Information Analysis by Using PrecisionTree

Kan Shao, Mitchell J. Small
Department of Civil and Environmental Engineering
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
March 29, 2007

1. Introduction

A broad view of uncertainty in carcinogenic risk assessment points out that the key steps in cancer risk assessment includes a) a determination of whether the agent is carcinogenic in humans; b) estimation of the agent potency within the range of dosage used in an animal study; c) quantitative extrapolation of risk from the test species to humans; and d) high- to low-dose extrapolation to estimate risks for the dose range experienced in the exposed human population. [1, 2]

The distributional approach, which has also been referred to as information analysis, builds event trees that allow to take into account all of the available relevant information and characterizes the dose-response relationship in terms of a probability distribution by assigning different weights to different information and assumptions.

One of the major benefits of information analysis is that this method does not focus on only one experimental data set, one dose scale, one dose-response model or another single factor [3]. Rather, this method reflects weight-of-evidence evaluations of all of the available dose-response information. This is response to recommendation from EPA that combining information from multiple studies increases the confidence and decreases the uncertainty of the final dose-response estimate.

Information analysis quantifies and reports attributes, and then summarizes them individually. Summaries can be frequency distributions indicating how often a particular attribute value occurs or, preferably, weighted frequency distributions indicating the proportion of scientific support for each combination of factor alternatives [3]. To support the enumeration of assumption combination and the calculation of associated probabilities, we used PrecisionTree which is an event/decision tree function in the Palisade's DecisionTools software package. This paper is intended to use PrecisionTree

as an example to illustrate the application of commercial software in information analysis for cancer risk assessment.

2. Background

One of the initial applications of the information analysis was for detailed assessments of low-dose cancer risk for formaldehyde [3-5], which is based primarily on the use of animal bioassay studies to fit dose-response relationships. A distributional assessment of cancer risk due to exposure to formaldehyde was also conducted by Fayerweather et al. [6] In this study, the event tree models for formaldehyde address uncertainty in six elements of dose-response characterization that is derived primarily from animal bioassay results. Table 1 enumerates these six factors and all the alternatives in each factor.

Table 1. The Alternatives for the Six Major Dose-Response Factors in Event Tree [3]

Factor 1: Human Carcinogenicity: Relevant Human Target Tissue

1. A Human Carcinogen
2. Not a Human Carcinogen

Factor 2: Mode of Action: Carcinogenic Mechanism

1. Cell Proliferation Involved But Not Genotoxicity
2. Genotoxicity Involved But Not Cell Proliferation
3. Cell Proliferation and Genotoxicity Both Involved

Factor 3: Dose Scale

1. Formaldehyde Concentration in the Air Inhale (ppm)
2. Daily Intake of Formaldehyde
3. Amount of Formaldehyde Induced DNA protein cross-links (DPX)

Factor 4: Dose-Response Model

1. Threshold: Probit
2. Sublinear: Multistage (5 Stage)
3. Sublinear and Low-Dose Linear: Multistage (5 Stage) and Linear below 1.0 ppm [Abbreviation = Sublinear (LDL)]
4. Linear: One-Hit

Factor 5: Experimental Data Set

1. Malignant Squamous Cell Carcinoma in the Rat Nasal Cavity (Abbreviation = M)
2. Malignant Squamous Cell Carcinoma and/or Benign Poly-poid Adenoma in the

Rat Nasal Cavity (Abbreviation = M&B)

Factor 6: Interspecies Extrapolation

1. Human and Rat Response Probabilities are Equal when the Dose in Humans Equals the Dose in Rats (Abbreviation = Same)
2. Human and Rat Response Probabilities are Equal when the Dose in Humans is {[Rat Body Weight] / [Human Body Weight]}$^{1/4}$ Times the Dose in Rats
3. Human and Rat Response Probabilities are Equal when the Dose in Humans is {[Rat Body Weight] / [Human Body Weight]}$^{1/3}$ Times the Dose in Rats

Now, we are going to give some details in each factor and specify the weights we use in the event tree.

