

《会員寄稿》

腫瘍発生時期の定量的検定法の模索

財団法人 食品農医薬品安全性評価センター 小林克己、大橋信之、廣内康彦  
岩田 聖、井上博之

はじめに

がん原性試験 [1] に於ける腫瘍発生数の統計処理は、一般的に腫瘍毎に試験期間に発生した動物数と供試動物数を用い、フィッシャーの確率計算法または  $X^2$  検定等が常用されている。しかし、対照群に比較して、より早期に発生した腫瘍か否かは不明である。途中死亡および切迫屠殺動物の検索による判定も難しいことがあり、腫瘍の発生時期を正確に把握することはできない。これらの問題点を考慮した統計法として Peto 検定 [2] (1980) が発表されている。この検定法は用量の要因、途中死亡および切迫屠殺動物について、発生腫瘍が死因か否か、死因に二次的な関与をしていたか、死因とは関係のない腫瘍かなどの判断を要求されている。このように病理学者の主観的、経験的判定を必要とし計算法も煩雑である。また、対照群に比べてどのくらい早期に腫瘍が発生したかを定量的に捉えることはこの検定法でもできず、最終的な統計検定の結果は対照群に対する各用量群との比較ではなく用量相関性をもって有意な増加と判断することが特徴である。しかし、この手法はフィッシャーの確率検定法または  $X^2$  検定に比べると、途中死亡および切迫屠殺動物の死因が腫瘍によるものか否かまた計画屠殺時に認められた場合などによって検定結果が異なる。癌 (悪性腫瘍) 発生が薬剤投与によって早期化されたか否かは癌原性試験における一つの重要な知見とされている [3]。たとえ腫瘍発生動物数が投薬群と対照群との間で有意差を示さない場合でも、投薬群で早期に腫瘍の発生を示している場合もありうる。しかし、癌発生時期を定量的に把握するための統計法の利用は見あたらない。この理由としては下記の点が考えられる。

①腫瘍発生の有無は試験途中で解剖しなく

ては把握できない。

- ②前癌病変および癌化早期化の判定法は完全ではない。
- ③従来は癌発生個数 (発生率) や発生動物数などが重要視されていた。
- ④データとして定量的な数値の大きさが十分でない。特に悪性腫瘍の発生が比較的少なく (0~10%)、通常1%内外のため、適当な手法が検討されなかった。
- ⑤データが正規分布しないため、パラメトリック検定ができない。
- ⑥計画屠殺動物と途中切迫・死亡動物に発生した該当腫瘍の取り扱いが難しい。

臨床観察で確認できる皮膚乳腺腫と死亡率の関係を含めて投薬の影響を統計学的に評価する方法については Mcknight [4] により幾つかの方法が紹介されている。

目 的

腫瘍発生動物数の統計処理法は一般にフィッシャーの確率検定法等のカテゴリック検定が最も常用されている手法である。これらの検定は試験終了時に解剖した動物中に何頭腫瘍発生が認められたかを吟味する。または途中死亡および切迫動物を含め試験期間中に腫瘍を発生した全ての動物数と供試動物数を用いて検定を実施している。また小林 [5] は良性・悪性および総腫瘍発生個数 (1頭の動物に複数の発生を示す) の検定には  $X^2$  検定の標本度数と理論度数との比較検定によって実施することを推奨している。しかし、対照群と投薬群の総腫瘍発生動物数が同じかまたは投薬群の方が少ない場合や、投薬群が対照群よりも早期に腫瘍が発生していても発生腫瘍数 (動物数) が重要視されるために、この早期発生の評価が十分でない可能性が存在する。

今回、著者らは実際の医薬品の癌原性試験成績を基に、対照群に比べてどの程度早い腫

瘍発生が有意性をもっているかについて有効な手法を模索した結果、ある知見を得たので報告する。

材料および方法

F344/DuCrjを用いた2年間(104週)の癌原性試験の成績について検索した。

1) 試験群と動物数

群の設定は雌雄それぞれ対照群を含めて低、中、高用量群の4用量群、1群50頭の計400頭を用い、104週間混餌投与を実施した。

2) 統計処理対象腫瘍

肺の肺胞/細気管支上皮腺腫、肺の肺胞/細気管支上皮癌およびリンパ節の悪性リンパ腫の良性および悪性腫瘍の3種各雌雄について検討した。

以上3種の腫瘍の発生時期(週齢)および基礎数値は、表-1、2および3に各々示した。

3) 統計モデルの選択理由

群によっては同数値のみで分散が0の場合もあり、データも正規分布を示さない。したがって、順位検定系のクラスカル・ワリス検定で有意差を認めた場合にはウイリアムズ・ウイルコクソン検定により2群間の検定を実施した。また、発生動物数の検定をフィッシャーの確率計算法によって実施した。なお、全群間を検定するクラスカル・ワリス検定を

