Regulatory Toxicology and Pharmacology 65 (2013) 311–315

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

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

A novel bottom-up approach to bounding low-dose human cancer risks from chemical exposures

Thomas B. Starr [a,b,*], James A. Swenberg [b]

[a] TBS Associates, 7500 Rainwater Road, Raleigh, NC 27615, USA
[b] Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, 166 Rosenau Hall, CB #7431, Chapel Hill, NC 27599-7431, USA

ARTICLE INFO

Article history:
Received 29 July 2012
Available online 23 January 2013

Keywords:
Bottom-up approach
Carcinogenic risk assessment
Unit risk
q_1^*
Upper 95% confidence bound risk estimate
DNA adducts
Formaldehyde

ABSTRACT

We propose a novel bottom-up approach to the bounding of low-dose human cancer risks from chemical exposures that does not rely at all upon high-dose data for human or animal cancers. This approach can thus be used to provide an independent "reality check" on low-dose risk estimates derived with dose-response models that are fit to high-dose cancer data. The approach (1) is consistent with the "additivity to background" concept, (2) yields central and upper-bound risk estimates that are linear at all doses, and (3) requires only information regarding background risk, background (endogenous) exposure, and the additional exogenous exposure of interest in order to be implemented. After describing the details of this bottom-up approach, we illustrate its application using formaldehyde as an example. Results indicate that recent top-down risk extrapolations from occupational cohort mortality data for workers exposed to formaldehyde are overly conservative by substantial margins.

1. Background

In 1976, Kenny Crump, David Hoel, Charles Langley, and Richard Peto published a landmark paper (Crump et al., 1976) showing that a non-decreasing dose–response relationship for cancer risk will be linear at sufficiently low doses as long as there is a non-zero background exposure to which the specific chemical exposures of interest simply add. This is the well-known "additivity to background" concept: at zero additional exposure, we are already somewhere up on the dose–response curve as a result of our non-zero background exposure, so the slope of the dose–response relationship at zero additional exposure will necessarily be non-zero and positive. Even a threshold dose–response relationship will have a non-zero slope at zero additional exposure if there are some individuals in the population of interest whose thresholds lie below their non-zero background exposure.

Then, in 1977, Crump, Harry Guess, and K.L. Deal published another landmark paper (Crump et al., 1977) that outlined the statistical and mathematical procedures for estimating and bounding the low-dose slope of the multistage dose–response model using constrained maximum likelihood methods and tumor data collected in laboratory animal bioassays conducted at very high exposure levels. It was in this paper that the now infamous "q_1^*", the upper 95% confidence bound on the coefficient of the linear term (i.e., the low-dose slope) of the presumed dose–response relationship, was created, and this value has dominated carcinogenic risk assessment ever since.

The dominance of q_1^* in risk assessment has been a consequence of two factors. First, there is the tyranny of small numbers, i.e., the small numbers of animals that have been utilized in laboratory animal carcinogenicity studies, typically, only about 50 animals per sex per dose group. This number is so small that even if the observed tumor incidence in a treated group is zero (0/50), the exact binomial upper 95% confidence bound on the true response probability is 0.0582, so true risks up to this value cannot be confidently ruled out. It is also not possible to distinguish statistically at the $p = 0.05$ level between a response as high as 0.08 (4/50 tumor-bearing animals) in a treated group and a null response (0/50) in a control group using Fisher's exact test. If the goal of risk assessment is to bound the dose of a chemical that is associated with an upper bound incremental cancer risk of only one per million (10^{-6}), then one can conservatively "guesstimate" the required dose, using the low-dose linear hypothesis, as being about 100,000-fold lower than the highest dose that produces no significant increase, compared to controls, in the probability of developing cancer. This is common knowledge among biostatisticians, and a source of frustration and heartburn among many toxicologists; it is, nevertheless, an irrefutable "fact of life".

The second factor behind the dominance of q_1^* is that until recently, the background exposures that may be responsible, at least in part, for our background cancer risks have not been quantified (two notable exceptions are radiation and our background body

* Corresponding author.
E-mail addresses: tbstarr@mindspring.com (T.B. Starr), jswenber@email.unc.edu (J.A. Swenberg).