Factor 1. Human Carcinogenicity: Relevant Human Target Tissue

In this event tree for formaldehyde, the alternatives for the human carcinogenicity are simply that in human exposure situations either formaldehyde is a human carcinogen or it is not. The probability that formaldehyde is a human carcinogen was assigned a value of 0.8, consistent with its EPA and IARC classifications as a probable human carcinogen.

Factor 2. Mode of Action: Carcinogenic Mechanism

There are three alternatives considered for formaldehyde: a) cell proliferation (inducing cell growth) only; b) genotoxicity only; and c) both cell proliferation and genotoxicity. According to the studies of Metcello and coworkers [7-10], the predominant probability (0.8) was assigned to cell proliferation only, with 0.195 probability assigned to both cell proliferation and genotoxicity [11], and only a probability of 0.005 assigned to genotoxicity only.

Factor 3. Dose Scale for Dose-Response Modeling

The dose scale used for formaldehyde's cancer dose-response modeling is an important factor and has been paid considerable attention. For the formaldehyde application, dosimetry options include: a) the concentration in the inhaled air; b) total daily intake; and c) a biomarker-based measurement of the covalent bonding of formaldehyde to DNA in respiratory tissues, as measured by the amount of formaldehyde induced DNA protein cross links (DPX) [12]. Based on the then-current

understanding of mechanisms in rats and humans, they are assigned weight of 0.1 (ppm inhaled), 0.3 (total daily intake), and 0.6 (DPX).

Factor 4. Dose-Response Model

The shape of the dose-response relationship is closely connected to the assumed carcinogenic mode of action in Factor 2, and is also affected by the dose scale in Factor 3. Four options are considered in both the formaldehyde assessment: a) a probit model representing highly nonlinear, threshold-like behavior at low dose; b) a five-stage multistage model that results in sublinear relationships at low dose; c) a five-stage multistage model above 1 ppm formaldehyde, with linear interpolation to zero below 1 ppm; and d) a one-stage (or one-hit) model that is essentially linear throughout the range of doses considered [1]. After considering the factor of mode of action and dose scale, the different assigned weights of dose-response model are shown in Table 2.

Table 2. Assigned Weights in Dose-Response Model

		4. Dose Response Model Prob[Dose Response Model | Mode of Action ⋂ Dose Scale]			
2. Mode of Action	**3. Dose Scale**	Probit nonlinear threshold	5-stage sublinear	5-stage linear	1-stage linear
Cell Proliferation	ppm inhaled	0.5	0.45	0.05	0
Genotoxicity	ppm inhaled	0	0.15	0.5	0.35
Both	ppm inhaled	0	0.5	0.5	0
Cell Proliferation	total daily intake	0.5	0.4	0.1	0
Genotoxicity	total daily intake	0	0.1	0.5	0.4
Both	total daily intake	0	0.4	0.6	0
Cell Proliferation	DNA protein Xlinks	0.2	0.4	0.4	0
Genotoxicity	DNA protein Xlinks	0	0.025	0.225	0.75
Both	DNA protein Xlinks	0	0.1875	0.8125	0

Factor 5. Experimental Data Set

In the formaldehyde assessment the estimates were based on a CIIT (Chemical Industry Institute of Technology) rat inhalation study data. The two options for the data sets included are: a) MSCC only (assigned an initial weight of 0.8); and b) both MSCC and benign polypoid adenoma (assigned a weight of 0.2).

Factor 6. Interspecies Extrapolation

The interspecies extrapolation method is linked to the dose scale, with the use of DPX yielding a strong preference for human-rat equivalence. The bodyweight-power extrapolations generally yield higher risk estimates than those obtained using the human-rate equivalence assumption. [1]

Table 3. Assigned Weights in Interspecies Extrapolation

	6. Interspecies Extrapolation Prob[Interspecies Extrapolation | Dose Scale]		
3. Dose Scale	Human same as Rat	BW to 3/4ths	BW to 2/3rds
ppm inhaled	0.2	0.4	0.4
total daily intake	0.5	0.25	0.25
DNA protein Xlinks	0.8	0.1	0.1

