

発癌性を評価するモデルと適用例

統計数理研究所 柳 本 武 美
" 鎌 倉 稔 成

(1981年2月 受付)

序

我々の生活環境中には極めて多くの化学物質があり、毎年増え続けている。人間は環境中から摂食、呼吸、接触を通して体内にとり込んでいる。それらの害作用の中で近年特に社会的な関心が高いのが発癌性である。ヒトの癌の多くが環境中の化学物質に由来していると思なされていることから発癌物質を環境中から取り除く努力がなされている。

化学物質の発癌性を評価するとき、従前はその有無を問題にしていた。しかしデータの蓄積と共に発癌性の強さが問題にされるようになってきた。実際の問題として発癌性がないことを確かめることは出来ないので、その強さを評価するというアプローチをとる方がより現実的である。発癌性の強さの評価は、環境中での使用の適否を判断する基礎資料となる。またサルモネラ菌等を用いた簡易試験を評価する材料を与える。しばしば変異原性試験は発癌性を評価する動物実験に代替できると考えられ、またそのようなデータも与えられている [8]。しかし、発癌性の強さが評価されていない以上、代替できるか否か、実際には良く分からないというのが実状である。

本稿では動物実験から発癌性の強さを評価するための統計モデルを検討し、日本で得られたデータを適用してその必要性、妥当性を調べる。

1. 統計モデル

今動物種と関心の対象となる腫瘍（あるいは前段階とみられる病変）は特定されているとしよう。動物実験では比較的良く管理がいき届くので、性を除けば主として化学物質の濃度と観察時のみが重要な要因であるとみなされる。試験をする化学物質の濃度を d として飼育し、 t 日経過したとき腫瘍が認められる (1) か、否か (0) を表す変数を $X(d, t; e)$ とする。悪性腫瘍のように不可避的なものであれば $X(d, t; e)$ は t に関して増加する。適当な t_0 を固定すると $X(d, t_0; e)$ となって用量-反応関係を表わすようになる。

データから $P(d, t) = \Pr(X(d, t; e) = 1)$ を推定するためには $P(d, t)$ に適当な制約が必要である。ここでは可能な限り少ない母数で表現するようにパラメトリックな場合を扱う。その場合分布型を仮定することが必要となるがここではワイブル分布を仮定する。最も基本的なモデルは

$$(1) \quad P(d, t) = 1 - e^{-\alpha t^\gamma}, \quad \alpha, \gamma > 0 \quad (\text{モデル 1})$$

である。ただし α と γ は d に依存する。Peto ら [10] はこれをより一般化して

$$(2) \quad P(d, t) = \begin{cases} 1 - e^{-\alpha(t-\delta)^\gamma} & \text{for } t > \delta (> 0) \\ 0 & \text{for } t \leq \delta \end{cases} \quad (\text{モデル 2})$$

を提案している。ここで δ は潜伏期間に対応する。即ち DNA の損傷から癌化 (Initiation) して病的に病変が認められるまでの増殖 (Promotion) の期間に対応する。次節の議論の準

備の為に (1), (2) 式で $\gamma \rightarrow 0$ とすると各々

$$(3) \quad P(d, t) = 1 - e^{-\alpha} \quad (=p) \quad (\text{モデル 3})$$

$$(4) \quad P(d, t) = \begin{cases} 1 - e^{-\alpha} \quad (=p) & \text{for } t > \delta \\ 0 & t \leq \delta \end{cases} \quad (\text{モデル 4})$$

となることを注意する.

(1) で α は d の関数だが具体的には例えば

$$(5) \quad P(d, t) = 1 - e^{-(\alpha_0 + \alpha d^h)t^\gamma} \quad (\text{モデル 5})$$

が考えられる. ここで $1 - e^{-\alpha_0 t_0^\gamma}$ は $d=0$, $t=t_0$ のときの腫瘍が認められる確率, 即ち $t=t_0$ での自然発癌率にあたる.

一方実験によって得られるデータはやはり動物種, 腫瘍の種類, 性を特定すれば, 主な情報は実験動物 i について (d_i, t_i, x_i) でまとめられる. ここで

d_i : 投与用量
 t_i : 剖検時
 x_i : 病理所見, 腫瘍があれば 1, なければ 0

日米での標準的なプロトコール [9], [14] では用量は 3 段階, その内に, コントロールを含み動物は各用量, 各性 50 匹が標準となっている. 腫瘍の存在の有無は解剖して病的に判断される. 試験は 2 年が標準であり, 終了時全例剖検に付される. 途中死んだり, 著しく弱った動物は可能な限り剖検される.

