UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

WASHINGTON, D.C. 20460

November 14, 2014

MEMORANDUM

SUBJECT: Additional information requested by Dr. Lianne Sheppard

FROM: Jennifer Jinot, Chemical Assessment Manager for ethylene oxide, NCEA

THRU: David Bussard, Director, Washington Division, NCEA

TO: Aaron Yeow, EPA Science Advisory Board

This memorandum provides some additional information requested by Dr. Lianne Sheppard, a member of the SAB panel reviewing the EPA draft assessment of ethylene oxide.

Please share this information with Dr. Sheppard, the augmented CAAC and with the public via the SAB website. The initial text in each bullet is a quote of the direct request as it was received from Dr. Sheppard. This is then followed by a response and information from EPA. For example, while EPA does not have the primary NIOSH data, we have provided additional information where possible in response to Dr. Sheppard's questions. We will be glad to respond to any further questions Dr. Sheppard or other members of the panel may have.

- "Please provide the data indicated in footnote 15 on page 4-3."

  EPA does not have the primary NIOSH data. EPA arranged for the additional analyses presented in the draft assessment to be conducted by Dr. Kyle Steenland, who was formerly one of the NIOSH investigators and analysts involved in the NIOSH EtO study. Dr. Steenland's submission to EPA is provided in Appendix D of the draft assessment. However, we provide below the additional information that we can in response to Dr. Sheppard's questions from available summaries or analyses of the NIOSH data.

- "Provide a "Table 1" for each of the analysis datasets giving summary statistics for demographic variables, key covariates, potential confounders, duration of observation, duration of employment, year of first employment, etc. Separate columns for cases and noncases."

  While EPA does not have the NIOSH data and therefore cannot provide these exact summary statistics, some summary statistics are available for the cancer mortality cohort as a whole from the NIOSH publications [Steenland et al. (1991, 2004); Stayner et al. (1993)], with some additional information provided by Dr. Steenland in Appendix D:

- the cohort is 45% male, 55% female; 79% white, 16% black, 5% other;
- the covariates considered in the mortality analyses were calendar year, age at risk, sex, and race.
- among the criteria considered in selection of the facilities for the study were known EtO exposure and the absence of any known confounding exposures;
- the average year of first exposure was 1970;
- the mean length of follow-up from first employment was 26.8 years (SD 8.5 years) at the end of the latest follow-up (through 1998);
- the mean duration of exposure of exposure at the end of the latest follow-up (through 1998) was 8.7 years (SD 9.3 years).

Some summary statistics are available for the breast cancer incidence cohort as a whole from the NIOSH publication [Steenland et al. (2003)], with some additional information provided by Dr. Steenland in Appendix D (note that the breast cancer incidence cohort is a subset of the mortality study cohort):

- cases and controls were matched on race;
- the covariates considered in the incidence analyses for the full breast cancer incidence cohort were calendar year, age, and race;
- the variables that were considered in the subcohort with interviews (5139 of the 7576 women in the full breast cancer incidence cohort) were body mass index, breast cancer in a first-degree relative, parity, age at menopause, age at menarche, socioeconomic status, and diet. Of these, only parity and breast cancer in a first-degree relative were important predictors for the cohort and only these were included in the final models. Information on some of these variables and on some exposure metrics for the cases and non-cases in the subcohort with interviews is provided in Table 2 of Steenland et al. (2003):

Table 2. Description of cases and non-cases with interview data[a]

Variable	Cases (n = 233)	Non-cases (n = 4906)
% Nulliparous	15.0%	11.6%
% With first-degree relative with breast cancer	16.3%	10.3%
% Pre-menopausal at diagnosis	14.4%	n.a.
Mean year of birth	1932 (s.d. 11.3)	1938 (s.d. 12.6)
Mean number of children	2.29 (s.d. 3.52)	2.36 (s.d. 3.34)
Mean BMI age 20	20.8 (1.6)	21.0 (1.6)
Median cumulative exposure	14.0 ppm-years	8.4 ppm-years
Means years exposed	13.0 (s.d. 9.2)	10.9 (s.d. 9.4)

[a] Based on those with complete interview data for parity and breast cancer in first degree relatives. Somewhat fewer subjects had complete data for menopausal status and BMI.

