獣医情報科学会1996年度講演会の報告

獣医生物統計学における最近の話題

米国メルク社統計解析部長 Carolyn P. Daurio

監訳 山根逸郎*

はじめに

1996年4月3日に、麻布大学にて獣医情報科学会の総会及び講演会が開催された。例年は総会の後に会員らによるシンポジウムを行うのが常であったが、本年度は米国メルク社で動物医薬品検査の統計解析を担当している Daurio 女史の来日に合わせて、獣医生物統計学の基礎と最近の話題について講演を依頼することになった。新しい動物医薬品が開発された際の臨床試験は非常に重要な意味を持つが、それらの方法論については良く理解されていない部分が多い。今回の講演では、薬剤の臨床試験を行うにあたり統計学的に留意すべき点、重要な点などについて約1時間半に亘り解説されたが、その内容を翻訳しここにまとめて報告する。この講演では、第1部に統計学の基礎が解説され、第2部に統計学の応用について話された。第1部は表1と照らし合わせて読んでいただくと理解しやすい。第2部で第1部の内容が反復している部分については、第1部の部分にまとめて書くか、紙面の都合のため一部割愛させて頂いた。

第1部 獣医生物統計学の基礎

動物薬研究とは、産業動物や愛玩動物に用いられる医薬品の発見や開発、そして認可に関わる一連のプロセスの事を指す。このプロセスとは科学的方法を用いて行われるべきものである(表1)。科学的な方法とは、試験の目的を明確にし研究に用いられる動物や処置について詳しく記述し、試験の順序について正確に表し、適切な統計解析の方法を示し研究者に科学文献上での新しい情報を提供するものである。プロトコールはこれら全てについて明確に記述されているもので、生物統計学者はプロトコール

作成上に非常に重要な役割を果たす。

まずプロトコールには、研究目的としての仮説が明確にされなければならない(表1)。仮説の検定とは、ある薬剤の処置を受けた群とコントロールの群の比較を行う事である。帰無仮説とは処置群とコントロール群の間に有意な差が無いとする仮説で、対立仮説とは研究を遂行するにあたり実際に調べたい仮説であり、帰無仮説が否定された時に取りうる仮説である。従って研究者は、帰無仮説を反証するために研究を行うことになる。

プロトコールには、研究に用いられる処置やその他の処置について詳しく述べられる(表1)。ここでの処置とは独立変数と考えられ、動物側の反応は対応変数と考えられる。与えられる薬量や投与方法などの詳しい記述に加え、試験に用いられる動物に併用して与えられる治療についても記述する。またどのような母集団より研究用の動物を得るのかについても充分考慮しなければならない。ここでの母集団とは、試験をしている薬剤がいずれ用いられるであろう動物集団を指す。

そして試験に用いられる動物に、反復やブロックを形成する。反復やブロックは、類似する形質を持つ動物のグループであり、通常一つの反復の中でのサンプルの数は処置の数と同じになる。反復の中の動物は、品種や性、年齢や生理的状态や他の要因など、処置への動物の反応に影響を与える要因において類似した形質を持つ必要がある。又、ブロックの作製方法についてもプロトコールに記述される。遺伝的な要因の影響を取り除くため、同腹の動物などが反復やブロックとしてしばしば用いられる。

プロトコールにはさらに動物を試験に用いるか、試験対象からはずすかの条件についても詳しく述べる。例えば、生産性に影響を与える薬剤の試験では、その薬剤がいずれ使われる動物を用いて試験をする。あるいは、蚤に対する薬剤の効果試験では、犬や猫

* 農林水産省家畜衛生試験場疫学研究室
〒305 茨城県つくば市観音台3-1-1

表1 科学的方法

仮 設	帰無仮説と対立仮説
材 料	薬剤 用いられる薬剤 併用して用いられる治療
サンプリングの方法	反復やブロックの形成 試験に入れる,あるいは排除するための条件
試験単位	動物単位, 飼育房の単位, 放牧地
方 法	研究デザインの決定 ランダム化、ブロックの作製 クロスオーバーデザイン 誤差の最小化 盲検法 ランダム化 対応変数の測定 スケジュール表の作製 連続変数 離散変数 名義変数, 順序変数, 間隔変数
分 析	記述統計学 中央傾向, 分散 分散統計学 仮説の検定
結 論	表とグラフによる表現 P値の解釈 信頼区間