2. 用量—反応モデル

簡単に解析するためにデータ (d_i, t_i, x_i) から t_i を無視することがある. 用量のステージを $d_1^\circ, \dots, d_k^\circ$ とするとこのデータは (d_j°, n_j, r_j) , $j=1, \dots, k$ で, まとめられる. ただし

$$(6) \quad \begin{aligned} n_j &= \# \{i | d_j^\circ = d_i\} \\ r_j &= \# \{i | d_j^\circ = d_i, x_i = 1\} \end{aligned}$$

で潜伏期を考慮することによって n_j の代わりに有効動物匹数 \tilde{n}_j を使うこともある. ここで \tilde{n}_j と \tilde{r}_j は

$$(7) \quad \begin{aligned} T_0 &= \text{Min} \{t_i; x_i = 1\} \\ \tilde{r}_j &= \# \{i | d_j^\circ = d_i, t_i \geq T_0\} \\ \tilde{n}_j &= n_j \end{aligned}$$

で定義される. データ $(d_j^\circ, \tilde{n}_j, \tilde{r}_j)$ $j=1, \dots, k$ から用量—反応解析がなされることも多い. ここで用量—反応解析は $k=3$ が多いことから, コントロール群と他の群と χ^2 -検定して発癌性の有無が判断されていることが多い.

明らかに次の命題が成り立つ.

命題 (i) データ (d_i, t_i, x_i) にモデル 3 を仮定すると $P(d)$ の最尤推定量 $\hat{P}(d)$ は

$$\hat{P}(d_j^\circ) = r_j/n_j$$

(ii) モデル 4 を仮定すると $P(d_j^\circ)$ の最尤推定量 $\hat{P}(d_j^\circ)$ と δ の最尤推定量 $\hat{\delta}$ は

$$\begin{aligned} \hat{P}(d_j^\circ) &= r_j/\tilde{n}_j \\ \hat{\delta} &= T_0 \end{aligned}$$

$X(d, t; e)$ が t について狭義増加とすればモデル 3, 4 は成立しないので, 現実的なモデルとはみなされない. 次の例は用量-反応解析に帰着させることの, まずさを示唆する.

例 塩化ビニールは労働衛生学の調査から発癌作用のあることが知られていた. 日本での研究は, [16] に報告されている. ラットを使って肝血管肉腫を調べた結果の要約が表 1 である [7]. この実験は他にも飼育期間を変えて行われていて, 高用量群では発癌作用以外の害作用によって寿命が短縮することが分っている. しかし原著では剖検には報告されていない.

この実験はデータの質が良く, 塩化ビニールが社会的に広く利用されていることから多くの研究者によって解析されている. しかしその引用は極めて恣意的に見え, 引用の規準がはっきりしない.

Rai ら [13] は, 全用量群を用いていた. アメリカの食品安全会議のレポート [15] では第 6 群迄を用いており, 筆者の 1 人もそのデータを引用して [18] 用量-反応モデルを比較した. Guess ら [3] は第 5 群までのデータを引用し, 更に Crump ら [2], Rai ら [12] は第 4 群までを引用して, 新しい解析法について吟味している.

実際的に考えて高用量群のデータを削除して解析することは妥当だと思われる. しかしある用量までは競合リスクが無視できて, それ以上は急に競合リスクの影響が大きくなるとは考え難い. 従ってどの群まで解析の対象とするかを定めることは困難である. Maltoni [7] によって報告されたこのデータは用量が高くなる割りに肝血管肉腫の発生が高頻度でないことに興味もたれている [17], [4]. その原因の本質は競合リスクにあると疑われている.

3. ワイブルモデルの巾乗母数

パラメトリックなモデルに限定してもモデルの設定には様々な問題を伴う. ワイブル分布以外の分布型を仮定することは分布の中央部を議論の対象にしているときには重要ではない. 計算が便利なことを考慮して選ぶ.

ワイブル分布を仮定したとき, その巾乗母数 γ を特定な値に固定することがある. $\gamma=3$, $\delta=0$ は経験的に $\gamma=3$ の近くであてはまることが多いという報告に, 由来している [8]. $\gamma=3$ とすると t を固定してもコントロール群と実験群を用いるとモデルの母数は一意に定まる. 推定すべき母数は実験群の用量を d とすると $\alpha(0)$, $\alpha(d)$ の 2 つになるからである. この事実に基づいて簡便法 (Simple Method) と呼ばれる手法が研究者の関心をひいていて多くの資料で利用されている. 簡便法では, (5) 式のモデルで $\beta=1$ を仮定する. 今剖検時 t が一定とすると, 標準時間 T での 50% の動物に腫瘍が発生する量, TD_{50} の点推定量は

$$TD_{50} = (\log 2) d (t/T)^{3/\gamma} \left[\log \left(1 - \frac{r_c}{n_c} \right) - \log \left(1 - \frac{r_e}{n_e} \right) \right]$$

ただし,

- d : 実験時の用量
- n_c, r_c : コントロール群の動物数, 腫瘍発生数
- n_e, r_e : 実験群

表 1 塩化ビニールをラットに 52 週間, 週 5 日, 1 日 4 時間吸入させて計 135 週間後に観察した結果

	塩ビ濃度 (ppm)	動物数	肝血管肉腫の発生数
1.	No treatment	58	0
2.	50	59	1
3.	250	59	4
4.	500	59	7
5.	2,500	59	13
6.	6,000	60	13
7.	10,000	61	9

