of stages. Taking k=2 gives the result:

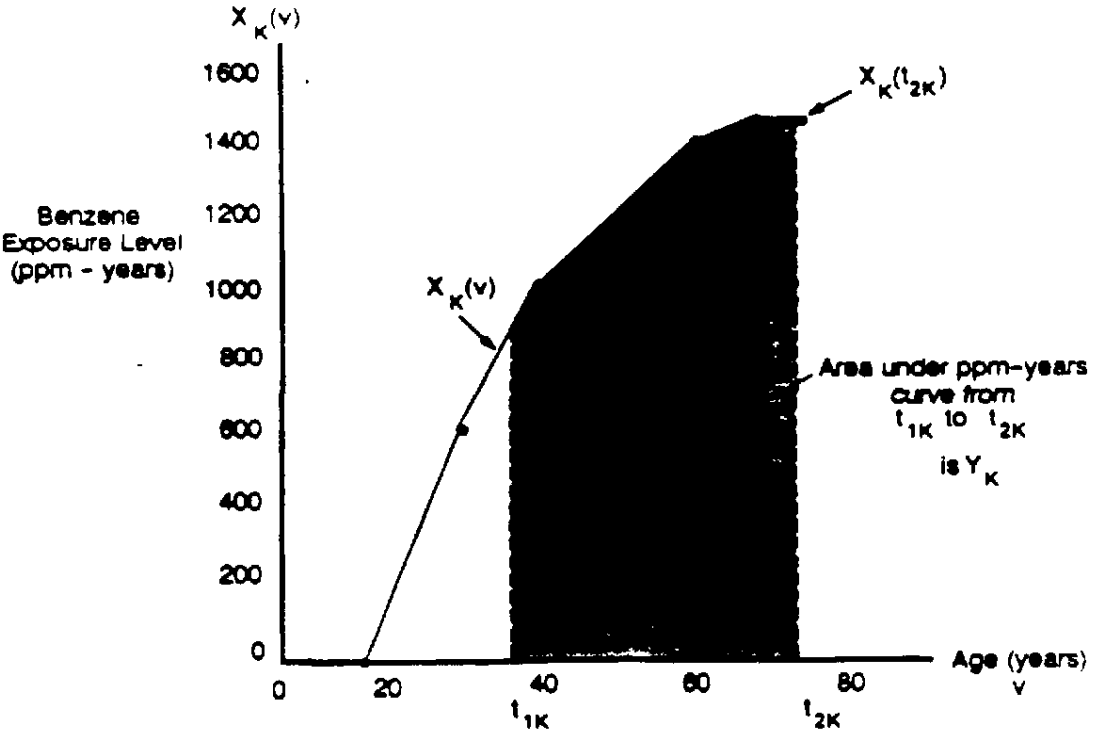
$$h(t) = \alpha t + \beta x(f-s) \qquad \text{(II-5}$$

where x(f-s) is the cumulative exposure. For continuous exposure, f = t and s = 0 so that h(t) = $\alpha t + \beta xt$.

Several observations are inconsistent with the relationships postulated in equations (II-4) and (II-5). First, background leukemia death rates are increasing over time much faster than the predicted level based on the linear equation (see Figure II-3). Second, h(t) does not remain proportional to the total past exposure [i.e., x(f-s)] as predicted, but declines over time following cessation of leukemogen exposure. This pattern is true for both radiation- and benzene-induced leukemia. Enterline (1987) has pointed out a

Typical Step Function for Exposure, Cumulative Exposure, and Area Under Cumulative Exposure Curve for Observation Period





SAL 000001484

number of additional factors indicating that cumulative exposure is not a biologically plausible exposure variable for carcinogens in general. As a result, little weight will be given to this approach.

Weighted Cumulative Dose

The weighted relative exposure assumption is reasonably consistent with observations of time-dependent changes in risk exhibited by the pattern of latency observed after exposure to the atomic bombs in Japan; however, the manner in which this observation has been represented by a weighting function is crude in nature since it employs a step function form that is inconsistent with biological processes. In the next section, an alternative weighting function is derived that is consistent with observations and theoretical biological considerations.

III. NEW METHODOLOGICAL DEVELOPMENT

A. Form of Weighting Function

The accumulated data on the time from leukemogen exposure to leukemia death and the preleukemia syndromes associated with the etiology of the disease are very consistent with the hypothesis that agents such as radiation, chemotherapeutic drugs, and benzene induce leukemogenesis by similar mechanisms (Rinsky et al. 1981, Yin et al. 1987, Radford 1983, Cadman et al. 1977, Foucar et al. 1979). It is reasonable to assume on this basis that once a cell has been transformed to malignancy, its behavior is leukemogen-independent and the probability distribution of the time from the cellular transformation to death due to leukemia, $w(t-v)$, that is obtained from different leukemogens are equivalent and can be used interchangeably.

