Uncertainties in the CIIT Model for Formaldehyde-Induced Carcinogenicity in the Rat: A Limited Sensitivity Analysis–I

Ravi P. Subramaniam,[1]* Kenny S. Crump,[2] Cynthia Van Landingham,[2] Paul White,[1] Chao Chen,[1] and Paul M. Schlosser[1]

Scientists at the CIIT Centers for Health Research (Conolly *et al.*, 2000, 2003; Kimbell *et al.*, 2001a, 2001b) developed a two-stage clonal expansion model of formaldehyde-induced nasal cancers in the F344 rat that made extensive use of mechanistic information. An inference of their modeling approach was that formaldehyde-induced tumorigenicity could be optimally explained without the role of formaldehyde's mutagenic action. In this article, we examine the strength of this result and modify select features to examine the sensitivity of the predicted dose response to select assumptions. We implement solutions to the two-stage cancer model that are valid for nonhomogeneous models (i.e., models with time-dependent parameters), thus accounting for time dependence in variables. In this reimplementation, we examine the sensitivity of model predictions to pooling historical and concurrent control data, and to lumping sacrificed animals in which tumors were discovered incidentally with those in which death was caused by the tumors. We found the CIIT model results were not significantly altered with the nonhomogeneous solutions. Dose-response predictions below the range of exposures where tumors occurred in the bioassays were highly sensitive to the choice of control data. In the range of exposures where tumors were observed, the model attributed up to 74% of the added tumor probability to formaldehyde's mutagenic action when our reanalysis restricted the use of the National Toxicology Program (NTP) historical control data to only those obtained from inhalation exposures. Model results were insensitive to hourly or daily temporal variations in DNA protein cross-link (DPX) concentration, a surrogate for the dose-metric linked to formaldehyde-induced mutations, prompting us to utilize weekly averages for this quantity. Various other biological and mathematical uncertainties in the model have been retained unmodified in this analysis. These include model specification of initiated cell division and death rates, and uncertainty and variability in the dose response for cell replication rates, issues that will be considered in a future paper.

KEY WORDS: Cell replication; DNA protein cross-links; formaldehyde; mutation; two-stage model

1. INTRODUCTION

Two modes of action are thought to be potentially relevant in the dose response for formaldehyde-induced cancer: 1) formaldehyde is thought to be directly mutagenic (Grafstrom *et al.*, 1985; Heck *et al.*, 1990; Speit & Merk, 2002); 2) at high exposures, formaldehyde is cytotoxic, resulting in cell injury and cell killing that induces increased cell replication (Monticello *et al.*, 1991, 1996). The relative importance of cytotoxicity compared with the direct mutagenicity to the observed tumorigenicity of the chemical is of considerable debate. The biologically-based modeling of formaldehyde carcinogenicity carried out

[1] NCEA, ORD, U.S. Environmental Protection Agency, Washington, DC, USA.
[2] ENVIRON International Corporation, Monroe, LA, USA.
* Address correspondence to Ravi P. Subramaniam, NCEA, ORD, U.S. Environmental Protection Agency, Mailcode 8623-D, 1200 Pennsylvania Ave. NW, Washington, DC 20460, USA; Subramaniam.Ravi@epa.gov.

by the CIIT Centers for Health Research (CIIT) further contributes to this debate (Conolly _et al._, 2003). In the CIIT model, DNA protein cross-links formed by formaldehyde (DPX) was considered a surrogate for an unknown mutagenic pathway (Conolly _et al._, 2003). The best fit of this model to the observed dose-response data was obtained when the probability of formaldehyde-induced mutations per cell generation was estimated to be zero, lending weight to the point of view that the mutagenic component is not important (Conolly _et al._, 2003, 2004; Slikker _et al._, 2004). In other words, the model explains the highly nonlinear tumor incidence seen in rats optimally by the effect on spontaneous mutations of the proliferative response induced at cytotoxic exposures. The authors of the CIIT model opted to make what they believed to be a conservative choice by 1) evaluating human risk at the upper statistical bound on the estimate for the proportionality constant relating the DPX concentration to the probability of mutation, and 2) by presenting results obtained with both a hockey-stick as well as a J-shaped dose-response curve for the cell replication rate (Conolly _et al._, 2004). The model predicted negative added risk at low dose with the J-shaped curve.

This article further advances the effort of Dr. Rory Conolly and other scientists at CIIT who have generously shared the bioassay data, other input files, and model code. In this article, we reimplement the CIIT modeling of the F344 rat bioassay data (Conolly _et al._, 2000, 2003; Kimbell _et al._, 2001a, 2001b) after modifying some of its features, including its use of data on historical control animals, while retaining other key assumptions and uncertainties, and examine the impact these changes have on the above inference. The changes we implement here serve as a base for carrying out further sensitivity analysis on other biological and mathematical uncertainties. The impact of other significant uncertainties in modeling the rat data and in extrapolation of these data to humans are being addressed separately.

The CIIT formaldehyde model for cancer is in the form of a two-stage clonal expansion model of cancer (Moolgavkar & Venzon, 1979; Moolgavkar & Knudson, 1981). The solution of the two-stage model equations employed in Conolly _et al._ (2003) assumed homogeneous (time-independent) parameters. However, several authors have developed solutions to the nonhomogeneous two-stage clonal expansion model as well as the more general multistage clonal expansion model (Crump _et al._, 2005a, 2006; Little, 1995; Little _et al._, 2002; Little & Wright, 2003; Moolgavkar

& Luebeck, 1990; Heidenreich _et al._, 1997; Portier _et al._, 2000, 1996). Crump _et al._ (2005a) determined that applying solutions derived for homogeneous models to nonhomogeneous models could, in general, lead to substantial error. Since parameters in the CIIT model are time-dependent, this article examines the impact of this assumption on the CIIT model; this issue was also examined in conference presentations by Crump _et al._ (2005b) and by Conolly _et al._ (2006).

The analysis by Conolly _et al._ (2003) treated tumors discovered incidentally in sacrificed animals the same as those found in animals that died naturally due to the tumor (which, in essence, assumes tumors are rapidly fatal). This was based on the consideration by these authors that tumors of this type were typically fatal within 1–2 weeks following detection in the laboratory (personal communication, R. Conolly). In the two formaldehyde bioassay data, however, animal sacrifices were scheduled at a number of times in addition to a final sacrifice of all surviving animals. In our reanalysis of these data, we found that 57 animals were observed to have tumors at the time of these sacrifices. In general, it is potentially problematic to assume that a tumor was fatal when it was discovered incidentally in an animal sacrificed at a scheduled time. Since we can readily differentiate between these two observations of tumor in a statistically rigorous manner, we expanded the CIIT model to eliminate this assumption.

In their analysis, Conolly _et al._ combined data on all historical control animals from National Toxicology Program (NTP) with the data from the two formaldehyde bioassays. We examine the sensitivity of model results to this procedure by 1) applying their model to only the concurrent data, and 2) including historical data only from NTP studies of inhalation exposures.

Other issues examined in this article are the extent to which uncertainty in estimating DPX half-life propagates to the two-stage modeling, and the sensitivity of the model to rapid temporal oscillations in DPX dynamics. In their physiologically-based pharmacokinetic model (PBPK) simulation of DPX concentrations, Conolly _et al._ (2000) assumed a value of 6.5×10^{-3} min^{-1} for the first-order rate constant for DPX clearance. This was the minimum value needed in the model to match the observation that virtually all of the DPX had cleared 18 hours following the cessation of exposure in a study in which rats were exposed to formaldehyde by inhalation (Casanova _et al._, 1994). Data from an _in vitro_ study indicate a slower clearance, with an average rate constant of

9.24×10^{-4} min^{-1} (Quievryn & Zhitkovitch, 2000). We discuss how this slower clearance rate could be consistent with the *in vivo* DPX data of Casanova *et al.*

The analyses presented here do not examine the sensitivity of the CIIT model results to uncertainties in the assumed dose response for initiated cell replication and death rates, to uncertainty and variability in the dose response for normal cell replication rates, and to the extrapolation of the normal cell replication dose response for beyond the range over which it was characterized from empirical data. The dose dependence of the growth advantage of initiated cell (I) division rates over those of normal (N) cells was given by the same functional form as assumed in Conolly *et al.* (2003).

2. METHODS

2.1. The CIIT Two-Stage Clonal Expansion Model

The CIIT two-stage clonal expansion model for formaldehyde (Conolly *et al.*, 2003) utilized a three-dimensional computational fluid dynamic (CFD) model (Kimbell *et al.*, 2001a, 2001b) to predict site-specific flux of formaldehyde from inhaled air into rat nasal tissue. The rat nasal airway surface was partitioned into 20 regions, termed "flux bins," each defined by a range of formaldehyde flux to the surface. The CFD model was used to predict, for each flux bin, the ratio of formaldehyde flux into the bin to the air concentration of formaldehyde (pmol/mm^2-h-ppm). The number of cells at risk for forming tumors in each flux bin of an adult animal was also estimated.

The two-stage clonal expansion model is defined in terms of the following parameters (the notation here is equivalent to, but differs from, that used by Conolly *et al.*):

N – number of normal cells that are eligible for progression to malignancy;
α_N – division rate of normal cells (h^{-1});
μ_N – rate at which an initiated cell is formed by mutation of a normal cell (per cell division of normal cells);
α_I – division rate of an initiated cell (h^{-1});
β_I – death rate of an initiated cell (h^{-1});
μ_I – rate at which a malignant cell is formed by mutation of an initiated cell (per cell division of initiated cells).

