

巻 頭 言

静岡県立大学薬学部
木村 良平

私は静岡薬科大学時代に当時会長でいらっしゃった林 栄一先生のお誘いを受けてこの静岡実験動物研究会の発足と同時に参加させていただき、その後幹事としても微力ながらお手伝いさせていただいておりましたが、この数年は第一線を離れ名ばかりの会員でした。このたび新たな気持ちでアクティブな会員として会に参加させていただきたいと考えております。どうぞよろしく願いいたします。

私のように薬物動態の研究に携わっている者にとっては実験動物は欠くべからざる存在です。実験動物に代わるいわゆる代替法ももちろん検討し、また、一部は*in vitro*で実験を行っていますが、その結果から生体内でおこる現象の全体の流れのごく限られた過程を説明できても、全体の流れを一部の過程のみで説明することはできません。いろいろな過程の*in vitro*実験の結果を寄せ集めてもやはり全体を正しく説明することは無理でしょう。また、*in vitro*実験で推論できる現象とは別の未知の要因が関わっていることはしばしばあることであり、*in vitro*の実験のみで最終的な結論を出すのはたいへん危険なことです。

私の教室の主要なテーマの一つに薬物受容体に関する研究があります。従来薬物受容体の研究では動物から取りだした臓器・組織標本を使い、種々の結合パラメータを求めていますが、*in vitro*で観察された薬物と受容体の結合がはたして生体内で同じようにおこるとは断定できません。投与された薬物が目的臓器・組織に運ばれ、いくつかの膜を透過し、蓄積し、その受容体と相互作用を持つことができるわけですから、*in vivo*での受容体結合パラメータを求め、ファルマコキネティクス、ファルマコダイナミクスの融合を図ることが重要と考えております。そこで、私の教室では実験動物を用いた*in vivo*実験が主となっています。

実験動物を用いることの是非についてはいろいろな立場の方がいろいろな意見を述べています。いずれももっともな意見です。実験動物を使うことが必須の者であってもいろいろな意見に耳を傾け、常に襟を正して行くべきであると思います。実験動物に対して単に感謝の気持ちを持つに留まらず畏敬の念をも持つべきです。実験動物を使うことが人類の幸福のため、地球全体の平和のためなどと申しながら、所詮は人間のエゴに端を発しているのですから。

《学術論文》

定量データに対する統計処理法の選択

小林克己, 北島省吾, 三浦大作, 謝明*, 安麗*, 女兆 宝玉*, 高崎興平**
 (財団法人 食品農医薬品安全性評価センター,
 *化工部農薬安全性評価監督検閲中心/沈陽, 中国, **東京農業大学)

【はじめに】

毒性試験を扱うものにとって最終判定の一助を担う統計処理の方法は、設定した群数、群構成および試験責任者の判断基準によってその手法が選択される。今回毒性試験に使用する統計手法の選択について解説したい。

【各検定法の解説】

1. パラメトリックおよびノンパラメトリック検定

パラメトリック検定は分布を利用した検定で、標準偏差などの表示ができる。ノンパラメトリック検定は分布を利用せず各定量値を大きさによって順位化し、順位差を検定する。したがって、平均順位などが表示できる。Table 1. に各パラメトリックに対応するノンパラメトリック検定を示した。

2. パラメトリック検定

2.1. t-検定(二群間の比較)

t-検定は一般的に分散比の大きさおよび各群の動物数によって三種類が常用されている。しかし、最近ではCochranは検出力が低いので使用を避けた方がよいとの指摘が専門家¹⁾によってなされている。したがって、等分散の場合はStudent, 不等分散の場合はAspin-Welchのt-検定が一般的に使用されている。

2.1.1. Studentのt-検定

両群の分散がほぼ同様な場合のみに実施する検定である。すなわち、両群の分散を算出し、分散比(分散の小さい群を分母に)をF表と比較し一様の分散か否かを有意水準5%で検定する。もし同様の分布を示した($p>0.05$, 有意差なし)場合はStudentのt-検定を用いて吟味する。

Table 1.

設定群数	パラメトリック検定	ノンパラメトリック検定 ¹⁾
二群のみ	Student, ◎ Aspin-Welch, Cochran などの t-検定	◎ Mann-Whitney の U 検定 Wilcoxon の検定
多群 (3 群以上)	分散分析(ANOVA)	Kruskal-Wallis の順位検定
	◎ Dunnett の多重比較検定 一般的な多重比較検定	Dunnett 型ノンバラ順位検定 ²⁾ ◎ Steel のノンバラ順位検定
	◎ Tukey の多重範囲検定法(群の大きさが同一) Tukey-Kramer の多重範囲検定(群の大きさが異なる)	Tukey 型ノンバラ順位検定 ²⁾ Tukey-Kramer の多重範囲検定
	◎ Duncan の多重範囲検定法	Duncan 型ノンバラ順位検定 ²⁾
	◎ Williams の多重比較検定(用量相関が前提)	Shirley-Williams の多重比較検定
	Scheffe の多重比較検定	Scheffe 型ノンバラ順位検定 ²⁾
	—	Jonckheere の傾向検定