である。通常 λ が一定でないときには、その算術平均が代用されている。 TD_{50} は発癌性の強さの典型的な測度であるが、これが上のようにデータの陽関数で表現できて大変に都合が良い。発癌性の強さが測れると簡易試験との比較が可能になる。簡易試験の代表として Ames によって開発されたサルモネラ菌を使った変異原性試験があり、Messelson ら [8] の比較研究、これに続く Ames, 杉村らの未発表の研究がなされている。

しかしながら $\gamma=3$ とか $\beta=1$ であるという根拠はどこにもない。疫学的アプローチから人癌の発生前年齢を調べると $\gamma=5\sim 6$ が得られている [1]。また用量-反応モデルでは化学物質によってあてはまる γ が違うことが経験的に知られている。発癌のプロセスとして DNA の損傷させる作用とその後の増殖を促進させる作用がかなり違った面があることを考えあわせれば γ が一定であるとは考え難いと思われる。従って簡便法はごく準備的な解析にのみ使えようである。

4. 腫瘍の発生と実験動物の死亡

腫瘍の発生の有無が動物の生存中から確認できる場合がある。古典的なコルタルを貼布して皮膚癌の発生を見るときには目で確認できる。人の場合には異常を訴えて入院する時点を確認することもある。いずれも腫瘍が発生してある程度の変化した段階で捉えられ、その段階も一応は客観的である。一方動物実験では剖検によって始めて腫瘍の発生を確認することが多い。実験中に死亡した動物を病理検索して腫瘍の存在が確認されたとすると両者の因果関係について2つの見方ができる。

(イ) 腫瘍の存在が死亡の原因である。

(ロ) 腫瘍の存在は死亡と全く関係がない。

おそらく腫瘍の存在は動物を弱らせるといった意味で関係しているが直接の死因は感染症など他の原因によると見なした方が妥当である場合が多い。そうした事実にも拘らず、実際の動物実験の規模から考えてモデルを簡単にするために(イ)か(ロ)のいずれか、出来れば一方、に分けた方が望ましいと考えられる。Peto [10] は、(ロ)の場合は偶然的 (incidental) と呼び、(イ)の場合、非偶然的 (nonincidental) であると呼んでいる。Peto [10] らを含め一般には(イ)が仮定される。

しかし我々は現実には、偶然的な腫瘍が多いと考える。特に、実験終了時に剖検に付した場合には動物の死の原因は実験の計画にあったから偶然的である筈である。

一方剖検による腫瘍が大きくて、飼育期間中の観察から腫瘍が非偶然的とみなした方が合理的な場合も当然ありうる。この場合動物 e が病的に腫瘍が見いだされる迄の $x_1(e)$ と、その後死に至るまでの $x_2(e)$ と、2段階に分ける必要がある。非偶然的な腫瘍が認められた剖検日 t_1 、偶然的な腫瘍が認められた剖検日 t_2 、腫瘍が認められなかった剖検日は各々

$$\begin{aligned} t_1 &= x_1(e) + x_2(e) \\ x_1(e) &\leq t_1 < x_1(e) + x_2(e) \\ t_2 &< x_1(e) \end{aligned}$$

と見なされる。モデルに新しい確率変数が入り、確率変数の和が入るので大変に複雑になってしまう。

5. 潜伏期間

細胞レベルで癌化してから病理検索によって確認されたり、あるいは臨床症状と示す迄には

増殖期がある。もしその増殖期間が一定であるとして実験期間がそれよりも短いとすれば、観察的には、発生がないことになる。用量-反応モデルにおいて、有効匹数 \tilde{n} が用いられる理由である。

潜伏期間の存在は明らかだけれども動物試験のデータ処理には必ずしも潜伏期間 δ を導入しなくてもよいと筆者らは考えている。その1つは δ は定数とは考え難いことによる。細胞を癌化させる作用は確率的であるが、増殖は一定である仮定は、モデルを簡単にはするが、現実的ではない。今 δ は確率変数であるとする、化学物質を与え始めてから、病理的症状を示すまでは2つの作用を表わす確率変数の和になる。この事情は潜伏期間を病理検索によって見いだされる時から死までの期間としても同様である。

発癌性の強さに主な関心があるときには敢えて潜伏期間を考える必要はない。もし潜伏期間の推定にも重要な関心があるとするればこれを考慮したモデルが必要になる。この場合は2式のモデルのように δ が定数であるモデルより確率変数と、みなした方が実際的である。確率変数とみなすとすれば、Gamma 分布のように再生性のある分布族の方が扱い易そうである。

潜伏期を推定するためには実験の計画の段階から配慮が必要である。通常の試験では計画的屠殺 (Scheduled Sacrifice) は試験終了時に限られる。初期の段階で発癌が認められるか否かは不明である。途中で計画的に屠殺する必要がある。一方試験は長期的な観察が必要だから必然的に実験動物の数の増加が求められる。