などが試験開始の時点で蚤が付着していることが確認されなくてはならない。また、疾病の発生を予防する目的で使われる薬剤の試験では、試験開始の時点で試験に使われる動物が目的とする病気を持っていないことを確認する。

試験を行うに当たり、試験単位を決定する。通常個々の動物が試験単位となるが、獣医分野の研究では、動物舎や牧区が試験単位となる場合もしばしばある。しかしながら、試験に用いられる動物は可能な限り別々に飼育され、飼育舍同士のあるいは囲いごとの交叉汚染が起きないように注意する。個々の

飼育舍ブロック1			飼育舍ブロック2		
群1	群2	群3	群1	群2	群3
処置2	処置3	処置1	処置3	処置1	処置2
反復1	反復1	反復1	反復1	反復1	反復1
反復2	反復2	反復2	反復2	反復2	反復2
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...
反復10	反復10	反復10	反復10	反復10	反復10

図1 ブロックデザインの例

動物が別々に飼育されている場合は、個々の動物そのものが試験単位となる。動物がそれぞれ別々に処置を受け、処置を受けたすべての動物が一緒に放牧されたような時は、すべての動物は同じ条件に飼育されているので、個々の動物が試験単位となる。鶏や豚などの飼料添加物の効果試験の場合は、飼育舎や囲いに入っている動物すべてに同時に薬剤を投与する事が多い。この場合、体重などの変数は個体別に測定されるものの、飼育舎や囲いが試験単位となる。もし動物がそれぞれ別々に処置を受け、それぞれの処置ごとに一緒に飼育された場合も、飼育舎の中での交叉汚染の可能性があるので、飼育舎が試験単位となる。

プロトコールにはさらにどのような研究デザインを作るか、どのような要因にてブロックを作製するかなどについても詳しく述べる(表1)。注射薬の試験に子豚を用いる場合、飼育舎をブロックとして形成する場合がある(図1)。もし3つの処置群がある場合、3つの飼育群が飼育舍ブロックの中に形成される。そしてそれぞれの飼育群の中にさらに反復が作られる。もし10頭の豚がひとつの飼育群に入れられる場合、図1の様に10頭ずつそれぞれの飼育群に入れていく。

通常試験に用いられる動物は処置ごとにランダムに振り分けられ、1頭の動物は1つの処置のみに用いられる。一方、1頭の動物がコントロールと試験群の両方に用いられるクロスオーバーデザインと呼ばれる研究手法もある。2つの期間に亘るクロスオーバーデザインでは、動物は2つの群に分けられ、処置Aと処置BをABの順あるいはBAの順に一定の期間をあけてそれぞれの処置を投与する。その際に、時間の経過や疾病の進行による影響が問題とな

るので、クロスオーバーデザインは通常健康な動物を用いた薬物動態の試験などに用いられる。なぜなら薬物動態の試験においては、家畜間の違いが非常に大きく、個体内での違いが比較的小さいため、処置を同じ動物に投与した方が処置による影響をより鋭敏に観察することが可能であるからである。

プロトコールには、試験中に発生する誤差を最小限にする方法についても述べる。この方法には、盲検法とランダム化と類似する形質を持つ動物を用いて反復を作製する方法などが挙げられる。盲検法とは、動物の反応を測定する際に生じる誤差を最小限に減らす方法である。もし変数を測定する者が、どの動物にどの処置が振り分けられているか知っている場合は、期待される結果に対して無意識に有利になる判定を出す可能性がある。それらに対処するために、コントロール群にも処置群と見た目で区別がつかないプラシーボを投与する方法と、投与する者と判定をする者を別々にする方法がある。