0273-2300
http://dx.doi.org/10.1016/j.yrtph.2013.01.004

burdens of dioxin-like compounds). Generally, little attention has been focused on quantifying background chemical exposures, and the exposures of interest have routinely been expressed as increments above whatever the background exposures might be. This is primarily due to the fact that human background exposures are complicated, uncontrolled, and usually unmeasured, while the animal studies that attempt to carefully control and minimize these background exposures have not routinely included measures of the corresponding internal (endogenous) doses that can arise via normal metabolism and other internal biochemical reactions.

Without knowing what background exposure is, expressed preferably as the concentration of a relevant exposure biomarker, e.g., DNA adducts, in the target tissue of interest, the only way to estimate the slope of the dose–response relationship at low doses has been via downward extrapolation from the observed tumor responses in small numbers of animals (or occupationally exposed people) at high external exposure levels, which forces us into the q_1^*conundrum. However, this situation has changed recently, and the change could profoundly alter carcinogenic risk assessment going forward, at least for those potentially carcinogenic substances that are always present in our bodies, even absent external exposure, because they are produced continuously by normal biochemical processes such as metabolism and biochemical synthesis and degradation. The key technological advance underpinning our novel "bottom-up" approach to risk assessment is the extraordinary ability to distinguish between and separately quantify the relevant internal exposures in target tissues that arise from internal background (endogenous) and external (exogenous) sources. In what follows, we outline this alternative approach to estimating and bounding low-dose cancer risks for such substances, and illustrate the potential for its application with the specific example of formaldehyde, an important commodity chemical that is currently under review by the US Environmental Protection Agency (USEPA).

2. The bottom-up approach

Let P_0 represent the background lifetime risk of a tissue-specific cancer in people, such as nasopharyngeal cancer or leukemia. Let C_0 represent the mean tissue-specific background steady-state concentration of a biomarker, such as a specific DNA adduct, that is presumed to be causally related to these cancers. Then the ratio P_0/C_0 provides an estimate of the low-dose slope of the relationship between the cancer risk and the corresponding tissue-specific DNA adduct concentration. Similarly, if C_{0L} represents the *lower* 95% confidence bound estimate for the same background adduct concentration, then the ratio P_0/C_{0L} provides an *upper* 95% confidence bound on the low-dose slope. This latter ratio is thus directly comparable to the q_1^* derived from high dose animal studies, as well as the upper bound slope estimates for the low-dose linear dose–response relationships that are typically inferred from epidemiologic analyses of occupational cohort cancer mortality, provided only that the dose metrics used in these two kinds of studies (animal bioassays and cohort mortality studies) are converted into the corresponding equivalent tissue-specific adduct concentrations.

The key elements of this bottom-up approach are illustrated in Fig. 1. What is most important to appreciate is that the central and upper bound slope estimates derived using this approach do not depend in any way on high-dose carcinogenicity data for laboratory animals or humans. The approach thus provides a completely independent "reality check" on low-dose slope estimates like q_1^* that are derived from analyses of high-dose laboratory animal tumor incidence data or occupational cancer mortality data.

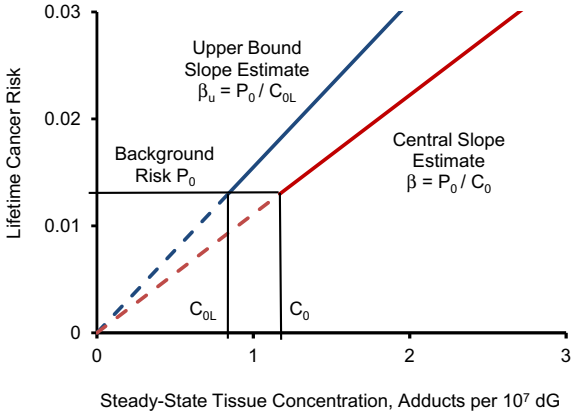

Fig. 1. Illustrating the "bottom-up" approach to bounding additional human cancer risks that may be associated with low level chemical exposures. P_0 is the background lifetime risk of a tissue-specific cancer. C_0 and C_{0L} are the central and lower 95% confidence bound estimates of the steady-state background concentration of specific DNA adducts linked to the cancer in the same tissue. β and β_u are the bottom-up central and upper 95% confidence bound estimates of the low-dose slope of the cancer risk-DNA adduct relationship.