3. PrecisionTree

The Palisade's DecisionTools is a software package can be purchased and subsequently embedded in the Microsoft Excel platform. PrecisionTree is one function in this software package. One of the advantages of PrecisionTree is that it is easy to build an event tree and assign different weights to each branch. At the end of each branch, the final probability will be calculated and presented automatically [13]. Figure 1 shows some end branches for calculating the cancer risk for inhalation exposure to formaldehyde, based on the information analysis in Quantitative Cancer Modeling and Risk Assessment by C. D. Holland and R. L. Sielken following the path of "Human Carcinogen —〉 Cell Proliferation Involved But Not Genotoxicity —〉 PPM —〉 Dosimetry Probit". We can easily define how many branches at each chance node (equivalent to an event node, indicated by a red circle) and give a name to this node, and

then input the weights for each branch. The weight on the outcome at the end of path in the event tree is equal to the product of the weights along that path. If there were many branches with different weights in the decision tree, we would need to input different weight for each branch. But in this study, there are many nodes with the same branches in the event tree. Therefore, we can copy and paste the same nodes to save time in the initial specification of model.

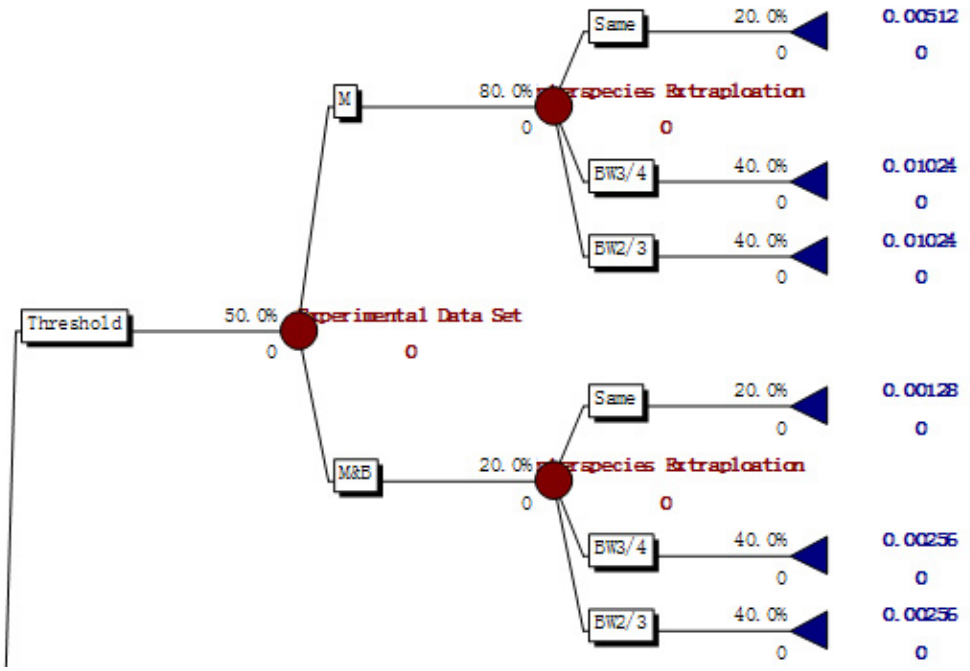

Figure 1. A Part of Branches in Event Tree

(Based on Portion of Tree for Cancer Risk from Formaldehyde, Lifetime Exposure to 1 ppb as Reported in Quantitative Cancer Modeling and Risk Assessment)

Another advantage of PrecisionTree is that this software is embedded in Excel, so we still can use the powerful functions and tools in Excel for statistical analysis. The following analysis is done by combining the results from the maximum likelihood estimate of cancer risk reported in Quantitative Cancer Modeling and Risk Assessment (Holland & Sielken) and the outcomes of the event tree.