Radford (1983) notes that post-radiation leukemia is the earliest observed radiogenic cancer, reaching a peak at five years following exposure. In contrast to the solid tumors induced by radiation, the leukemia effect virtually disappears after 25 or 30 years. A simple estimate of the average time between cell transformation and leukemia death can also be made on the basis of a recent study by Yin et al. (1987). These investigators described 30 cases of leukemia in a cohort of benzene exposed workers in China for whom exposure was continuous from the start of employment until death due to leukemia. Under the assumption that over time, each exposure had an equally likely chance of inducing the critical cellular transformation, an estimate of the modal time from that transformation until death is simply one-half of the observed latency period (i.e., time of first exposure until death). This value is 11.4/2 = 5.7 years, which is in very good agreement with the atomic bomb radiation data.

A weighting function that is consistent with these observations is a simple function of the form:

$$w(t-v) = K^2(t-v)\exp-[K(t-v)] \qquad \text{(III-1)}$$

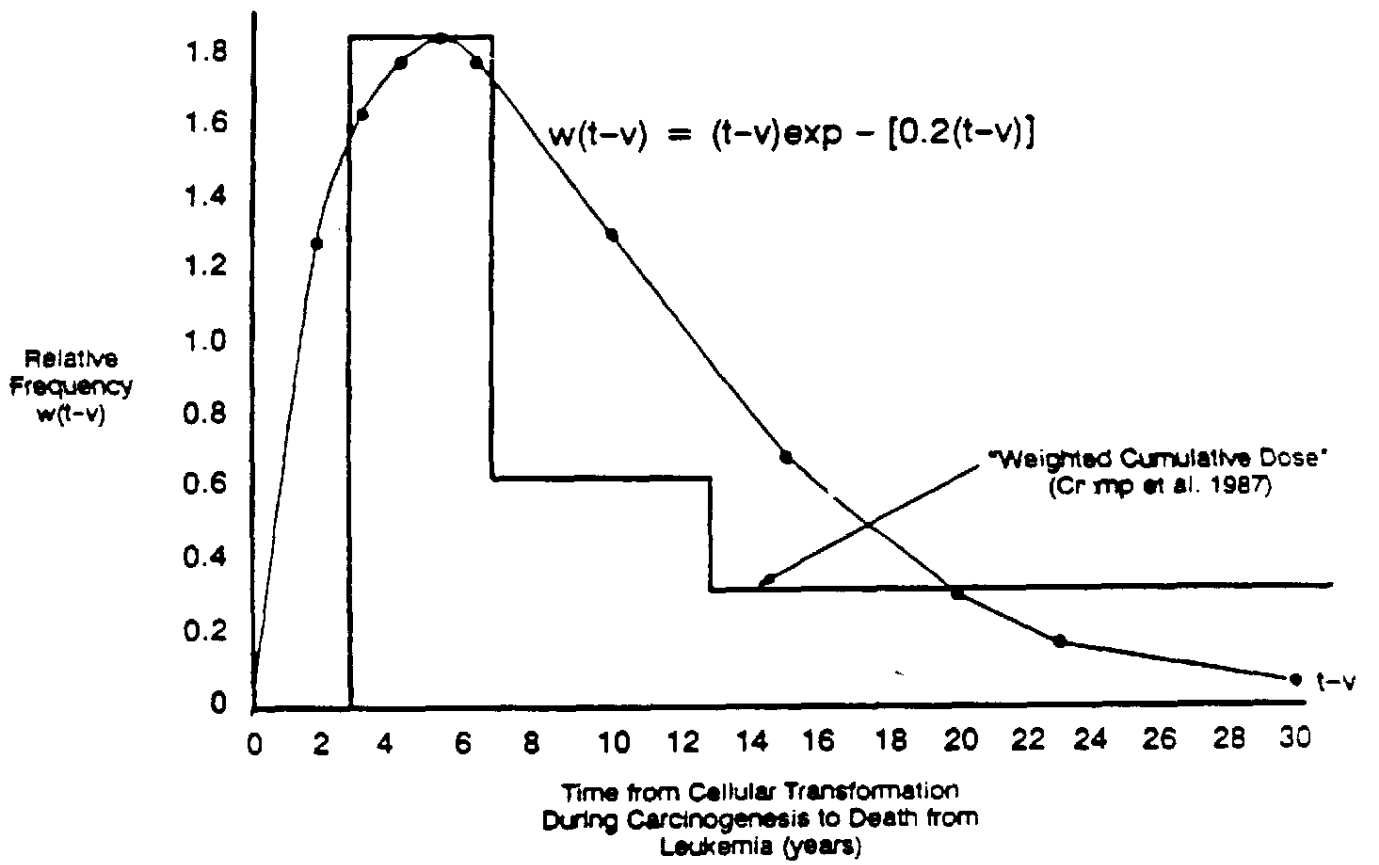
where $(t-v)\geq 0$ and $K = 0.2$. This relationship results in a function that predicts a maximum leukemia death rate five years after transformation and only about 4% of the maximum response at 30 years. The latter would be non-detectable using available data bases due to their lack of power to detect such a small increase. This functional form is illustrated in Figure III-1 and is contrasted to Crump's "weighted cumulative dose". In all subsequent analyses equation (III-1), with K=0.2, shall be used as the weighting function in the risk models that are developed.

