A comprehensive Review of Occupational and General Population Cancer Risk: 1,3-Butadiene Exposure-Response Modeling for

All Leukemia, Acute Myelogenous Leukemia, Chronic Lymphocytic Leukemia, Chronic Myelogenous Leukemia, Myeloid Neoplasm and Lymphoid Neoplasm

Robert L. Sielken Jr., Ph.D., and Ciriaco Valdez-Flores, Ph.D., P.E.
Sielken & Associates Consulting, Inc.
3833 Texas Avenue, Suite 230, Bryan, TX, 77802
Tel: 979-846-5175; Fax: 979-846-2671; Email: SielkenAssoc@aol.com

Slides Prepared for the 2014 Symposium on Understanding the Health Risks of Lower Olefins

November 5, 2014

The exposure-response modeling for butadiene has evolved considerably in the last 30 years.

The earliest modeling was based on animal studies

(e.g., USEPA 1985, IARC 1986, 1992, NIOSH 1991, CARB 1991, OSHA 1996, and Cagen et al. 1996).

Although human data are generally considered
more relevant to humans,

the exposure-response modeling based on human
epidemiological studies had to wait for
sufficient quantitative exposure information.

Starting in 1995,
the needed quantitative exposure data were
provided by the University of Alabama at Birmingham (UAB)
epidemiological study of North American workers
in the styrene-butadiene rubber (SBR) industry.

The most recent EPA risk assessment of BD (USEPA 2002) concluded that the UAB SBR epidemiological study provided the best published set of data to evaluate human cancer risk from 1,3-butadiene (BD) inhalation exposure.

Similarly, the Texas Commission on Environmental Quality (TCEQ) reviewed the scientific literature and concluded that there were no other epidemiology studies that would be appropriate to evaluate human cancer risk from BD inhalation exposure (TCEQ 2008).

With the **availability of quantitative exposure data**
for BD epidemiology studies,
the exposure-response models evolved quickly.

The earliest models (USEPA 1998 and 2002)
were primarily restricted to “regulatory” models;
that is, linear models
with the hazard rate for a broad group of cancer
endpoints (e.g., leukemia)
modeled as a linear function of one predictive variable
(specifically, cumulative BD ppm-years).

In general, there are five components in the evolution of the models:

(1) the specific cancer endpoint,

(2) the exposure variables chosen to be the relevant predictors of cancer,

(3) the representation of an individual's exposure history,

(4) the shape of the exposure-response relationship (e.g., restricted to be linear or not), and

(5) the statistical techniques used to do the modeling.

The initial reports regarding the UAB study suggested that there was a noticeable increase in leukemia among workers exposed to butadiene and other chemicals

who performed tasks in certain job categories (maintenance, laboratory, etc.)

(Delzell et al. 1995 and 1996).

Leukemia was the initial regulatory cancer endpoint.

However, leukemia is a GROUP of cancer endpoints involving multiple cell types, modes of action, biological processes, etc.

For butadiene, the target cancer endpoint has evolved from this problematic endpoint (All Leukemia) to more specific endpoints:

Acute Myelogenous Leukemia (AML),
Chronic Lymphocytic Leukemia (CLL),
Chronic Myelogenous Leukemia (CML),
Myeloid Neoplasm, and
Lymphoid Neoplasm.

EPA (1998 and 2002) selected
cumulative butadiene exposure
(specifically cumulative BD ppm-years)
as the single exposure variable to use to predict
All Leukemia.

EPA's risk assessments have never been updated

beyond their use of the initial study data

even though UAB added seven more years of follow-up

and substantially improved their exposure estimates

(Delzell et al., 2000, 2001, Macaluso et al., 2004; and

Sathiakumar et al., 2005).

While cumulative BD ppm-years may be a convenient variable to use as a basis for regulations,
cumulative BD ppm-years is
only one of the exposure variables considered by UAB and is not necessarily the best one.

UAB epidemiological data
include five non-exposure covariates
and nine exposure variables:

UAB epidemiological data include

five non-exposure covariates:

1. age,
2. years since hire,
3. calendar year,
4. race, and
5. plant.

UAB epidemiological data include 9 exposure variables:

1. cumulative BD ppm-years,
2. cumulative exposure to STY ppm-years,
3. cumulative exposure to DMDTC,

4. cumulative number of BD high-intensity tasks (HITs) (i.e., tasks with exposures above 100 ppm BD),
5. cumulative number of STY HITs (i.e., tasks with exposures above 50 ppm STY),

6. cumulative exposure to BD ppm-years concentrations below 100 ppm (concentrations > 100 ppm set equal to 100 ppm),
7. cumulative exposure to BD ppm-years concentrations above 100 ppm (concentrations ≤ 100 ppm set equal to 0 ppm),

8. cumulative exposure to STY ppm-years concentrations below 50 ppm (concentrations > 50 ppm set equal to 50 ppm), and
9. cumulative exposure to STY ppm-years concentrations above 50 ppm (concentrations ≤ 50 ppm set equal to 0 ppm).

Cumulative number of BD HITs
is a better predictor of All Leukemia
than cumulative BD ppm-years.

This is also true for CML, AML, and Myeloid Neoplams.

In fact, cumulative BD ppm-years is not the best predictor of any one of the six endpoints we evaluated:

All Leukemia, CLL, CML, AML,
Lymphoid Neoplasms, and Myeloid Neoplasms

Table 2. Statistical significance (p-value) of the effect of adding one of the exposure variables to the Cox proportional hazards model with no exposure variables[1]

Exposure Variable	Leukemia	CLL	CML	AML	Lymphoid	Myeloid
BD (cumulative ppm-years)	0.0287*,c	0.0239*,b	0.5261	0.9490	0.0565	0.4222
STY (cumulative ppm-years)	0.1438	0.0990	0.7047	0.9654	0.0406*,c	0.8291
DMDTC (cumulative mg/cm-years)	0.5964	0.9027	0.6549	0.7544c	0.9733	0.5043
Cumulative # of BD HITs	0.0083**,b	0.6071	0.0843b	0.9065	0.3525	0.0647b
Cumulative # of STY HITs	0.2815	0.9584	0.4219	0.7938	0.4027	0.2638
BD ≤ 100 ppm (cumulative ppm-years)	0.0071**,a	0.0190*,a	0.0281*,a	0.3052a	0.0113*,a	0.0552a
BD > 100 ppm (cumulative ppm-years)	0.0593	0.0401*,c	0.8652	0.7975	0.1120	0.6433
STY ≤ 50 ppm (cumulative ppm-years)	0.0531	0.1188	0.1461c	0.6512b	0.0277*,b	0.3064c
STY > 50 ppm (cumulative ppm-years)	0.2958	0.1331	0.6697	0.7904	0.0865	0.7668

The representation of an individual's exposure history has evolved from cumulative exposures GROUPED into categories to more continuous exposure variables based on the plant- and calendar-year-specific job-exposure-matrices (JEMs) provided by UAB.