3. An illustration of the bottom-up approach using currently available data for formaldehyde

Formaldehyde is a highly reactive chemical and an essential metabolic intermediate that is generated endogenously in all living cells, and N^2-hydroxymethyl-deoxyguanosine (dG) adducts have been detected and quantified in various tissues of rats (Lu et al., 2010 and 2011) and cynomolgus macaques (Moeller et al., 2011) exposed to various concentrations of stable isotope-labelled [$^{13}CD_2$]-formaldehyde by inhalation. These formaldehyde-DNA adducts are potentially promutagenic because adduction takes place on the amino groups participating in Watson–Crick base pairing, and adduct formation is widely considered to be a key event in the initiation of mutations that lead to carcinogenesis (Swenberg et al., 2011). Thus, the tissue-specific concentration of these adducts provides an excellent internal dose metric with which to illustrate the bottom-up approach to bounding the low-dose slope of dose–response relationships for human cancer risk.

The use of [$^{13}CD_2$]-formaldehyde permits the simultaneous measurement of both endogenous and exogenous formaldehyde-DNA adducts with sensitive Liquid Chromatography–Electrospray Ionization-Tandem Mass Spectrometry-Selected Reaction Monitoring (LC–ESI-MS/MS-SRM) methods. While endogenous dG adducts were detected in all of the examined tissues, exogenous dG adducts formed with inhaled [$^{13}CD_2$]-formaldehyde were detected only in the tissues taken from the site of initial contact with exogenous formaldehyde, i.e., rat and monkey nasal respiratory epithelium (Swenberg et al., 2011).

Because no exogenous dG adducts were detected in these studies in any distant site tissues, including bone marrow and the blood, we can state with confidence that if such exogenous adducts were present in these tissues, then their amounts would necessarily have been smaller than the LC–ESI-MS/MS-SRM method's detection limit (DL). We have therefore used the method's DL (reported in Moeller et al. (2011) as 20×10^{-18} mol) as a worst case upper bound on the level of exogenous dG adducts that could be present and yet remain undetected in the bone marrow of [$^{13}CD_2$]-formaldehyde-exposed monkeys. The above molar DL was converted to an equivalent DL expressed in terms of the number of adducts, namely, 1.03×10^{-3} per 10^7dG, using the average amount of monkey DNA collected in the bone marrow samples (Moeller et al., 2011), and the amount of guanine (0.20, expressed as a fraction) that is present in monkey DNA (Casanova et al., 1991).

The formaldehyde-DNA adduct data utilized in our bottom-up slope calculations are provided in Table 1. These values are the mean ± standard error of the number of endogenous and exogenous dG adducts per 10^7dG in nasal respiratory epithelium (2.49 ± 0.23 and 0.25 ± 0.020, respectively), and the bone marrow (17.5 ± 1.31, endogenous dG adducts only) as determined in monkeys following two 6 h exposures to 2 ppm [$^{13}CD_2$]-formaldehyde (data taken from Table 3 in Swenberg et al., 2011). Also presented are the lower 95% confidence bound estimates for endogenous dG adducts in both tissues, i.e., the mean values minus 1.645 times their respective standard errors.

