

## 《会員寄稿》

ビーグル犬を用いた毒性試験から得られる定量データの取り扱い  
第一報：血液生化学的検査値の統計処理法の検索財団法人 食品農医薬品安全性評価センター 小林克己・大橋信之  
大川原一弘・井上博之

## 要 旨

イヌを用いた毒性試験は、ラット・マウス等のゲッ歯類に比較して個体差が大きく、加えて供試動物数が極めて少ないこと、一方動物自体が大きく、採血量が十分なため、投薬前から試験期間を通して時系列的に同一動物からの測定値が入手できるという点に大きな特徴をもっている。このため、試験群設定時の多くの場合、同腹を避け、体重に偏りが無いように動物の分配をおこなうと、投薬前の各群の血液生化学的検査値にしばしば、大きな差が散見され、臨床検査値を参考にした群構成への修正を迫られる事が多い。これらの事例から、従来一般的に実施されてきたように実測値をそのままt-検定、分散分析系またはノンパラメトリック系で投薬の影響を吟味する統計処理を犬試験にあてはめることには、問題が多い。従って、イヌの血液化学的検査値の時系列データの取扱いには、投薬前の値（初回値）から次の測定値の増減を算出しt-検定または、分散分析系の方法を用い、対照群と各投薬群間の差あるいは、各群相互間の差を吟味する方法が推奨される。この方法によって、薬剤の影響をより正確に考察に反映できるばかりでなく、投薬前の値も一層有効に活用できる。

## 1. はじめに

ビーグル犬（以下イヌと略す）を用いた毒性試験は、医薬品毒性試験法ガイドライン（薬審1 第24号）[1] および農薬の登録申請に係る毒性試験の取り扱いについて（59農蚕第4200号）[2] に定められた方法で実施されている。これらイヌを用いたガイドラインでは、ゲッ歯類のラット・マウスの場合と異なり血液生化学的検査には投薬前の検査を必要

とした解剖時および途中の測定を含め3～4回の検査が同一個体で実施される。しかし、供試匹数は一群3・4匹とゲッ歯類の場合に比べて極めて少ないのが特徴である（表1）。

表1 イヌを用いた毒性試験の概要  
—血液検査回数および供試匹数—

ガイドライン	医薬品	農 薬	
試 験	反復投与試験	亜急性毒性試験	慢性毒性試験
期 間	1, 3, 6, 12ヵ月の4種	90日間	12ヵ月間
測 定 回 数	投薬前・途中1回以上	投薬前・途中1～2回以上	6ヵ月毎
	解剖時	解剖時	
供 試 匹 数	各群3匹以上	各群少なくとも4匹以上	

一般的にラット・マウスの試験開始時の群分け（grouping）には、体重を指標に層別連続的無作為化 [3]（stratified sequenced randomization）を用い各群平均体重値間に有意差の出ないよう動物を分配する。しかし、イヌの試験群設定では、体重測定と同時に採血も実施し体重によって群分けを行った数日後に血液検査の成績を参照して群構成を行なう。しかしながら、投与前の全調査項目の平均値に有意差が無いように群構成をすることは、限られた数の動物を用いた条件下に於て極めてむずかしい。こうして投薬前に設定された各群の血液検査値は、平均値および分散が必ずしも同様ではなく、時には有意差を示すこともある。一般的に、イヌのこれら血液生化学的検査値に対する統計処理法としては、体重・飼料摂取量・飲水量・血液学的検査・臓器重量等と同様に決定樹方式によるt-検定法 [4, 5] および分散分析（ANOVA）から成る多重比較（範囲）検定法 [4, 6] 等が用いられ、投薬群と対照群との比較および相互群間の比較が実施されている。

本研究ではこれら血液化学的検査値の中から代表例をあげ、従来から常用されている上記各決定樹法を用い、各測定週毎に対照群に対する差を吟味した。同時に時系列を含めた二元配置の分散分析法、投与前（初回値）の値との比較および投薬前に対する増減を算出し、各週毎に対照群との比較等を併せて比較検討した。これら少数標本では群間変動が大きい上に、各群間の投薬前の平均値のバラツキ等も加わるため、時系列データを活用することによって、イヌのこれらパラメータのより適切な統計処理方法を検索した。