Sielken & Associates Consulting Inc. has had a unique opportunity (in addition to UAB and Health Canada) to have access to the individual data for the purpose of exposure-response modeling.

Cumulative exposures put equal emphasis on exposure magnitude and duration.

Although other exposure metrics based on unequal emphases have been explored, no simple alternative

(e.g., cumulating [$\text{magnitude}^j \times \text{duration}^k$] for different powers j and k)

has emerged as necessarily more predictive.

Lags and windows of exposure have been considered but not found to have much impact.

Non-cumulative exposure metrics have not been extensively considered.

Whereas nonlinear shapes for the exposure-response model have been considered for other chemicals,

for butadiene,

the shape of the regulatory model has predominately kept to a simplified assumption of linearity.

The statistical techniques used in the exposure-response model have evolved from Poisson regression techniques to Cox proportional hazards model techniques.

The first step in Poisson regressions is to create exposure categories which categorize the person-years in an individual's exposure history into specified intervals.

This categorization discards much of the information in the individual exposure histories that UAB has captured in their job-exposure matrices.

Cox proportional hazards models allow
the exposure in the model to be whatever specific values
each worker had in his work history
and how that exposure changed over time.

Because UAB made the individual exposure histories
available to Sielken & Associates
(under a confidentiality agreement)
and our exposure-response modeling had evolved from
Poisson to Cox modeling,

Sielken & Associates did not have to rely on reported
categorical results and Poisson regression.

The individual exposure histories

in the most recent UAB data

are fully incorporated

into the exposure-response Cox modeling reported in

Sielken and Valdez-Flores (2011) and (2013) and

herein.

The University of Alabama at Birmingham (UAB)

butadiene exposure data have evolved

1995 – First Estimates of Exposure Available

2000 – Revised Estimates of Exposure

2004 – Seven More Years of Follow-up and

Most Recent Exposure Estimates

The 2004 epidemiological data has more decedents, more leukemia decedents, and more years of followup than previously

	2004	1995 and 2000
Decedents	5,593	4,227
Leukemia	81	58
Years of follow up	500,400	426,000
Average Years of follow up	30.0	25.7

In the UAB data set, there are a total of 16,585 workers with exposure estimates for BD, STY and DMDTC.

Exposure estimates are available for six of the eight plants, and the Cox Proportional Hazards regression modeling is based on these six plants.

Recent Risk Assessment Publications

of Cancer Risks for

1,3-butadiene (BD)

Using UAB Newest 2004 Exposure Data:

Sielken et al. (2007),

TCEQ (2008),

Sielken and Valdez-Flores (2011), and

Sielken and Valdez-Flores (2013)

Although Leukemia and other Lymphohematopoietic Cancers had been the primary endpoints of interest,

more recent analyses have focused on

Leukemia Subtypes

Lymphoid Neoplasms

and

Myeloid Neoplasms

and the results for these endpoints are reported herein.

Breakdown of Leukemia Subtypes (ICD 9)

Category	Description	Number	ICD9
1	Acute lymphocytic leukemia	3	204.0
2	Acute myelogenous or monocytic leukemia	26	205.0, 206.0
3	Acute leukemia – other/unknown	4	207.0
4	Chronic lymphocytic leukemia	25	204.1
5	Chronic myelogenous leukemia	16	205.1
6	Chronic leukemia – other/Unknown	1	207.1
7	Non-AML – unspecified lymphocytic	2	204.9
8	Non-AML – unspecified myelogenous	3	205.9
9	Non-AML – other non-AML/unknown	1	207.8

6 Endpoints Analyzed and Reported

Description	Number	ICD9
Leukemia Subtypes		
All leukemia subtypes combined	81	204 - 208
Acute myelogenous or monocytic leukemia (AML)	26	205.0, 206.0
Chronic lymphocytic leukemia (CLL)	25	204.1
Chronic myelogenous leukemia (CML)	16	205.1
Additional Endpoints Analyzed with a Larger Number of Decedents than the three Leukemia Subtypes		
Lymphoid neoplasm	120	200 - 204
Myeloid neoplasm	56	205 - 206

Exposure-Response Modeling

Lack of Sensitivity to the Assignment of Uncertain Deaths to Subtypes

Sielken and Valdez-Flores (2013) showed that

the results for CLL, CML and AML are essentially unchanged

by either including or excluding

the “other/unknown” and “unspecified” leukemia types

in CLL, CML, and AML deaths.

USEPA 2002 Risk Assessment

Based on the animal data, EPA estimates a unit risk of 0.3 per ppm.

Based on the outdated UAB data (including outdated exposure estimates) and flawed modeling, EPA estimates a Final Unit risk of 0.08 per ppm.

TCEQ (2008) and Grant et al. (2009)

Based on most recent update of UAB study
and most recent exposure estimates

TCEQ's unit risk factor (URF) = 0.0011 per ppm

1. All leukemia
2. Upper Bounds (instead of best estimates, MLEs)
3. Data restricted to lower 95% of exposure range
4. Model based on only cumulative BD ppm-years
   and not any other exposure variables or covariates

In TCEQ's discussion of uncertainty, they acknowledge that "The inclusion of age and number of HITs > 100 ppm BD as covariates in the Cox regression modeling may result in URFs that are more relevant to BD exposures experienced by the general population. As mentioned previously, once age is in the model, inclusion of number of BD HITs results in a significant improvement in the fit (likelihood)."