We have also estimated the corresponding steady-state exogenous dG adduct levels that would result from continuous 24 h/day, 7 days/week exposure. To accomplish this, we used the adduct levels measured in monkeys by Moeller et al. (2011) immediately after the two 6 h exposures (30 h after the onset of the first exposure), together with a simple one compartment linear kinetic model of adduct buildup and elimination with a 63 h elimination half-life (mean adduct lifetime $T = 63/\ln(2) = 90.9$ h) as has been determined in rats (Swenberg et al., 2012). For example, if C_{x30} represents the measured exogenous DNA adduct concentration after two 6 h exposures on consecutive days to a given airborne formaldehyde concentration, and C_{xS-S} represents the model-predicted asymptotic steady-state adduct concentration that would result from continuous exposure to the same airborne formaldehyde concentration, then $C_{xS-S} = C_{x30}/\{[1-\exp(-6/T)] \times [1 + \exp(-24/T)]\} = 8.85^{*}C_{x30}$. The steady-state adduct concentrations that are predicted by this formula to arise from continuous lifetime exposure to 2 ppm [$^{13}CD_2$]-formaldehyde are also provided in Table 1.

At present, we do not have estimates of endogenous or exogenous dG adduct concentrations in human tissues, so we have made the simple assumption that the DNA adduct data collected by Moeller et al. (2011) in cynomolgus macaques are directly relevant to humans without any interspecies scaling adjustments. For the background lifetime risks of developing nasopharyngeal cancer (NPC) and leukemia (LEU), we have relied on two different sources. For NPC, we have taken the estimate of 7.25×10^{-4} that is provided in USEPA's 2 June 2010 draft formaldehyde assessment (see Table C-1, p C-3 and Section 5.2.2). For leukemia, we used the Both Sexes, All Race lifetime risk estimate of 1.3×10^{-2} from Table 1.14 of *SEER Cancer Statistics Review 1975–2007* (Altekruse et al., 2010).

Table 2 presents the results from using the bottom-up approach with these data and assumptions to calculate upper bound estimates of human nasopharyngeal cancer and leukemia risk from lifetime continuous exposure to 1 ppm formaldehyde. To obtain bottom-up estimates corresponding to 1 ppm formaldehyde, we first calculated bottom-up estimates for 2 ppm (the lowest exposure level used by Moeller et al. (2011), and then simply divided those estimates by a factor of two, since the bottom-up approach assumes linearity of the dose–response relationship. We chose 1 ppm so as to be able to compare our risk estimates simply and directly with those derived by USEPA from epidemiologic data using cumulative formaldehyde exposure as the dose metric, namely 0.011 ppm^{-1} for NPC and 0.057 ppm^{-1} for leukemia (see Table 6–3, pp 6–41- 6–42 of the Agency's draft assessment).

For nasopharyngeal cancer (NPC), the bottom-up upper bound risk estimate (0.038×10^{-2}) is nearly 29-fold lower than USEPA's "plausible upper bound" estimate of 1.1×10^{-2}, i.e., about 1%. In contrast, the bottom-up upper bound estimate of leukemia risk ($<3.9 \times 10^{-6}$) is more than 14,000-fold lower than the corresponding USEPA estimate of 5.7×10^{-2}, i.e., about 6%. The marked disparity between these estimated cancer risks for this distant site suggests strongly that the excess risk of leukemia that has been reported in association with workplace formaldehyde exposures is not due to those exposures. If our plausible assumption that formaldehyde dG adducts provide a valid molecular dosimeter for relating potential human cancer risks to formaldehyde exposure is correct, then the much larger risks derived by USEPA from the adult human cancer data are overly conservative.

4. Strengths and limitations of the bottom-up approach

We are confident that the estimates obtained from this simple approach to bounding low-dose human cancer risks are conservative for several reasons. Most importantly, the bottom-up approach attributes *all* of the background risk of specific cancers to the endogenous formaldehyde dG adducts that are found in the corresponding tissues. If only a fraction f of the total background risk P_0 were due to the background endogenous adduct concentration C_0, then it is only this fraction of the total background risk P_0 to which our assumed linear dose–response relationship should apply. The slope estimate P_0/C_0 and the associated upper 95% confidence bound described herein would therefore exaggerate the actual slope of the low-dose response and its upper estimated 95% confidence bound by the factor $1/f$.