科学的な方法の基本として、ランダム化の概念がある。臨床試験の統計解析を行うに当たり、ランダム化は非常に重要な意味を持つ。ランダム化を行うことにより、既知の他の要因による結果への影響を除くだけでなく、未知の要因による結果への影響も取り除くことができる。ランダム化を行うにあたり、まず何らかの方法で試験に用いられる動物を並べ直す必要がある。それにより、最初の動物がXと呼ばれる処置を受ける、という事が決まった際に、どの動物が最初の動物であるかが確認できる。もしブロックが用いられる場合は、体重順などによりすでに動物はある種の法則に基づいて順番に並んでいると言えよう。

動物をある一定の法則に並べ替えた後、硬貨の裏表を用いるかサイコロを振るか乱数表を用いるなどの方法を使い処置群へのランダムな割り振りを行う。乱数表の例として、図2を挙げた。上の左端より右に向かって乱数表を読むと、最初に3の数字があるので、反復1の最初の動物は3の処置を受ける。次に1の数字が出てくるので、反復1の2頭目は1の処置を受ける。残りは2の処置のみなので、反復1の最後の動物は2の処置を受ける。同様に反復2の動物は最初に2の数字が出ていたので2の処置を受ける。その次の乱数は1が出ていたので1の処置を与える。この様にして、反復ごとにそれぞれの3つの処置をランダムに振り分けていく。

プロトコールにはさらに対応変数をどの様に測定

	3190	9278	1816	7697
	6032	7574	7685	5258
	9286	0449	5051	4772
反復 1	3 頭の処置	3—1—2 の順		
反復 2	3 頭の処置	2—1—3 の順		
反復 3	3 頭の処置	1—3—2 の順		
反復 4	3 頭の処置	2—1—3 の順		

図2 乱数表

するかを詳しく述べる。ここで言う対応変数とは、帰無仮説を評価するために用いられるものである。そして、どのような頻度で対応変数を測定するかは、スケジュール表などを作製すると分かりやすい。

独立変数や対応変数は連続変数か離散変数として測定される。連続変数の例としては、体重や体温や肝臓の酵素値などが挙げられる。離散変数は、さらに名義変数や順序変数や間隔変数などに分けられる。名義変数とは、観察される結果がカテゴリーで表される変数であるが、その観察値に一定の順序は存在しない。動物の品種や下痢の有・無などの記録も名義変数の一種である。順序変数とは、観察される判定に一定の順序がある場合を呼ぶ。病状が無い・軽い・重いなどの例や、ダニの数が、100未満、100-1000、1000以上などの順序のあるカテゴリーに分けられるのも順序変数である。間隔変数とは、足し算とか引き算が可能である非連続的な数値のことである。蚤や内部寄生虫卵数、生まれてきた子豚の数、尿中の細菌の数などが間隔変数の良い例である。統計解析を行うにあたっては、間隔変数は連続変数として扱われる場合もしばしばある。

統計的な分析とは、記述統計手法を用いて結果を表し、分析統計手法を用いてデータが帰無仮説に反しているか検証する(表1)。前述したように、もし帰無仮説が棄却された場合、対立仮説が採用され研究の目的は達せられたことになる。

記述統計学とは、それぞれの処置による対応変数の反応の結果を要約するものである。この統計手法では、それぞれの処置群ごとに中心傾向の値と分散度を測定し、処置の影響について表現する。中心傾向の最も一般的なものは算術平均である。左右対称の分布をする生物学的データの中心傾向を表すには、この算術平均が最も適している。図3の例では算術平均は16となる。中央値はデータの半分の数が中央