In the Phase II report, actual time-to-tumor data may be used to estimate the parameters of the time-to-tumor distribution from studies in which benzene, radiation, and chemotherapy have been shown to induce leukemia. Techniques can be developed to make efficient use of these data to estimate K or the parameters in such models as the Weibull, log-normal, or log-logistic. Efficient mathematical methods of using the time-to-tumor information from several epidemiological studies to estimate the parameters in the risk models will be developed in the future if resources permit.

Based upon the previous analysis, a linear absolute risk model will be assumed for benzene with a continuous weighting function that simulates the observed latency period for leukemia among Japanese post-atomic bomb radiation, benzene-exposed Chinese workers, and benzene-exposed Pliofilm workers in Ohio. The main improvement afforded by the new model is its ability to utilize exposure data from individuals in the benzene-exposed cohort. The availability

FIGURE III-1

Relative Leukemia Latency Period



of individual exposure and survival data allows the development of an estimation procedure that is 100% efficient in its use of the observed data (i.e., maximum likelihood estimation of the time to benzene-induced leukemia death). The development of this approach is discussed in Section III.B.

B. Maximum Likelihood Method of Leukemia Potency Parameter Estimation Using Individual Exposure and Time-to-Tumor Data

The method of estimation employed by EPA for benzene and leukemia used approximations and grouped data to obtain estimates of the unknown dose-response model parameters (EPA 1985). The basic data was person years observed and expected number of cases within exposure intervals. The type of information that was used is shown in Table III-1. The use of such summarized data resulted in the incorporation of potential biases into the method of estimation and precluded a logical method for introducing non-linear relationships into the dose-response effect submodel $G[x(v)]$. The extent of bias that is introduced by using summary data is discussed by Thomas et al (1985), who concluded that it is a major factor in the distortion of the shape of dose-response relationships and results in risk estimates that are unreliable. In contrast, using the exact exposure pattern and vital status determination for each individual in the cohort eliminates the defects associated with EPA's approach. To use individual data and the methods developed in this section, the following information is needed:

1. Sex
2. Race
3. Date of birth
4. Age follow-up begins
5. Age follow-up ends
6. Status at end of follow-up (alive or dead)
7. Cause of death if dead
8. Exposure history

TABLE III-1

OBSERVED AND EXPECTED LEUKEMIAS
IN RINSKY ET AL. COHORT BY WEIGHTED
CUMULATIVE DOSE; 1940-1978 FOLLOWUP

X_j Weighted Cumulative Exposure in ppm-years (average)	O_j Observed	E_j Expected	SMR	W_j (person-years)
0-3 (0.54)	2	1.55	129	24633
3-10 (5.9)	0	0.46	0	5843
10-30 (17.7)	0	0.36	0	4544
30-100 (56.6)	1	0.33	305	3947
100-300 (163.5)	3	0.204	1470	2286
300+ (554.8)	2	0.047	4250	633

1. General Approach

In this section, the maximum likelihood method for estimating the unknown parameters in the age-specific cancer rate function for benzene-induced leukemia is derived.

The benzene-exposed cohort is assumed to contain a total of N individuals For the kth individual in the cohort, the vector t_k is defined as

$$t_k = (t_{0k}, t_{1k}, t_{2k})$$

where t_{0k} = age at which an individual was first exposed to benzene;

t_{1k} = age at which an individual first entered the cohort for observation, and

t_{2k} = age at which an individual was last observed.