There are six major factors in the quantitative dose-response characterizations to build the event tree. These six factors and all of the alternatives in each factor are listed

in Table 1. Because there are over 200 branches in the event tree, the entire tree and the weight assigned to each path will not be displayed here. A way of communicating the state of knowledge concerning the added risk at 1 ppb of formaldehyde is the weighted frequencies of the event tree outcomes. Alternative risk attributes include maximum likelihood estimates, and upper bound, and lower bound estimates. Here, I will communicate the weights with the maximum likelihood estimates and upper bounds, and then concentrate on showing the results and introducing the methods briefly.

4. Results

There are 216 values of the maximum likelihood cancer risk estimate which are linked with different weights at the end of branches. We first calculate the log 10 of the MLE risk estimate, i.e., the $\log_{10}$(MLE), sort these values into 19 categories, and then sum up the weights in the same category. The probability distribution and the cumulative distribution are show in Figures 2 and 3, which include a weight of 20% on Not a Human Carcinogen.

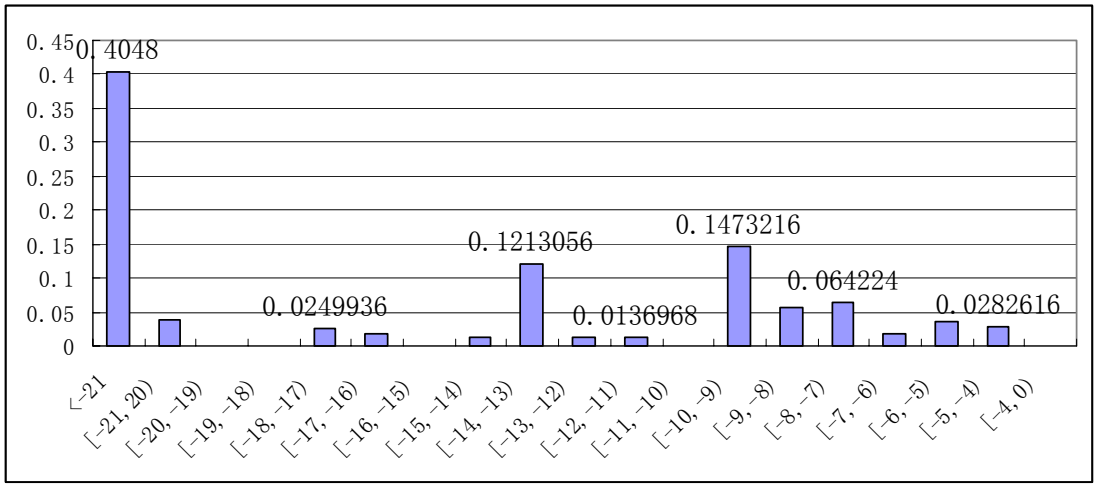

Figure 2. Probability Distribution of Log10 (MLE) including Not a Human Carcinogen