Table 1
Endogenous and exogenous DNA adduct concentrations (per 10^7dG) in nasal epithelial tissue and bone marrow of cynomolgus macaques exposed via inhalation for 6 h on two consecutive days to 2 ppm [$^{13}CD_2$]-formaldehyde (data taken from Moeller et al. (2011)). Also shown are the 8.85-fold higher steady-state exogenous adduct concentrations that are expected to result from lifetime continuous inhalation exposure to 2 ppm [$^{13}CD_2$]-formaldehyde (see text for details).

Tissue	Endogenous adducts at 30 h	Exogenous adducts at 30 h	Exogenous adducts at steady-state
Nasal epithelium mean ± se	2.49 ± 0.23	0.250 ± 0.020	2.21 ± 0.18
lower 95% bound	2.11		
Bone marrow mean ± se	17.5 ± 1.31	<0.00103[a]	<0.00912[a]
lower 95% bound	15.34		

[a] No exogenous adducts were detected in bone marrow; upper limits estimate based on the detection limit reported in Moeller et al. (2011).

Table 2
Comparison of estimated lifetime risks of developing nasopharyngeal cancer (NPC) and leukemia (LEU) from continuous lifetime inhalation exposure to 1 ppm formaldehyde, as estimated with the bottom-up approach and, alternatively, by USEPA using top-down linear extrapolation from epidemiologic data (as taken from Table 6–3, pp 6–41-6–42 of the Agency's 2 June 2010 draft assessment).

Cancer	Background risk, P_0	Bottom-up slope, P_0/C_{0L}[a]	Bottom-up risk at 1 ppm[b]	USEPA risk at 1 ppm
NPC	7.25×10^{-4}	3.44×10^{-4}	0.038×10^{-2}	1.1×10^{-2}
LEU	1.30×10^{-2}	8.50×10^{-4}	$<3.9 \times 10^{-6}$	5.7×10^{-2}

[a] for NPC, $3.44 \times 10^{-4} = 7.25 \times 10^{-4}/2.11$
for LEU, $8.50 \times 10^{-4} = 1.30 \times 10^{-2}/15.3$.
[b] for NPC, $0.038 \times 10^{-2} = 3.44 \times 10^{-4} \times (2.21/2)$
for LEU, $<3.9 \times 10^{-6} = 8.50 \times 10^{-4} \times (<0.00912/2)$.

Second, the approach is linear at low doses simply because it assumes linearity at all doses. This can create problems if we extrapolate the bounding bottom-up risk estimates to very high exogenous exposure levels, such as those producing statistically significant increases in tumor incidence in exposed laboratory animals or humans. At such levels, it is expected that the dose–response relationship for tumor incidence may well be highly nonlinear due to a variety of factors that become important only at high doses, such as cytotoxicity, tissue damage, and enhanced cell proliferation that markedly increases the probability of mutations. Such non-genotoxic high dose phenomena are not accounted for in our simple linear model, so the confidence bounds that it generates should not be expected to hold at the high exogenous exposures where these phenomena take place.

Third, we have used *lower* 95% confidence bounds on the estimated mean endogenous DNA adduct levels (C_0) to generate, by simple inversion, the corresponding *upper* 95% confidence bounds on the slope (P_0/C_0) of the linear relationship that has been assumed between cancer risks and steady-state adduct concentrations. This follows directly from a Taylor series expansion of the ratio P_0/C_0 about its expected value when P_0 is taken to be constant, as we have assumed herein, and only the estimated mean background DNA adduct concentration C_0 has uncertainty associated with it. In this special case, the variance of the ratio P_0/C_0 is given approximately by (c.f., **Stuart and Ord, 1994**, p. 351):

$$\mathrm{Var}(P_0/C_0)/(P_0/C_0)^2 \approx \mathrm{Var}(C_0)/C_0^2 \quad (1)$$

If uncertainty in the estimate of the mean background risk P_0 is also characterized and it is independent of, i.e., uncorrelated with, the uncertainty in C_0, as is the case herein, because the estimates of P_0 and C_0 are derived from two completely different data sets, then the additional uncertainty in the estimated ratio P_0/C_0 that is due to the uncertainty in P_0 can also be readily accommodated (ibid.):

$$\mathrm{Var}(P_0/C_0)/(P_0/C_0)^2 \approx \mathrm{Var}(C_0)/C_0^2 + \mathrm{Var}(P_0)/P_0^2. \quad (2)$$

The upper 95% confidence bound on the slope estimate P_0/C_0 is thus increased as a result of incorporating the additional uncertainty in P_0. Additional discussion of uncertainty in the estimate of the ratio P_0/C_0 in a linear regression context is provided in the Appendix.