- for the full breast cancer incidence cohort, the mean duration of exposure at the end of follow-up (through 1998) was 10.7 years (SD 9.2 years) (median 7.4 years);
- non-respondents had a median birth year of 1937 and a median cumulative exposure of 8.0 ppm -years; the corresponding figures for respondents were 1938 and 8.6 ppm-years.

• "Summarize the exposure distribution for the analysis datasets (e.g. nested case-control). Give descriptive statistics and consider showing as two histograms, one for cases and one for all controls selected into all risk sets."

Some summary statistics are available for the cancer mortality cohort as a whole at the end of follow-up (through 1998) (with no lag) from the NIOSH publication by Steenland et al. (2004):

- mean cumulative exposure was 26.9 ppm-years (SD 65.7 ppm-years) (median 5.6 ppm-years);
- cumulative exposure was higher for males (mean 37.8, SD 87.6, median 7.6) than for females (mean 18.2, SD 38.2, median 4.6).

Some summary statistics are available for the breast cancer incidence cohort as a whole at the end of follow-up (through 1998) (with no lag) from the NIOSH publication [Steenland et al. (2003)], with some additional information provided by Dr. Steenland in Appendix D (note that the breast cancer incidence cohort is a subset of the mortality study cohort):

- information on some exposure metrics for the cases and non-cases in the subcohort with interviews is provided in Table 2 of Steenland et al. (2003), which is reproduced in the above response;
- cumulative exposure at the end of follow-up, with no lag, had a mean of 13,524 ppm-days (37.0 ppm-years), with a standard deviation of 13,254 ppm-days. These data are highly skewed, with a range from 5 to 253,848 ppm-days. The 25th percentile is 926 ppm-days, while the 75th is 10,206 ppm-days;
- some information on the cumulative exposure ranges for the cumulative exposure deciles with a 15-year lag is provided in Table D-1a of Appendix D:

Cumulative exposure, 15-year lag	**Mean cumulative exposure (ppm-days)**
0 (Lagged out)	
>0–355 ppm-days	157
356–842 ppm-days	580
843–1361 ppm-days	1097
1362–2187 ppm-days	1725
2188–3772 ppm-days	2899
3773–5522 ppm-days	4546

5523–7891 ppm-days	6554
7892–14483 ppm-days	14384
14484–25112 ppm-days	18859
>25112 ppm-days	48807

- "Provide the SEs of the linear combination of coefficients for the spline models after the knot
  - P. D-58: Need the SE of spl1+spl2
  - Table 4-3: add in the slope after the know with its SE"

    For the maximum likelihood estimate for exposures above the knot, RR = 1 + (β1 × exp + β2 × (exp – knot)).

    For the 95% upper confidence limit for exposures above the knot, RR = 1 + (β1 × exp + β2 × (exp-knot) + 1.645 × sqrt(exp^2 × var1 + (exp-knot)^2 × var2 + 2 × exp × (exp-knot) × covar)), where exp = cumulative exposure, var = variance, covar = covariance.

    Parameters and equations for the different two-piece spline models are summarized in the attached EXCEL spline model spreadsheet.

- "Add to the tables (Table 4-2, 4-7, others) the RR estimates and their 95% CIs at the exposure means. Also calculate and present for the exposure medians. Make sure to include the category means and medians in the footnotes, not just the ranges."

    EPA does not have the lagged cumulative exposure means and medians. Approximate RR estimates from the various models for different cumulative exposure levels can be seen in the figures in the draft assessment (e.g., Figure 4-1 on page 4-9 for lymphoid cancer; Figure 4-3 on page 4-25 for breast cancer mortality; and Figure 4-5 on page 4-34 for breast cancer incidence [in the subcohort with interviews]). The 95% CIs at a given exposure level can be calculated from the relevant model equations.

    The category means and medians for the categories in Table 4-2 are:

    > Mean exposures for both sexes combined with a 15-year lag for the categorical exposure quartiles in Table 4-2 were 446; 2,143; 7,335; and 39,927 ppm × days. Median values were 374; 1,985; 6,755; and 26,373 ppm × days. These values are for the full cohort, not just the risk sets.