In the CIIT model, each of these parameters takes on a different value in each flux bin, resulting effec-

tively in 20 two-stage models. Definitions of these parameters in the Conolly *et al.* model and their modifications in our analysis are described below.

2.2. Cell Division and Death Rates

Several steps were involved in the construction of a dose-response relationship for normal cell replication rates (α_N) in Conolly *et al.* (2003) using cell labeling index data (Monticello *et al.*, 1991, 1996). In these studies, F344 rats were exposed to formaldehyde at the same six air concentrations as in the formaldehyde bioassays, and unit labeling indexes were measured at six different sites in the rat nose and at eight different times after the beginning of exposure (1, 4, and 9 days, and 6, 13, 26, 52, and 78 weeks). First, the data from separate pulse labeling (Monticello *et al.*, 1991) and continuous labeling experiments (Monticello *et al.*, 1996) were averaged over replicate animals and pooled. For a given formaldehyde exposure concentration, these unit-length labeling indexes were time-weighted and averaged over the six sites. Second, the unit-length labeling index was related to the labeling index, using data from an experiment where both quantities were measured (Monticello *et al.*, 1990). Third, the labeling index values so derived were related to cell replication rates using a formula due to Moolgavkar and Luebeck (1992) derived for continuous labeled data. Fourth, formaldehyde fluxes computed using steady state CFD simulations were also averaged over the six sites. The modeling in Conolly *et al.* was based on these averages even though the labeling index varied both over anatomical site and with time at each air concentration, and flux and the number of cells at risk varied over site.[3] Cell replication rate (α_N) versus flux then showed a "J-shaped" curve; for averaged flux $\leq 1.24 \times 10^3$ pmol/mm^2-h (corresponding to exposures at 0.7 and 2 ppm), α_N was less than its value in unexposed animals. For this portion of the dose-response curve, Conolly *et al.* (2003) also considered a "hockey-stick" shaped dose-response curve with a threshold at 2 ppm, and progressively increasing rates at higher exposures. We refer the reader to Conolly *et al.* (2003) for further details of their model and to Table 1 of that paper for the derived time-weighted average of cell division rates as a function of flux averaged over the sites. Details

---

[3] Reasons given for this decision included uncertainty in how to extrapolate temporal variations in cell division rates to humans, and site-to-site variations in cell division rates that were not consistent with predicted flux at these sites.

pertaining to the partitioning of the rat nasal lining into flux bins are provided in Table 1 of Kimbell *et al.* (2001a).

Whereas the flux values in the tabulation of flux versus cell division rates in Conolly *et al.* (2003) only went up to 9.3×10^3 pmol/mm^2-h, the predicted fluxes in the flux bins went up to 39.3×10^3 pmol/mm^2-h. Therefore, Conolly *et al.* had to extrapolate the flux-cell division rate relationship to higher flux values. This was accomplished by assuming a maximum division rate (α_{max}, an estimated parameter) at the highest flux value of 39.3×10^3 pmol/mm^2-h and linearly interpolating between this point and the flux-cell division rate point corresponding to a flux of 9.3×10^3 pmol/mm^2-h. This upward extrapolation has been retained in the current reimplementation.

There are no separate data on the cell division rate of initiated cells (α_I), and Conolly *et al.* (2003) assumed this rate to be the following function of α_N, the division rate of normal cells:

$$\alpha_I = \alpha_N \times \left(multb - multfc \times [\alpha_N - \alpha_{Nbasal}]^+\right),$$

where α_{Nbasal} is the empirically obtained average cell division rate in unexposed normal cells, multb and multfc are estimated parameters, and the "$[\;]^+$" stands for the maximum of the quantity in the brackets and zero. The value of α_{Nbasal} was equal to 3.39×10^{-4} h^{-1} as determined by Conolly *et al.* (2003) from the raw averaged labeling index data. Conolly *et al.* (2003) state that this formulation provided the best fit of the model to the tumor data. The death rate for initiated cells was set equal to the cell division rate of normal cells for all formaldehyde flux values, i.e., $\beta_I = \alpha_N$. These formulations for α_I and β_I were retained in our reimplementation; alternatives to these will be studied in future work. In our reimplementation, we ensured $\alpha_I > 0$ by constraining the parameter optimization so that $1 < multb < 10$ and $0 < multfc < 10$.

2.3. Mutation Rates and DPX

Conolly *et al.* (2003) suggested that the mutagenic role for DPX in their current model could be interpreted as a surrogate for some other unknown mutagenic pathway, and provided a discussion of uncertainty about the role of DPX formation in mutation. DPX concentrations were estimated from a study by Casanova *et al.* (1994) in which rats were exposed for 11 weeks (5 days/week) + 4 days, 6 hours/day, to filtered air (naive) or to 0.7, 2, 6, or 15 ppm formaldehyde (preexposed). On the 5th day of the 12th week, the rats were then exposed for 3 hours to 0, 0.7, 2,

6, or 15 ppm ^{14}C-labeled formaldehyde (with preexposed animals exposed at the same concentration as the preceding 12 weeks and 4 days), the animals were sacrificed, and DPX concentrations were determined at two sites in the nasal mucosa. Conolly *et al.* (2000) used the naive[4] rat data to develop a PBPK model that predicted the time-course of DPX concentrations as a function of formaldehyde flux at these sites (estimated in the CFD modeling). In Conolly *et al.* (2003), this PBPK model was then used to predict DPX concentrations in each flux bin by hour over a week, at the bioassay exposure pattern of 6 hours per day, 5 days per week. DPX concentrations thus calculated were incorporated into the two-stage clonal expansion model by defining the mutation rate of normal and initiated cells as the same linear function of DPX concentration,

$$\mu_N = \mu_I = \mu_{Nbasal} + KMU \times DPX,$$

where the unknown constants μ_{Nbasal} and KMU were estimated by fitting to the tumor bioassay data.

In the development of the PBPK model for DPX, Conolly *et al.* assumed a value of 6.5×10^{-3} min^{-1} for the first-order rate constant for the clearance of DPX, such that the DPX predicted at the end of a 6-hour exposure to 15 ppm was reduced to exactly the detection limit for DPX in 18 hours.[5] This determination of rapid clearance was based on an observation by Casanova *et al.* (1994) that the DPX concentrations observed in the preexposed animals were not significantly higher than those in naive animals (in which there was no significant DPX accumulation). However, *in vitro* data (Quievryn & Zhitkovitch, 2000) indicate a slower clearance, with an average rate constant of 9.24×10^{-4} min^{-1}. Given that there was considerable thickening of the nasal tissues in the more highly exposed preexposed animals and possible changes in the metabolic capacity of those tissues, we believe that the small amount of day-to-day DPX accumulation predicted using this slower *in vitro* clearance rate constant is not inconsistent with the data of Casanova *et al.* (1994). This motivated the need for examining the sensitivity of the cancer model to uncertainty in the DPX clearance rate constant. The DPX

---

[4] Conolly *et al.* state that they used the preexposed data. However, we compared their values to those given by Casanova *et al.* (1994), applying the DNA concentration of 4.1 μg DNA/mg tissue that they used, and determined clearly that they in fact used the naive rat data.

[5] Eighteen hours is the period between the end of one day's 6-hour exposure during the preexposure period and the beginning of the next.

concentrations used in our modeling were obtained by refitting the PBPK model to the naive rat DPX data using this *in vitro* clearance rate; the fit we obtained was as good as that in Conolly *et al.* (2000). It should be noted that this slower DPX repair rate was obtained in the *in vitro* study using transformed, immortalized, human cell lines. However, it appears that DPX repair in normal cells would only be even slower. When nontransformed freshly purified human peripheral lymphocytes were used instead, the half-life for DPX repair was about 50% longer than in the cultured cells (Quievryn & Zhitkovitch, 2000). The revised PBPK model and the above issues are described further in Appendix A. We next compared our cancer model results with those we obtained using DPX concentrations generated with the Conolly *et al.* (2000) PBPK model parameters.

The experimental data on DPX concentrations, to which the Conolly *et al.* (2000) PBPK model was fit, was collected at a single time point (3 hours into one day's exposure), along with the observation of no accumulation of DPX in animals preexposed to formaldehyde. The ability of PBPK models with different DPX half-lives to fit the same experimental data without time-course measurements is not particularly surprising, but the wide range in the above values for the DPX clearance rate led us to question the appropriateness of accounting for hourly variations in the PBPK model predicted DPX dynamics in the cancer modeling. These temporal oscillations were rapid relative to the variations of other time-dependent parameters in the CIIT two-stage cancer model, such as body weight. Further, the measured long-term trends in cell division rates were not incorporated into the CIIT model. Moreover, these oscillations in DPX concentrations translated to similar high frequency oscillations in the tumor hazard (see Section 3); it seemed unlikely to us that the true probability of tumor (as opposed to the modeled probability) would be sensitive to high frequency DPX oscillations. These considerations prompted us to utilize weekly averages of the DPX data, after first determining that the CIIT model predictions were not sensitive to these variations. The large daily fluctuations in predicted DPX concentrations greatly increased the computational time required in our implementation of the two-stage clonal expansion model.