潜伏期間を考慮しないモデルの採用を示唆するもう1つの理由は、一般に潜伏期間の存在を印象づけている原因が単なる見かけである疑いがあることである。実際に初めて腫瘍が発見される時 T_0 は比較的大きいことが多い。しかし、その事実のみから正の δ を仮定する必然性はない。ワイブルモデル $1 - e^{-\alpha t^\beta}$ をあてはめて $\hat{\alpha}$, $\hat{\beta}$ を推定して $1 - \exp(-\hat{\alpha} T_0^{\hat{\beta}})$ を計算すると小さい値を得ることが多い。 T_0 より短いときには、確率的にみて腫瘍が発生しないようにみえる。データを表にプロットしたとき一見して正の潜伏期間が必然的にみえることと似たことは我々は別の場面で経験している。実際用量-反応モデルの解析において正の閾値を仮定する解析が提案されている。見かけ上閾値があるように見えるデータも正の閾値の仮定のないモデルで十分説明できる [19]。

以上の我々の経験と考察から簡単に標準的なモデルとして $P(t, d) = 1 - \exp(-\alpha(d)t^\beta)$ が第一選択のモデルとの見解を得た。実際の実験の動物数、実験の計画に応じて適宜改良すべきである。

以下の節では実際に日本で得られたデータの解析をする。

6. 過酸化水素

過酸化水素は殺菌剤、漂白剤として使われていた食品添加物であった。試験は広島大学原医研で行なわれた。マウスに 0.4, 0.1, 0 (コントロール) %の過酸化水素を含んだ水を自由摂取させ、試験期間は2年間観察している。雌雄各群 48 匹~51 匹が用いられ、死亡動物は腐敗が進行していたものを除いて剖検された。終了時には全例が剖検に付されている。試験は [9] に準拠して行なわれている。十二指腸に発生した Carcinoma, Adenoma, Hyperplasia, Erosion が注目された。詳しくは、論文として [6] で報告される予定であ

表 2 過酸化水素を経口投与したマウスの十二指腸における発生

H ₂ O ₂ 濃度 (%)	性	匹数	Er	Hyp	Ad	Ca
0.4	M	50	2	30	2	1
	F	49	3	31	0	4
0.1	M	51	0	16	2	1
	F	50	1	24	4	0
0 (Control)	M	48	0	2	0	0
	F	50	2	7	1	0

る。表2で各群の発生数を与える。

先ず、1節で述べたモデル1, 2, 3, 4をあてはめる。モデルに含まれる母数を推定するために最尤推定法を用いる。モデル3, 4については、2節にあるように容易に推定される。モデル1については $\log \alpha$ を改めて α とおいて ($-\infty < \alpha < \infty$)

$$(1') \quad P(d, t) = 1 - \exp(-\exp(\alpha + \gamma \log t))$$

と書き換えた方が計算が易いようである。モデル2も同様の工夫をほどこすがそれでも計算が複雑なので δ を固定して α と γ を求め、 δ を動かした。 δ は (7) 式の T_0 より小さいから

表3 過酸化水素を含んだ飲料水を飲んだマウスに認められた腫瘍の出現時間についての4つのモデルのあてはめ

用量	腫瘍	モデル	母数の数	対数尤数	推定された母数				
					δ	α	γ		
雄 の 場 合	0.4% AD-CA	1	2	-11.202	490.5	-8.729	.953		
		2	3	-9.702		-1.257	-.298		
		3	1	-11.348					
		4	2	-10.326					
	HY-CA	1	2	-18.907	410.5	-23.107	3.749		
		2	3	-9.812		.436	.112		
		3	1	-32.052					
		4	2	-10.326					
雌 の 場 合	0.1% AD-CA	1	2	-11.156	366.5	-9.206	1.019		
		2	3	-9.028		.339	-.619		
		3	1	-11.410					
		4	2	-10.881					
	HY-CA	1	2	-31.898	366.5	-6.814	.967		
		2	3	-28.631		1.017	-.294		
		3	1	-33.675					
		4	2	-29.514					
雌 の 場 合	0.4% AD-CA	1	2	-13.376	467.5	-11.974	1.506		
		2	3	-10.482		-.312	-.525		
		3	1	-13.768					
		4	2	-12.057					
		HY-CA	1	2		-14.606	311.1	-21.377	3.483
			2	3		-12.906		-5.420	1.136
			3	1		-28.036			
			4	2		-20.659			
	0.1% AD-CA	1	2	-13.030	456.5	-23.531	3.267		
		2	3	-11.873		-.938	-.285		
		3	1	-13.938					
		4	2	-12.438					
		HY-CA	1	2		-27.120	456.5	-19.088	2.959
			2	3		-24.121		1.145	-.187
			3	1		-34.296			
			4	2		-24.436			
0% HY-CA	1	2	-18.979	559.5	-35.075	5.150			
	2	3	-17.642		.256	-.334			
	3	1	-21.983						
	4	2	-18.277						