The exposure function for the kth individual is denoted as

$$x_k(v) \qquad 0 \le v \le t_{2k}.$$

The general form of the age-specific leukemia death rate for the kth individual at age $t \ge t_{1k}$ was developed in the previous section and is defined for the kth individual as

$$h_k(t) = \alpha(t) + \int_0^t G[x_k(v)]w(t-v)dv \qquad \text{(III-2}$$

which is also known as the individual hazard rate. In the absence of competing mortality the probability of survival until age t for the kth individual is b definition

$$S_k(t) = \exp-[\Lambda_k(t)], \text{ where} \qquad \text{(III-3}$$

$$\Lambda_k(t) = \int_{t_{1k}}^{t_{2k}} h_k(v)dv \qquad \text{(III-4}$$

which is referred to as the accumulated hazard.

Two types of observations for an individual within the study cohort are possible:

1. death at age t_{2k} from a form of leukemia that is suspected to be benzene-induced; and

2. no death due to leukemia by age t_{2k}.

By definition, the probability density function for those who died of leukemia given that they survived until age t_{1k} is

$$f_k(t) = h_k(t)S_k(t). \qquad \text{(III-5)}$$

where $h(\cdot)$ and $S(\cdot)$ are the hazard and probability of survival functions.

The M individuals that died of leukemia in the cohort of size N are given the first M subscripts, so that the likelihood of the observed leukemia cases occurring in the cohort can be expressed as

$$L = \left[\prod_{k=1}^{M} f_k(t_k)\right]\left[\prod_{k=M+1}^{N} S_k(t_k)\right] = \left[\prod_{k=1}^{M} h_k(t_k)\right]\left[\prod_{k=1}^{N} S_k(t_k)\right] \qquad \text{(III-6}$$

Alternatively, using the log-likelihood function, the expression can be written as

$$\ln L = \sum_{k=1}^{M} \log h_k(t_k) - \sum_{k=1}^{N} \Lambda_k(t_k) \qquad \text{(III-7}$$

where $\Lambda(\cdot)$ is the cumulative hazard function defined in equation (III-4).

Estimates for the unknown parameters in the hazard function $h(\cdot)$ can be obtained from the likelihood equation by solving the set of n non-linear equations of the form

$$\frac{\partial \ln L}{\partial \theta_s} = 0 \qquad s = 1,2,\ldots,n \qquad \text{(III-8}$$

where n is the total number of unknown parameters for which estimates are desired. The development of these estimation procedures in a slightly less general way is described in Gart et al. (1986, sec. 6.3. ii).

2. Parameter Estimation for Absolute Risk Model

The general form for the absolute risk model may be expressed as

$$h_k(t) = \alpha(t) + \beta \int_0^t G[x_k(v)]w(t-v)dv, \qquad \text{(III-9)}$$

which may also be written

$$h_k(t) = \alpha(t) + \beta X_k(t) \qquad \text{(III-10)}$$

where $X_k(t) = \int_0^t G[x_k(v)]w(t-v)dv,$ (III-11)

which represents the effective benzene exposure level. In this case, the log-likelihood function has the form

$$\ln L = \sum_{k=1}^{M} \log [\alpha(t_k) + \beta X_k(t_k)] - \beta \sum_{k=1}^{N} \int_{t_{1k}}^{t_{2k}} X_k(v)dv - \int_0^{t_{2k}} \alpha(v)dv. \qquad \text{(III-12)}$$

If the weighting function is assumed to be completely specified and $\alpha(t)$ is known from life tables, the only unknown parameter in the likelihood function is the term β (i.e., the cancer potency). To estimate β, the derivative of the log likelihood with respect to β is set equal to zero, which results in the relationship

$$\frac{d\ln L}{d\beta} = \sum_{k=1}^{M} \frac{X_k(t_k)}{\alpha(t_k)+\beta X_k(t_k)} - \sum_{k=1}^{N} \int_{t_{1k}}^{t_{2k}} X_k(v)dv = 0 \qquad \text{(III-13)}$$

This expression may be solved for λ using iterative non-linear techniques. Such a solution requires that the values are known for $X_k(v)$ and its integral.