    The category means and medians for the categories in Table 4-7 are:

    > Mean exposures for females with a 20-year lag for the categorical exposure quartiles in Table 8 of Steenland et al. (2004) were 276; 1,453; 5,869; and 26,391 ppm × days. Median values

were 250; 1,340; 5,300; and 26,676 ppm × days. These values are for the risk sets but should provide a good approximation to the full cohort values.

The category means and medians for the categories in Table 4-10 are:

Mean exposures for females with a 15-year lag for the exposure categories in Table 3 of Steenland et al. (2003) were 280; 1,241; 3,304; 8,423; and 36,022 ppm × days. Median values were 253; 1,193; 3,241; 7,741; and 26,597 ppm × days. These values are for the risk sets but should provide a good approximation to the full cohort values.

- "Add additional modeling results in Table 4-3 to facilitate comparisons. In particular, the log-cumulative exposure results should be included. Also the estimated slope after the knot and its SE. (These are obtained as a linear combination of B1 and B2)."

  The additional modeling results for comparison can be found in Table 4-2 on page 4-6 of the draft assessment. Spline results in Section 4.1 focus on the low-exposure spline segments because the Section is focused on deriving (low-exposure) unit risk estimates. As noted, the slopes above the knots can be calculated from the beta1 and beta2 estimates, which can all be found in Appendix D.

- "Clarify calculations, and information used in the calculations, most likely in an appendix:
  - Go through all the details of one set of unit risk estimate calculations. Provide the rationale for each adjustment factor, etc. Show an approximate "back of the envelope" estimate using overall (i.e. age-adjusted rather than age-specific) estimates of background rates."

    The sample life-table calculation provided in Appendix E (see also attached EXCEL life-table spreadsheet) for lymphoid cancer incidence goes through the steps for calculating the LEC_{01} (different exposure values can be tried in column K to zero in on the one that yields an extra risk of 0.01). From that, the unit risk is calculated as $0.01/LEC_{01}$.

    For a sample "back-of-the-envelope" calculation, please see the attached EXCEL calculation spreadsheet.

  - "Please connect the extra risk formula (4-2) on p 4-7 with the unit risk calculations. Please define the terms more clearly and present in the context of examples."

    The calculation of Rx, Ro, and the extra risk can be seen in the life-table example provided in Appendix E.

  - "Provide a table of sample ppm-day to ppm conversions for use in the EC01 calculations. Give conversion values at the knots and also for some of the results. (A few of these conversions are in table footnotes, e.g. Table 4-5.)"

The conversions are done within the life-table calculation (see example in Appendix E). The conversions provided in the footnotes to Table 4-5 are crude calculations assuming an average lifespan for the purposes of gauging whether or not calculated EC_{01} estimates for the two-piece spline models are appropriately in the range of the lower spline segments, since only those parts of the spline models were used to derive the EC_{01} estimates.

- "Please provide a table of overall estimates of background rates (i.e. crude, age-adjusted) for all cancer outcomes considered in this document."

The age-adjusted background rates (per year) are as follows (2007-2011 rates from SEER website, except where noted):

lymphoid cancer mortality in both sexes: 11.7/100,000[1]

lymphoid cancer incidence in both sexes: 31.9/100,000[2]

breast cancer mortality in females: 22.2/100,000

breast cancer incidence in females: 124.6/100,000

[1] Sum of non-Hodgkin lymphoma and myeloma 2007-2011 rates from SEER plus lymphoid leukemia 2007 rate from NCHS.
[2] Sum of non-Hodgkin lymphoma, chronic lymphocytic leukemia, acute lymphocytic leukemia, and myeloma 2007-2011 rates from SEER.