### 2. 検討材料および調査方法

イヌの血液化学的検査値の多くは、加齢によって変化することが知られている。五十嵐ら [7] は、加齢と共にT-Bil, D-Bil, TP, Alb, Cre, TTT, ZTTおよびNaが増加しALP, GluおよびIPが減少することを報告している。これらのことから、標準的な12ヵ月間反復投与試験（慢性毒性試験）から得られたデータ（表2および3）を基にして、加齢によって変化するALP（アルカリホスファターゼ）活性値と加齢によって変化の少ないGPT（グルタミン酸ピルビン酸トランスアミナーゼ）活性値の二つの測定項目を選んだ。対照群および高薬量群は各5、低薬量群および中薬量群は各4匹の計4群18匹（雄）、測定週は、投薬前1週（約5ヵ月齢）、投薬後26および52週の3点とした。

表2 臨床検査個体値

群	ALP活性値 (U/l)			GPT活性値 (U/l)		
	-1週	26週	52週	-1週	26週	52週
対照群	181	88	90	5.0	4.2	2.8
	480	140	105	3.8	3.7	2.8
	198	76	67	3.5	3.7	3.8
	174	80	81	3.8	4.1	4.1
	235	144	128	3.7	3.6	3.0
低薬量群	195	87	73	3.5	5.0	2.9
	184	105	92	4.7	4.4	4.2
	201	100	104	3.3	3.3	3.2
	124	55	51	3.0	4.0	3.9
中薬量群	111	72	53	2.9	3.9	4.7
	170	57	54	3.4	4.7	6.2
	173	59	57	2.5	3.9	3.5
	167	61	60	2.3	4.4	5.4
高薬量群	222	108	109	3.6	4.6	5.3
	162	121	94	3.6	3.7	6.0
	211	151	128	3.6	4.6	5.7
	170	79	67	4.3	6.6	8.0
	191	89	71	2.6	4.1	4.1

表3 各統計値の平均値および標準偏差

群	ALP活性値 (U/l)			GPT活性値 (U/l)		
	-1週	26週	52週	-1週	26週	52週
対照群	254 ± 129	106 ± 34	94 ± 23	40 ± 6	39 ± 3	33 ± 6
低薬量群	176 ± 35	87 ± 22	80 ± 23	36 ± 7	42 ± 7	6 ± 6
中薬量群	155 ± 30	62 ± 7	56 ± 3	28 ± 5	42 ± 4	50 ± 11
高薬量群	191 ± 26	110 ± 28	94 ± 26	35 ± 6	47 ± 11	58 ± 14

統計処理法は、t-検定（両側検定、図1）、多重比較（範囲）検定法（図2）の決定樹およびボンフェローニの傾向検定法 [8] を用いた。パッケージソフトはMuscot [ユックムス(株)] を用いた。二元配置の共分散分析法の基礎的数値の算出は、SEIKO S-500 基本プログラムライブラリー [株精工舎] で実施し群間・週齢間の差および有意差が認められた場合は、同一誤差項を用いて対照群と各投薬群間の差を手計算で吟味した。

有意差の表示は、対照群に対する各投薬群および投薬前に対する投薬後の帰無仮説の信頼の確率を示した（\* p ≤ 0.05, \*\* p ≤ 0.01）。薬量相関性の検定は、ボンフェローニの傾向検定法を用い有意差マーク（#）を付した。

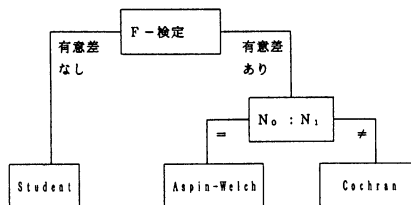


図1 t-検定の決定樹

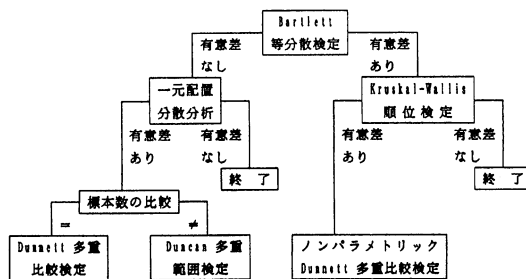


図2 多重比較（範囲）検定の決定樹

### 3. 結果

各週齢毎に於ける対照群と各投薬群の差の検定 [t-検定および多重比較（範囲）検定法の決定樹による検討] の結果は表4に示した。

ALP 活性値は、投薬後26および52週で対照群と中薬量群間にt-検定で有意差が認められたが、多重範囲検定法では、各群とも有意差が認められなかった。

GPT 活性値は、投薬前1週でt-検定によって対照群と中薬量群間に有意差が認められ

たが、分散分析法では、有意差が認められなかった。投薬後 26 および 52 週では、t-検定および多重範囲検定法とも同一の群間で有意差が認められた。

表 4 各週齢毎に於ける対照群と各投薬群の差の検定 [t-検定および多重比較 (範囲) 検定法の決定樹による検討]