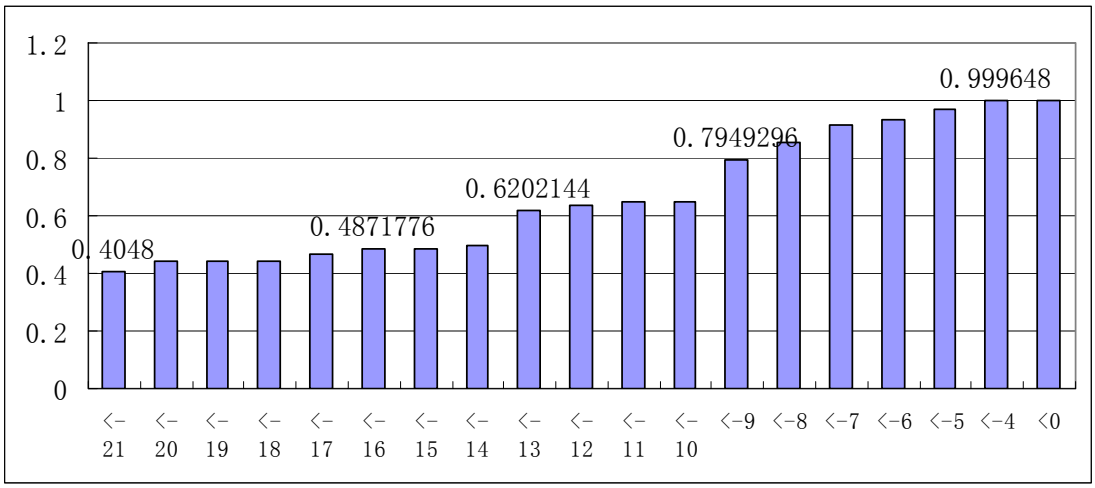

Figure 3. Cumulative Prob. Distribution of Log10 (MLE) including Not a Human Carcinogen

If we only consider the Human Carcinogen condition, we can obtain the results in Figures 4 and 5.