用いて有意差が認められた場合には、2群間の差を検定するこのウイリアムズ・ウイルコクソンの他に、ノンパラメトリックなダンネット、チュキーおよびシェフェの検定の3法が一般に使用されている。ウイリアムズ・ウイルコクソン検定を含めこれらの4検定法は、ほぼ同様の検出力を示す。したがって、今回はウイリアムズ・ウイルコクソン検定を使用し検討した。

4) 統計モデル

a: フィッシャーの確率計算法 (Fisher's Exact Test) は2群間の検定。

b: クラスカル・ワリス (Kruskal-Wallis) 検定は全群間の検定。

c: ウイリアムズ・ウイルコクソン (Williams-Wilcoxon) 検定は2群間の検定。

d: ヨンキー (Jonckheere) の傾向検定は全群間の検定。この傾向検定は全群間を検定し、薬剤投与による影響が投薬量と用量相関性を示して発生したかどうかを検討するものである。

以上の統計ソフトはマスコットNo.1 (ユックムス株式会社、東京) を用いた。

表-1 肺の肺胞/細気管支上皮腺腫の発生週齢の生データ

途中解剖および計画解剖によって発生が認められた週齢							
対照群	低用量群	中用量群	高用量群	対照群	低用量群	中用量群	高用量群
104	77 104	84 104	79 104	104	78	66 101	32 71
104	89 104	90 104	82 104	104	78	77 101	47 75
104	97 104	95 104	86 104	104	96	81 104	63 76
104	104 104	99 104	90 104		102	83 104	64 82
104	104 104	100 104	97 104		102	85 104	66 83
104	104	101 104	100 104		104	86 104	67 83
104	104	103 104	101		104	89 104	68 85
	104	104 104	104		104	94 104	68
	104	104 104	104		104	95	69
	104	104 104	104		104	97	70
動物数	7	15	20	16	3	10	18
平均順位	38.5	30.2	29.2	25.4	41.5	33.7	29.5
平均値 ± S. D.	104 ± 0	100 ± 8	101 ± 5	98 ± 9	104 ± 0	98 ± 11	93 ± 11

表-2 肺の肺胞/細気管支上皮癌の発生週齢の生データ

途中解剖および計画解剖によって発生が認められた週齢							
対照群	低用量群	中用量群	高用量群	対照群	低用量群	中用量群	高用量群
95	85 104	72 100	67 103	103	78	74 103	70
100	90 104	77 103	82	104	86	78 104	72
	92	92 104	92	104		85	77
	104	97	100			96	83
動物数	2	6	7	5	3	2	6
平均順位	11.5	12.4	10.1	8.3	13.2	7.3	8.7
平均値 ± S. D.	98 ± 4	97 ± 9	92 ± 13	89 ± 15	104 ± 1	82 ± 6	90 ± 13

表—3 リンパ節の悪性リンパ腫の発生週齢の生データ

	途中解剖および計画解剖によって発生が認められた週齢							
	対照群	低用量群	中用量群	高用量群	対照群	低用量群	中用量群	高用量群
動物数	91	84	60	74	95	74	77	70
	97	90	90	104	104	104	97	71
	104	102	94	104	96	81	81	75
	104	104	96	90	103	104	85	
	104	104	100	100	104	104	91	
動物数	5	5	8	9	7	6	8	3
平均順位	17.7	14.7	13.8	11.7	17.0	13.3	11.8	2.3
平均値±S.D.	100±6	97±9	94±15	93±11	100±5	90±15	90±10	72±3

結果ならびに考察

肺の肺胞／細気管支上皮腺腫、肺の肺胞／細気管支上皮癌およびリンパ節の悪性リンパ腫の病変をクラスカル・ワリス、ウイリアムズ・ウイルコクソンおよびヨンキーの各順位和検定を用いて検索した結果を表—4に示した。各データとも平均順位（早期発生順位、mean of rank）は対照群が最下位（数値は大）であり、一見して用量が増加するにしたがい早期発生のあったことがうかがえる。これらの検定結果、下記のことを判明した。

1) 例1；肺の肺胞／細気管支上皮腺腫  
雄の発生数はフィッシャーの確率計算法によると対照群に対して各投薬群で有意差を示したが、早期発生に対する有意差は認められなかった。雌の発生数はフィッシャーの確率計算法によると、雄同様に対照群に対して各投薬群で有意差を示した。早期発生癌の指標としての平均順位は対照群（41.5）に対し、高用量群の（10.8）のみに有意差（ $p < 0.01$ ）が示された。したがって、対照群の平均発生週齢104週齢に対して高用量群は35週間早く発生し、加えてヨンキーの傾向検定で用量相関性（ $p < 0.001$ ）が認められた。