TCEQ (2008) and Grant et al. (2009)

Based on most recent update of UAB study and most recent exposure estimates

TCEQ's unit risk factor (URF) = 0.0011 per ppm

TCEQ assessments present results for several exposure-response models.

When these Cox proportional hazards models match those used by Sielken & Associates the results agree.

Choices (2) to (4), namely,

2. Upper Bounds (instead of best estimates, MLEs),
3. Data restricted to lower 95% of exposure range, and
4. Model based on only cumulative BD ppm-years and not any other exposure variables or covariates,

cause TCEQ's upper bounds on the all leukemia risk to be about 10 fold greater than the best estimates for all leukemia herein (i.e., in the slides to follow).

Sielken & Associates Results

The following slides contain our best estimates of

1. Added Risks at 1 ppm (environmental & occupational)
    a. All Leukemia, CLL, CML, AML,
       Lymphoid Neoplasms, and Myeloid Neoplasms
    b. Models based on
        (i) cumulative BD ppm-years, and
        (ii) cumulative BD ppm-years +
            statistically significant exposure covariate (1%)

2. EC(1/Million), EC(1/100,000), and EC(1/10,000)
    (environmental & occupational), (70 years and 85 years)

All Leukemia

cumulative DMDTC mg/cm-years,
cumulative # of BD HITs,
cumulative # of STY HITs, and
cumulative STY > 50 ppm-years)

significantly (at the 1% significance level) improve the maximum likelihood over the model with cumulative BD ppm-years alone.

The slope (per cumulative BD ppm-year) of the Cox proportional-hazards model for leukemia is statistically significantly different than zero before adding any of the eight exposure covariates; however, the slope is not statistically significantly different than zero after any one of the eight exposure covariates is added to the Cox proportional hazards model.

All Leukemia
(Continued)

The slope (per cumulative BD ppm-years)
is NOT statistically significantly different than zero
when the exposure-response modeling
is based on the person-years
with cumulative BD ppm-years less than or equal

200 ppm-years for All Leukemia

CLL

None of the covariates significantly improves the maximum likelihood at the 5% significance level.

The slope is statistically significantly different than zero when the model includes BD ppm-years alone.

The slope (per cumulative BD ppm-years) is NOT statistically significantly different than zero when the exposure-response modeling is based on the person-years with cumulative BD ppm-years less than or equal

300 ppm-years for CLL.

Lymphoid Neoplasms
(which includes CLL)

None of the covariates significantly improves the maximum likelihood at the 5% significance level.

The slope is statistically significantly different than zero when the model includes BD ppm-years alone.

The slope (per cumulative BD ppm-years) is NOT statistically significantly different than zero when the exposure-response modeling is based on the person-years with cumulative BD ppm-years less than or equal

400 ppm-years for Lymphoid Neoplasms.

CML, AML, and Myeloid Neoplasms

The slopes are NOT statistically significantly different than zero when the model includes cumulative BD ppm-years alone.

For CML and AML, when any exposure covariate that makes a SS improvement at either 1% or 5% significance level is added, the estimated slope for cumulative BD ppm-years is negative.

For Myeloid Neoplasms, when any exposure covariate that makes a SS improvement at either 1% or 5% significance level is added, the estimated slope for cumulative BD ppm-years is either negative or not significantly different than Zero at the 5% significance level.

The slope (per cumulative BD ppm-years)
is NOT statistically significantly different than zero
when the exposure-response modeling
is based on the person-years
with cumulative BD ppm-years less than or equal

(1) 200 ppm-years for All Leukemia
(2) 300 ppm-years for CLL
(3) 400 ppm-years for Lymphoid Neoplams

True even though there are plenty of person-years
and cancer deaths observed below these ppm-levels.

Sielken and Valdez-Flores (2011) and (2013)

Added Risk – Leukemia

Lifetime added risk by age 70 years
for an environmental and occupational exposure concentration of 1 ppm

Endpoint	Covariate	Slope (MLE)	Slope (Std. Dev.)	Statistical Significance of Slope	Lag	Added Risk	
						Environmental	Occupational
Leukemia	STY > 50 ppm (cumulative ppm-yrs)	0.000159	0.000140	NS	0	6.5×10^{-5}	1.3×10^{-5}
Leukemia	Cumulative # of BD HITs	0.000201	0.000130	NS	0	8.2×10^{-5}	1.7×10^{-5}
Leukemia	Cumulative # of STY HITs	0.000113	0.000140	NS	0	4.6×10^{-5}	9.5×10^{-6}
Leukemia	DMDTC (cumulative mg/cm-yrs)	0.000179	0.000123	NS	0	7.3×10^{-5}	1.5×10^{-5}
Leukemia	None	0.000291	0.000103	SS(1%)	0	1.2×10^{-4}	2.5×10^{-5}

Added Risk – Leukemia Types

Lifetime added risk by age 70 years
for an environmental and occupational exposure concentration of 1 ppm

Endpoint	Covariate	Slope (MLE)	Slope (Std. Dev.)	Statistical Significance of Slope	Lag	Added Risk	
						Environmental	Occupational
CLL	None	0.000417	0.000128	SS(1%)	0	2.7×10^{-5}	5.8×10^{-6}
CML	None	0.000213	0.000286	NS	0	4.3×10^{-6}	9.0×10^{-7}
AML	Cumulative # of STY HITs	-0.000801	0.000798	NS	0	0	0
AML	STY > 50 ppm (cumulative ppm-yrs)	-0.000795	0.000820	NS	0	0	0
AML	DMDTC (cumulative mg/cm-yrs)	-0.000275	0.000601	NS	0	0	0
AML	None	-0.000010	0.000411	NS	0	0	0

Added Risk - Lymphoid and Myeloid Neoplasm

Lifetime added risk by age 70 years
for an environmental and occupational exposure concentration of 1 ppm