¹⁾ 順位 (和) 検定(ranked sum test). ²⁾ 日本でのみ使用, その他は世界的に使用されている。

◎ 推奨できる手法(後述)。

2.1.2. Aspin-Welchのt-検定

両群の分散が異なるが、標本数が同一(最近では異なっても実施する)の場合に実施する検定である。各群の例数がほぼ等しい場合にはStudentのt-検定結果と概ね一致する。最近では、不等分散性を考慮して、はじめからAspin-Welchの手法を応用している例が多い。したがって、二群間の手法では、本手法で分析するのが最良と思える。

2.1.3. Cochranのt-検定

両群の分散および標本数が異なる場合に実施する検定である。我が国では最近検出力が低いこと理由に使用されなくなりつつあるが、諸外国では使用されている。

2.2. 一元配置の分散分析(ANOVA, One-way Analysis of Variance)

三群以上の設定の場合に用いる。一般的には、分散分析で有意差が検出された場合、どの群間に差が存在するのかわからない。分散分析法は特に農学分野を含めた畜産獣医関係の論文に多い。計算は1)全平方和の算出、2)群間平方和の算出、3)誤差項の平方和を計算、全平方和-群間平方和、4)平方和を自由度で割って分散を算出、5)分散比を算出してF-分布表の値と比較して有意差を判定する。その他に二元(用量と測定時期)および三元配置(用量、測定時期、性差)の分散分析も使用されている。したがって、有意差が認められれば、どの群間に有意差があるかは下記のいくつかの検定法で吟味する。しかし、分散分析とDunnnettの検定によって2回の多重検定(5%+5%=10%水準)を実施したことになる。また分散分析で有意差が認められなくてもDunnnettの検定で有意差が認められることも散見する。したがって、分散分析を実施せずにDunnnettなどの検定を直接使用することを推奨する。

2.3. Dunnnettの多重比較検定

- ・ 対照群との比較のみに使用する。
- ・ 各群の分散がほぼ等しく標本数もほぼ同一の場合。
- ・ 分散分析表の誤差項の標準誤差を用いて

対照群と各用量群間の差を吟味する。
・ 通常5および1%水準で判定する。

標本数が一定の場合に適用すると言われているが、極端な標本数の減少が認められない場合は使用できると言われている²⁾。現在では、標本数が変化しても対応できるように“確率ダネット”(パラメトリック・ノンパラメトリックとも)としてプログラム化されている^{3, 4)}。

2.4. TukeyおよびDuncanの多重範囲検定法

Table 2.

群	平均値
1	15.51 ^a
2	8.02 ^b
3	8.75 ^b
4	8.96 ^b
5	13.75 ^a

異符号間は5%水準で有意差を示す

- ・ 全群間の比較に使用し、各群の分散がほぼ等しくTukeyは標本数が同一の場合、Duncanは標本数が異なる場合に使用する。
- ・ 分散分析表の誤差項の標準誤差を用いて全群間の平均値の差を吟味する。
- ・ 通常5および1%水準で判定する。しかし、一般的に表中ではTable 2.の様に5%のみの有意水準で表示異符号間は5%水準で有意差を示す。している文献が多い。

Duncanの多重範囲検定法は外国で実施した毒性試験および関連試験で比較的使用されている。我が国では検出力が低いとのことで使用されない傾向にある。God and Weil⁵⁾によると標本数が同一の場合と同一でない場合に異なった計算式が与えられている。本検定法は、対照薬剤を設定した場合に有用である。

2.5. Scheffeの多重比較検定

Scheffeの多重比較検定は対比較以外の場合

も対象となる。各群の分散がほぼ等しく群の大きさが異なる場合に使用する。しかし、群の大きさが等しい場合は、検出力が少し低くなるといわれている。通常5および1%水準で判定する。

有意差の検出力が低いことから最近ではあまり使用されていない。対照群に対して40%の差があっても有意差が検出できない場合もある。

2.6. Williamsの多重比較検定

対照群と一薬剤をいくつかの用量に設定し、各群の平均値に用量相関性(対照群に対して用量群の値が増加あるいは減少)が見られる場合に使用する。Williams⁵⁾はWilliamsの検定を毒性試験のみに適用することを提言している。死亡によって群内動物数が極めて変化した群が存在する場合は適用できない。したがって、各群の大きさが等しいときに用

い、等分散性が前提となっている。Williamsの検定は、対照群との比較を高用量群から低用量群の順に一つずつ行い、ある用量群で有意差がないと、それより低い用量群は自動的に有意差なしとする。したがって、毒性試験では、用量相関性を前提としている場合が多いので最良と思われる。しかし、我が国では余り使用されていない。

2.7. 多重比較・範囲検定のまとめ

Scheffeの多重比較検定の検出力が低い理由は、本検定は対比の組み合わせがDunnnett, TukeyおよびDuncanの検定に比較して多いためである⁶⁾。各検定による5%水準は各検定の対比数を保証している。したがって、対比が多い検定法は棄却限界値(各分布表中の値、各検定による計算値がこの値より大きい場合は有意差とする)が大きい。

橋本⁶⁾をTable 3.に転載し一部加筆した。

Table 3.