Parameters and equations for EtO two-piece spline models; cumulative exposure units of ppm × days

breast cancer incidence models (subcohort with interviews; with 15-year lag)

model	knot	beta1	SE1	beta2	SE2	var1	var2	covar
log-linear	5800	0.0000770	0.0000317	-0.0000724	0.0000334	1.00E-09	1.12E-09	-1.E-09
linear	5800	0.000119	0.0000677	-0.000105	0.0000705	4.58E-09	4.97E-09	-4.64E-09

equations and sample calculations (using the **linear** two-piece spline model) for the cumulative exposure of 3772 ppm-days, which is the cutpoint between the bottom 5 deciles and the top 5 deciles in terms of the # of cases, with 15-year lag (Table D-1a):

exp	RR	95% UCL
3772	1.45	1.87

and for a cumulative exposure above the knot (7544 ppm-days, i.e., 2 x 3772):

exp	RR	95% UCL
7544	1.71	2.36

breast cancer mortality models (with 20-year lag)

model	knot	beta1	SE1	beta2	SE2	var1	var2	covar
log-linear	700	0.000688	0.000417	-0.000678	0.000419	1.74E-07	1.76E-07	-1.75E-07
log-linear	13000	0.0000607	0.0000309	-0.0000583	0.0000371	9.55E-10	1.38E-09	NR
linear	700	0.000830	0.000614	-0.000807	0.000619	3.77E-07	3.83E-07	-3.80E-07

NR = not reported

no sample calculations for breast cancer mortality are presented here, because EPA did not use any of these spline models to develop risk estimates.

lymphoid cancer mortality models (with 15-year lag)

model	knot	beta1	SE1	beta2	SE2	var1	var2	covar
log-linear	100	0.01010	0.00493	-0.01010	0.00493	2.43E-05	2.43E-05	NR
log-linear	1600	0.000489	0.000255	-0.000486	0.000256	6.50E-08	6.55E-08	NR
linear	100	0.010090	0.004458	-0.01009	0.004458	1.99E-05	1.99E-05	-2.52E-05

NR = not reported

equations and sample calculations (using the **log-linear** two-piece spline model with the knot at 1600) for the cumulative exposure of 1227 ppm-days, which is 1/3 the value used in the above-the-knot example below.

exp	RR	95% UCL
1227	1.82	3.05

and for a cumulative exposure above the knot: 3680 ppm-days, which is the cutpoint between the bottom 2 quartiles and the top 2 quartiles in terms of the # of cases, with 15-year lag (Table D-3a):

exp	RR	95% UCL
3680	2.20	4.29

assuming covar of -6.53E-08

all lymphohematopoietic cancer mortality models (with 15-year lag)

model	knot	beta1	SE1	beta2	SE2	var1	var2	covar
log-linear	500	0.002010	0.000773	-0.00201	0.000774	5.98E-07	5.99E-07	NR
linear	500	0.003673	0.002345	-0.003668	0.002345	5.50E-06	5.50E-06	-5.70E-06

no sample calculations for all lymphohematopoietic cancer mortality are presented here, because EPA did not use any of these spline models to develop risk estimates.

NR = not reported

@ 0.0114 ppm env exp (LEC01 for lymphoid cancer incidence in M&F; Rothman linear regression (w/o highest exp gp); cum exp; 15y lag)

age interval	all cause mort (x E5/year)	lymphoid inc (x E5/year)	all cause hazard rate (h*)	prob of surv interval (q)	prob of surv up to interval (S)	lymphoid hazard rate (h)	cond prob of lymph incid in int (R0)
<1	685.2	1.9	0.0069	0.9932	1	0.0000	0.00002
1-4	29.9	8.1	0.0012	0.9988	0.993171421	0.0003	0.00032
5-9	14.7	4.2	0.0007	0.9993	0.991984298	0.0002	0.00021
10-14	18.7	3.2	0.0009	0.9991	0.991255458	0.0002	0.00016
15-19	66.1	3.5	0.0033	0.9967	0.990329067	0.0002	0.00017
20-24	94	3.2	0.0047	0.9953	0.987061432	0.0002	0.00016
25-29	96	4.1	0.0048	0.9952	0.982433129	0.0002	0.00020
30-34	107.9	6	0.0054	0.9946	0.977728749	0.0003	0.00029
35-39	151.7	9	0.0076	0.9924	0.972468106	0.0005	0.00044
40-44	231.7	13.2	0.0116	0.9885	0.965119839	0.0007	0.00063
45-49	352.3	20.9	0.0176	0.9825	0.954003442	0.0010	0.00099
50-54	511.7	32.5	0.0256	0.9747	0.937345814	0.0016	0.00150
55-59	734.8	49.2	0.0367	0.9639	0.913668011	0.0025	0.00221
60-64	1140.1	70.1	0.0570	0.9446	0.880709012	0.0035	0.00300
65-69	1727.4	101.1	0.0864	0.9173	0.83190835	0.0051	0.00403
70-74	2676.4	128.7	0.1338	0.8747	0.763071914	0.0064	0.00460
75-59	4193.2	163	0.2097	0.8109	0.667495253	0.0082	0.00491
80-84	6717.2	179.8	0.3359	0.7147	0.541245129	0.0090	0.00413
85+	13823.5	168.9			0.386840429		
							0.027965459
	2004 data	2000-2004 SEER					