検 定 法	群	ALP 活性値 (U/l)			GPT 活性値 (U/l)		
		-1 週	26 週	52 週	-1 週	26 週	52 週
t-検定法	対 照 群	-	-	-	-	-	-
	低薬量群	NS	NS	NS	NS	NS	NS
	中薬量群	NS	*	*	*	NS	*
	高薬量群	NS	NS	NS	NS	NS	**
パートレットの等分散検定		**	NS	*	NS	NS	NS
	分散分析法	-	NS	-	NS	NS	S+
クラスカル・ワリスの順位検定		NS	-	NS	-	-	-
	ボンフェローニの傾向検定	-	-	-	-	-	**

\*p<0.05, \*\*p<0.01, \*\*\*p<0.001.  
+ ダンカンの多重範囲検定法で吟味した結果、t-検定と同一の結果を示した。

二元配置の分散分析法の結果は表 5 に示した。単に -1, 26 および 52 週の週齢間の差を吟味する場合、-1 週のデータを統計値として使用できるが、群間をも含めて吟味する場合は、投薬を実施していない -1 週は統計値として使用できない。従って、週間は 26 および 52 週の 2 ポイント、群間は 4 群の 4 ポイント (2×4) の二元配置が応用できる。ALP 活性値は、26 および 52 週間を通して対照群に対して中薬量群のみに有意差を示した。又、各薬量間を通して 26 と 52 週との間に有意差は認められず、26 週に比べて 52 週の活性値の低下は認められなかった。ボンフェローニの傾向検定の結果、薬量相関性は認められなかった。GPT 活性値は、ALP 活性値と同様に 26 および 52 週との間にいずれの薬量間にも統計的な有意差は認められなかった。26 および 52 週間を通して対照群に対して中および高薬量群に有意差が認められて、またボンフェローニの傾向検定の結果、薬量相関性も併せて認められた。

表 5 二元配置の分散分析表

ALP 活性値					
項 目	平方和	自由度	分散	分散比	確率
合 計	27076	35			
週齢間	982	1	982	-	-
群 間	10130	3	3377	6.5	p ≤ 0.01
対照群対低薬量群	1	1	1213	2.3	-
対照群対中薬量群	1	1	7389	14.3	p ≤ 0.01
対照群対高薬量群	1	1	16	-	-
誤 差	15964	31	515		
GPT 活性値					
項 目	平方和	自由度	分散	分散比	確率
合 計	4418	35			
週齢間	26	1	26	-	-
群 間	1674	3	558	6.3	p ≤ 0.01
対照群対低薬量群	1	1	36	-	-
対照群対中薬量群	1	1	452	5.1	p ≤ 0.05
対照群対高薬量群	1	1	1427	16.2	p ≤ 0.01
誤 差	2718	31	88		

各群毎に前回値との比較結果は表 6 に示した。表 3 で示した平均活性値を対照群および各群の投薬前の値に対する % 値として各々下表のカッコ内に示した。ALP 活性値の投薬前の値は、対照群に対して各投薬群が 61~75% と低く投薬前の値は必ずしも近似ではない。GPT 活性値の投薬前の値は、ALP 活性値のそれとほぼ同様で、70~90% を示している。

ALP 活性値は、-1 週に対して各群とも時系列で同様に減少していることが分かる。GPT 活性値は、対照群で若干減少の推移を示しているが、中および高薬量群は、-1 週に対して増加の傾向を示していることが分った。従って、各群共、投薬前の値に対する 26 および 52 週の値を各々多重比較 (範囲) 検定法の決定樹法 (図 2) によって吟味した。しかしながら、ALP 活性値のように、対照群の投薬前の値に対して 26 および 52 週で有意差を確かめた場合は、他の投薬群で減少が認められていても薬剤の影響を正確に判断できない。

GPT 活性値は、中薬量群で 26 および 52 週に、高薬量群で 52 週に有意に増加を示すものと判断された。

表 6 各群毎に前回値との比較

群	ALP 活性値 (U/l)			GPT 活性値 (U/l)		
	-1 週	26 週	52 週	-1 週	26 週	52 週
対 照 群	254 (100) N (100)	105 (42) * (41)	94 (37) * (37)	40 (100) (100)	39 (98) (98)	33 (83) (83)
低薬量群	176 (100) (69)	87 (49) ** (49)	80 (45) ** (45)	36 (100) (90)	42 (117) (117)	36 (100) (100)
中薬量群	155 (100) N (61)	62 (40) (40)	56 (36) ** (36)	28 (100) (70)	42 (150) * (150)	50 (179) ** (179)
高薬量群	191 (100) (75)	110 (58) ** (58)	94 (49) ** (49)	35 (100) (88)	47 (134) (134)	58 (166) * (166)

N: ノンパラメトリック検定。  
\*p<0.05, \*\*p<0.01.