群構成	Dunnnettの多重比較検定	TukeyおよびDuncanの多重範囲検定	Scheffeの多重比較検定
対照群	┌──┐	┌──┐	┌──┐
低用量群	┌──┐	┌──┐	┌──┐
高用量群	┌──┐	┌──┐	┌──┐

各検定による棄却限界値(各分布表中の値、各検定による計算値がこの値より大きい場合は有意差とする)をTable 4.に示した。二群間でのみの検定に用いるt検定の棄却限界値がもっとも小さく、ついでDunnnett, Williams, Scheffe, DuncanおよびTukeyの順になる。この表と上の表を対比することによって、多群を設定した場合にt検定を使用することの間違いが理解できる。Scheffeの多重比較検定では有意差の検出力が弱いといわれている割りに棄却限界値が小さい理由は、計算法がその他の3法と異りまたF分布表を使用しているためである。

各検定法の使用上の注意点をTable 5.にまと

Table 4.

検定法	有意水準	棄却限界値
t-test for 2 group	5%	2.10
Dunnnettの多重比較検定	5%	2.30
Williamsの多重比較検定	5%	2.50
Tukeyの多重範囲検定	5%	3.80
Duncanの多重範囲検定	5%	3.11
Scheffeの多重比較検定	5%	2.63

群数は4、一群内標本数は10で両側検定による。

めた。

Table 5.

単に群間差の検定は両側検定で、対照群に対し強弱の問いかけは片側検定を使用する		
群構成	比較対象	使用統計処理名
二群のみの設定	二群間一回のみ	Aspin-Welch の t -検定
対照, 低用量, 中用量, 高用量	対照群と各用量群間のみの差の検定 (検定回数は 3 回)	Dunnnett の多重比較検定
対照, 低用量, 中用量, 高用量	対照群と各用量群間のみの差の検定で用量相関性が前提 (検定回数は 3 回)	Williams の多重比較検定
対照, 低用量, 中用量, 高用量 対照薬剤	対照群と各用量群間および対照薬剤群と各用量群で各群の標本数がほぼ一定又は異なっている場合 (検定回数は 10 回)	Tukey の多重範囲検定(標本数がほぼ一定), Duncan の多重範囲検定(標本数が異なる)
対照, 低用量, 中用量, 高用量 対照薬剤	1 対の検定から 1 群対 2 群, 群対 1 群対 1 群など任意の組み合わせが可能 (検定回数はかなり多い)	Scheffe の多重範囲検定 (標本数が異なる)

Williams の多重比較検定以外は対照群, 薬剤 A, 薬剤 B, 薬剤 C の設定でも使用できる。

統計手法の選択

3. ノンパラメトリック検定

3.1. Wilcoxon の順位和検定および Mann-Whitney の U 検定

二群のみの設定による試験の比較手法である。我が国では、一般毒性試験の検査項目では、この検定法の使用例が少ない(生殖毒性試験では一部のデータに対して常用されている)。その理由に毒性試験が多群を前提にしているからで、 t -検定が不適当とする考えと同様である。しかし本検定法は我が国で常用されている Dunnnett, Scheffe などの順位和検定に比べて標本数が 3 で有意差が検出できることから感度が極めてよい。上記二つの検定の検出力は同一である。

3.2. Dunnnett 型ノンパラメトリック順位検定

3 群以上の多群を設定した場合に用いる。Kruskal-Wallis の順位検定で有意差を示した場合に用いられている。対照群と各群間に有意差があるか順位を検定⁷⁾する。

3.3. Steel の順位和検定

3 群以上の多群を設定した場合に用いる。本検定法は米国保健局 (NIH) および諸外国で

広く使用されている。我が国では極めて少ないが本検定法は Mann-Whitney の U 検定とほぼ同様な検出力を持つことから今後は使用例が増加すると考えられる。

3.4. 順位和検定に対する注意点

順位和検定を使用する場合、あらかじめ一群内の標本数が幾つあれば有意差が検出できるか把握する必要がある。各順位和検定で有意差の検出できる群内最小標本数は Table 6. の通りである。Mann-Whitney の U 検定を除いて

Table 6.

検定法	4 群設定	5 群設定
Scheffe type	22	40
Hollander-Wolfe	19	30
Turky type	18	32
Dunnnett type	15	26
Wilcoxon	8	12
Steel type	4	6
Mann-Whitney U ¹⁾	3	-

¹⁾ 2 群間検定。

は、多群間検定である。

Table 6. の順位和検定の中で国際的に使用さ