extra risk = 0.010009375

A	B	C	D	E	F	G	H
Interval number (i)	**Age interval**	**All cause mortality**	**lymphoid cancer incidence (×10⁵/yr)**	**All cause hazard rate**	**Prob of surviving interval (q)**	**Prob of surviving up to**	**lymphoid cancer hazard rate (h)**
1	<1	685.2	1.9	0.0069	0.9932	1.0000	0.0000
2	1-4	29.9	8.1	0.0012	0.9988	0.9932	0.0003
3	5-9	14.7	4.2	0.0007	0.9993	0.9920	0.0002
4	10-14	18.7	3.2	0.0009	0.9991	0.9913	0.0002
5	15-19	66.1	3.5	0.0033	0.9967	0.9903	0.0002

6	20-24	94	3.2	0.0047	0.9953	0.9871	0.0002
7	25-29	96	4.1	0.0048	0.9952	0.9824	0.0002
8	30-34	107.9	6	0.0054	0.9946	0.9777	0.0003
9	35-39	151.7	9	0.0076	0.9924	0.9725	0.0005
10	40-44	231.7	13.2	0.0116	0.9885	0.9651	0.0007
11	45-49	352.3	20.9	0.0176	0.9825	0.9540	0.0010
12	50-54	511.7	32.5	0.0256	0.9747	0.9373	0.0016
13	55-59	734.8	49.2	0.0367	0.9639	0.9137	0.0025
14	60-64	1140.1	70.1	0.0570	0.9446	0.8807	0.0035
15	65-69	1727.4	101.1	0.0864	0.9173	0.8319	0.0051
16	70-74	2676.4	128.7	0.1338	0.8747	0.7631	0.0064
17	75-59	4193.2	163	0.2097	0.8109	0.6675	0.0082
18	80-84	6717.2	179.8	0.3359	0.7147	0.5412	0.0090
							Ro =

extra risk = (Rx−Ro)/(1−Ro) = 0.01001

exp duration mid int (XTime)	cum exp mid int (Xdose)	exp lymphoid haz rate (hx)	exp all cause haz rate (h*x)	exp prob of surv int (qx)	exp prob surv up to int (Sx)	exp cond prob of lymphoid in int (Rx)
0	0.00	0.00002	0.0069	0.9932	1.0000	0.00002
0	0.00	0.00032	0.0012	0.9988	0.9932	0.00032
0	0.00	0.00021	0.0007	0.9993	0.9920	0.00021
0	0.00	0.00016	0.0009	0.9991	0.9913	0.00016
2.5	31.64	0.00018	0.0033	0.9967	0.9903	0.00018
7.5	94.92	0.00017	0.0047	0.9953	0.9871	0.00017
12.5	158.20	0.00022	0.0048	0.9952	0.9824	0.00022
17.5	221.49	0.00034	0.0054	0.9946	0.9777	0.00033
22.5	284.77	0.00052	0.0077	0.9924	0.9724	0.00050
27.5	348.05	0.00079	0.0117	0.9884	0.9650	0.00075
32.5	411.33	0.00128	0.0179	0.9823	0.9538	0.00121
37.5	474.61	0.00205	0.0260	0.9743	0.9369	0.00190
42.5	537.90	0.00319	0.0375	0.9632	0.9128	0.00286
47.5	601.18	0.00467	0.0582	0.9435	0.8793	0.00399
52.5	664.46	0.00691	0.0882	0.9156	0.8296	0.00549
57.5	727.74	0.00902	0.1364	0.8725	0.7595	0.00640
62.5	791.02	0.01171	0.2132	0.8080	0.6627	0.00699
67.5	854.31	0.01323	0.3401	0.7117	0.5354	0.00601
						0.03769