2) 例2；肺の肺胞／細気管支上皮癌  
雄では対照群に比較して癌の発生数および発生時期に有意差が認められなかった。一方、雌では対照群に比較して各投薬群の癌の発生数に有意差が認められなかったが、発生時期の指標としての平均順位は対照群（13.2）に対して高用量群（3.5）のみに有意差（ $p < 0.01$ ）を示した。したがって、対照群の平均発生週齢の104週に対して高用量群は28週間早く発生し、加えてヨンキーの傾向検定で用量相関性（ $p < 0.01$ ）が認められた。

3) 例3；リンパ節の悪性リンパ腫

雄では対照群に比較して投与群で、癌の発生数および発生の早期化のいずれにも有意差が認められなかった。一方、雌では対照群に比較して各投薬群の癌の発生数に有意差が認められず、発生の早期化の指標としての平均順位は対照群（17.0）に対して高用量群（2.3）のみに有意差（ $p < 0.01$ ）を示した。したがって、対照群の平均発生週齢である100週齢に対し高用量群は28週間早く発生し、加えてヨンキーの傾向検定で用量相関性（ $p < 0.01$ ）が認められた。

以上のことから、例2の雌および例3の雌のように発生動物数に有意差が認められなくても早期発生がウイリアムズ・ウイルコクソン法で示され、併せて用量相関性も認められた。これら3要素である、発生動物数、早期発生および用量相関性の相互関係を表—5に示した。たとえ発生動物数に差がなくても早期発生がうかがえる場合も重要な知見である。また例1の雌が示すように対照群に対して用量群が発生動物数、早期発生および用量相関で有意差が認められれば、薬剤の明らかな影響が認められたことになる。

農薬の発がん性試験 [6] では一般に試験途中で中間屠殺動物用のサテライト群を設定しているが、医薬のがん原性試験ではサテライト群を設定しないため、最終計画屠殺を除き、切迫および死亡動物のみからは腫瘍発生の早期化が把握できない。また各群の動物数がきわめて小さい場合、順位検定では検出力が低く問題があるといえる。その他、各群の発生動物数に大きな差がある場合は検出力に問題がないか、これらの問題点を含め今後も長期反復投与試験における腫瘍発生の統計的解説分野の発展が望まれる。

表—4 各種腫瘍の発動物数、発生時期および用量相関性の相互関係を示す統計処理結果

試験群	雄			雌		
	発生頭数	検定法 K-W <sup>1)</sup> W-W <sup>2)</sup>	ヨンキーの傾向検定	発生頭数	検定法 K-W <sup>1)</sup> W-W <sup>2)</sup>	ヨンキーの傾向検定
<b>肺の肺胞/細気管支上皮腺腫</b>						
対照群	7	} NS	} NS	3	} **	} ***
低薬量群	15*					
中薬量群	20*					
高薬量群	16*			17***	**	
<b>肺の肺胞/上細気管支皮膚腫</b>						
対照群	2	} NS	} NS	3	} *	} **
低薬量群	6					
中薬量群	7					
高薬量群	5			4	**	
<b>リンパ節の悪性リンパ腫</b>						
対照群	5	} NS	} NS	7	} *	} **
低薬量群	5					
中薬量群	8					
高薬量群	9			3	**	

<sup>1)</sup>クラスカル・ワイリス検定。<sup>2)</sup>ウイリアムズ・ウイルクソン検定。  
 各群は50頭。発生頭数はフィッシャーの確率計算法で実施した。  
 NS:5%水準で有意差なし。\*p<0.05, \*\*p<0.01および\*\*\*p<0.001で有意差を示す。

表—5 癌原性試験の3要素、腫瘍発生数動物数、発生時期および用量相関性の相互関係

対照群	対各用量群の有意差		薬量相関性	判 定
発生頭数	発生時期			
あり	あり	あり	あり	極めて重要
あり	あり	なし	なし	重要
あり	なし	あり	あり	極めて重要
なし	あり	なし	なし	要注意

有意差は5%水準による。

参考文献

- 1) 厚生省薬務局審査課, 医薬品の製造(輸入)承認に必要な毒性試験のガイドラインについて, 薬審1第24号, 平成元年9月11日.
- 2) Peto, R., M. C. Pike, N. E. Day, R.G. Gray, P. N. Lee, S. Parish, J. Peto, S. Richardls and J. Wahrendorf.(1980). Long-term and short-term screening assay for carcinogens: A critical Appraisal, IARC, Lyon.
- 3) 青木繁伸(1989). 長期動物発癌試験成績の評価法について, 第38回医薬安全性研究会.
- 4) Barbara Mcknight (1991). STATISTICS IN TOXICOLOGY: Chapter 22, Survival Analysis, pp471—518, GORDEN AND BREACH SCIENCE PUBLISHERS, TOKYO.
- 5) 小林克己(1989). Categorical検定による腫瘍発生数の検討. 医薬安全性研究会, No.29, 11—14.
- 6) 農林水産省農蚕園芸局植物防疫課, 農薬の登録申請に係る毒性試験成績の取扱いについて, 59農蚕第4200号, 昭和60年1月28日.