$0 \leq \delta \leq T_0 - 1/2$ として区間を 100 等分した値に限定した。厳密な意味での最尤法ではないが、我々の目的の為には十分であると考えた。実際 δ が T_0 に非常に近いモデルは受け入れ難い。生物学的にみて $\gamma \geq 0$ と見なす方が合理的であるが $\infty > \gamma > -\infty$ の範囲で計算した。 $\hat{\gamma} < 0$ が推定されればそれ自体モデルの悪さを示すものと考えた。

腫瘍としては Carcinoma か Adenoma のいずれかが+のとき+とみなした場合、腫瘍 CA-AD とそれに Hyperplasia を加えたいずれかが+のとき、+とみなした場合、腫瘍 CA-HY について計算した。報告された腫瘍は偶然的であるとみなした。性別、用量毎に計算した結果とそのモデルの尤度を表 3 に与えた。またモデル 5 についても計算して表 4 に与えた。雄、コントロール群では、実験終了時のみに腫瘍が認められていることがわかる。モデル 5 ではこの群を除外した。

表 4 過酸化水素によるマウスの腫瘍発生データに巾乗母数一定の同時ワイブルモデルをあてはめた結果。最右欄は巾乗母数が変化したときの尤度

腫瘍	性	対数尤数	α_1	α_2	α_3	γ	
AD-CA	M	-22.358	—	- 8.996	- 8.939	.986	-22.358
	F	-31.036	-18.985	-17.554	-17.382	2.348	-29.753
HY-CA	M	-62.260	-19.467	-17.674	-16.328	2.662	-56.914
	F	-60.958	-23.608	-21.910	-20.823	3.394	-60.705

表 3 でモデル 2 の巾乗母数 γ の推定値が負になっている場合が 9 例中 7 例となっており、このときの δ の推定量は最初の腫瘍が認められた日 T_0 に近い。我々は計算の簡単のため $\delta \leq T_0 - 1/2$ としたがその端点を取っていることが多い。これは供試動物数が少なく、従って腫瘍の認められたマウスが少ない場合、母数 δ を導入すると T_0 時点で腫瘍が発生した 1 匹のみのデータに強く影響されるためと考えられる。一方モデル 3 は、モデル自体受け入れ難いけれども、見かけ上の適合は他のモデルに比べて良い。0.4% 群の雌で腫瘍 HY-CA の場合、モデル 1, 2 に比べて有意に適合が悪いのが例外的になっている。この場合にはモデル 2 の巾乗母数 γ の推定値は 1.136 で他の例により、大きい値になっている、見かけ上の適合が良い事実は有効動物数 \tilde{n}_j と用いる伝統的な手法の妥当性を支持している。

モデル 1 はモデル 4 と同じく、モデルに 2 つの母数を含んでいるが、上に述べた場合を除いてあてはまりが相対的に悪い。得られた推定量を概観すると γ は [.953 ~ 7.929] の範囲に入って尤もらしく、モデル 2 のような破綻もなく、またモデル 4 のように極端に 1 点に強く影響をうけることもない。

モデル 1 がモデル 4 より適合が悪いことは、腫瘍の発生数が多くない場合には構造的に起る疑いがある。最初の腫瘍発見例のデータに影響されると考えるからである。この直観的な疑いをチェックするために次のような小規模なシミュレーションを行なった。モデルとして $1 - \exp(-e^{-\alpha+2 \log t})$, $\alpha = 15, 16$ とおいて $n = 50$ として t_i としては最大 730, 最小 130 とした。剖検匹数として $t = 730$ で 20 匹とし、他は等間隔に各 1 匹を布置した。ほぼ今の実験の結果に似せた。データを、乱数を用いて発生させモデル 1 ~ 4 をあてはめてそのあてはまりの良さを尤度でみたところ上に述べた直観と一致していた。

以上の考察では用量を固定して別個に議論した。次に各用量同時に 1 つのモデルにあてはめる。本来モデル 5 のように用量の効果も関数で表現をしておくことが望ましいが、上の場合用量が 3 群に限定されているので

$$P(d_i, t) = 1 - e^{-\alpha_i t^\gamma}, \quad i = 1, \dots, k \quad (\text{モデル 6})$$

ここで k は用量群の数 (この例では 3) とする. このモデルでは γ が d に依存しない制約がある. データにあてはめて得られた結果が表 4 である. 腫瘍 AD-CA で雄の場合発生が 0 なので α_1 が計算できない. 但し実験終了時のみ腫瘍が認められた場合でも他の用量で γ が推定できればこの場合, 用量を固定した場合とちがって, 推定が可能である. 推定された γ の値は .986~3.394 で通常常識的に考えられている範囲である. Messelson ら [8] は, 3 に近いとみなしている. γ 一定の仮定を調べてみる. この仮定は常識的にみてやや無理があるが, モデルを簡単にするために近似的に成り立って欲しい仮定である. 表 4 でモデル 6 の下での尤度を与える. これは表 3 モデル 1 の尤度を用量についての和と比べる. この場合実験終了時のみ腫瘍がみられた例ではモデル 1 で $\gamma \rightarrow 0$ としてモデル 3 に近づけられるので, モデル 3 の尤度を用いた. 尤度の和は表 4 の最右欄に与えた. モデル 6 の方が推定した母数が 2 だけ少ない (腫瘍 AD-CA, 雄では 1) ことから, 腫瘍 HY-CA, の雄のみ, 有意であるとみなされる. 腫瘍 AD-CA, 雌の場合各用量により γ の推定値が違うが, 実験のサイズが小さく腫瘍の発生数が小さいことが関与していると考えられる.