I	J	K	L	M	N	O	P
Cond prob of lymphoid	**Exp duration mid interval**	**Cum exp mid interval**	**Exposed lymphoid**	**Exposed all cause hazard**	**Exposed prob of surviving**	**Exposed prob of surviving up to**	**Exposed cond prob**
0.00002	0	0.00	0.00002	0.0069	0.9932	1.0000	0.00002
0.00032	0	0.00	0.00032	0.0012	0.9988	0.9932	0.00032
0.00021	0	0.00	0.00021	0.0007	0.9993	0.9920	0.00021
0.00016	0	0.00	0.00016	0.0009	0.9991	0.9913	0.00016
0.00017	2.5	31.64	0.00018	0.0033	0.9967	0.9903	0.00018

0.00016	7.5	94.92	0.00017	0.0047	0.9953	0.9871	0.00017
0.00020	12.5	158.20	0.00022	0.0048	0.9952	0.9824	0.00022
0.00029	17.5	221.49	0.00034	0.0054	0.9946	0.9777	0.00033
0.00044	22.5	284.77	0.00052	0.0077	0.9924	0.9724	0.00050
0.00063	27.5	348.05	0.00079	0.0117	0.9884	0.9650	0.00075
0.00099	32.5	411.33	0.00128	0.0179	0.9823	0.9538	0.00121
0.00150	37.5	474.61	0.00205	0.0260	0.9743	0.9369	0.00190
0.00221	42.5	537.90	0.00319	0.0375	0.9632	0.9128	0.00286
0.00300	47.5	601.18	0.00467	0.0582	0.9435	0.8793	0.00399
0.00403	52.5	664.46	0.00691	0.0882	0.9156	0.8296	0.00549
0.00460	57.5	727.74	0.00902	0.1364	0.8725	0.7595	0.00640
0.00491	62.5	791.02	0.01171	0.2132	0.8080	0.6627	0.00699
0.00413	67.5	854.31	0.01323	0.3401	0.7117	0.5354	0.00601
0.02797						**Rx =**	0.03769

"back-of-the-envelope" calculation for breast cancer incidence EC01

EC01 (ppm lifetime constant exposure) (from Table 4-13 of assessment)

0.0112 ppm

conversion to ppm-days of occupational exposure:

10 m3/day 8-hour occ breathing rate
20 m3/day 24-hour non-occ breathing rate
240 working days per year
365 days per year
75 years approximate average lifetime in life-table analysis
60 years of exposure in lifetime, with 15-year lag (75-15 =60)

746.06 ppm-days (lagged)

for the linear two-piece spline model:

beta1 = 0.000119 per ppm-day (from Table 4-11 of assessment) (RR = 1 + beta*exposure)

at 746.06 ppm-days RR = 1.088781

Ro = background lifetime risk = background annual rate * 75 years = 124.6/100,000/year * 75 years = 0.09345

Rx = (total) lifetime risk in exposed = Rx = RR * Ro = 0.101747

extra risk = (Rx - Ro)/(1 - Ro) = 0.009152

thus, this crude calculation roughly recreates the extra risk of 0.01 for which the EC01 was derived.

or, working in the other direction:

for 0.01 extra risk, (Rx-Ro)/(1-Ro) = 0.01, so Rx = (0.01* (1-Ro)) + Ro = 0.102516

RR = Rx/Ro = 1.097009

occ exp = (RR-1)/beta = 815.2025 ppm-days

env exp = 0.012238 ppm

which is close to the derived EC01 of 0.0112 ppm.

for the unit risk:

for the LEC01, the same calculation applies but using the model: RR = 1 + exposure* (beta1 + 1.645*SE1)
(using the 95% upper bound on risk to derive the 95% lower bound on exposure)

then unit risk = 0.01/LEC01 (per ppm)