

A comparison of one-and two-sided tests for judging significant differences in quantitative data obtained in toxicological bioassay of laboratory animals

メタデータ	言語: Japanese 出版者: 浜松医科大学 公開日: 2014-10-31 キーワード (Ja): キーワード (En): 作成者: 小林, 克己 メールアドレス: 所属:
URL	http://hdl.handle.net/10271/1667

学位論文の内容の要旨及び論文審査の結果の要旨

学位記番号	医博論第 390号	学位授与年月日	平成16年 3月 9日
氏名	小林克己		
論文題目	A comparison of one-and two-sided tests for judging significant differences in quantitative data obtained in toxicological bioassay of laboratory animals (実験動物の毒性試験で得た定量的データの有意差検定に用いる片側検定と両側検定の比較)		

論文題目

A comparison of one- and two-sided tests for judging significant differences in quantitative data obtained in toxicological bioassay of laboratory animals (実験動物の毒性試験で得た定量的データの有意差検定に用いる片側検定と両側検定の比較)

論文の内容の要旨

[はじめに]

医薬品および農薬の開発に必要な実験動物を用いた毒性試験法は、そのガイドラインに示されているが、統計解析法については明確に述べられていない。毒性試験のうちげっ歯類を用いた78週(1.5年)から104週(2年)にわたる長期慢性毒性試験は重要である。通常は無投薬の対照群が1群と、2~4群の薬剤投与群を設定する。測定結果が出たら順次有意差検定をするのであるが、通常検査は8項目で検定すべき回数が800から900におよぶことがある。検定法によって対照群と各薬剤投与群間の差を5%水準の両側または片側検定で吟味する。両側検定は、①単に群間に差があるかどうかの問いかけの場合 ②群間の差が事前に想定できない場合 ③試験責任者が正負の両方同時の結果を望む場合に使用する。片側検定は、①強弱を扱う問いかけの場合 ②群間の差が事前に想定できる場合 ③試験責任者が一方のみの重要性を指摘する場合に使用する。

ところが、この両側と片側の検定法の使用についてもガイドラインには定められてはいない。そこで、どちらが適切な方法かを探究した。

さらに、毒性試験における統計解析法はこれのみでなく等分散性の検討などいくつかの段階を経て最終的に評価する。これらを総称して決定樹(decision tree)という。副論文ではここで探求した検定法を含む新しい決定樹を提案した。

[材料ならびに方法]

測定項目と測定回数が多い慢性毒性試験/がん原性試験併合試験(1群80頭×5群設定、26、52、78週で各10頭を計画解剖、残りは104週で解剖)の試験成績を選び、まず検出力の高いt検定で有意差($P < 0.05$)を示す項目を選んだ。これらを対象として、t検定およびダンネットの多重比較検定を用いて両側検定および片側検定で吟味し、有意差($P < 0.05$)検出割合を比較検討した。

なお、主な測定項目は、飼育管理データ(体重、摂餌量、飼料効率など)、血液学的検査(HCT、HGB、RBC、MCV、MCH、MCHC、PLT、WBC、NEUT、LYMPH、RCなど)、生化学的検査(Glu、T-CHO、TG、PL、NEFA、BUN、T-Bil、Crea、TP、Alb、Na、K、Cl、Ca、IP、AST、ALT、ALPなど)、尿検査(尿量、比重など)および器官重量(脳、心、肝、腎、脾、副腎、精巣、卵巣など)などである。

一方、文献に掲載されているげっ歯類を用いた毒性試験を対象に検定方法について調査した。対象誌は、医薬品研究(Vol 22(1)~Vol 24(3))、応用薬理(Vol 47(1)~Vol 48(2))、日本癌学会誌(Vol 83(1)~Vol 84(12))およびNational Toxicology Program(NTP)(1987~1992)、Toxicology and Applied pharmacology(TAP)(Vol 112(1)~Vol 119(2))、Food and Chemical Toxicology(FCT)(Vol 29(6)~Vol 31(4))である。

決定樹については、その変遷をたどり、それぞれの手法の特徴を分析して、上述したように検討した検定法を含めて適切な統計法を考案した。

[結果]