2.4. Number of Cells at Risk

We followed the procedure in Conolly *et al.* (2003) to estimate N, the number of cells in each flux bin at

risk of malignant tumor formation as a function of time. This involved fitting a Gompertz curve to the body weight data from the Monticello *et al.* (1996) bioassay of formaldehyde, provided to us by CIIT.[6]

2.5. Tumor Data

Conolly *et al.* (2003) fit their model to the combined data on squamous cell carcinoma (SCC) from the Kerns *et al.* (1983) and Monticello *et al.* (1996) formaldehyde bioassays, along with historical control data from all NTP bioassays as of 1999. The ages of the animals at death were obtained by adding 63 days (the age of the animals in the Monticello *et al.* study at the beginning of exposure) to the nominal days on study. Conolly *et al.* augmented the data from these two bioassays with data on 7,684 control animals from NTP bioassays, 13 of which were determined to have a nasal SCC. The actual ages at death were used by Conolly *et al.* in the modeling for animals with nasal SCC, and a common age at death of 794 days was used for the 7,671 control animals that died without a nasal SCC.

Tumor rates in control groups from different NTP studies are known to vary due to genetic drift in animals over time and differences in laboratory procedures, such as diet, housing, and pathological procedures (Haseman, 1995). In order to minimize extra variability when using historical control data, the current NTP practice is to limit the historical control data, as far as possible, to studies involving the same route of exposure and to use historical control data from the most recent studies (Peddada *et al.*, 2007). Bickis and Krewski (1989) analyzed 49 NTP long-term rodent cancer bioassays, and found a large difference in determinations of carcinogenicity depending on the use of historical controls with concurrent control animals. To investigate the effect of including historical controls in the CIIT model, we conducted analyses using the following sets of data for controls (fraction of animals with SCC is denoted in parentheses):

1. only concurrent controls (0/347);
2. concurrent controls plus all the NTP historical control data used by Conolly *et al*. (13/8,031); and

---

[6] The growth curve we used was substantially similar to that in Conolly *et al*., except that the maximum fractional body weight is 0.9 in the Conolly *et al.* work, while it is 1.0 in ours. This did not make a noticeable difference in the predictions of the clonal growth model.

Table I. Summary of F344 Rat Data on Squamous Cell Carcinoma (SCC)

Formaldehyde Exposure (ppm)	Number of Animals	Number with SCC	Percent with SCC
All NTP historical controls (Conolly *et al.*, 2003)			
0	7,684	13	0.17
NTP inhalation historical controls			
0	4,602	1	0.02
Concurrent data[a]			
0	341	0	0
0.07	107	0	0
2	353	0	0
6.01	343	3	0.87
9.93	103	22	21.4
14.96	386	162	42.0

[a]Combined data from Kerns *et al.* (1983) and Monticello *et al.* (1996).

3. concurrent controls plus data from historical controls we obtained from NTP inhalation studies (1/4,949) (NTP, 2005). This inhalation database contained information on 4,602 animals, of which only three were diagnosed with a nasal SCC. Of these, two of the tumors in the inhalation historical controls were determined to have originated in tissues other than the nasal cavity upon further review[7] (Dr. Kevin Morgan and Ms. Betsy Gross Bermudez, personal communication). These two tumors were therefore not included.

To reduce computational time in our implementation of the CIIT model, animals without SCC were placed in one of 10 groups depending upon their age at death, and the average age at death for a group was applied to the entire group. The animals with SCC were assigned their actual ages. The ages at death of these control animals were calculated by adding the study-specific average at the beginning of the study to the days on study. As in Conolly *et al.* (2003), we also combined data from the Kerns *et al.* (1983) and Monticello *et al.* (1996) inhalation bioassays for formaldehyde. Both sets of historical control data, along with the data from the two formaldehyde bioassays, are summarized in Table I.

2.6. Tumor Probability and Corrections to Likelihood Function

The unknown parameters in the model were estimated by maximizing the log-likelihood of the tumor

incidence data (Cox & Hinkley, 1974). The rat bioassay data exhibit significant site-specificity in nasal cancer risk, but in many cases the tumors were too large to exactly identify the site or flux bin in which they originated (Monticello *et al.*, 1996). Conolly *et al.* did not partition the computed risk across these sites but instead combined results from the individual flux bins to calculate the overall probability of a tumor (in any flux bin).[8] To compute the combined hazard (instantaneous or integrated) at a given time point, the individual (instantaneous or integrated) hazards from the bin-specific two-stage models were summed over the 20 flux bins. The resulting hazards were for the occurrence of the first malignant cell.

We modified the likelihood expression constructed by Conolly *et al.* (2003) by assuming that tumors discovered in animals that died naturally (i.e., animals that were not part of a scheduled sacrifice) were fatal, and tumors found in animals at a scheduled sacrifice were incidental. Two time delays, as suggested by Dr. Christopher Portier (personal communication), are required to implement the approach:

D_O = the delay from the occurrence of the first malignant cell until the resulting tumor is large enough to be observed at necropsy,

and

D_{OF} = the additional delay from when the tumor resulting from the occurrence of the first malignant cell is large enough to be observed at necropsy until it causes the death of the animal.

There are three cases to consider: (1) an animal that dies without a tumor; (2) an animal that dies with an incidental tumor; and (3) an animal that dies with a fatal tumor.

1) If an animal dies at time t (either naturally or by sacrifice) and no tumor is observed, the contribution to the likelihood is
$1 - P(t - D_O)$ = Probability [a tumor has not become observable by time t], = Probability [no malignant cell has appeared by time $t - D_O$];

2) If a tumor is observed in an animal subject to a scheduled sacrifice at time t, the tumor is assumed to be incidental, and the contribution to the likelihood is

---

[7] Two control animals with nasal SCCs from the NTP inhalation study on methyl methacrylate (NTP TR314) were determined to have tumors that originated in the posterior palate (Animal 12) and the buccal cavity (Animal 26).

[8] However, since the model is effectively composed of distinct two-stage models for each flux bin, it allows one to calculate site-specific tumor risk.

$P(t - D_O) - P(t - D_O - D_F) =$ Probability [a tumor has become observable by time t, but has not yet resulted in the death of the animal];

3) It a tumor is observed in an animal that dies naturally at time t, the tumor is assumed to be fatal, and the contribution to the likelihood is

$p(t - D_O - D_F) = P'(t - D_O - D_F) =$ Probability density of [death from a tumor at time t].

2.7. Statistical and Computational Methods

Our reimplementation of the CIIT two-stage clonal expansion model involved seven adjustable parameters (α_{max}, multb, multfc, μ_{Nbasal}, KMU, D_O, and D_{OF}), which includes all those in the Conolly *et al.* (2003) implementation, except that we required an additional time delay in order to handle both fatal and incidental tumors. These parameters were estimated by maximizing the log-likelihood of the tumor bioassay data. Statistical confidence bounds were computed using the profile likelihood method (Cox & Hinkley, 1974; Cox & Oakes, 1984; Crump, 2002).

Our computer code was written in Microsoft Excel 2002 SP3 Visual Basic. The Excel Solver was used to make the required function optimizations. Our implementation of the two-stage model (i.e., computation of tumor hazard and tumor probability) involved numerically solving the differential equations for computing the functions in Crump *et al.* (2005a) labeled A (eq. 9), B (eq. 10), E (eq. 14, which yielded the bin-specific instantaneous hazards), and G (below eq. 14, which yielded the bin-specific cumulative hazards). In all results presented herein, these systems of differential equations were solved using a Cash-Karp variable step size Runge-Kutta routine (Press *et al.*, 1992) with a tolerance of 10^{-10}.

3. RESULTS

We first attempted to reproduce the Conolly *et al.* (2003) results under similar conditions and assumptions. The results of this comparison are provided in Appendix B.

3.1. Evaluation of the Hoogenveen *et al.* Formulation

As noted earlier, the differential equation in Hoogenveen *et al.* (1999), which was solved in the calculations reported by Conolly *et al.* (2003), is not generally valid for a two-stage model having time-dependent parameters. To evaluate the effect of this assumption, the hazard and the probability of tumor in the 15 ppm exposure group were calculated by applying the same set of parameter values in the Hoogenveen *et al.* equation and the Crump *et al.* (2005a) equations. Graphically, the two solutions were very similar.

3.2. Revision of the Conolly *et al.* (2000) DPX Model

The weekly average DPX concentrations predicted with our revision to the Conolly *et al.* (2000) PBPK model (as described earlier and in Appendix A) were larger than the Conolly *et al.* concentrations by essentially a constant ratio equal to 4.21 (range of 4.12–4.36) when averaged over flux bin and exposure concentrations. Accordingly, cancer model fits using the two sets of DPX concentrations provided very similar parameter estimates, except that the parameter KMU was 4.23 times larger with the Conolly *et al.* DPX concentrations. The two optimization runs were identical except for the DPX concentrations. In the remainder of this article, only results based upon our revised DPX parameters will be presented.