Fourth, for the case of leukemia, exogenous DNA adducts were never detected in monkey bone marrow even though the methodology has sufficient statistical power to detect a single N^2-hydroxymethyl-deoxyguanosine adduct in 10 billion deoxyguanosine molecules. We have therefore assumed, as a worst case, that exogenous DNA adducts could have been present at a level just barely below the detection limit of the ultrasensitive LC–ESI-MS/MS-SRM methodology. One could reduce this estimate substantially using less conservative assumptions regarding the sampling distribution of non-detected exogenous DNA adduct concentrations (c.f., Ginevan and Splitstone, 2003, pp 123–125).

Fifth, we have made reasonable assumptions in converting adduct concentrations measured in monkey tissues after two 6 h/day exposures to 2 ppm airborne formaldehyde on consecutive days to the higher steady-state adduct levels that would arise from continuous exposures to the same airborne formaldehyde concentration for a lifetime, but our extrapolation using a simple linear pharmacokinetic model has not yet been validated. Data from longer term studies out to 28 consecutive days of exposure are currently being analyzed, so the remaining uncertainty regarding the half-life of formaldehyde DNA adducts should be better resolved in due course.

Even so, the cross-species extrapolation from DNA adduct data obtained in monkeys to human formaldehyde exposures remains unvalidated. However, unvalidated assumptions can be replaced at some point with data-driven alternatives. For example, in the near future, we expect to obtain data regarding endogenous formaldehyde dG adducts in human tissues. Human blood samples are readily obtainable, and opportunistic sampling of other critical tissues such as nasal tissue and bone marrow is certainly possible. Such data could be used to confirm and/or replace our plausible dosimetric assumption that endogenous formaldehyde dG adduct amounts in monkey and human tissues are directly comparable. Obtaining exogenous adduct concentrations in humans may be more problematic. However, the extraordinary sensitivity of the Lu et al. (2010, 2011) and Moeller et al. (2011) methodology may offer the prospect of detecting such adducts using short-term voluntary human exposures.

Finally, another important limitation of the bottom-up approach is its reliance on the assumption of a linear dose–response relationship between cancer risks and DNA adduct concentrations in target tissues, even though the true dose–response may be highly nonlinear at sufficiently high exogenous exposure levels. For this reason, we advocate the bottom-up approach only as a potential means for generating tighter upper bounds on low-dose human cancer risks than it may be possible to achieve with top-down approaches. The bottom-up approach may not be appropriate for developing "best" or central estimates of low-dose human cancer risks which, at least in our view, can best be accomplished through a comprehensive and deep mechanistic understanding of how chemical exposures give rise to human cancer.

5. Summary

The Lu et al. (2010, 2011) and Moeller et al. (2011) LC–ESI-MS/MS-SRM methodology differentiates clearly between DNA adducts formed with formaldehyde molecules of endogenous and exogenous (inhaled) origin. This remarkable technological achievement has made it possible to develop upper-bound estimates of potential cancer risk with a unique bottom-up approach that extrapolates upward from background (endogenous) exposures and background cancer risks, as opposed to the typical top-down extrapolations from cancer incidence in laboratory animals or human workers subjected to very high exposure levels.

While we have illustrated the bottom-up approach with the example of formaldehyde, we expect it to be readily generalizable to other chemicals. For example, vinyl chloride, ethylene oxide, methanol, and acetaldehyde are all known to produce specific DNA adducts from endogenous and exogenous sources, and other chemicals are likely to be added to this list in the near future. The target tissue dose concept can also be generalized to include other forms of endogenous DNA damage, such as abasic sites, lesions arising from oxidative stress, and also biomarkers of effect, such as mutations (Swenberg et al., 2008, 2011). The potential of the bottom-up approach to impact human cancer risk assessment appears great.