7. AF-2

AF-2 はハム・ソーセージ, 豆腐など加工食品の防腐剤として巾広く使用されていた食品添加物であった. その変異原性は極めて強力である.

国立がんセンターで発癌性試験がハムスターを用いて行われた. 試験期間は 660 日で前胃に発生した乳頭腫と癌が実験者の関心をひいた. AF-2 投与群は 0.08% 群と 0.16% 群より成り各群雌雄 25 匹, 対象群は各 20 匹より成っている. ここでの性の区別はせず併せて解析する. 444 日に死亡したマウスに乳頭腫が見つかったのが最初で, それ以前の剖検日は入手できなかった. 対象群には腫瘍は発見されず, AF-2 投与群のみに現われた. 図 1, 2 は投与群での剖検日 (444 日以後) と腫瘍の有意を図示した.

乳頭腫と癌は共に偶然的であると見なした. 癌は殆んど実験終了時に観察されており, 癌は乳頭腫より重篤な病変であるからいずれかが認められるとき腫瘍が認められるとして解析し

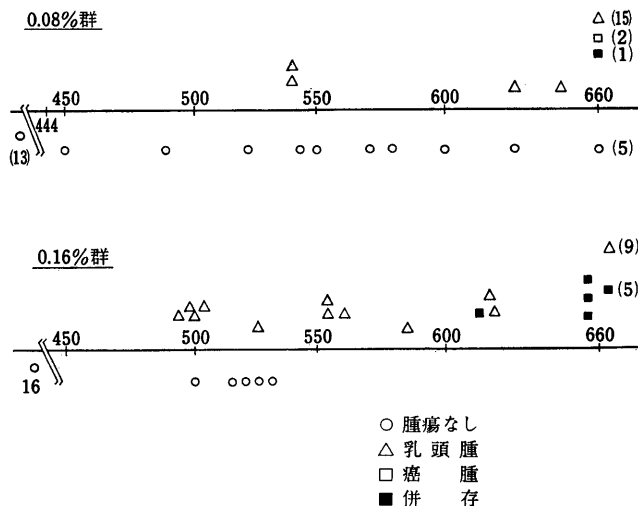


図 1 AF-2 を含んだ飼料をハムスターに投与した実験での剖検時と腫瘍
 図中 () 内の数字は重複数を示す

表 5 AF-2 をハムスターに投与して腫瘍の発生を観察したデータへのモデルのあてはめ
腫瘍は乳頭腫か癌

用 量	モデル	多数尤度	推定された母数				
			δ	α	α_1	α_2	γ
0.08%	1	-22.336		-47.178			7.325
	2	-22.336	0	-47.178			7.325
	3	-24.541					
	4	-34.497					
0.16%	1	-8.316	417.9	-98.630			15.782
	2	-8.001		-10.014			
	3	-12.315					
	4	-33.651					
0.08 U 0.16%	6	-31.779			-61.661	-59.829	9.569

た。前節に倣って解析した結果が表5にまとめられる。但し443日以前の正確な剖検日は得られなかったので仮りに330日と444日の間に等間隔に配置した。

モデル1~4の内ではモデル1のあてはまりの良さが顕著であり、表4と対照的である。 γ の推定値は7.325, 15.782と大変に大きい。実際実験終了時の近くでの剖検例では殆んど腫瘍が認められている。

モデル6をあてはめると共通の γ の推定値として、9.569が得られる。両用量群のモデル1の尤度の和からモデル6の尤度を引いて2倍すると2.254になる。両群の γ の値が共通とみなして良さを調べる規準として検定法を用いる時、上記の値は漸近的に自由度1の χ^2 -分布の実現値と見なされる。従って統計的に有意水準5%、有意でなく、 γ が共通と見なしておいてデータが一応説明できることになる。

8. OPP-ナトリウム

OPP-Naは輸入柑橘類にかび防止剤として現在も使用が認可されている食品添加物である。柑橘類に10ppm以下となっていて、WHOの1日摂取許容量は体重1kg当り1mgである。