3.3. Comparison of Results Using Hourly DPX Data Versus Weekly Average DPX Values

Table II compares results obtained using weekly-averaged or hourly DPX concentrations in one case. These calculations utilized the J-shape cell replication model and the NTP inhalation controls along with the concurrent CIIT controls. The optimized value of the parameter KMU obtained using hourly-varying DPX values is twice that obtained using weekly-averaged values and the optimized value of μ_{Nbasal} from the use of hourly-varying values is only 64% of that obtained when weekly averages of DPX were used. Despite these differences in parameter estimates, there is good agreement between the log-likelihoods (difference of only 0.43), and the tumor probabilities (Fig. 1A). Although the log-likelihoods obtained from these two fits are similar, there are some differences in the parameter values. However, the instantaneous hazard obtained using the hourly DPX values (Fig. 1B)

Parameter	DPX Data		Ratio Hourly/Weekly
	Hourly	Weekly	
Log-likelihood	−1,518.96	−1,519.39	
μ_{Nbasal}	$7.61 \cdot 10^{-7}$	$1.20 \cdot 10^{-6}$	0.64
KMU	$1.03 \cdot 10^{-6}$	$4.93 \cdot 10^{-7}$	2.09
D_O	244.6	236.9	1.03
D_{OF}	72.6	71.3	1.02
multb	1.069	1.060	1.01
multfc	3.174	2.332	1.36
α_{max}	0.042	0.045	0.94

Note: Both analyses used inhalation historical controls and J-shaped cell replication data. The SCC of two animals in the historical controls were determined to have originated outside the nasal cavity. In contrast to the rest of this article, these were not excluded here because the purpose was only to evaluate the need for using hourly-varying DPX concentrations.

exhibited very rapid oscillations around the hazard obtained using weekly-average DPX predictions. These oscillations closely mimicked those predicted for the DPX concentrations, but are not apparent in the probability graph because the integration process tends to smooth them out. Weekly averages of DPX concentrations are used in results presented in the remainder of the article.
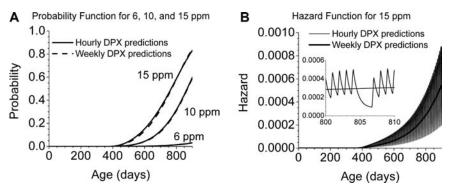
3.4. Reimplementations of the CIIT Model

Table III compares results from runs using either the J-shape or hockey stick cell replication model, in conjunction with modeling the three choices for control data discussed earlier. All of these runs used our revised DPX parameters, computed the likelihood in a manner that distinguished between tumors found at a scheduled sacrifice and those found in animals that died naturally, and used the time inhomogeneous

solutions to the two-stage model. In each analysis, α_{max} was restricted to be no larger than 0.045 h^{-1}, the largest value considered by Conolly *et al.* (2003) to be biologically plausible.

In all cases, the J-shape and the hockey stick cell replication models provide similar fits (i.e., similar log-likelihoods). In particular, the two maximum likelihood estimate (MLE) fits that use only concurrent tumor data are essentially identical. This is expected because both cases predict a zero probability of background SCC so that the reduced cell replication rate predicted at the lower doses has no effect, and at higher exposures the cell replication rates predicted by the J-shape and hockey stick cell replication models are identical (Table 1 in Conolly *et al.* (2003)).

Estimates of D_O, D_{OF}, multb, and multfc from different analyses are very similar, and in each case α_{max} is estimated at its assumed maximum biologically plausible value of 0.045 h^{-1} (Table III). All of these estimates are roughly similar to those obtained by Conolly *et al.* (2003) (see Appendix B for details), even though the analyses differ in the definition of the likelihood and in the DPX model. Note that one must add D_O and D_{OF} in order to compare with the time delay obtained in Appendix B. The MLE estimates of μ_{Nbasal} are zero for concurrent control data, which makes the ratio, KMX = KMU/μ_{Nbasal}, undefined or infinite.

Conolly *et al.* (2003) estimated KMU to be zero for both the hockey stick and J-shape cell replication models. However, our estimate for the coefficient KMU (obtained using the solution of Crump *et al.*, 2005a) is zero only for the case of the model with the hockey stick curve for cell replication and with control data as used by Conolly *et al.* It is positive in all other cases, and statistically significantly so in all cases in which either inhalation control data or concurrent controls were used. As a further exercise on the sensitivity of the dose-response modeling on the control data, we also examined the case

Fig. 1. Comparison of probabilities and hazards estimated from fits in Table II using hourly or weekly-averaged DPX concentrations. (A) Probability function for 6, 10, and 15 ppm exposure concentration. (B) Hazard function for 15 ppm exposure concentration.
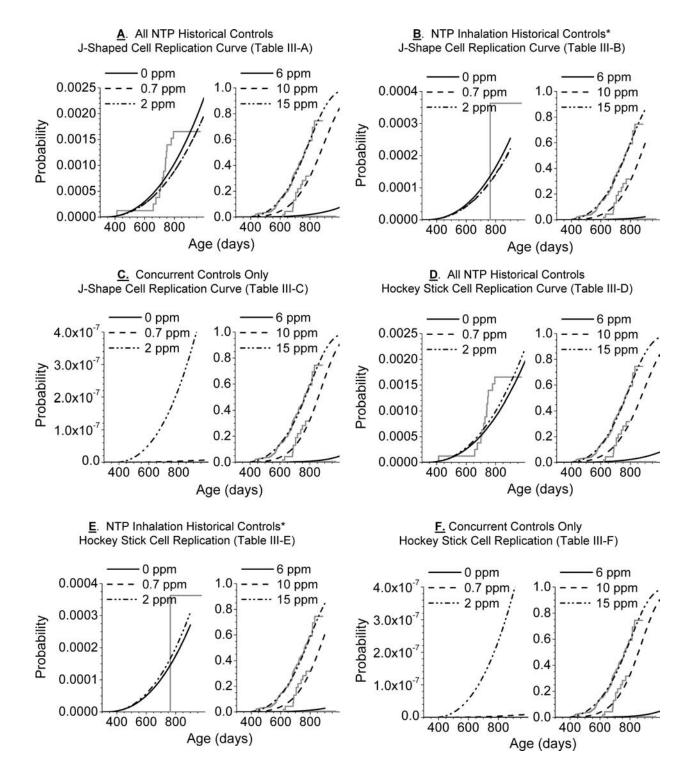
Table III. Results from Different Control Animals and Cell Replication Models

	A	D	B	E	C	F
Control Animals (Combined with Concurrent Controls)	All NTP	All NTP	NTP Inhalation	NTP Inhalation	Concurrent Only	Concurrent Only
Cell Replication Dose Response	Historical J-Shaped	Historical Hockey	Historical J-Shaped	Historical Hockey	J-Shaped	Hockey
Log-likelihood	−1,692.65	−1,693.68	−1,493.21	−1,493.35	−1,474.29	−1,474.29
μ_{Nbasal}	$1.87 \cdot 10^{-6}$	$2.12 \cdot 10^{-6}$	$7.32 \cdot 10^{-7}$	$9.32 \cdot 10^{-7}$	0.0	0.0
KMU	$1.12 \cdot 10^{-7}$	0.0	$6.84 \cdot 10^{-7}$	$6.18 \cdot 10^{-7}$	$1.20 \cdot 10^{-6}$	$1.20 \cdot 10^{-6}$
KMX (KMU/μ_{Nbasal})	0.06	0.0	0.94	0.66	∞	∞
KMX 90% CI	(0.0, 0.40)	(0.0, 0.25)	(0.26, 6.20)	(0.2, 5.20)	(0.42, ∞)	(0.41, ∞)
D_O	238	238	237	237	243	243
D_{OF}	67.4	67.3	67.3	70.9	68.8	68.8
multb	1.05	1.05	1.06	1.06	1.08	1.08
multfc	1.87	1.77	2.44	2.54	3.35	3.35
α_{max}	0.045	0.045	0.045	0.045	0.045	0.045
α_{max} 90% CI	(0.029, 0.045)	(0.029, 0.045)	(0.026, 0.045)	(0.027, 0.045)	(0.027, 0.045)	(0.027, 0.045)

(but do not include in the tabulated results) when the two nasal tumors in the historical NTP inhalation controls determined to have originated elsewhere (posterior palate and buccal cavity) were also included as nasal SCC. KMU varies considerably depending on the NTP historical controls lumped with the concurrent controls. It progressively increases in the following sequence: 1) including all NTP historical controls; 2) only NTP inhalation controls with the two tumors from non-nasal tissue; 3) only NTP inhalation controls without these two tumors; 4) using only concurrent controls. With concurrent controls only and the J-shape cell replication model, the MLE estimate for KMU (1.2×10^{-6}) is larger than the statistical upper bound obtained by Conolly *et al.* (2003) (8.2×10^{-7}). It should also be kept in mind that our estimate would be about 4.2 times larger still had the Conolly *et al.* DPX model been used. Excluding the two nonnasal SCC from the inhalation historical controls increases the ratio KMU/μ_{Nbasal} by 1.7- or 2.3-fold, depending on whether the hockey stick or J-shape, respectively, is used for the cell replication dose response, and that of the upper confidence bound on this ratio by 4.4- or 4.6-fold.