検定回数700回を片側 t -検定で実施した結果、578回が対照群に対して一定方向に有意差を示し、39回が対照群を挟んで高いか低い両方向で有意差を認め、その比は15:1であった。つぎに、片側検定と両側検定による t -検定およびダンネットの多重比較検定による有意差検出率を計算すると、片側検定の検出率は、両側検定に比較してそれぞれ85%および86%で各棄却限界値(各分布表)の比率0.78と0.85に類似していた。

文献上の傾向；(1)国内3誌の103論文中90論文が検定法の記載なし、残り13論文が両側検定の使用であった。(2)TAP誌の103論文中87論文が記載なし、12論文が両側検定で4論文が片側検定の使用であった。(3)FCT誌の51論文中33論文が記載なし、15論文が両側検定で3論文が片側検定の使用であった。(4)NTPの報告書では、幾つかの試験でダンネットの多重比較検定およびウィリアムズ検定(用量相関性を前提とした検定)が使用され、体重および臓器重量は両側検定で、血液学的検査、生化学的検査および尿検査は順位和検定(片側検定)で実施していた。

既存の決定樹は、複雑または簡単な経路を辿る手法の二つに、また等分散検定を採用している手法と採用しない手法に分類できた。出来るだけ簡潔で検出力の高い決定樹が適切であると判断した。

[結論]

毒性試験から得られたデータは、医薬品ではヒトに対する投薬の安全量を算定する基礎に、また農薬では一日最大摂取許容量等の決定の基盤となるものであるから、厳密な毒性チェックが必要である。実験動物を用いた毒性試験から得られる定量データの有意差検定は、従来どちらを選択するかは明確ではなかったが、検出力の高い片側検定を使用すべきである。決定樹については、著者らが提案した簡潔な経路を辿りかつ検出力の高い手法、すなわちまず等分散検定をして、有意でなければダンネットの多重比較検定(片側検定)を、有意であればスチールの検定(片側検定)を採用すべきである。

論文審査の結果の要旨

我々が日常暴露されうる種々の化学物質、医薬品、農薬、化粧品などが、生物特にヒトにどのような影響を与えるかについての研究は、医薬品ではヒトに対する投薬の安全量を算定する基礎に、また農薬では一日最大摂取許容量等の決定の基盤となるものであるから、厳密な毒性チェックが必要である。伝統的に、げっ歯類を多数用いた毒性試験がおこなわれており、その方法論(実験計画、対象動物種、系統、病理所見を含めた指標)の標準化がおこなわれている。一方、その結果をどのような統計処理をおこなって、どの基準を有意とするのが一番安全かつ科学的であるかについては十分検討されていない。申請者は長年にわたり、本邦の医薬品、農薬の長期(1-2年)毒性試験の6割のシェアをもつ財団法人、医薬品農薬食品安全評価センターにおける実際の大規模毒性試験に携わりつつ、検定方法の違い、特に片側検定と両側検定による結果の違いにより、どのような解釈が生まれるかを検証し、また、文献上の同規模毒性試験結果に見られる検定方法も再検討した。そもそも両側検定は、①単に群間に差があるかどうかの問いかけの場合 ②群間の差が事前に想定

できない場合 ③試験責任者が正負の両方同時の結果を望む場合に使用する。片側検定は、①強弱を扱う問いかけの場合 ②群間の差が事前に想定できる場合 ③試験責任者が一方のみの重要性を指摘する場合に使用するものとされている。

申請者は片側検定のほうが特に低用量領域でのその毒性検出力(有意差検出力)の増大することを見だし、低暴露量の生物学的影響或いは人体に対する影響を厳密に評価するにはより有用で、さらに副論文において低暴露量(低用量)での差を検出しやすい効率的な統計手法の決定樹を提案したものである。

[材料ならびに方法]