実験は東京都衛生研究所においてラットを用いて行なわれた。ラットは7群に分けられ、各群21匹が割りつけられた。0, 0.125, 0.25, 0.5, 1.0, 2.0, 4.0%のOPP-Naを含んだ飼料が投与された。当初は優性致死試験として開始されたが、55週で死亡したラットに膀胱腫瘍が見られ、その後も続いたことから91週で生存動物が解剖された。同時に各群10匹を用いて13週の悪急性試験がなされ、尿路系の組織学的検査が行なわれた。実験の結果から発癌性を示すとみなす研究者と実験の不備を指摘する研究者とあって議論が分かれているが、実験の結果のみから判断すると、前者の方が尤もらしいので、ここでは議論に立入らず発癌性を認めた上で考慮する。

実験中に死亡して剖検により認められた腫瘍は非偶然的であるとみなされている。91週の剖検時に認められた腫瘍は非偶然的と考え難いから、死亡との関係からみて偶然的、非偶然的な2種類の癌があることになる。各群の動物数も少ないこと、また実験終了まで生存していた動物数が多いことから、用量-反応モデルにより解析を主とした。この場合剖検日が偏りとして働くことに注意すると、9週目に死亡した0.5%群の1例が他の例(55週日以降)に比べて早いことから除外することが考えられるが、結果に影響がないので除外しない。

表6 OPP-Naを含んだ飼料を91週投与した雄ラットに出現した尿路系腫瘍

飼料中濃度	剖検匹数	尿路系腫瘍 ^{a)}	腎の移行皮癌	膀胱の移行上皮癌・癌肉腫	途中剖検匹数 ^{b)}
0	20	0	0	0	1(2)
0.125	20	0	0	0	1(2)
0.25	20	0	0	0	0(1)
0.5	21	1	1	0	2 ^{c)}
1.0	21	7(8)	1	5(6)	2
2.0	21	20	1	19	7
4.0	20	17(20)	13(18)	8	3(4)

a) ()内は Hyperplasia を含む b) ()内は剖検不能を含む c) 内1例は9週目

膀胱の腫瘍のデータをみると、2.0%群では19/21の割合で発生しているのに対して、4.0%群では8/20にすぎず、プロビットモデルなど通常のモデルには適合しない。しかし高濃度でデータがゆらぐことは生物学的にみてもあり得ることである。次に腎臓では、4.0%群で急激に増加している。プロビットモデルをあてはめると、4.0%群での増加が急激であてはまりは良くない。2.0%群では膀胱に、4.0%では腎臓にと異った臓器に腫瘍が認められている。両臓器は尿路系器管としてみなすことができるから、改めて尿路系の発癌を調べると、

$$\hat{P}(d) = \Phi(1.501 \log d - 0.370)$$

となる。 χ^2 -値は通常の手段により低用量群をまとめて10.489(自由度2)となって適合は良くないが、膀胱あるいは腎臓のみに比べて、あてはまりが良くなる。更に良性腫瘍であるHyperplasiaを含めると、4.0%群では20例中全てに病変が認められてあてはまりが良くなる。この事実は尿路系器管としてまとめて扱う1つの根拠とみなされる。

表7 OPP-Naを含んだ飼料を13週投与したラットに出現した膀胱腫瘍

飼料濃度	供試匹数	剖検匹数	腫瘍の認められた匹数
0	10	7	0
0.125	10	10	0
0.25	10	10	0
0.5	10	9	0
1.0	10	8	1
2.0	10	10	9
4.0	10	10	1

次に剖検時を考慮に入れた解析について議論する。表7はまず同一の条件でなされた別の実験であるが、13週で生存動物がすべて解剖されている。結果は91週飼育した場合と良く似ている。2.0%群では9/10に腫瘍が認められ、4.0%群では1/10と少ない一方6匹に腎不全が認められている。早い時期に腫瘍が発生(病理検索で見つかるまで大きくなっている)していることと再現性があることが認められる。この場合単純に simple method を用いて発癌性の強さを評価すれば2つの実験で大きく食い違う。

この食い違いは simple method のモデルの不適切さと共に長期試験の途中死亡ラットに認められた腫瘍は非偶然的とみなされていることにも依存すると考えられる。

実験打ち切り後の剖検で認められた癌は偶然的とみなされる。用量 d を固定すればその尤度は

$$L = \prod_{i=1}^{n_0} p(d, t_i)^{x_i} (1 - P(d, t_i))^{1-x_i} \cdot \prod_{j=n_0+1}^{n_0} P(d, T)^{x_j} (1 - P(d, T))^{1-x_j}$$

ここで $p(d, t) = \frac{\partial}{\partial t} P(d, t)$, T は実験終了時, $t_i < T$, $i=1, \dots, n_0$ となる。1%, 2%,

4%群について尿路系の腫瘍についてデータをあてはめる。用量を固定したとき $P(t)$ はモデル 1 と同様に Weibull 分布関数と仮定すると

$$\begin{aligned} d = 1.0\% \text{ のとき} \\ \hat{\alpha} = -97.590 \quad \hat{\gamma} = 15.000 \\ d = 2.0\% \text{ のとき} \\ \hat{\alpha} = -52.811 \quad \hat{\gamma} = 8.327 \\ d = 4.0\% \text{ のとき} \\ \hat{\alpha} = -87.990 \quad \hat{\gamma} = 13.764 \end{aligned}$$