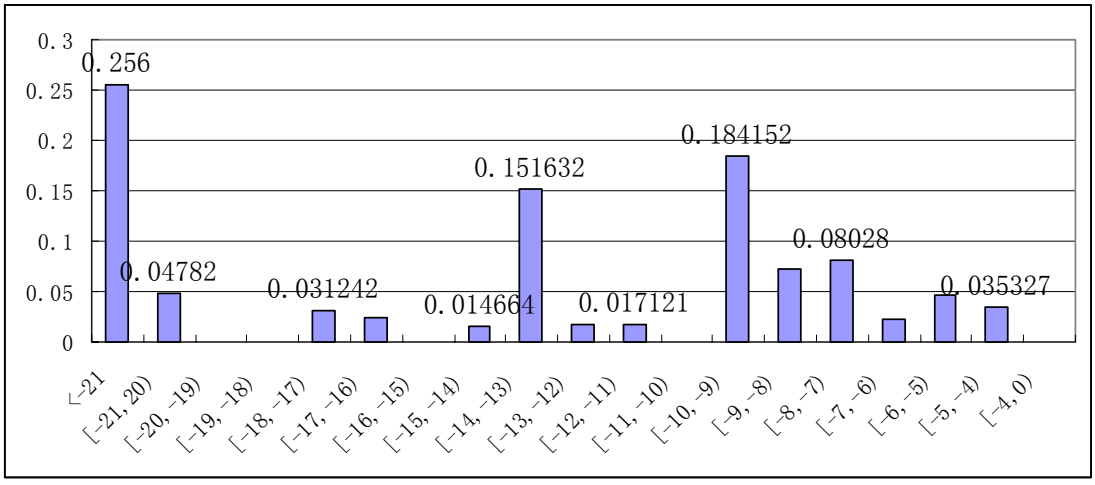

Figure 4. Probability Distribution of Log10 (MLE) only Human Carcinogen

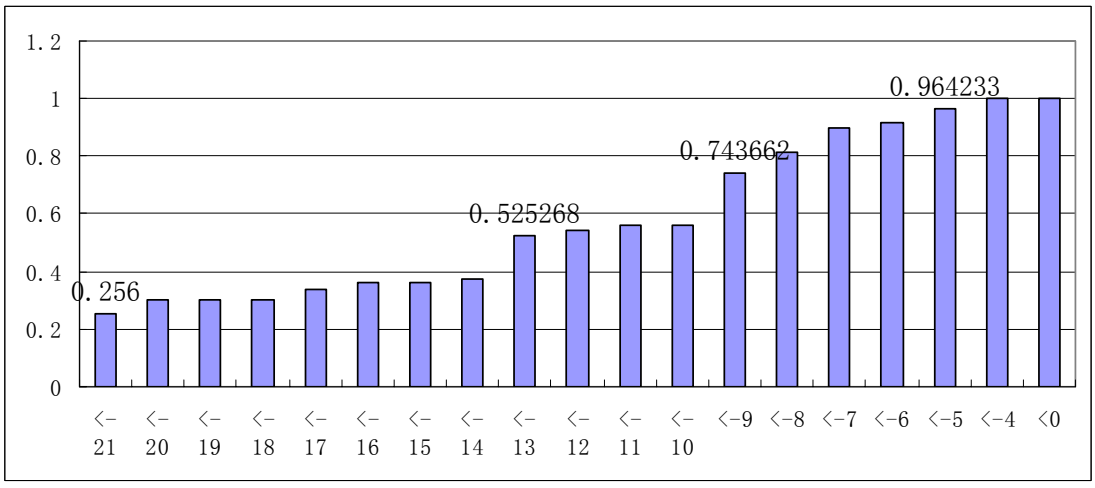

Figure 5. Cumulative Prob. Distribution of Log10 (MLE) only Human Carcinogen

Similarly, there are 216 values of 95th upper confidence limit which are linked with different weights at the end of branches. We calculate the log 10 of the 95th Upper Confidence Limit risk estimate, i.e., the $\log_{10}$ (95th UCL), sort these values into 6 categories and then sum up the weights in the same category. The probability distribution and the cumulative distribution are show in Figures 6 and 7, which include a weight of 20% on Not a Human Carcinogen. Figures 8 and 9 show the results that assume that formaldehyde is a Human Carcinogen.

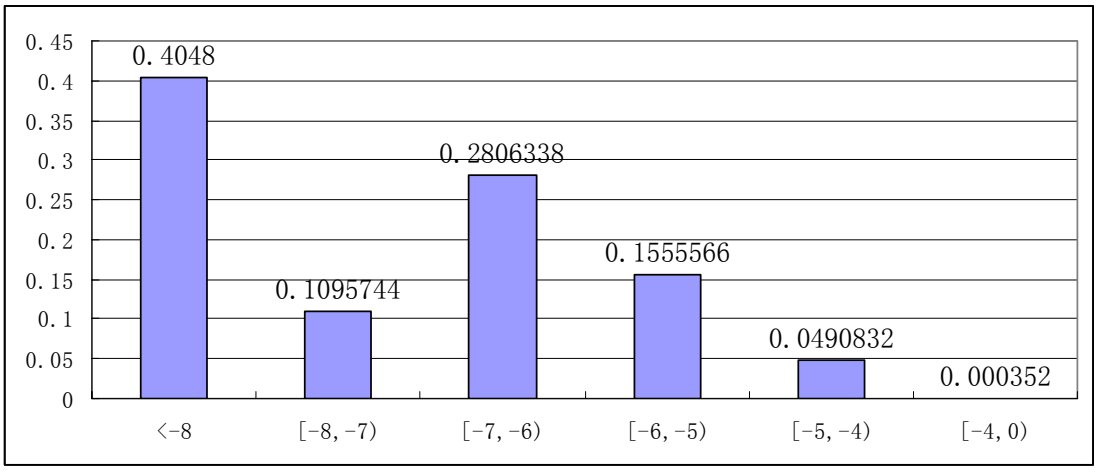

Figure 6. Probability Distribution of Log10 (95th UCL) including Not a Human Carcinogen