Endpoint	Covariate	Slope (MLE)	Slope (Std. Dev.)	Statistical Significance of Slope[3]	Lag	Added Risk	
						Environmental	Occupational
Lymphoid	None	0.000227	0.000098	SS(5%)	0	1.7×10^{-4}	3.7×10^{-5}
Myeloid	Cumulative # of STY HITs	-0.000230	0.000327	NS	0	0	0
Myeloid	Cumulative # of BD HITs	-0.000084	0.000327	NS	0	0	0
Myeloid	DMDTC (cumulative mg/cm-yrs)	0.000036	0.000227	NS	0	7.6×10^{-6}	1.6×10^{-6}
Myeloid	STY > 50 ppm (cumulative ppm-yrs)	-0.000051	0.000276	NS	0	0	0
Myeloid	None	0.000155	0.00018	NS	0	3.3×10^{-5}	6.9×10^{-6}

Environmental EC – Leukemia

Average **environmental** BD exposure concentrations (ppm) for a lifetime of exposure corresponding to specified added risks and ages (70 and 85 years)

Environmental							
Response	Covariate	Average Butadiene Exposure Concentration (ppm)					
		By Age 70 Years			By Age 85 Years		
		Added Risk			Added Risk		
		1 in 1,000,000 or 0.000001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1,000,000 or 0.000001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001
All Leukemia	STY > 50 ppm (cumulative ppm-yrs)	0.016	0.157	1.538	0.005	0.048	0.473
	Cumulative # of BD HITs	0.012	0.124	1.216	0.004	0.038	0.374
	Cumulative # of STY HITs	0.022	0.221	2.164	0.007	0.067	0.666
	DMDTC (cumulative mg/cm-yrs)	0.014	0.139	1.366	0.004	0.042	0.420
	None	0.009	0.086	0.840	0.003	0.026	0.259

Environmental EC – Leukemia Types

Average **environmental** BD exposure concentrations (ppm) for a lifetime of exposure corresponding to specified added risks and ages (70 and 85 years)

Environmental							
Response	Covariate	Average Butadiene Exposure Concentration (ppm)					
		By Age 70 Years			By Age 85 Years		
		Added Risk			Added Risk		
		1 in 1,000,000 or 0.000001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1,000,000 or 0.000001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001
CLL	None	0.038	0.375	3.319	0.009	0.089	0.858
CML	None	0.236	2.269	16.825	0.077	0.758	6.582
AML	Cumulative # of STY HITs	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	STY > 50 ppm (cumulative ppm-yrs)	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	DMDTC (cumulative mg/cm-years)	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	None	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated

Environmental EC – Lymphoid and Myeloid Neoplasm

Average **environmental** BD exposure concentrations (ppm) for a lifetime of exposure corresponding to specified added risks and ages (70 and 85 years)

Environmental							
Response	Covariate	Average Butadiene Exposure Concentration (ppm)					
		By Age 70 Years			By Age 85 Years		
		Added Risk			Added Risk		
		1 in 1,000,000 or 0.000001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1,000,000 or 0.000001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001
Lymphoid	None	0.006	0.059	0.580	0.002	0.018	0.176
Myeloid	Cumulative # of STY HITs	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	Cumulative # of BD HITs	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	DMDTC (cumulative mg/cm-years)	0.132	1.313	12.642	0.043	0.434	4.272
	STY > 50 ppm (cumulative ppm-yrs)	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	None	0.031	0.305	2.936	0.010	0.101	0.992

Occupational EC – Leukemia

Average **occupational** BD exposure concentrations (ppm) for a lifetime of exposure corresponding to specified added risks and ages (70 and 85 years)

Occupational							
Response	Covariate	Average Butadiene Exposure Concentration (ppm)					
		By Age 70 Years			By Age 85 Years		
		Added Risk			Added Risk		
		1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1000 or 0.001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1000 or 0.001
All Leukemia	STY > 50 ppm (cumulative ppm-yrs)	0.748	7.324	61.174	0.242	2.401	22.398
	Cumulative # of BD HITs	0.592	5.793	48.397	0.191	1.900	17.718
	Cumulative # of STY HITs	1.052	10.305	86.080	0.340	3.379	31.517
	DMDTC (cumulative mg/cm-yrs)	0.664	6.505	54.341	0.215	2.133	19.896
	None	0.409	4.001	33.428	0.132	1.312	12.238

Occupational EC – Leukemia Types

Average **occupational** BD exposure concentrations (ppm) for a lifetime of exposure corresponding to specified added risks and ages (70 and 85 years)

Occupational							
Response	Covariate	Average Butadiene Exposure Concentration (ppm)					
		By Age 70 Years			By Age 85 Years		
		Added Risk			Added Risk		
		1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1000 or 0.001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1000 or 0.001
CLL	None	1.702	15.073	78.961	0.450	4.341	32.863
CML	None	10.711	78.972	275.369	3.823	33.286	165.300
AML	Cumulative # of STY HITs	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	STY > 50 ppm (cumulative ppm-yrs)	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	DMDTC (cumulative mg/cm-years)	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	None	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated

Occupational EC – Lymphoid and Myeloid Neoplasm

Average **occupational** BD exposure concentrations (ppm) for a lifetime of exposure corresponding to specified added risks and ages (70 and 85 years)

Occupational							
Response	Covariate	Average Butadiene Exposure Concentration (ppm)					
		By Age 70 Years			By Age 85 Years		
		Added Risk			Added Risk		
		1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1000 or 0.001	1 in 100,000 or 0.00001	1 in 10,000 or 0.0001	1 in 1000 or 0.001
Lymphoid	None	0.273	2.701	24.376	0.089	0.889	8.561
Myeloid	Cumulative # of STY HITs	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	Cumulative # of BD HITs	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	DMDTC (cumulative mg/cm-years)	6.200	59.625	443.048	2.183	21.503	188.220
	STY > 50 ppm (cumulative ppm-yrs)	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated	No Limit Estimated
	None	1.440	13.848	102.911	0.507	4.995	43.717

THANK YOU

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

Delzell et al. (1995) and Macaluso et al. (1996)

-- Outdated Version of UAB Data Set

-- Delzell et al. (2000 and 2001)
and Sathiakumar et al. (2005)
Added 7 more years of follow-up

-- Macaluso et al. (2004) substantially improved
UAB's exposure estimates and also changed
the definition of peaks from
≥ # ppm over 15 minutes
to ≥ # ppm over any length of time

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

EPA had UAB data on the number of peaks

≥ 100 ppm for any amount of time for BD

≥ 50 ppm for any amount of time for STY

EPA did NOT include number of peaks

in their final exposure-response models.