Fig. 2 compares the predicted probability of death from a SCC as a function of age to the corresponding graphs of Kaplan-Meier nonparametric estimates of this probability in the rat bioassays, for the cases listed in Table III. These six fits are all virtually indistinguishable for formaldehyde concentrations ≥ 6 ppm. The model appears to provide an excellent fit to the 15 ppm data, where most of the SCC occurred. The fits to the 10 ppm data are less satisfactory and the

models predict a somewhat larger response at older ages in the 6 ppm group than observed. The curves for 0.7 ppm and 2 ppm are indistinguishable from each other in Panels A, B, D, and E. The models incorporating either all NTP historical control data or the NTP inhalation historical control data also fit the control data reasonably well, although it must be kept in mind that there were very few SCC in these groups. In the fits that used only concurrent controls, no SCCs were present and the predicted MLE probability of a SCC was zero. It may be noted, however, that absence of tumors in the limited number of concurrent animals does not imply that the calculation will necessarily predict a zero background probability of tumor.

Table IV examines the contribution of the DPX component, which in this model represents the directly mutagenic potential of formaldehyde, to the calculated tumor probability. This is demonstrated for the optimized models that use the NTP inhalation historical control data (cases B, E in Table III) by expressing the tumor probability resulting from the DPX term as a fraction of the absolute overall added tumor probability, |P(d) – P(0)|, at exposure concentration d. For a given optimized model, the DPX-related tumor probability was obtained by subtracting from the overall tumor probability the value calculated by setting the DPX concentration to zero. In the range of exposures where tumors were observed (6.0 ppm – 15.0 ppm), the DPX term is responsible for 58–74% of the added tumor probability. Below 6.0 ppm, the DPX contribution varies between 2% and 80% depending on the dose-response curve used for cell replication. It must be noted that in the case

Fig. 2. Graphs of probability of a fatal SCC predicted by best fitting model, compared to Kaplan-Meier curves for fatal SCC. Dark gray line is Kaplan-Meier curve and is zero for 0.7 and 2 ppm. (A) All NTP historical controls, J-shaped cell replication curve (Table III-A). (B) NTP inhalation historical controls,* J-shaped cell replication curve (Table III-B). (C) Concurrent controls only. J-shaped cell replication curve (Table III-C). (D) All NTP historical controls, hockey stick cell replication curve (Table III-D). (E) NTP inhalation historical controls,* hockey stick cell replication curve (Table III-E). (F) Concurrent controls only, hockey stick cell replication curve (Table III-F).* SCCs in controls originating from nonnasal tumors were excluded.

Table IV. Contribution of DPX to Tumor Probability (for Models in Table III with Inhalation Historical Controls)

Model	Exposure, d(ppm)	Tumor Probability[a] P(d)	Tumor Probability Due to DPX Term[b] $P_{DPX}(d)$	$\dfrac{P_{DPX}(d) \times 100}{|P(d) - P(0)|}$
B	0.0	$2.54 \cdot 10^{-4}$	0	
(J-shape α)	0.7	$2.17 \cdot 10^{-4}$	$6 \cdot 10^{-7}$	2%
	2.0	$2.21 \cdot 10^{-4}$	$5 \cdot 10^{-6}$	14%
	6.01	0.023	0.016	68%
	9.93	0.581	0.431	74%
	14.96	0.845	0.600	71%
E	0.0	$2.71 \cdot 10^{-4}$	0	
(Hockey shape α)	0.7	$2.71 \cdot 10^{-4}$	$7 \cdot 10^{-7}$	80%
	2.0	$3.10 \cdot 10^{-4}$	$6 \cdot 10^{-6}$	14%
	6.01	0.028	0.016	58%
	9.93	0.590	0.377	64%
	14.96	0.833	0.501	60%

[a]Cumulative probability at 900 days of fatal SCC.
[b]Calculated subtracting from P(d) the probability obtained setting the DPX concentration to zero.

of 0.7 ppm and 2.0 ppm exposure concentrations, the numerator and denominator of the ratio in the last column of Table IV are both obtained by subtracting two very small quantities, each of which is rather imprecise. Therefore, numerical error in estimating this ratio is to be expected when the added tumor probability is very small.

4. DISCUSSION

The aim of this article was to examine the sensitivity of the two-stage clonal expansion model in Conolly *et al.* (2003) for predicting formaldehyde carcinogenicity in the F344 rat to some key assumptions and use of control data. In this article, we implemented solutions to this model that are valid for nonhomogeneous models. The mathematical methods used here for the solution account for time-dependence in variables, while the solution implemented by Conolly *et al.* (Hoogenveen *et al.*, 1999) is valid only for homogeneous models, that is, when variables do not change in time. For the cases examined here, our solutions were not very different from those obtained with the Hoogenveen *et al.* method. Thus, the use of the Hoogenveen *et al.* formalism is not likely to have resulted in any significant error in the Conolly *et al.* solutions. However, we present this conclusion with the caveat that the analyses presented in this article retained the time-independent form of the cell replication rates (time-weighted averages) as used in

Conolly *et al.* (2003). On the other hand, the labeling index data of Monticello *et al.* (1991, 1996) from which these rates were derived clearly vary with time. Crump *et al.* (2005a) showed that the error in using the Hoogenveen *et al.* equation can be substantial when there is considerable time-dependence in cell replication rates. Therefore, the formulation that is valid for nonhomogeneous models should be preferred because of the generally unknown amount of error associated with applying the Hoogenveen *et al.* formulation to a nonhomogeneous model. The Hoogenveen *et al.* solution used by Conolly *et al.* has a significant advantage in computing speed. However, simulations with our approach did not require a large investment of computational resources, even though a single calculation of the likelihood required solving more than 8,000 sets of four coupled ordinary differential equations,[9] with hundreds of likelihood calculations often necessary for parameter optimization.

The CIIT cancer risk assessment for formaldehyde was based on pooling the large number of all NTP historical controls with controls that were concurrent to the Kerns *et al.* (1983) and Monticello *et al.* (1996) formaldehyde bioassays. As seen from Table I, the incidence of nasal SCC is substantially different in

[9] The two formaldehyde bioassays involved more than 400 unique ages at death, not counting the additional death times in the historical control data. Furthermore, the hazard calculation for a single time point required solving a two-stage clonal expansion model for each of the 20 flux bins.

all NTP historical controls when compared with only the NTP inhalation historical controls. In our calculations, inclusion of historical controls had a strong impact on the tumor probability curve below the range of exposures over which tumors were observed in the formaldehyde bioassays. This can be seen in Fig. 2. The MLE probabilities of a fatal tumor for exposure concentrations below 6 ppm were roughly an order of magnitude higher when all the NTP historical controls were used (Panels A and D) than that predicted when historical controls were drawn only from inhalation bioassays (Panels B and E, correspondingly; these had lower tumor incidence), and by many orders of magnitude than that predicted when only concurrent controls were used in the analysis (Panels C and F, correspondingly). Note that this comparison should not be inferred to apply to upper bound risk estimates as there were many fewer concurrent than historical controls, so error bounds could be much larger in the case where concurrent controls were used. This will be addressed in a future paper. However, model fits to the tumor data in the 6–15 ppm exposure concentration range (that is, exposures where tumors were observed) were qualitatively indifferent to which of these control data sets were used. This observation emphasizes the statistical aspect of the CIIT modeling—that significant interplay among the various adjustable parameters allows the model to achieve a good fit to the tumor incidence data independent of the control data used. However, our results show that changes in the control data affect KMU, the proportionality constant relating DPX with the mutational probability, resulting in significantly different tumor predictions at lower exposure concentrations. Our analysis indicates that detailed evaluation of bioassays from which control animals are drawn is necessary if historical controls are to be combined with concurrent controls. The bioassays in consideration need to be similar, and comparable to the concurrent control group, with regard to the many factors that are known to affect tumor and survival rates of the control animals (Haseman, 1995).

Induction of a proliferative response to cytotoxicity plays a critical role in the carcinogenesis of many compounds that are also mutagenic. For formaldehyde and vinyl acetate, two-stage cancer modeling has been used to reflect upon the relative importance of cytotoxicity-induced cell proliferation versus the chemically-induced direct mutations in explaining the observed tumorigenicity in rodent bioassays (Bogdanffy *et al.*, 1999, 2001; Conolly, 1995; Conolly *et al.*, 2003, 2004; Slikker *et al.*, 2004). The reanalyses presented in this article (in particular, the conclusions from Table IV) indicate that, based on existing mechanistic data, a large contribution from formaldehyde's mutagenic potential may be needed in the mathematical model to explain formaldehyde carcinogenicity. The choice of control tumor data had substantial bearing on the biological inferences one might draw from the two-stage cancer modeling. In the Conolly *et al.* (2003) model where all NTP historical controls were included, the constant of proportionality relating the probability of mutation per cell generation to DPX concentrations (KMU) was estimated to be zero when either J-shape or hockey-stick shaped curves were considered for the cell replication dose response. In our reimplementation of this model, KMU was estimated to be zero only when the hockey stick shape was used in conjunction with the inclusion of all NTP historical controls. In all cases that included only the NTP inhalation historical controls with the concurrent or used only the concurrent controls, KMU was estimated to be statistically significantly positive. When only concurrent controls were used, the MLE estimate of KMU is roughly twice that obtained when inhalation historical controls were included. With concurrent controls only and the J-shaped cell replication curve, the MLE estimate for KMU is about 1.5-fold higher than the statistical upper bound obtained for it by Conolly *et al.* (2003), and would have been about 6 times higher had the Conolly *et al.* DPX model parameters been used.