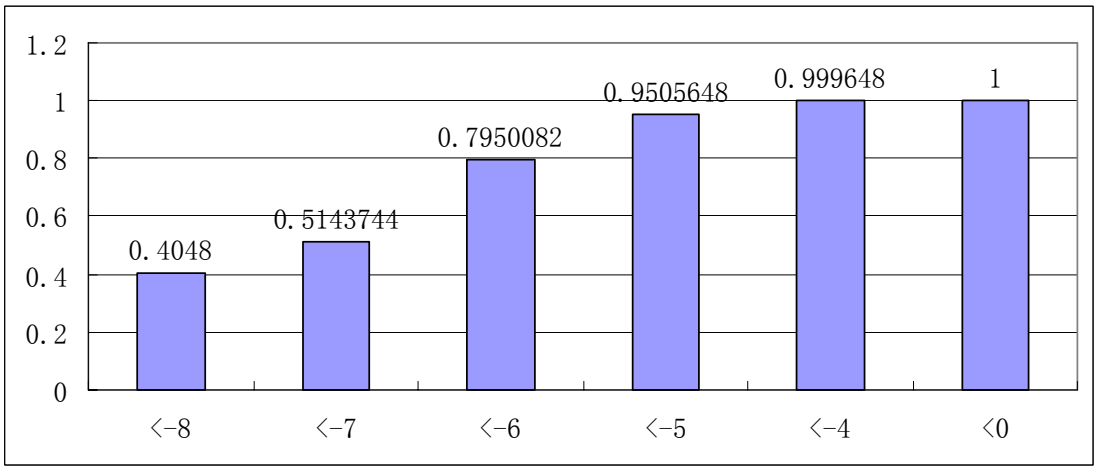

Figure 7. Cumulative Prob. Distribution of Log10 (95th UCL) including Not a Human Carcinogen

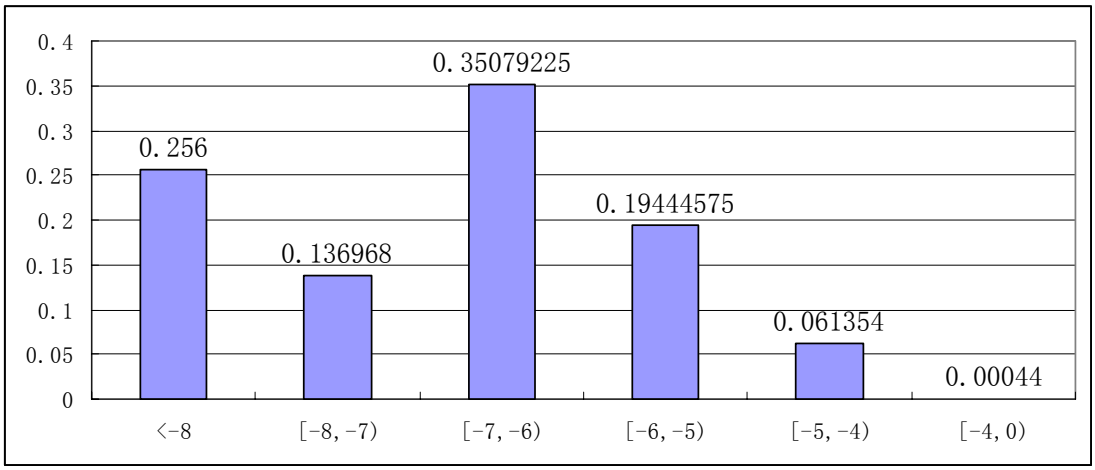

Figure 8. Probability Distribution of Log10 (95th UCL) only Human Carcinogen

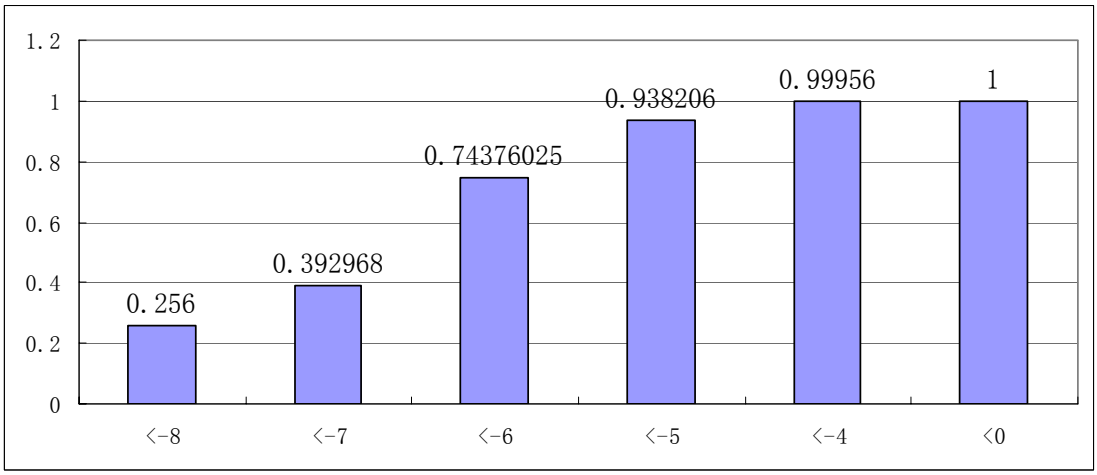

Figure 9. Cumulative Prob. Distribution of Log10 (95th UCL) only Human Carcinogen