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

EPA used Poisson Regression Modeling
(rather than preferred Cox Proportion Hazards Modeling)

-- Poisson regression modeling
forces exposure to be categorized
into a few categories

-- rather than continuous exposure values
with each individual's exposure values
being included exactly

USEPA 2002 Risk Assessment
Epidemiologically Based Cancer Risk Assessment

EPA did not obtain the individual worker data

and did not do their own exposure-response modeling

but relied on the modeling done by Health Canada (1998).

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

Flaws in Health Canada's Modeling:

Flaws in Health Canada's Approach for Addressing Covariates and Confounders

Health Canada did not explicitly consider any confounding exposures other than styrene ppm-years. BD and STY HITs were not included. DMDTC was not included.

Health Canada's approach to estimating the role of covariates is flawed. In essence, Health Canada's approach to estimating the role of covariates and ultimately the effect of butadiene exposure on the rate ratios ignores groups of person-years in which there was not a leukemia response.

Furthermore, Health Canada's approach estimates the effect of combinations of covariates on insufficient sample sizes. The insufficient sample sizes result from over-partitioning the data into 29,403 categories. The over-partitioning causes the estimate of each combination of covariates to be roughly based on a single person-year.

Sielken & Associates Presentation at Butadiene Risk Assessment Meeting with USEPA March 29, 2005

USEPA 2002 Risk Assessment
Epidemiologically Based Cancer Risk Assessment

Delzell et al. (1995) models:

(1) Log-linear: $RR = \exp(\beta \times X)$

(2) Power: $RR = \exp(\beta \times \ln[1+X])$

(3) Linear: $RR = 1 + \beta \times X$

(4) Polynomial: $RR = 1 + \beta_1 \times X^{p} + \beta_2 \times X^{q}$

(5) Square Root: $RR = 1 + \beta \times X^{1/2}$

X = exposure

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

Health Canada (1998) models:

(1) Log-linear: $RR = \exp(\beta \times X)$

(2) Power: $RR = \exp(\beta \times \ln[1+X])$

(3) Linear: $RR = 1 + \beta \times X$

(4) Polynomial: $RR = 1 + \beta_1 \times X^p + \beta_2 \times X^q$

(5) Square Root: $RR = 1 + \beta \times X^{\alpha}$

X = exposure

Health Canada used a much finer stratification (29,403 categories) for combinations of age, calendar year, and years since hire than Delzell et al.

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

Problems with Excess Risk Calculation:

Estimated excess risks corresponding to a potential of 85 years of exposure

- well beyond the workers' exposure duration in the UAB study
- years beyond age 70 have much higher background leukemia mortality rates and much greater impact on risks than years before age 70
- that is, the ages with least data have the greatest impact on calculated risks

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

Problems with Excess Risk Calculation:

Estimated excess risks corresponding to leukemia INCIDENCE:

- used an exposure-response model
  fit to leukemia **MORTALITY** data
  as a **substitute** for an exposure-response model
  fit to leukemia **INCIDENCE** data

USEPA 2002 Risk Assessment

Epidemiologically Based Cancer Risk Assessment

Problems with Excess Risk Calculation:

EPA historically had used MLEs
for cancer risk estimates from human data
rather than upper bounds as used with animal data.

For butadiene, EPA used upper bounds.

USEPA 2002 Risk Assessment
Epidemiologically Based Cancer Risk Assessment

Problems with Excess Risk Calculation:

EPA uses an “adjusment factor of 2”

as an unsupported fudge factor

to increase estimated risks 2 fold

Final Unit risk estimate of 0.08 per ppm.

The U.S. EPA Science Advisory Board (SAB) recommendations

of using the most recent data

and

giving consideration to peak exposures to BD have been followed

by Sielken & Associates herein.

The U.S. EPA Science Advisory Board (SAB) recommendations
of using the most recent data
and
giving consideration to peak exposures to BD have been followed.

Macaluso *et al*., (2004) defined the number of BD peaks (i.e., the number of High-Intensity Tasks or **HITs**) for a worker
as the number of tasks (regardless of the duration)
performed in the workplace
that entailed exposures above 100 ppm.

Macaluso *et al*., established the threshold of 100 ppm
as a point that is
"sensitive enough
to capture significant excursions in exposure intensity
and specific enough
to exclude small excursions compatible with the
imprecision of our estimation procedures."

Separate Exposure-Response Models were fit to each of the six endpoints analyzed:

All Leukemia
Acute Myeloid Leukemia (AML)
Chronic Lymphocytic Leukemia (CLL)
Chronic Myeloid Leukemia (CML)
Lymphoid Neoplasms
and
Myeloid Neoplasms

Exposure-Response Modeling - Model

Cox Proportional Hazards Regression was used in the analyses by Sielken and Valdez-Flores (2011 and 2013)

Hazard Rate = BHR × NECV × OECV × f(BDppm-years)

where

BHR = Background Hazard Rate

NECV = categorical effect of Non-Exposure CoVariates

OECV = categorical effect of Other Exposure CoVariates (other than cumulative BD ppm-years)

f(BDppm-years) = continuous function of cumulative BD ppm-years

Exposure-Response Modeling - Model

The **Cox Proportional Hazards Regression** assumes that the relationship between the rate ratio (RR) and the cumulative BD ppm-years is log-linear

$$\ln(RR) = \beta \times \text{cumulative BDppm-years}.$$

This log-linear model is the closest model to the linear model in the low-dose region that can be fit using standard Cox proportional hazards software.

Age is the time index used in Cox regression to construct the likelihood function, so that, the effect of age is implicitly included in the Cox models.

Exposure-Response Modeling - Covariates

The effects of

non-exposure covariates
and
other exposure covariates

are modeled as **categorical effects**
(rather than continuous effects)
because it is less restrictive and
does not require that the covariate effects be smooth or monotone or known to have a specific functional form.