We used the new molecular dosimetry information for formaldehyde DNA adducts in the bottom-up approach to estimate upper-bound lifetime human nasopharyngeal cancer and leukemia risks that might arise from continuous inhalation exposure to 1 ppm formaldehyde. This provides a totally independent "reality check" on estimates derived with the conventional top-down approach to human cancer risk assessment. Comparison of the resulting bottom-up risk estimates with corresponding top-down estimates derived by USEPA from epidemiologic data for exposed workers show the latter to be markedly higher. The large discrepancies between the results we obtained with molecular dosimetry data incorporated into the bottom-up approach and those that relied on worker cancer mortality and uncertain retrospective occupational exposure reconstructions call into serious question the credibility of attributing large increases in human mortality from these cancers to occupational formaldehyde exposure.

Conflict of interest statement

TBS has served as a consultant on risk assessment issues related to formaldehyde for the American Chemistry Council. The formaldehyde research conducted by JAS has been funded in part by the NIEHS, the American Chemistry Council, Formacare, and the Texas Commission for Environmental Quality. The sponsors do not have access to research results until they have been accepted for publication. JAS has also served as a formaldehyde consultant to ENVIRON International.

Acknowledgements

TBS received financial support for this work from the Research Foundation for Health and Environmental Effects, Washington DC. We also thank an anonymous reviewer for pointing out that estimating an upper bound on the slope of the dose-response relationship at and near the background exposure level can be considered in the context of linear regression through the origin.

Appendix A

An anonymous reviewer pointed out to us that estimation of an upper percentile confidence bound on the slope of the dose–response curve near the background exposure level C_0 can also be considered in the context of linear regression, with the regression line forced through the origin (c.f., Neter et al., 1996, pp 159–163). The situation we are concerned with here is a special one, because there is only one Y value, namely P_0, which is derived from national cancer statistics, i.e., the Y_i (P_{0i}) values that would be associated in a regression context with the n individual X_i measurements of the background endogenous DNA adduct concentration (C_{0i}) are actually all equal to the single mean background risk estimate P_0. Given this constraint, it is not difficult to show algebraically that the slope estimate b_1 resulting from linear regression through the origin of the equal Y_i on the individual C_{0i} is given by:

$$b_1 = P_0/C_0 \times /\{1 + ((n-1)/n) \times Var(C_0)/C_0^2\}, \quad (3)$$

where $Var(C_0)$ is the estimated variance of the C_{0i}. So long as $Var(C_0)$ is small compared to ${C_0}^2$, the term in curly brackets in Equation 3 will be close to unity in value, and the regression slope estimate b_1 will be approximately equal to P_0/C_0, but in any case, b_1 will always be *smaller* than the ratio P_0/C_0. Thus P_0/C_0 will be a conservative, i.e., larger, estimate of the low-dose slope than the linear regression estimate b_1 given in Equation 3.

In addition, the estimated variance of the regression slope estimate b_1 for this special case of equal Y_i can be shown algebraically to be given by:

$$Var(b_1) = b_1^2 \times [(Var(C_0)/n)/C_0^2]. \quad (4)$$

Thus, the percentiles of the sampling distribution for b_1 in this special case are exactly complementary to those of the sampling distribution of the mean background DNA adduct concentration C_0, i.e., an upper 95% confidence bound on the regression slope estimate b_1 corresponds exactly to a lower 95% confidence bound on the mean background DNA adduct concentration.

When the additional uncertainty associated with individual background risk estimates P_{0i} is considered, so long as there is no covariance between the P_{0i} and C_{0i}, as is the case here, the regression estimate of b_1 remains the same as that given by equation A1. However, the estimated variance of b_1 changes to:

$$Var(b_1)/b_1^2 = (Var(C_0)/n)/C_0^2 + (Var(P_0)/n)/P_0^2) \times [1 + ((n-1)/n) \times Var(C_0)/C_0^2], \quad (5)$$

where $Var(P_0)$ represents the variance of the individual background risk measurements P_{0i}. Thus, the upper 95% confidence bound estimate for the regression slope estimate b_1 is increased as a result of incorporating the additional uncertainty associated with the individual background risk estimates P_{0i}.