値より低い値であり、半分の数がその中央値より高い値となるように設定される中心傾向の値である。図3においては、中央値は15となる。中央値はサンプルサイズが大きく分布が歪んでいる場合や、スコアとして記録されたデータなどに用いられ、小さなサンプルサイズのデータにはあまり信頼できる値ではない。幾何平均とは、対数変換されたデータの算術平均をさらに逆対数をとって得られる中央傾向の値である。図3の例では幾何平均は14.96となる。データの数値が0である場合は対数をとることはできないので、対数をとる前に1を加え逆対数を取った後で1をさらに減ずるという方法を取る。生物分野においては、対数変換した値を用いた分析統計手法がしばしば用いられる。

分散度を表す数値として最も一般的なのが範囲である。これは単純にデータの拡がり方を表すもので、最高値から最低値を引き算したものである。図4の例では範囲は $25-10=15$ である。分散は分散度をより正式に表す数値である。図4の例では42.5となる。標準偏差とは分散の平方根である。複数の処置群が設定された場合、それぞれの処置群ごとの分散が等しいかどうかは非常に重要な意味を持つ。もしそれぞれの処置群ごとの分散が等しい場合は、グループの分散は均質 (Homogeneous) と呼び、そうでない場合は異質 (heteroscedasticity) と呼ぶ。これを検定するために、Hartley's F-max test や Bartlett test などを用いる。

分析統計学は処置群ごとの比較を行い、帰無仮説が正しいかを検定する重要な方法である。パラメトリック統計学を用いるには、対応変数が釣り鐘型の正規分布をし、それぞれの処置ごとの分散が等しい

ことが前提条件として必要である。もしこれらの前提条件が満たされた場合は、パラメトリック統計は最も検出力の高い統計手法、すなわち帰無仮説が間違っているときに帰無仮説を最も棄却しやすい統計手法となる。

パラメトリック統計の最も一般的な手法としてt検定と分散分析が挙げられる。t検定は2つの処置群に有意な差があるか検定するものであり、分散分析は2つ以上の処置群間に有意な差があるか検定するものである。分散分析では、すべての分散をブロックによる分散、処置による分散、どの要因にもよらない誤差による分散に分ける。通常ブロックを導入することにより、誤差の分散を小さくすることが可能で、処置間の有意差を見いだす感受性を増すことができる。2つ以上の処置を比較し、もし処置の効果が有意であっても、どの処置とどの処置の間に有意な差があるかは不明である。それらを明らかにするためには、応用の部分で述べられる多重比較を行う必要がある。

さらに科学的方法とは、研究者に何らかの結論や推論を与えるものでなくてはならない(表1)。生物統計学者は、処置ごとの算術平均や幾何平均の値を用いて表などを作製し結果を要約する。必要とあればグラフなどを使い、結果を視覚に訴える方法も用いられる。p値は結果の解釈に利用される。p値とはデータが表す事象が帰無仮説と同じである確率を表し、もしp値が0.05より低い場合は帰無仮説は棄却され、薬剤の効果試験の場合は処置は効果があったと結論づける。結果の平均値の信頼区間を示すことにより、結果の解釈に客観性を持たせる。通常、信頼区間とは、

点推定の値 $\pm$ 係数 $\times$ 標準誤差

という数式で表される。ここでの点推定の値は、通常平均値であり、係数はStudentのtの値などが用いられる。信頼区間は単なる平均値と異なり、分散度も考慮に入れて母集団の平均がどの程度の精度で測定されたかを示すものである。

第2部 獣医生物統計学の応用

分散分析を行う際の予測変数(独立変数)は離散変数である。これは処置などの固定された変数もあれば、反復やブロックなどのようにランダムな変数もある。分散分析においては、しばしば一つ以上の相互作用の要因が導入される。相互作用とは、一つ

サンプル=10, 10, 15, 20, 25
平均 =16
中央値 =15
幾何平均=14.96

図3 中央傾向

サンプル=10, 10, 15, 20, 25
範囲 =15
分散 =42.5
標準偏差=6.52
標準誤差=2.92

図4 分散度

の要因による結果がもう一つの要因に作用されることを言う。従って、処置と性別の間に相互作用が存在する場合は、それぞれの性別ごとに処置の影響を調べる必要がある。共分散分析とは、一つの予測変数が離散変数で、共分散と呼ばれる変数が連続変数である場合に用いられる。獣医分野においては、研究開始時における初期データ、例えばヘマトクリット値の初期値を共分散として用いることが多い。