投薬前の値に対する増減値を算出し対照群との比較検討の結果は表 7 に示した。前項で述べたように群構成時の値 (投薬前の値) は、各群の平均値が近似しない場合が時々ある。もし、投薬前の値に有意差が認められた場合や、有意差はないが対照群との間に実数値上めだつた差がある様な場合にはそれ以降に測定した 26 および 52 週までの値を通常の方法で吟味したならば検定結果の信頼性は低くなる。従って、投薬前の値が各群間で大きく差を示す場合、投薬前の値に対する各測定

週の増減 ( $\Delta$ ) を算出し,  $t$ -検定および多重比較 (範囲) 検定法の決定樹を応用できる. 表に多重比較 (範囲) 検定法の決定樹 (図 2) で吟味した.

ALP 活性値は,  $-1 \sim 26$  週および  $-1 \sim 52$  週でそれぞれ等分散が認められずノンパラメトリック検定の結果いずれも有意差が認められなかった.

GPT 活性値は,  $-1 \sim 26$  週および  $-1 \sim 52$  週でそれぞれ同様の傾向を示した. 即ち, 対照群では,  $-1$  と  $26$  週値が同様の値,  $-1$  および  $52$  週値では, 若干の増加が認められた. しかしながら, 中薬量および高薬量群では  $-1 \sim 26$  週および  $-1 \sim 52$  週でそれぞれ対照群に対して減少が認められ, いずれも薬量相関性を示した.

表7 投薬前の値に対する増減値を比較

群	ALP 活性値 (U/l)		GPT 活性値 (U/l)	
	-1~26週	-1~52週	-1~26週	-1~52週
対照群	148 ± 108	160 ± 122	1 ± 4**	7 ± 10**
低薬量群	98 ± 18	96 ± 20	-6 ± 8	1 ± 7
中薬量群	93 ± 36	99 ± 28	-15 ± 5**	-22 ± 10**
高薬量群	82 ± 30	97 ± 22	-12 ± 8*	-23 ± 9**

\* $p < 0.05$ , \*\* $or$ \*\*\* $p < 0.01$ .

#### 4. 考察および結論

イヌを用いた毒性試験の特徴としてのデメリットは, ラット・マウス等に比較して供試動物数が少ない点と薬剤に対する感受性に個体差の大きい点である. 一方メリットとしては, 投薬前から試験期間を通して時系列に同一動物について測定値が入手できることである. 本研究では, 従来一般的に実施されてきたような実測値をそのまま  $t$ -検定や分散分析系又は, ノンパラメトリック系で薬剤間の差を吟味することをやめ, イヌの血液化学的検査値を含めた他のパラメータ (体重・血液検査値・尿検査値) を時系列データとして取扱い, 投薬前の値 (初回値) から次の測定値

の値の増減を算出し  $t$ -検定または分散分析系の方法を用いて対照群と各投薬群間の差あるいは, 各群相互間の差を吟味する方法を試みた. その結果, 本法が薬剤の影響をより正確に考察に反映させたと考えられる. またこの方法によって, 投薬前の値もより有効に活用できる.

#### 5. 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] 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.
- [3] 日本製薬工業協会・医薬品評価委員会・基礎研究会 (1984), 実験動物技術者のための教育資料, Section: ELT/ACA/00701, 割つけ方法.
- [4] Gad Shayne and Weil S. Carrol (1986). STATISTICS AND EXPERIMENTAL DESIGN FORTOXICOLOGISTS, The Telford Press, U.S.A.
- [5] 吉村功編著 (1987), 毒性・薬効データの統計解析, サイエンティスト社.
- [6] 佐久間昭 (1981), 薬効評価—計画と評価—II, 東京大学出版会.
- [7] 五十嵐功ら (1991), ビーグル犬における血液学および血液生化学的検査値の統計的解析 (C-53), 第 38 回日本実験動物学会総会.
- [8] Jonckheere, A.R. (1954). A distribution-free  $k$ -sample test against ordered alternatives. *Biometrika*., 41, 133 - 145.