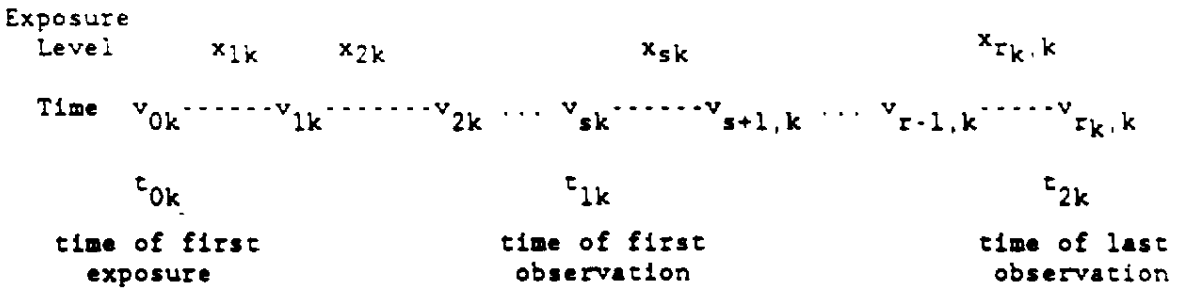
$$Y_k = \int_{t_{1k}}^{t_{2k}} X_k(v)dv, \qquad \text{(III-13a)}$$

over the observation period. These values (i.e., $X_k(v)$ and Y_k) are derived in the next section for a general exposure scenario and two weighting functions.

3. Exposure Pattern

The time-dependent exposure function x(v) is usually based upon a series of analytical measurements taken over time. On occasion, an underlying structural form for x(v) may be assumed, such as $x(v) = x(0)\exp-\lambda t$ (i.e., first order decay). However, in most situations, the measured value will define a step function for which it is assumed that exposure is constant over some defined interval such as a working year.

A step function may be viewed in the following manner. Consider the exposure pattern for the kth individual who has r_k exposure intervals represented below:



where the v_{jk} are the times at which the exposure level changed for the kth individual except for the cases v_{0k}, v_{sk}, and $v_{r_k,k}$, which are defined in the

representation. The schematic representation may also be expressed by the mathematical function

$$x_k(v) = x_{jk} \qquad v_{j-1,k} \leq v < v_{jk} \qquad j = 1,2,\ldots,r_k \qquad \text{(III-14)}$$

A typical depiction of this type of exposure pattern was displayed in Figure II-4.

In the next subsections, the effective exposure levels for the benzene-exposed cohort are derived assuming a step function for exposure and two different types of weighting functions.

Equal Weights

For the case in which all times in the future are given equal weight (i.e., $w(t-v) = dv$, $0 \leq v \leq t$), the weighted cumulative exposure (effective exposure) for a step function until times $v_s \leq v \leq v_{s+1}$ may be expressed as:

$$X_k(v) = \sum_{j=1}^{s} x_{jk}(v_{jk}-v_{j-1,k}) + x_{s+1,k}(v-v_{sk}) \qquad \text{(III-15}$$

At the point of change in exposure level, it may also be expressed as:

$$X_k(v_{sk}) = \sum_{j=1}^{s} x_{jk}(v_{jk}-v_{j-1,k}) \qquad \text{(III-1}$$

The total effective exposure at the end of the observation period is described as

$$X_k(t_{2k}) = \sum_{j=1}^{r_k} x_{jk}(v_{jk}-v_{j-1,k}) \qquad \text{(III-1}$$

which is time-dependent total exposure in ppm-years. These expressions (i.e., III-15, III-16, and III-17) are derived by substituting equation (III-14) into equation (III-11) and integrating. The likelihood equation solution also

requires that the area under the ppm-years curve during the observation period, t_{1k} to t_{2k}, be obtained. For the step function, this area may be expressed as:

$$Y_k = \sum_{m=s_k}^{r_{k-1}} (v_{m+1,k}-v_{mk}) \sum_{j=1}^{m} x_{jk}(v_{jk}-v_{j-1,k}) + x_{m+1,k} (v_{m+1,k}-v_{mk})^2/2 \qquad \text{(III-18)}$$

which is obtained by substituting the definition for $X_k(v)$ given in equation (III-15) into equation (III-13a) and integrating. The step function exposure pattern, the cumulative exposure (ppm-years) function, and the area under the cumulative exposure function during the observation period were all shown in Figure II-4.

The numerical values for $X_k(t_k)$, $\alpha(t_k)$ and Y_k can be calculated from the derived equations and substituted into the derivative of the log-likelihood equation to obtain a numerical solution for β.

It is informative to see how the estimator behaves in a simple case. Consider the case where exposure is constant over an individual's entire lifetime. In this case, the hazard may be expressed as

$$h(t) = \alpha(t) + \beta x t \qquad \text{(III-19)}$$

and the cumulative hazard as

$$\Lambda(t) = \int_0^t \alpha(v)dv + \frac{\beta x t^2}{2} \qquad \text{(III-20)}$$

where $\int_0^t \alpha(v)dv$ is the expected spontaneous (background) number of leukemia cases by age t.