変換とは、対応変数が正規分布しない場合や、処置群間の反応の分散が等しくない場合に用いられる。変換された値はより正規分布に近づき、処置群間の分散は等しくなることがしばしばある。従って変換することにより、t検定や分散分析を行うための前提条件を満たすことになる。対数変換は処置群間の比率が問題になる場合や、各群の分散が平均に比例して増大する場合や、寄生虫数のように対応変数が実際に対数正規分布する場合などに用いられる。データが途中で切れているような場合にはランク変換を用い、後に述べるノンパラメトリック統計を用いて分析される。ランク変換された値の中央傾向を表す数値は中央値である。さらに対応変数が比率などの場合には arcsin 変換、それ以外には平方根変換などの方法がある。

2つ以上の処置があり、すべての処置間に有意な差があるとするだけでは十分でない場合には、多重比較という方法を用いて、どの処置とどの処置の間に有意な差があるかを検定する。生物統計学者は前もってどの比較をするか考慮しておく必要がある、それらはプロトコールに記されていないからではない。通常、各処置間の比較は5%の有意水準で行われる。もし前もってある特定の2処置の比較をするのではない場合、2処置間比較 (pairwise comparison) が通常用いられる。これらの2処置間比較には、Tukey's test や Scheffe の方法や Dunnett の方法や Duncan の方法などがある。

試験に用いられる薬剤の量と反応との関係を調べるのが、量反応関係の試験である。成長期の豚に投与する飼料添加物の量と増体量との関係などがその良い例である。これらの関係が線形であるか、プラトーと呼ばれるそれ以上添加物の量を増やしても反応が変わらない点に達したかなどについて注意を払う必要がある。図5は駆虫薬と内部寄生虫卵の数との量反応関係を調べたものである。x軸は薬剤の量、Y軸には反応の内部寄生虫卵数の値が示されている。中程度の薬剤の量でプラトーが生じ、それ以上に薬

図5 量反応関係

剤の量を増やしても寄生虫卵数の減少は観察されていない。

欠損値が処置と無関係にランダムに発生する場合は、欠損値の与える薬剤効果の判定に対する影響は最小限である。例えば犬や猫を用いた試験で、犬や猫が畜主の引越などによる理由により試験群からはずれる場合は、ランダムに欠損値が生じていると言える。薬剤の処置と関係のある理由により欠損値が生じる場合は、それらの影響について充分考慮する必要がある。薬剤の処置の効果があり畜主が動物を病院に連れて来なくなるとか、副作用の影響により動物が死亡するなどが、ランダムに欠損値が生じない例である。

順序変数の値など正規分布の理論が応用できない対応変数が得られた時は、対応変数に対する前提条件が少なくすむノンパラメトリックの統計手法が用いられる。対応変数の種類や処置の数などに応じて、ノンパラメトリックの手法を選択する。例えば、コントロール群と処置群の影響を、疾病の有・無として2×2分割表にて比較する場合、Fisherの直接確率計算法が用いられる。順序変数として記録されたデータか、連続変数として記録され正規分布の前提条件を満たさない場合、あるいは各群の分散が等しくない場合には、対のある2つのサンプルに対してWilcoxonのランクテストが用いられる。2つ以上の処置があり、ブロックなどを用いている試験の場合には、Friedmanのテストが用いられる場合もある。これは、パラメトリック統計における乱塊法に類似する方法である。

終わりに

以上、Daurio 先生が講演で話された内容を、講演

原稿を基に翻訳した。講演自体の内容は話し言葉で行われたため、報告書としての全体の構成を考えた上で若干の意識を用いた部分は御了承頂きたい。又、今回の講演の実施に当たって、数々の労を担って下さった日本 MSD 社の木下現氏にこの場を借りて深謝いたします。