測定項目と測定回数が多い慢性毒性試験/がん原性試験併合試験(1群80頭×5群設定、26、52、78週で各10頭を計画解剖、残りは104週で解剖)の試験成績を選び、まず検出力の高い t -検定で有意差($P < 0.05$)を示す項目を選んだ。これらを対象として、 t -検定およびダンネットの多重比較検定を用いて両側検定および片側検定で吟味し、有意差($P < 0.05$)検出割合を比較検討した。なお、主な測定項目は、飼育管理データ(体重、摂餌量、飼料効率など)、血液学的検査(HCT、HGB、RBC、MCV、MCH、MCHC、PLT、WBC、NEUT、LYMPH、RCなど)、生化学的検査(Glu、T-CHO、TG、PL、NEFA、BUN、T-Bil、Crea、TP、Alb、Na、K、Cl、Ca、IP、AST、ALT、ALPなど)、尿検査(尿量、比重など)および器官重量(脳、心、肝、腎、脾、副腎、精巣、卵巣など)などである。文献に発表しているげっ歯類を用いた毒性試験を対象に検定方法について調査したのは、医薬品研究(Vol 22(1)~Vol 24(3))、応用薬理(Vol 47(1)~Vol 48(2))、日本癌学会誌(Vol 83(1)~Vol 84(12))およびNational Toxicology Program(NTP)(1987~1992)、Toxicology and Applied Pharmacology(TAP)(Vol 112(1)~Vol 119(2))、Food and Chemical Toxicology(FCT)(Vol 29(6)~Vol 31(4))に発表されたものを対象にした。

[結果]

検定回数700回を片側 t -検定で実施した結果、578回が対照群に対して一定方向に有意差を示し、39回が対照群を挟んで高いか低いかの両方向で有意差を認め、その比は15:1であった。つぎに、片側検定と両側検定による t -検定およびダンネットの多重比較検定による有意差検出率を計算すると、片側検定の検出率は、両側検定に比較してそれぞれ85%および86%で各棄却限界値(各分布表)の比率0.78と0.85に類似していた。一方従来の文献では、(1)国内3誌の103論文中90論文が検定法の記載なし、残り13論文が両側検定の使用であった。(2)TAP誌の103論文中87論文が記載なし、12論文が両側検定で4論文が片側検定の使用であった。(3)FCT誌の51論文中33論文が記載なし、15論文が両側検定で3論文が片側検定の使用であった。(4)National Toxicology Program(NTP)の報告書では、幾つかの試験でダンネットの多重比較検定およびウィリアムズ検定(用量相関性を前提とした検定)が使用され、体重および臓器重量は両側検定で、血液学的検査、生化学的検査および尿検査は順位和検定(片側検定)で実施していた。また、これらの決定樹 decision tree は複雑および簡単な経路を辿る手法の二つに、また等分散検定を採用している手法と採用しない手法に分類できた。

[結論]

申請者はこれらの検討から以下のように結論づけた。実験動物を用いた毒性試験から得られる定量データの有意差検定は片側検定のほうが、検出力が高く、とくに生物学的およ

び医学的に重要な、低用量領域の微妙な差をのがすことがない。

従って、片側検定を使用すべきである。さらに簡潔かつ合理的に統計学的解析を行う決定樹については、まず等分散検定をして、有意でなければダンネットの多重比較検定(片側検定)を、有意であればスチールの検定(片側検定)をおこなうという手法により、短手数で感度のすぐれた検定を行うことができ、有用性の高いものであった。

本研究について審査委員会では以下の試問を行った

- 1) 対象とした自験毒性試験はいくつか
- 2) 毒性試験のガイドラインはどんなものか
- 3) 受託機関は本邦にどのくらいあるか
- 4) 大規模長期毒性試験はその施設で年間どのくらいの数おこなわれているか
- 5) 一群何匹使用するのが一般的か
- 6) 動物への投与法は
- 7) 投与法、解剖といった手技に資格認定があるか
- 8) 試験計画責任者のトレーニングあるいは資格認定をおこなっているのか
- 9) 毒性試験ではベータ値はどのようにあつかわれているか
- 10) 5%を有意差の基準にする根拠はなにか、追跡調査の実例はあるか
- 11) この研究の新規性と医学分野における重要性を簡潔にいうとどういうことか

これらに対し、申請者の回答は適切であり、問題点も十分理解しており、博士(医学)の学位論文にふさわしいと審査員全員一致で評価した。

論文審査担当者	主査	梶村春彦	副査	青木伸雄	副査	木村通男
---------	----	------	----	------	----	------