APPENDIX H

Lifetable Analysis and Weighted Linear Regression based on Results from Charbotel et al. (2006)

CONTENTS—Appendix H: Lifetable Analysis and Weighted Linear Regression based on Results from Charbotel et al. (2006)

APPENDIX H: LIFETABLE ANALYSIS AND WEIGHTED LINEAR REGRESSION BASED ON RESULTS FROM CHARBOTEL ET AL. (2006) H-1
H.1. LIFETABLE ANALYSIS H-1
H.2. EQUATIONS USED FOR WEIGHTED LINEAR REGRESSION OF RESULTS FROM CHARBOTEL ET AL. (2006) H-1
H.3. REFERENCES H-4

APPENDIX H: LIFETABLE ANALYSIS AND WEIGHTED LINEAR REGRESSION BASED ON RESULTS FROM CHARBOTEL ET AL. (2006)

H.1. LIFETABLE ANALYSIS

A spreadsheet illustrating the extra-risk calculation for the derivation of the lower 95% bound on the effective concentration associated with a 1% extra risk (LEC_{01}) for renal cell carcinoma (RCC) incidence is presented in Table H-1.

H.2. EQUATIONS USED FOR WEIGHTED LINEAR REGRESSION OF RESULTS FROM CHARBOTEL ET AL. (2006) (source: Rothman [1986], p. 343-344)

Linear model: $RR = 1 + bX$

where RR = risk ratio, X = exposure, and b = slope

b can be estimated from the following equation:

$$\hat{b} = \frac{\sum_{j=2}^{n} w_j x_j R\hat{R}_j - \sum_{j=2}^{n} w_j x_j}{\sum_{j=2}^{n} w_j x_j^{\,2}} \qquad \text{(Eq. H-1)}$$

where j specifies the exposure category level and the reference category ($j = 1$) is ignored.

The standard error of the slope can be estimated as follows:

$$SE(\hat{b}) \approx \sqrt{\frac{1}{\sum_{j=2}^{n} w_j x_j^{\,2}}} \quad . \qquad \text{(Eq. H-2)}$$

The weights, w_j, are estimated from the confidence intervals (as the inverse of the variance):

$$Var(R\hat{R}_j) \approx R\hat{R}_j^{\,2} Var[\ln(R\hat{R}_j)] \approx R\hat{R}_j^{\,2} \times \left[\frac{\ln(\overline{RR}_j) - \ln(\underline{RR}_j)}{2 \times 1.96}\right]^2 \qquad \text{(Eq. H-3)}$$

where $\overline{RR}_j$ is the 95% upper bound on the RR_j estimate (for the jth exposure category) and $\underline{RR_j}$ is the 95% lower bound on the RR_j estimate.

Table H-1. Extra-risk calculation[a] for environmental exposure to 1.82 ppm TCE (the LEC_{01} for RCC incidence)[b] using a linear exposure-response model based on the categorical cumulative exposure results of Charbotel et al. (2006), as described in Section 5.2.2.1.2.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Interval number (*i*)	**Age interval**	**All cause mortality ($\times 10^5$/yr)**	**RCC incidence ($\times 10^5$/yr)**	**All cause hazard rate (*h**)**	**Prob. of surviving interval (*q*)**	**Prob. of surviving up to interval (*S*)**	**RCC cancer hazard rate (*h*)**	**Cond. prob. of RCC incidence in interval (*Ro*)**	**Exp. duration mid interval (*x*time)**	**Cum. exp. mid interval (*x*dose)**	**Exposed RCC hazard rate (*hx*)**	**Exposed all cause hazard rate (*h*x*)**	**Exposed prob. of surviving interval (*qx*)**	**Exposed prob. of surviving up to interval (*Sx*)**	**Exposed cond. prob. of RCC in interval (*Rx*)**
1	<1	685.2	0	0.0069	0.9932	1.0000	0.000000	0.000000	0.5	2.77	0.000000	0.0069	0.9932	1.0000	0.000000
2	1-4	29.9	0	0.0012	0.9988	0.9932	0.000000	0.000000	3	16.61	0.000000	0.0012	0.9988	0.9932	0.000000
3	5-9	14.7	0	0.0007	0.9993	0.9920	0.000000	0.000000	7.5	41.52	0.000000	0.0007	0.9993	0.9920	0.000000
4	10-14	18.7	0.1	0.0009	0.9991	0.9913	0.000005	0.000005	12.5	69.20	0.000006	0.0009	0.9991	0.9913	0.000006
5	15-19	66.1	0.1	0.0033	0.9967	0.9903	0.000005	0.000005	17.5	96.88	0.000006	0.0033	0.9967	0.9903	0.000006
6	20-24	94	0.2	0.0047	0.9953	0.9871	0.000010	0.000010	22.5	124.56	0.000013	0.0047	0.9953	0.9871	0.000013
7	25-29	96	0.7	0.0048	0.9952	0.9824	0.000035	0.000034	27.5	152.24	0.000049	0.0048	0.9952	0.9824	0.000048
8	30-34	107.9	1.6	0.0054	0.9946	0.9777	0.000080	0.000078	32.5	179.91	0.000117	0.0054	0.9946	0.9777	0.000114
9	35-39	151.7	3.2	0.0076	0.9924	0.9725	0.000160	0.000155	37.5	207.59	0.000245	0.0077	0.9924	0.9724	0.000237
10	40-44	231.7	6.3	0.0116	0.9885	0.9651	0.000315	0.000302	42.5	235.27	0.000504	0.0118	0.9883	0.9650	0.000484
11	45-49	352.3	11	0.0176	0.9825	0.9540	0.000550	0.000520	47.5	262.95	0.000919	0.0180	0.9822	0.9537	0.000869
12	50-54	511.7	17.3	0.0256	0.9747	0.9373	0.000865	0.000801	52.5	290.63	0.001507	0.0262	0.9741	0.9367	0.001393
13	55-59	734.8	26.2	0.0367	0.9639	0.9137	0.001310	0.001175	57.5	318.31	0.002375	0.0378	0.9629	0.9124	0.002127
14	60-64	1140.1	36.2	0.0570	0.9446	0.8807	0.001810	0.001549	62.5	345.99	0.003409	0.0586	0.9431	0.8786	0.002909
15	65-69	1727.4	44.6	0.0864	0.9173	0.8319	0.002230	0.001777	67.5	373.67	0.004358	0.0885	0.9153	0.8286	0.003456
16	70-74	2676.4	49	0.1338	0.8747	0.7631	0.002450	0.001750	72.5	401.35	0.004961	0.1363	0.8726	0.7584	0.003518
17	75-59	4193.2	51.6	0.2097	0.8109	0.6675	0.002580	0.001554	77.5	429.03	0.005407	0.2125	0.8086	0.6617	0.003223
18	80-84	6717.2	44.4	0.3359	0.7147	0.5412	0.002220	0.001021	82.5	456.71	0.004809	0.3384	0.7129	0.5351	0.002183
							***Ro* =**	0.010736						***Rx* =**	0.020586
Extra risk = (*Rx* − *Ro*)/(1 − *Ro*) = 0.00996															