Exposure-Response Modeling - Covariates

None of the Non-Exposure covariates (years since hire, calendar year, race, or plant) make a statistically significant improvement in the maximum likelihood for **any** the endpoints analyzed

Four of the Exposure covariates (three for AML)

[cumulative # of BD HITs,
STY > 50 ppm (cumulative ppm-years),
cumulative # of STY HITs, and
DMDTC (cumulative mg/cm-years)]

make a significant improvement on the model's fit to **leukemia**, **AML**, or **myeloid neoplasm** at the 1% significance level

None of the eight Exposure covariates makes a significant improvement at the 1% significance level for **CLL**, **CML**, or **lymphoid neoplasm.**

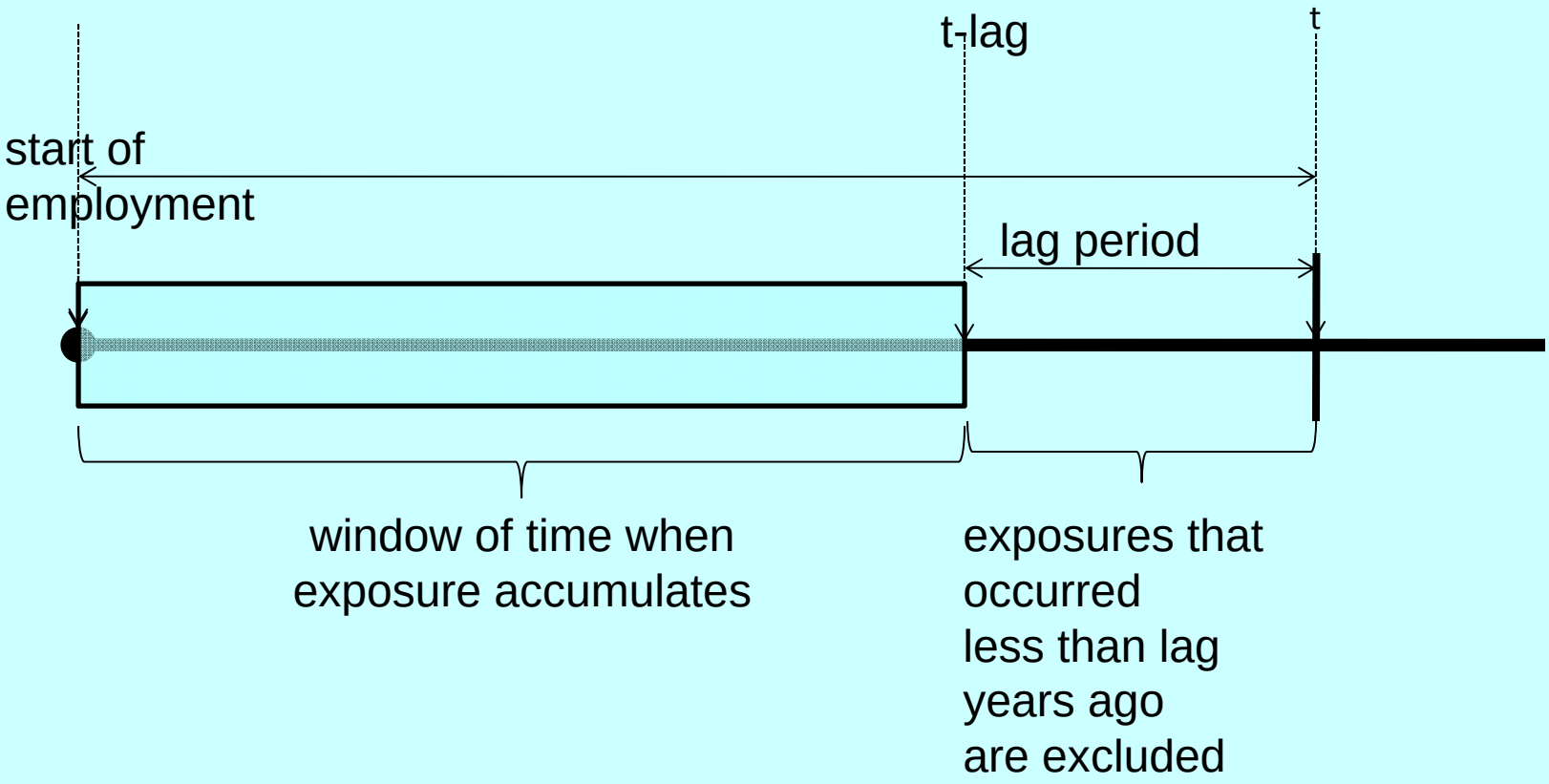
Exposure-Response Modeling - Covariates
Statistical Significance (p-value) of Non-Exposure and Exposure Covariates the Model's fit to the UAB data

Covariate	Leukemia	CLL	CML	AML	Lymphoid	Myeloid
Years since hire	0.3591	0.8696	0.2738	0.7785	0.8725	0.4532
Calendar year	0.2672	0.1544	0.3484	-0.3223	0.5723	0.4733
Race	0.5286	0.1960	0.9383	-0.3836	0.2502	0.7280
Plant	0.1399	0.0809	0.6164	0.0802	0.3659	0.5610
STY (cumulative ppm-years)	0.2019	0.4096	0.3567	0.3019	0.2117	0.1372
DMDTC (cumulative mg/cm-years)	0.0007**	0.1196	0.0823	-0.0065**	0.4832	0.0027**
Cumulative # of BD HITs	0.0001**	0.0901	-0.0287*	-0.0977	0.2595	-0.0001**
Cumulative # of STY HITs	0.0002**	0.5195	-0.0716	-0.0038**	0.0408*	-0.0001**
BD ≤ 100 ppm (cumulative ppm-years)	0.2078	0.2084	-0.3779	0.2989	0.3488	0.1249
BD > 100 ppm (cumulative ppm-years)	0.0108*	0.3542	-0.2155	-0.0924	0.1450	-0.0256*
STY ≤ 50 ppm (cumulative ppm-years)	0.0301*	0.3277	-0.4270	0.2837	0.1809	0.1847
STY > 50 ppm (cumulative ppm-years)	6.02×10^{-5}**	0.2326	0.2721	-0.0057**	0.5358	-0.0036**