The ratio of this constant to the probability of spontaneous mutation per cell generation ($KMX = KMU/\mu_{Nbasal}$) is of particular interest since KMU/μ_{Nbasal} was assumed to be invariant between rodent and human in the extrapolation carried out in the CIIT human model (Conolly *et al.*, 2004). While the MLE estimate of KMX is zero in the CIIT animal model (Conolly *et al.*, 2003), it takes a range of values from zero to 0.9 mm^3/pmol and undefined (or infinite, when $\mu_{Nbasal} = 0$) in the various cases examined in this article. The 95% upper confidence bound on this ratio ranges from 0.25 to 6.2 (values that would be four times larger had the Conolly *et al.* DPX concentrations been used) to infinite. Thus, the extrapolation to human risk using the approach in Conolly *et al.* (2004) becomes particularly problematic when only concurrent controls are used.

The bioassay data at high exposure concentrations have maximum impact on model calibration, and thereby, on model inferences. On the other hand, the strong influence of using all the NTP historical controls on the low-dose region of the time-to-tumor

curves presented in Fig. 2 suggests that large uncertainties may arise in extrapolating to human from such considerations alone.

We examined the sensitivity of the CIIT two-stage model to the temporal variations in DPX as predicted by PBPK modeling. Model results remained substantially unchanged, prompting us to utilize weekly averages of the DPX data. It is possible that the time-dependence of DPX concentrations could become relevant when temporal variations in cell replication rate is incorporated in the clonal growth modeling. Our decision to ignore the DPX time-dependence was also prompted by other considerations. Whereas the cumulative probability of tumor occurrence remained relatively unaffected when weekly averages of DPX were considered, the hourly variation in DPX predicted by Conolly et al. (2000) resulted in rapid variations in the tumor hazard predicted by the model. This is because the formation of malignant cells depends on the DPX level, and tumors are modeled to occur following a fixed time delay after the first malignant cell. However, because of these rapid variations, the exact time of tumor occurrence could impact significantly upon the likelihood function, which does not make sense biologically. Thus, we believe that the use of hourly-varying DPX levels in two-stage modeling seems inappropriate as it could result in artifactual variation in predicted cancer rates. The variations may also point to inherent limitations in the two-stage clonal expansion model in handling such rapid temporal oscillations in input data. It may be possible to overcome this difficulty by modeling the growth kinetics of malignant clones as implemented by Sherman and Portier (2000). Alternatively, the time from the formation of a malignant cell until the time the resulting malignant clone is large enough to produce an observable response can be considered a random variable. The time to observable response in the whole animal would be the minimum of such times. Hazelton et al. (2001) used a gamma distribution to represent the lag time between generation of the first malignant cell and death from lung cancer.

In addition, it is uncertain as to whether DPX is indeed completely removed in 18 hours as modeled in Conolly et al. (2000). As explained in Section 2, such an interpretation of the Casanova et al. (1994) data assumes no significant induction of formaldehyde or aldehyde dehydrogenase (enzymes that metabolize formaldehyde) during the 12-week exposures, and ignores the possible reduction of DPX concentration levels due to thickening of tissue upon exposure. These factors could, however, readily offset the small amounts of DPX accumulation that is predicted when the experimental in vitro rate constant for DPX removal (Quievryn & Zhitkovitch, 2000) is used in the model. Compared to the value in Conolly et al. (2000), this removal rate constant corresponds to a 7-fold longer mean half-life for DPX. Given these uncertainties in the PBPK model predicted DPX dynamics, weekly-averaged DPX concentrations are more reliable.

DPX concentrations obtained with a revised PBPK model that fixed DPX half-life at the 7-fold higher in vitro value resulted in weekly-average DPX values that were generally 4-fold higher than those predicted by Conolly et al. (2000). These revised higher DPX values were offset by a 4-fold lower value for the linear coefficient, KMU, that was estimated by likelihood optimization against the tumor data. Other than this parameter, the revised DPX values did not affect the quality of model fits or the value of any other parameter. However, revised DPX profiles could affect human risk estimates obtained using the CIIT model.

Our construction of the likelihood function distinguished between tumors discovered incidentally in sacrificed animals and tumors found in animals that died naturally due to the tumor. This allowed us to examine whether the assumption of rapidly fatal tumors was reasonable in the context of the formaldehyde two-stage cancer model. In our reimplementation of the model, the estimates for the constant D_{OF} (the lag between the time when a tumor is large enough to be detectable by necropsy and the time when the tumor kills the animal) ranged from 67 to 71 days and estimates for D_O (the time lag for a detectable tumor to manifest from the occurrence of the first malignant cell) ranged from 237 to 243 days. That is, D_{OF} was found to be a significant fraction of $(D_O + D_{OF})$. It must be noted that the use of these constant time delays represents a considerable approximation as it assumes that a malignant cell will lead to a tumor after a fixed time. We are in the process of modifying the approach to include a distribution in delays.

The use of weight of evidence in considering mode of action of a putative carcinogen is emphasized in the U.S. EPA's cancer guidelines (USEPA, 2005). While generally not perceived as being central to the weight of evidence process, mathematical dose-response modeling has played a role in the scientific literature in debating the relative weights to be placed on different possibilities in the case of carcinogens with multiple modes of action. The results of this article point to some caveats, and illustrate the

importance of detailed sensitivity analyses of parameter uncertainties and model specification as a prerequisite to such use. The modifications made in this article provide the basic changes we perceived necessary to enable further sensitivity analyses on assumptions in the CIIT model. These additional issues include examining the strength of biological assumptions on the model structure for initiated cell division and death rates, uncertainty and variability in the dose response for cell replication rates in the rodent, and its relevance to humans, and uncertainty in the extrapolation of modeled DPX concentrations to humans. These are substantial questions not examined here that have the potential to impact the risk characterization presented in the CIIT human health risk assessment for formaldehyde.

APPENDIX A: REVISION TO CONOLLY *ET AL.* (2000) DPX MODEL

The PBPK DPX model of Conolly *et al.* (2000) was refit to the regional DPX data of Casanova *et al.* (1994),[10] using the DPX removal rate measured by Quievryn and Zhitkovitch (2000) *in vitro* in three human cell lines. Quievryn and Zhitkovitch obtained a mean half-life for DPX removal of 12.5 hours (range 11.6–13.0 hours). Since DPX repair rates were not directly measured in rat tissues, we used this value for the half-life in our model, yielding a rate constant, $k_{loss} = 9.2 \cdot 10^{-4}$ min^{-1}. Because we did not scale the DPX data to a per volume basis (i.e., we kept them as pmol DPX/mg DNA rather than scaling to pmol/mm^3 tissue), the value of the binding constant we used was $1.56 \cdot 10^{-3}$ mm^3/mg DNA/min rather than $6.4 \cdot 10^{-6}$ min^{-1} as used by Conolly *et al.* but the values are equivalent based on that linear scaling.

The CFD model for the rat was used to predict the average flux to the "whole-nose" region of the rat nasal passages. The CFD model predicts flux to be proportional to the exposure level:

$$\text{Flux(region)} = f_{HCHO}(\text{region}) \cdot \text{ppm}_{exp},$$

where ppm_{exp} is the exposure level in ppm, Flux(region) is the amount of formaldehyde delivered to a given region the nasal mucosa per unit surface area per unit time, and $f_{HCHO}(\text{region})$ is the proportionality constant relating the two (flux at 1 ppm for a region). The model equations for formaldehyde dosimetry and DPX formation in the nasal lining were then:

$$\frac{d}{dt}[\text{HCHO}] = (f_{HCHO} \cdot \text{ppm}_{exp}/\delta(\text{resion}))$$

$$- \frac{V_{max} \cdot [\text{HCHO}]}{K_m + [\text{HCHO}]_j} - k_f[\text{HCHO}]$$

$$- k_b \cdot [\text{HCHO}]$$

$$\frac{d}{dt}[\text{DPX}] = k_b \cdot [\text{HCHO}] - k_{loss} \cdot [\text{HCHO}].$$

---

[10] For which the individual values are listed by Georgieva *et al.* (2003).

Table AI. Parameter Values for PBPK DPX Model

Parameter	Revised Value	Conolly *et al.* (2000)
V_{max} (pmol/min-mm^3)	967	1,008
K_m (pmol/mm^3)	42.6	70.8
k_f (min^{-1})	1.76	1.08
k_{loss} (min^{-1})[a]	$9.2 \cdot 10^{-4}$	$6.5 \cdot 10^{-3}$
k_b[b]	$1.56 \cdot 10^{-3}$ mm^3/mg DNA-min	
δ(low tumor region)	0.095 mm	
δ(high tumor region)	0.067 mm	
f_{HCHO}(low tumor region)	7.0 pmol/mm^2-min-ppm	
f_{HCHO}(high tumor region)	12.5 pmol/mm^2-min-ppm	

[a]Revised value from Quievryn and Zhitkovitch (2000); Conolly *et al.* (2000) value interpreted from naive versus preexposed rat DPX data.
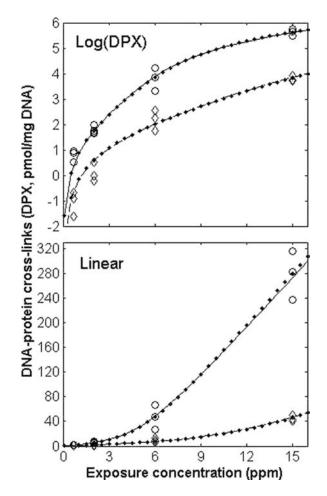[b]Different units from those used by Conolly *et al.* (2000).