Column A: interval index number (i).
Column B: 5-year age interval (except <1 and 1−4) up to age 85.
Column C: all-cause mortality rate for interval i ($\times 10^5$/year) (2004 data from NCHS [2007]).
Column D: RCC incidence rate for interval i ($\times 10^5$/year) (2001-2005 SEER data [http://seer.cancer.gov]).
Column E: all-cause hazard rate for interval i (h^*_i) [= all-cause mortality rate × number of years in age interval].[c]
Column F: probability of surviving interval i without being diagnosed with RCC (q_i) [= $\exp(-h^*_i)$].
Column G: probability of surviving up to interval i without having been diagnosed with RCC (S_i) [$S_1 = 1$; $S_i = S_{i-1} \times q_{i-1}$, for i > 1].
Column H: RCC incidence hazard rate for interval i (h_i) [= RCC incidence rate × number of years in interval].
Column I: conditional probability of being diagnosed with RCC in interval i [= $(h_i/h^*_i) \times S_i \times (1-q_i)$], i.e., conditional upon surviving up to interval i without having been diagnosed with RCC [Ro, the background lifetime probability of being diagnosed with RCC = the sum of the conditional probabilities across the intervals].
Column J: exposure duration (in years) at mid-interval ($x$time).
Column K: cumulative exposure mid-interval ($x$dose) [= exposure level (i.e., 1.82 ppm) × 365/240 × 20/10 × $x$time] (365/240 × 20/10 converts continuous environmental exposures to corresponding occupational exposures).
Column L: RCC incidence hazard rate in exposed people for interval i (hx_i) [= $h_i \times (1 + \beta \times x\text{dose})$, where $\beta = 0.001205 + (1.645 \times 0.0008195) = 0.002554$] [0.001205 per ppm × year is the regression coefficient obtained from the weighted linear regression of the categorical results (see Section 5.2.2.1.2). To estimate the LEC_{01}, i.e., the 95% lower bound on the continuous exposure giving an extra risk of 1%, the 95% upper bound on the regression coefficient is used, i.e., MLE + 1.645 × SE].
Column M: all-cause hazard rate in exposed people for interval i (h^*x_i) [= $h^*_i + (hx_i - h_i)$].
Column N: probability of surviving interval i without being diagnosed with RCC for exposed people (qx_i) [= $\exp(-h^*x_i)$].
Column O: probability of surviving up to interval i without having been diagnosed with RCC for exposed people (Sx_i) [$Sx_1 = 1$; $Sx_i = Sx_{i-1} \times qx_{i-1}$, for i> 1].
Column P: conditional probability of being diagnosed with RCC in interval i for exposed people [= $(hx_i/h^*x_i) \times Sx_i \times (1-qx_i)$] (Rx, the lifetime probability of being diagnosed with RCC for exposed people = the sum of the conditional probabilities across the intervals).

[a] Using the methodology of BEIR IV (1988).
[b] The estimated 95% lower bound on the continuous exposure level of TCE that gives a 1% extra lifetime risk of RCC.
[c] For the cancer incidence calculation, the all-cause hazard rate for interval i should technically be the rate of either dying of any cause or being diagnosed with the specific cancer during the interval, i.e., (the all-cause mortality rate for the interval + the cancer-specific incidence rate for the interval—the cancer-specific mortality rate for the interval [so that a cancer case isn't counted twice, i.e., upon diagnosis and upon death]) × number of years in interval. This adjustment was ignored here because the RCC incidence rates are small compared with the all-cause mortality rates.

MLE = maximum likelihood estimate, SE = standard error.

H.3. REFERENCES

BEIR ([Committee on the] Biological Effects of Ionizing Radiation). (1988) Health risks of radon and other internally deposited alpha emitters. BEIR IV. Washington, DC: National Academy Press.

Charbotel B, Fevotte J, Hours M, Martin J-L, Bergeret A. (2006) Case-control study on renal cell cancer and occupational exposure to trichloroethylene. Part II: epidemiological aspects. Ann Occup Hyg 50: 777-787.

NCHS (National Center for Health Statistics). (2007) National Vital Statistics Reports, vol. 55, no. 19; August 21, 2007, Table 3. National Center for Health Statistics, Hyattsville, MD.

Rothman KJ. (1986) *Modern Epidemiolgy*. Little, Brown and Company, Boston.