Exposure-Response Modeling - Lag

The BD cumulative exposure metric considered simple lags

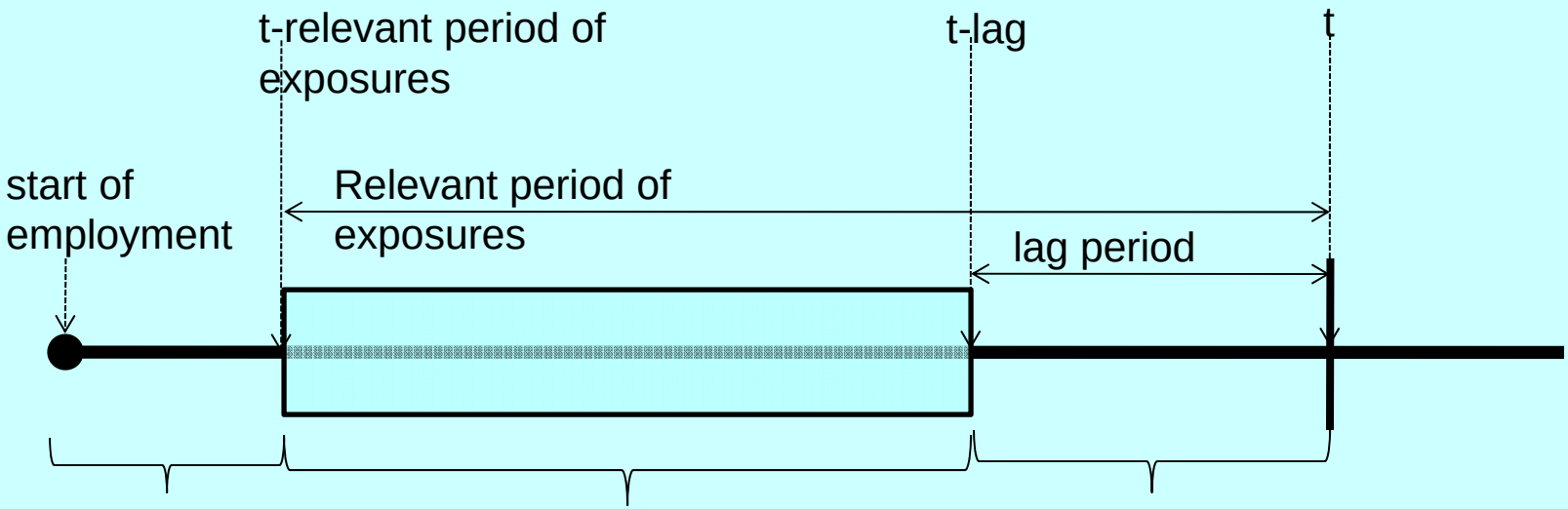
in which exposures in the preceding specified number of years are excluded



Exposure-Response Modeling - Lag

The BD cumulative exposure metric also considered simple exposure windows

in which exposures in the preceding specified number of years are excluded in addition to excluding exposures more than a specified number of years into the past



Exposure-Response Modeling - Lag

The Cox model fit with exposure lags of 5, 10, …, 35 years

does NOT significantly (at the 1% significance level)

improve the Cox model fit with no exposure lag

for any of the six endpoints analyzed

Table 4
The variability in the risk characterizations (exposure–response model slope and its significance) when the exposure–response modeling is restricted to different percentages of the exposure range.

Cox proportional hazards model with adjustment for HITs Restricted exposure range				Slope (Beta) of exposure–response model				Statistical significance of slope		
Person – years	Percentage among 71 leukemia mortalities with positive exposures (%)	Percentage among all workers (%)	# of leukemia mortalities	Maximum likelihood estimate	Standard error	95% Lower confidence limit	95% Upper confidence limit	Likelihood ratio test	Wald test	SS[a] or not SS
ALL (⩽9274)	100	100	81	0.00020	0.00013	−0.00001	0.00041	0.1931	0.1232	Not SS
⩽7750	100	99.9	81	0.00028	0.00014	0.00004	0.00052	0.1067	0.0529	Not SS
⩽2015	99	99	80	0.00050	0.00040	−0.00015	0.00115	0.2282	0.2057	Not SS
⩽1123	95	98	77	0.00058	0.00058	−0.00037	0.00153	0.3292	0.3137	Not SS
⩽401	75	91	63	0.00027	0.00162	−0.00240	0.00293	0.8695	0.8698	Not SS
⩽185	50	81	46	−0.00912	0.00483	−0.01706	−0.00118	0.0437	0.0593	SS and negative
⩽34	25	55	28	−0.02507	0.02667	−0.06894	0.01880	0.3349	0.3472	Not SS

[a] Statistically significant at 5% significance level.

Table 5
Maximum likelihood estimate of the slope for cumulative BD ppm-years (in the Cox proportional hazards model with cumulative BD ppm-years as a continuous variable and the only exposure variable), standard error, confidence interval, and statistical significance for different maximum levels of cumulative BD ppm-years for deaths in which leukemia is the primary cause of death or a contributing cause of death: Total leukemia.

Cumulative Butadiene ppm-years intervals included in the estimation	Maximum likelihood estimate of the slope	Standard error of the estimate of the slope	95% Confidence interval on the slope		Likelihood ratio test that slope = 0 (*p*-value)
All	0.00029	0.00010	0.00009	0.00049	0.0263*
⩽1338 ppm-years	0.00121	0.00036	0.00051	0.00191	0.0024**
⩽1000 ppm-years	0.00145	0.00047	0.00054	0.00236	0.0045**
⩽500 ppm-years	0.00296	0.00086	0.00127	0.00465	0.0014**
⩽400 ppm-years	0.00283	0.00111	0.00065	0.00501	0.0156*
⩽300 ppm-years	0.00305	0.00155	0.00001	0.00609	0.0591
⩽200 ppm-years	0.00089	0.00267	−0.00434	0.00612	0.7415
⩽100 ppm-years	0.00224	0.00536	−0.00827	0.01275	0.6793

* Statistically significant at the 5% significance level.
** Statistically significant at the 1% significance level.