Here, [HCHO] is the concentration of formaldehyde in the tissue, δ(region) is the tissue thickness of a region, V_{max} and K_m are the maximum metabolic removal rate and saturation constant, respectively, k_f is the rate constant for first-order removal processes, k_b is the rate constant for DPX formation (binding), and k_{loss} is the rate of DPX removal or loss.

The resulting parameter values, as compared to those of Conolly *et al.*, are listed in Table AI, while the resulting fits to the DPX data of Casanova *et al.* (1994) are shown in Fig. A1, with simulations obtained using the Conolly *et al.* (2000) parameters shown for comparison. As can be seen in Fig. A1, the revised parameter values give virtually identical fits to those obtained with the parameters of Conolly *et al.* Model simulations and parameter optimization were performed using Matlab software (version 7.3, The MathWorks, Inc., Natick, MA), in particular using the *fminsearch* (Nelder-Mead) optimization algorithm. Goodness of fit was quantified using a log-likelihood function (LLF; for a derivation, see the appendix of Cole *et al.*, 2001) with the objective function written as:

$$-LLF = n \cdot [\log(2\pi) + 1] + \gamma \sum_{i=1}^{n} \log(f_i)$$

$$+ n \cdot \log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{(z_i - f_i)^2}{f_i^{\gamma}}\right),$$

where n is the number of measurements, γ is a heteroscedasticity parameter estimated simultaneously

Fig. A1. PBPK model fits to DPX data of Casanova *et al.* (1994). Symbols are data from high (circles) and low (diamonds) tumor regions of the F344 rat nasal epithelium. Smooth lines (solid and dashed) are simulations with the revised parameter values listed in Table AI. Dotted lines are simulations with parameter values of Conolly *et al.* (2000) and cannot be distinguished from our curves.

with other parameters, f_i is the model prediction of the ith data point, and z_i is the ith measurement.

The key difference in these parameter values is that Conolly *et al.* (2000) selected their value of the DPX removal rate constant, k_{loss}, by interpreting the data of Casanova *et al.* (1994) for DPX formation in naive versus preexposed rats (air-exposed controls vs. those exposed for 11 weeks + 4 days) as indicating that DPX removal must be high enough to completely eliminate the DPX formed in a single 6-hour exposure by the beginning of the next day; i.e., in 18 hours. Their k_{loss} value is the smallest value, which would reduce the highest measured level of DPX in those experiments to just below the limit of detection in 18 hours.

However, examination of the Casanova *et al.* (1994) data shows that there was a significantly decreased (~40%) level of the DPX in high tumor regions of preexposed animals versus naive animals at 6 and 15 ppm, and that the weight of the tissues dissected from those regions increased substantially, indicating a thickening of the tissues. After testing the result of changing the tissue thickness in the PBPK model for DPX, we found that such a change alone could not account for the dramatic reduction in DPX levels after preexposure, even with the higher value of k_{loss} used by Conolly *et al.* (2000). Therefore, we concluded that in addition to the gross increase in tissue weight, these data indicate either an induction in the activity of enzymes that remove formaldehyde (aldehyde and formaldehyde dehydrogenase) or other changes in the biochemical properties of the tissue. Given such a change, it is entirely possible that those experimental results are consistent with the smaller experimental value of k_{loss} used here. In particular, if V_{max} increases with exposure (in a tissue region- and dose-specific manner), then it is possible to explain the naive versus preexposed data of Casanova *et al.* with the value of k_{loss} effectively measured *in vitro* by Quievryn and Zhitkovitch (2000). Given that this value was measured directly, rather than obtained by indirect interpretation of measurements made at only two time points where significant changes in the tissue had occurred, and that the fit we then obtained to the acute DPX data (Fig. A1) was excellent, we believe that use of this lower value for k_{loss} and corresponding values for the remaining parameters is justified.

APPENDIX B: REPRODUCING RESULTS FROM THE CIIT MODEL

We first attempted to reproduce the results of Conolly *et al.* (2003) by maximizing the log-likelihood for both the J-shape and hockey stick cell replication models under conditions similar to those employed by Conolly *et al.* To do so, we assumed all tumors were fatal, used the Hoogenveen *et al.* (1999) solution, and used the Conolly *et al.* values for DPX and body weight. However, we used average weekly values of DPX, whereas Conolly *et al.* used hourly values. Table BI compares the parameter estimates we obtained with those obtained by Conolly *et al.* Although the two sets of parameters are similar, there are some differences. The Conolly *et al.* log-likelihoods are considerably smaller than

Table BI. Comparison of Conolly *et al.* (2003) Model Fits with Those We Obtained Using the Same Computational Methods and Assumptions

Parameter	J-Shape		Hockey Stick	
	Conolly	This Analysis	Conolly	This Analysis
Log-likelihood	−2,131.5	−2,053.3	−2,133.1	−2,055.1
μ_{Nbasal}	$1.35 \cdot 10^{-6}$	$1.84 \cdot 10^{-6}$	$1.47 \cdot 10^{-6}$	$2.08 \cdot 10^{-6}$
KMU	0	0	0	0
D_O	290.9	299.4	297.4	301.6
multb	1.072	1.079	1.070	1.084
multfc	2.583	3.128	2.515	3.665
α_{max}	0.0435	0.0403	0.0435	0.0355

ours. In addition, the parameter "multfc" (which determines the dose-dependence of the initiated cell growth advantage over normal cells) is considerably higher in our analysis, and in the case of our analysis with the hockey stick curve, our estimate for α_{max} is lower.

To further explore these differences, we compared in Fig. B1 the predicted probabilities obtained by Conolly *et al.* and those obtained using our best-fitting model for the case of a J-shaped dose response for the cell replication (Table BI), with the corresponding Kaplan-Meier curves. Our fit seems better for 15 ppm, where most of the tumors occurred, as the Conolly *et al.* fit lies below the Kaplan-Meier curve. Our fit also appears better for 0 ppm, and possibly for 10 ppm. The Conolly *et al.* fit appears to be better for 6 ppm, where three tumors occurred.

The discrepancy in likelihood and fit may be primarily due to an error in the likelihood calculation in the computer program used to implement the CIIT model (Liao, 2006), wherein the log-likelihood for the survivor data was double-counted. It is possible that this error may not impact on predictions of risk at low dose significantly (the Conolly *et al.* model predicts a two-fold lower tumor probability for the controls). After correcting this error, Dr. Liao (2006) found that the likelihood in the CIIT model improved from −2,131.5 to −2,067.7 (J-shape cell replication), which is much closer to our value of −2,053.3. It is possible that the remaining discrepancy would likely be further reduced if the CIIT model were reoptimized after correcting this error.
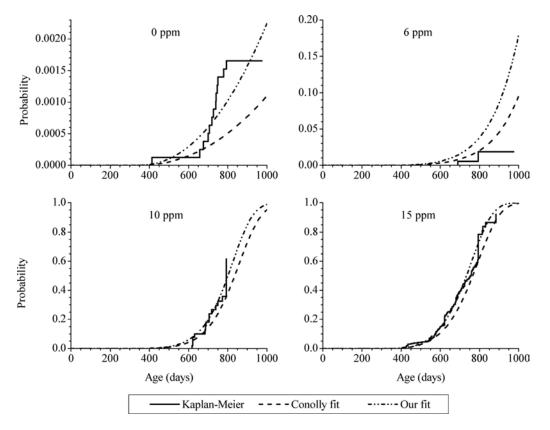
Fig. BI. Comparison of Conolly *et al.* (2003) fit and our fit to the Kaplan-Meier curves for probability of a fatal SCC, using the Conolly *et al.* assumption that all SCC were fatal.

REFERENCES

Bickis, M., & Krewski, D. (1989). Statistical issues in the analysis of the long-term carcinogenicity bioassay in small rodents: An empirical evaluation of statistical decision rules. *Fundamental and Applied Toxicology*, *12*, 202–221.

Bogdanffy, M., Plowchalk, D., Sarangapani, R., Starr, T., & Andersen, M. (2001). Mode-of-action-based dosimeters for interspecies extrapolation of vinyl acetateinhalation risk. *Inhalation Toxicology*, *13*(5), 377–396.

Bogdanffy, M., Sarangapani, R., Plowchalk, D., Jarabek, A., & Andersen, M. (1999). A biologically based risk assessment for vinyl acetate-induced cancer and noncancer inhalation toxicity. *Toxicological Sciences*, *51*(1), 19–35.

Casanova, M., Morgan, K., Gross, E., Moss, O., & Heck, H. A. (1994). DNA-protien cross-links and cell replication at specific sites in the nose of F344 rats exposed subchronically to formaldehyde. *Fundamental and Applied Toxicology*, *23*, 525–536.

Cole, C. E., Tran, H. T., & Schlosser, P. M. (2001). Physiologically based pharmacokinetic modeling of benzene metabolism in mice through extrapolation from in vitro to in vivo. *Journal of Toxicology and Environmental Health*, Part A *62*, 439–465.

Conolly, R. B. (1995). Cancer and non-cancer risk assessment: Not so different if you consider mechanisms. *Toxicology*, *102*, 179–188.