となる。このような解析も将来研究の余地があるが、そのためには実験の段階で動物の死因をどう評価するかという困難な問題の解決をまたなければならない。

謝 辞

この研究は「人がんの原因としての発癌物質の短期検索法の開発と評価に関する研究」(厚生省がん研究助成金(55-指)代表者 河内卓氏)より、補助をうけた。河内卓、伊藤明弘、平賀興吉の各氏にはデータを解析する機会を与えて頂き、また討論にも応じて頂いた。高橋淑子さんは数多くのプログラムを手際よく提供してくれた。併せて厚くお礼致します。

参 考 文 献

- [1] Armitage, P. and Doll, R. (1954). The age distribution of cancer and a multistage theory of carcinogenesis, *British Jour. Cancer*, **8**, 1-12.
- [2] Crump, K.S., Guess, H.A. and Deal, K.L. (1977). Confidence intervals and test of hypotheses concerning dose response relations inferred animal carcinogenicity data, *Biometrics*, **34**, 437-451.
- [3] Guess, H., Crump, K. and Peto, R. (1977). Uncertainly estimates for low-dose-rate extrapolations of animal carcinogenicity data, *Cancer Research*, **37**, 3475-3489.
- [4] Haseman, J.K., Hoel, D.G. and Jennrich, R.I. Some practical problems arising from use of the gamma multi-hit model for risk estimation, unpublished.
- [5] Haward, P.H. and Durkin, P.R. (1978). Human health effects: Holoethers, A draft prepared for Env. Criteria and Assess. Office, US EPA.
- [6] Ito, A., Watanabe, H. and Naito, M. Prevalence of gastric erosions and duodenal tumors with continuous oral administration of hydrogen peroxide in C57BL/6J mice, submitted to *Cancer Research*.
- [7] Maltoni, C. (1975). The value of predictive experimental bioassays in occupational and environmental carcinogenesis, an example: vinyl chloride, *Ambio*, **4**, 18-23.
- [8] Messelson, M. and Russell, K. (1977). Comparisons of carcinogenic and mutagenic potency, *Cold Spring Harbor Laboratory*, 1473-1481.
- [9] 小田嶋成和 (1976). 化学物質の癌原性検索法指針, 厚生省がん研究助成金に依る研究班資料.
- [10] Peto, R. and Lee, P. (1973). Weibull distributions for continuous carcinogenesis experiments, *Biometrics*, **29**, 457-470.
- [11] Pike, M.C. (1966). A method of analysis of a certain class of experiments on carcinogenesis, *Biometrics*, **22**, 142-161.
- [12] Rai, K. and van Ryzin, J. (1979). Risk assessment of toxic environmental substances using a generalized multi-hit dose-response model, *Energy and Health, SIAM*, 99-177.
- [13] ——— (1980). The use of quantal response data to make predictions, *Proc. of Symp. on Scientific of Toxicity Assesment*, (ed. Witshi, H.R.) North-Holland, 273-290.
- [14] Sontag, J.M., Page, N.P. and Saffiotti, U. Guidelines for carcinogen bioassay in small rodents, *DHEW Publication*, 76-801, U.S.A.
- [15] Scientific Committee, Food Safety Council (1977). Proposed system for safety assessment.
- [16] 田中豊穂, 吉村 功 (1977). 塩化ビニール製造作業者の死因と中毒の可能性について, I. II. 科学 47

- 巻, 342-350, 413-421.
- [17] 柳本武美 (1980). 用量-反応モデルにおける許容分布とその母数の制約, 京大数理解析研究所講究録, 384, 48-58.
- [18] Yanagimoto, T. and Hoel, D.G. (1980). Comparisons of models for estimation of safe doses using measures of the heaviness of tail of a distribution, *Ann. Inst. Statist. Math.*, **32**, B, 465-480.
- [19] Yanagimoto, Y. and Yamamoto, E. (1979). Estimation of safe doses—critical review of the hockey stick regression method, *Env. Health Persp.*, **32**, 193-199.

Statistical Models to Assess Carcinogenic Potency
with their Applications

Takemi Yanagimoto and Toshinari Kamakura
(The Institute of Statistical Mathematics)

Models to assess animal experiments for carcinogenesis are studied with special reference to the assumption on latency of observable tumors and the evaluation of tumors found in sacrificed animals. Time to tumor occurrence models based on the Weibull distribution seem to be preferable to either binary response models or the simple model which are often used in carcinogenic potential risk assessment of chemical compounds.

Practical data sets analysed here include those of hydrogen peroxide, AF-2 and OPP-NA, which are (were) used as food additives. The experiences of the model fitting suggest that the assumption on positive latency may be unnecessary in some cases, and tumors found in sacrificed animals are regarded to be incidental.