Table 6
Characteristics of the data sets evaluated in Table 5.

Cumulative butadiene ppm-years intervals included in the estimation	# of Person years	Observed # of leukemia deaths: leukemia is primary or contributing cause of death	Observed # of leukemia deaths: leukemia is primary cause of death	Expected[a] # of leukemia deaths: leukemia is primary cause of death	SMR[b]: observed/expected: leukemia is primary cause of death	Power[c] to detect a slope of 0.005	Power to detect a slope of 0.0099[d]
All	500,378 (100%)	81 (100%)	68 (100%)	55	124	0.97	1.00
⩽1338 ppm-years	494,348 (98.8%)	78 (96%)	65 (96%)	54	121	0.82	1.00
⩽1000 ppm-years	489,694 (97.9%)	75 (93%)	62 (91%)	53	117	0.78	1.00
⩽500 ppm-years	469,876 (93.9%)	68 (84%)	58 (85%)	50	116	0.41	0.97
⩽400 ppm-years	461,234 (92.2%)	63 (78%)	54 (79%)	49	111	0.44	0.96
⩽300 ppm-years	445,524 (89.0%)	57 (70%)	50 (74%)	46	108	0.37	0.89
⩽200 ppm-years	419,682 (83.9%)	48 (59%)	45 (66%)	43	105	0.26	0.70
⩽100 ppm-years	364,411 (73.8%)	41 (51%)	38 (56%)	36	106	0.10	0.25

[a] Using US male mortality rates.
[b] Standardized mortality ratio (SMR).
[c] Using the Beaumont and Breslow (1981) approach, the statistical power to detect increases in leukemia deaths in the UAB cohorts, when the expected number of cancer deaths is based on a specified slope (coefficient of BD cumulative ppm-years in the linear rate ratio model) is given by the probability of a standard normal random variable exceeding $1.645 - 2 \times (RR^{0.5} - 1) \times (Expected)^{0.5}$ where RR is the ratio of the number of leukemia deaths predicted to the number of leukemia deaths observed and *Expected* is the number of leukemia deaths expected to occur due to background leukemia rates.
[d] 0.0099 is the value of the slope used in USEPA (2002).

Table 5
Maximum likelihood estimate of the slope for cumulative BD ppm-years (in the Cox proportional hazards model with cumulative BD ppm-years as a continuous variable and the only exposure variable), standard error, confidence interval, and statistical significance for different maximum levels of cumulative BD ppm-years for deaths in which all leukemia, CLL, or lymphoid neoplasm is the primary cause of death or a contributing cause of death.

Cumulative BD ppm-years included in the estimation	Maximum likelihood estimate of the slope	Standard error of the estimate of the slope	95% Confidence interval on the slope		*p*-Value in test that slope = 0	
			Lower limit	Upper limit	Likelihood ratio test	Wald's test
All leukemia[b]						
All	0.00029	0.00010	0.00009	0.00049	0.0263*	0.0047**
⩽1338 ppm-years[e]	0.00121	0.00036	0.00051	0.00191	0.0024**	0.0007**
⩽1000 ppm-years	0.00145	0.00047	0.00054	0.00236	0.0045**	0.0018**
⩽500 ppm-years	0.00296	0.00086	0.00127	0.00465	0.0014**	0.0006**
⩽400 ppm-years	0.00283	0.00111	0.00065	0.00501	0.0156*	0.0105*
⩽300 ppm-years	0.00305	0.00155	0.00001	0.00609	0.0591	0.0482*
⩽200 ppm-years	0.00089	0.00267	−0.00434	0.00612	0.7415	0.7391
⩽100 ppm-years	0.00224	0.00536	−0.00827	0.01275	0.6793	0.6759
CLL[c]						
All	0.00042	0.00013	0.00017	0.00067	0.0195*	0.0011**
⩽1338 ppm-years	0.00160	0.00057	0.00048	0.00272	0.0145*	0.0052**
⩽1000 ppm-years	0.00210	0.00073	0.00067	0.00353	0.0110*	0.0040**
⩽500 ppm-years	0.00411	0.00147	0.00123	0.00699	0.0101*	0.0051**
⩽400 ppm-years	0.00403	0.00192	0.00027	0.00779	0.0505	0.0361*
⩽300 ppm-years	0.00033	0.00352	−0.00657	0.00723	0.9253	0.9249
⩽200 ppm-years	−0.00560	0.00661	−0.01856	0.00736	0.3625	0.3969
⩽100 ppm-years	−0.00195	0.01090	−0.02331	0.01941	0.8568	0.8583
Lymphoid neoplasms[d]						
All	0.00023	0.00010	0.00004	0.00042	0.0546	0.0199*
⩽1338 ppm-years	0.00084	0.00033	0.00021	0.00148	0.0164*	0.0094*
⩽1000 ppm-years	0.00107	0.00041	0.00027	0.00187	0.0140*	0.0082*
⩽500 ppm-years	0.00157	0.00078	0.00005	0.00309	0.0524	0.0427*
⩽400 ppm-years	0.00171	0.00094	−0.00014	0.00356	0.0813	0.0704
⩽300 ppm-years	−0.00061	0.00149	−0.00353	0.00231	0.6787	0.6821
⩽200 ppm-years	−0.00259	0.00232	−0.00714	0.00196	0.2498	0.2649
⩽100 ppm-years	−0.00862	0.00498	−0.01838	0.00114	0.0679	0.0836

Shaded rows have both *p*-values > 0.05.
[a]Exposure–response models include all exposures.
[b]Originally reported in Table 5 in Sielken and Valdez-Flores (2011) based on likelihood ratio test.
[c]Originally reported in Table 8 in Sielken and Valdez-Flores (2011) based on likelihood ratio test.
[d]Analogous results using Poisson modeling reported in Sielken et al. (2007).
[e]1,338 is the 97.5-th percentile of BD ppm-years among the all leukemia cases.
*Statistically significant at the 5% significance level.
**Statistically significant at the 1% significance level.