Conolly, R., Crump, K. S., & Clewell, H. J. (2006). Stochastic simulation to obtain the exact simulation of the two-stage model equations. Society of Toxicology, conference presentation.

Conolly, R., Kimbell, J., Janszen, D., Schlosser, P., Kalisak, D., Preston, J., & Miller, F. (2003). Biologically motivated computational modeling of formaldehyde carcinogenicity in the F344 rat. *Toxicological Sciences*, *75*, 432–447.

Conolly, R., Kimbell, J., Janszen, D., Schlosser, P., Kalisak, D., Preston, J., & Miller, F. (2004). Human respiratory tract cancer risks of inhaled formaldehyde: Dose-response predictions derived from biologically-motivated computational modeling of a combined rodent and human dataset. *Toxicological Sciences*, *82*(1), 279–296.

Conolly, R., Lilly, P., & Kimbell, J. (2000). Simulation modeling of the tissue disposition of formaldehyde to predict nasal DNA-protein cross-links in fischer 344 rats, rhesus monkeys, and humans. *Environmental Health Perspectives*, *108*(Suppl 5), 919–924.

Cox, D., & Hinkley, D. (1974). *Theoretical Statistics*. London: Chapman and Hall.

Cox, D., & Oakes, D. (1984). *Analysis of Survival Data*. New York: Chapman and Hall.

Crump, K. (2002). Benchmark analysis. In A. El-Shaarawi & W. W. Piegorsch (Eds.), *Encyclopedia of Environmetrics* (pp. 163–170). West Sussex, UK: Wiley.

Crump, K., Subramaniam, R., & Van Landingham, C. (2005a). A numerical solution to the non-homogeneous two-stage MVK model of cancer. *Risk Analysis*, *25*, 921–926.

Crump, K., Subramaniam, R., Van Landingham, C., Chen, C., & White, P. (2005b). Alternatives and refinements to the two-stage model for formaldehyde carcinogenicity in the rat. Society for Risk Analysis, conference presentation.

Crump, K., Subramaniam, R., & Van Landingham, C. (2006). Acknowledgment of prior solution to MVK cancer model. *Risk Analysis*, *26*, 3–4.

Grafstrom, R. C., Curren, R. D., Yang, L. L., & Harris, C. C. (1985). Genotoxicity of formaldehyde in cultured human bronchial fibroblasts. *Science*, *228*(4695), 89–91.

Haseman, J. K. (1995). Data analysis: Statistical analysis and use of historical control data. *Regulatory Toxicology and Pharmacology*, *21*, 52–59.

Hazelton, W. D., Luebeck, E. G., Heidenreich, W. F., & Moolgavkar, S. H. (2001). Analysis of a historical cohort of Chinese tin miners with arsenic, radon, cigarette smoke, and pipe smoke exposures using the biologically based two-stage clonal expansion model. *Radiation Research*, *156*(1), 78–94.

Heck, H., Casanova, M., & Starr, T. (1990). Formaldehyde toxicity—New understanding. *Critical Reviews in Toxicology*, *20*, 397–420.

Heidenreich W., Luebeck E. G., Moolgavkar S. H. (1997). Some properties of the hazard function of the two-mutation clonal expansion model. *Risk Analysis*, *17*, 391–399.

Hoogenveen, R., Clewell, H., III, Andersen, M., & Slob, W. (1999). An alternative exact solution of the two-stage clonal growth model of cancer. *Risk Analysis*, *19*(1), 9–14.

Kerns, W., Pavkov, K., Donofrio, D., Grlla, E., & Swenberg, J. (1983). Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Research*, *43*, 4382–4392.

Kimbell, J., Overton, J., Subramaniam, R., Schlosser, P., Morgan, K., Conolly, R., & Miller, F. (2001a). Dosimetry modeling of inhaled formaldehyde: Binning nasal flux predictions for quantitative risk assessment. *Toxicological Sciences*, *64*, 111–121.

Kimbell, J., Subramaniam, R., Gross, E., Schlosser, P., & Morgan, K. (2001b). Dosimetry modeling of inhaled formaldehyde: Comparisons of local flux predictions in the rat, monkey, and human nasal passages. *Toxicological Sciences*, *64*, 100–110.

Liao, D. K. (2006). RE: Error in the likelihood calculation in the computer program used to implement the CIIT model. Personal Communication to Dr. Kenny Crump.

Little, M. (1995). Are two mutations sufficient to cause cancer? Some generalizations of the two-mutation model of carcinogenesis of Moolgavkar, Venzon, and Knudson, and of the multistage model of Armitage and Doll. *Biometrics*, *51*, 1278–1291.

Little, M., Haylock, R., & Muirhead, C. (2002). Modelling lung tumour risk in radon-exposed uranium miners using generalizations of the two-mutation model of Moolgavkar, Venzon and Knudson. *International Journal of Radiation Biology*, *74*, 49–63.

Little, M., & Wright, E. (2003) A stochastic carcinogenesis model incorporating genomic instability fitted to colon cancer data. *Mathematical Biosciences*, *183*, 111–134.

Monticello, T., Miller, F., & Morgan, K. (1991). Regional increases in rat nasal epithelial cell proliferation following acute and subchronic inhalation of formaldehyde. *Toxicology and Applied Pharmacology*, *111*, 409–421.

Monticello, T. M., Morgan, K. T., & Hurtt, M. E. (1990). Unit length as the denominator for quantitation of cell proliferation in nasal epithelia. *Toxicologic Pathology*, *18*, 24–31.

Monticello, T., Swenberg, J., Gross, E., Leininger, J., Kimbell, J., Seilkop, S., Starr, T., Gibson, J., & Morgan, K. (1996). Correlation of regional and nonlinear formaldehyde-induced nasal cancer with proliferating populations of cells. *Cancer Research*, *56*(5), 1012–1022.

Moolgavkar, S., & Knudson, A. J. (1981). Mutation and cancer: A model for human carcinogenesis. *Journal of the National Cancer Institute*, *66*(6), 1037–1052.

Moolgavkar, S. H., & Luebeck, G. (1990). Two-event model for carcinogenesis: Biological, mathematical and statistical considerations. *Risk Analysis*, *10*, 323–341.

Moolgavkar, S. H., & Luebeck, E. G. (1992). Interpretation of labeling indices in the presence of cell death. *Carcinogenesis*, *13*, 1007–1010.

Moolgavkar, S., & Venzon, D. (1979). Two-event model for carcinogenesis: Incidence curves for childhood and adult tumors. *Mathematical Biosciences*, *47*, 55–77.

National Toxicology Program (NTP). (2005). *Control Data from NTP Technical Report Series Nos. 210, 267, 272, 306, 311, 314, 329, 346, 363, 371, 375, 377, 379, 380, 386, 421, 437, 440, 447, 448, 450, 451, 453, 454, 461, 462, 466, 467, 468, 471, 472, 475, 482, 484, 486, 487, 490, 492, 499, 500, 507, 513, 515, 519 and Control Data from the Wollastonite Calcium Silicates Study*. U.S. Department of Health and Human Services, NTP, Research Triangle Park, NC. Available at http://ntp-apps.niehs.nih.gov/ntp˙tox/index.cfm (accessed in July and August 2005).

Peddada, S., Dinse, G., & Kissling, G. (2007). Incorporating historical control data when comparing tumor incidence rates. *Journal of the American Statistical Association*, In Press.

Portier, C. J., KoppSchneider, A., & Sherman, C. D. (1996). Calculating tumor incidence rates in stochastic models of carcinogenesis. *Mathematical Biosciences*, *135*, 129–146.

Portier, C. J., Sherman, C. D., & KoppSchneider, A. (2000). Multistage, stochastic models of the cancer process: A general theory for calculating tumor incidence. *Stochastic Environmental Research and Risk Assessment*, *14*, 159–210.

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (1992). *Numerical Recipes in Fortran*. New York: Cambridge University Press.

Quievryn, G., & Zhitkovich, A. (2000) Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. *Carcinogenesis*, *21*(8), 1573–1580.

Rao, G. N., Plegorsh, W. W., & Haseman, J. K. (1987). Influence of body weight on the incidence of spontaneous neoplasms in rats and mice of long-term studies, *American Journal of Clinical Nutrition*, *45*, 252–260.

Sherman, C. D., & Portier, C. J. (2000). Calculation of the cumulative distribution function of the time to a small observable tumor. *Bulletin of Mathematical Biology*, *62*(2), 229–240.

Slikker, W. J., Andersen, M., Bogdanffy, M., Bus, J., Cohen, S., Conolly, R., David, R., Doerrer, N., Dorman, D., Gaylor, D., Hattis, D., ROgers, J., Setzer, R., Swenberg, J., & Wallace, K. (2004). Dose-dependent transitions in mechanisms of toxicity: Case studies. *Toxicology and Applied Pharmacology*, *201*(3), 226–294.

Speit, G., & Merk, O. (2002). Evaluation of mutagenic effects of formaldehyde in vitro: Detection of crosslinks and mutations in mouse lymphoma cells. *Mutagenesis*, *17*(3). 183–187.

USEPA. (2005). *Guidelines for Carcinogen Risk Assessment*. United States Environmental Protection Agency Risk Assessment Forum. Technical Report Number EPA/630/P-03/001b NCEA-F-0644b.