5. Discussion

Using the PrecisionTree, we also did the cancer risk assessment well. Comparing to other professional software such as the “Bayesian Belief Networks” (BBNs), one of the advantages of PrecisionTree is that it can provide the exact value of each possible outcome. However, the most inconvenient feature of PrecisionTree is that it is not easy to change the assigned weights and also difficult to simulate different situations.

Reference:

[1] Small, M.J. A comprehensive look at uncertainty analysis of carcinogenic risk assessments: part 2 tools for assessing uncertainty in fundamental biological assumptions and associated models. U.S. EPA White Paper, March 2007.

[2] Cohen, J.T. A comprehensive look at uncertainty analysis of carcinogenic risk assessments: Taking a broad view. U.S. EPA White Paper, March 2007.

[3] Holland, Charles D. and Robert L. Sielken Jr. (1993). Quantitative Cancer Modeling and Risk Assessment, Prentice Hall, Englewood Cliffs, New Jersey. (Chapter 7).

[4] Evans, John S., John D. Graham, George M. Gray, and Robert L. Sielken Jr. (1994) "A Distributional Approach to Characterizing Low-Dose Cancer Risk," Risk Analysis, Vol. 14, No. 1, pp. 25-34.

[5] Evans, John S., John D. Graham, George M. Gray, and Robert L. Sielken Jr. (1995) "A Distributional Approach to Characterizing Low-Dose Cancer Risk," Low-Dose Extrapolation of Cancer Risks, pp. 253-274, ed. by Stephen Olin, William Farland, Colin Park, Lorenz Rhomberg, Robert Scheuplein, Thomas Starr, and James Wilson, ILSI Press, Washington, D.C.

[6] Fayerweather, W.E., J.J. Collins, A.R. Schnatter, F.T. Hearn, R.A. Menning and D.P. Reyner. (1999). Quantifying uncertainty in a risk assessment using human data. Risk Analysis, 19(6): 1077-1090.

[7] Morgan KT, Kimbell JS, Monticello TM, Patra AL, Fleishman A. (1991). Studies of inspiratory airflow patterns in the nasal passages of the F344 rat and rhesus monkey using nasal molds: relevance to formaldehyde toxicity. Toxicol Appl Pharmacol., 110(2): 223-240.

[8] Monticello TM, Miller FJ, Morgan KT. (1991). Regional increases in rat nasal epithelial cell proliferation following acute and subchronic inhalation of formaldehyde. Toxicol Appl Pharmacol., 111(3): 409-421.

[9] Monticello TM, Morgan KT. (1994). Cell proliferation and formaldehyde-induced respiratory carcinogenesis. Risk Analysis, 14(3): 313-319.

[10] Monticello TM, Swenberg JA, Gross EA, Leininger JR, Kimbell JS, Seilkop S, Starr TB, Gibson JE, Morgan KT. (1996). Correlation of regional and nonlinear formaldehyde-induced nasal cancer with proliferating populations of cells. Cancer Research, 56(5): 1012-1022.

[11] Goldmacher, VS; Thilly, WG. (1983). Formaldehyde is mutagenic for cultured human cells. Mutation Research, 116(3-4): 417-422.

[12] Schlosser, P.M., P.D. Lilly, R.B. Conolly, D.B. Janszen, J.S. Kimbell (2003). Benchmark Dose Risk Assessment for Formaldehyde Using Airflow Modeling and a Single-Compartment, DNA-Protein Cross-Link Dosimetry Model to Estimate Human Equivalent Doses. Risk Analysis, 23(3), 473–487.

[13] Clemen, R.T., Reolly T. (2001) Making Hard Decisions, Duxbury, Thomson Learning, California. (Chapter 3).