References

Altekruse, S.F., Kosary, C.L., Krapcho, M., Neyman, N., Aminou, R., Waldron, W., Ruhl, J., Howlader, N.,Tatalovich, Z., Cho, H., Mariotto, A., Eisner, M.P., Lewis, D.R., Cronin, K., Chen, H.S., Feuer, E.J., Stinchcomb, D.G., Edwards, B.K. (eds). 2010. *SEER Cancer Statistics Review, 1975-2007*. National Cancer Institute, Bethesda, MD. <http://seer.cancer.gov/csr/1975_2007/>, based on November 2009 SEER data submission, posted to the SEER web site, 2010.

Casanova, M., Morgan, K.T., Steinhagen, W.H., Everitt, J.I., Popp, J.A., Heck, H.d'A., 1991. Covalent binding of inhaled formaldehyde to DNA in the respiratory tract of rhesus monkeys: pharmacokinetics, rat-to-monkey interspecies scaling, and extrapolation to man. Fundam. Appl. Toxicol. 17, 409–428.

Crump, K.S., Guess, H.A., Deal, K.L., 1977. Confidence intervals and test of hypotheses concerning dose response relations inferred from animal carcinogenicity data. Biometrics 33, 43–451.

Crump, K.S., Hoel, D.G., Langley, H., Peto, R., 1976. Fundamental carcinogenic processes and their implications for low dose risk assessment. Cancer Res. 36, 2973–2979.

Ginevan, M.E., Splitstone, D.E., 2003. Statistical Tools for Environmental Quality Measurement. Chapman & Hall/CRC Applied Environmental Statistics, Boca Raton, FL.

Lu, K., Collins, L.B., Ru, H., Bermudez, E., Swenberg, J.A., 2010. Distribution of DNA adducts caused by inhaled formaldehyde is consistent with induction of nasal carcinoma but not leukemia. Toxicol. Sci. 116 (2), 441–451.

Lu, K., Moeller, B., Doyle-Eisele, M., McDonald, J., Swenberg, J.A., 2011. Molecular dosimetry of N(2)-hydroxymethyl-dG DNA adducts in rats exposed to formaldehyde. Chem. Res. Toxicol. 24 (2), 159–161.

Moeller, B.C., Lu, K., Doyle-Eisele, M., McDonald, J., Gigliotti, A., Swenberg, J.A., 2011. Determination of N^2-hydroxymethyl-dG adducts in the nasal epithelium and bone marrow of non-human primates following $^{13}CD_2$- formaldehyde inhalation exposure. Chem. Res. Toxicol. 24 (2), 162–164.

Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W., 1996. Applied Linear Statistical Models, 4th ed. WCB McGraw-Hill, Boston, MA, see pp. 159–163.

Stuart, A., Ord, J.K., 1994. Kendall's Advanced Theory of Statistics, Volume 1, Distribution Theory, 6th ed. Halsted Press, John Wiley and Sons, New York.

Swenberg, J.A., Fryar-Tita, E., Jeong, Y.C., Boysen, G., Starr, T.B., Walker, V.E., Albertini, R.J., 2008. Biomarkers in toxicology and risk assessment: informing critical dose-response relationships. Chem. Res. Toxicol. 21, 253–265.

Swenberg, J.A., Lu, K., Moeller, B.C., Gao, L., Upton, P.B., Nakamura, J., Starr, T.B., 2011. Endogenous versus exogenous DNA adducts their role in carcinogenesis epidemiology and risk assessment. Toxicol. Sci. 120 (S1), S130–S145.

Swenberg, J.A., Moeller B.C., Lu K., Rager, J.E., Fry, R., Starr, T.B. 2012. Formaldehyde carcinogenicity research: 30 years and counting for mode of action, epidemiology, and cancer risk assessment. *Toxicol. Path.* http://tpx.sagepub.com/content/early/2012/11/16/0192623312466459.