

経時データの解析

第 24 回 高橋セミナー

2006 年 4 月 15 日

高橋 行雄

2006/04/17 10:17

表紙裏

2006年4月11日，新規作成，
ファイル名：E:\¥Q8¥Q8_ANZ_S23_共変量¥S23_共変量.doc

経時データの解析

目 次

1. はじめに.....	1
2. JPMA 浦狩ら (1996) の研究報告の概要.....	2
2.1. 抗炎症剤の抗浮腫作用のスクリーニング試験のデータ.....	2
反復測定分散分析の適用.....	3
曲線のあてはめ.....	4
2.2. HPD 投与前後の DA 遊離量のデータ.....	6
2.3. 解析の留意点.....	9
3. 経時データの解析について安全研での議論の集大成.....	10
3.1. はじめに.....	10
3.2. 単回, 反復, および逐次投与の区別.....	13
3.3. 経時データの分類.....	13
3.4. 検定の多重性の問題.....	14
3.5. 全ての時点を用いた検定.....	14
3.6. 良くある質問.....	15
3.7. 投与前値の調整法の比較.....	17
4. 解析の考え方, 実験計画の再考.....	20
4.1. 差または比をとるべきか, 投与前値を共変量とすべきなのかの判定.....	20
浮腫の容積.....	21
ドパミンの遊離量.....	22
4.2. 陽性対照と同程度の反応を起こす用量の推定.....	25
浮腫の容積.....	25
直線のあてはめ.....	25
ロジスティック曲線のあてはめ.....	26
4.3. 実験計画の再考.....	27
浮腫の容積.....	27

ドパミンの遊離量.....	27
5. JMP による線形混合モデルの解析.....	28

図表目次

表 2.1 JMPのデータ（浮腫の容積_浦狩JPAM.JMP）.....	4
表 2.2 JMPのデータ（ドパミン遊離量_浦狩JPMA.JMP）.....	7
表 3.1 時点間に相関がある場合の分散の比較.....	18
表 4.1 経時データの性質.....	20
図 2.1 JMPによる生データのプロット（浮腫の容積_浦狩JPAM_T.JMP）.....	5
図 2.2 JMPによる平均値のプロット（浮腫の容積_浦狩JPAM_Tの要約.JMP）.....	5
図 2.3 生データのプロット.....	7
図 2.4 JMPによる平均値のプロット（ドパミン遊離量_浦狩JPMA_Tの要約.JMP）.....	8
図 2.5 変化率の平均値の経時変化.....	9
図 3.1 剤形別薬物濃度の経時推移.....	15
図 3.2 相関係数の変化による分散の比較.....	19
図 4.1 2時間目（左）と5時間目（右）の浮腫容積についての用量反応関係.....	21
図 4.2 2時間目と5時間目の浮腫容積の用量.....	22
図 4.3 投与前3時点の平均値の分布.....	22
図 4.4 60分値のドパミン遊離量と投与前3ポイントの平均値の関連.....	23
図 4.5 60分から120分の平均ドパミン遊離量.....	24
図 4.6 240分のドパミン遊離量の比較.....	24

1. はじめに

皆様へ、

2006年4月15日(土)の106回安全研の前に同じ会場で午前10時から第24回のセミナーを行います。

テーマは、「経時データの解析」です。1996年に製薬協の「薬理と毒性の統計学的諸問題」についてシンポジウムがあり、その中で、塩野義製薬の浦狩さんたちが、「経時測定データ」について報告しました。ここで用いられた「抗炎症剤の抗浮腫作用のスクリーニング試験のデータ」、および「HPD投与前後のDA遊離量のデータ」を用いて、経時測定データの解析の基本的な考え方と実践を示したいと思います。

JMPのバージョン6から、線形混合モデルの解析が安心して行えるようになったこともあり、これまで安全研で取り上げられてきた経時データについても、解析の「考え方と実践」を示します。

高橋行雄

2. JPMA 浦狩ら (1996) の研究報告の概要

シンポジウムの報告集(73-81)は、本文2ページ+パワーポイント配布資料からなっている。そこで取り上げられている2つの事例について概観する。

経時測定データ

臨床評価部会第二分科会Dグループ

浦狩保則 塩野義製薬(株)
瓦林良和 (株)大塚製薬工場
小泉貴央 生化学工業(株)
松本正人 明治製菓(株)

0. 化合物投与時の反応値の経時的推移に基づいて、用量反応関係を検討する薬理試験が数多く実施されている。そのような試験では、定められた数個の時点で反応値が測定されるのが通常である。すべての実験動物(個体)で測定時点を揃えるため、欠測がなければ、各個体の経時測定値から同じ時点のデータを抜き出すことが可能である。そして、群間比較あるいは用量上の回帰分析を時点ごとに行い、化合物の用量反応関係を評価する。これが、わが国の製薬企業における経時測定データの解析の一つの典型である。ここで注意したいことは、そのような解析だけでは**個体内の変化**への視点が欠けているために、薬理作用あるいは毒性の見逃しが生じ得ることである。

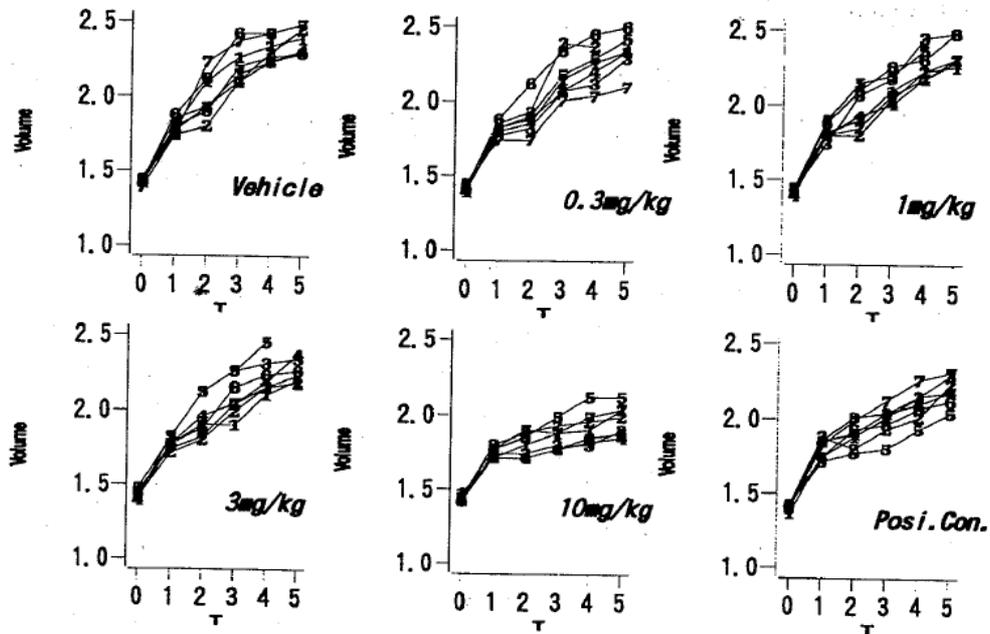
1. 個体内の反応値の変化に目を向けて、その用量反応関係を検討するには、経時測定データを意味のある(なるべく少数個の)指標に要約するのが有効である。本報告では、二つの薬効薬理試験の経時測定データの解析の中で、そのような要約指標として「時間の1次傾向」、「時間の2次傾向」、「全時点を通じた平均」、「曲線下面積(AUC)」をとりあげる。測定時点が等間隔であれば「時間の1次傾向」は、より馴染みのある「傾き」として表現することができる。

2.1. 抗炎症剤の抗浮腫作用のスクリーニング試験のデータ

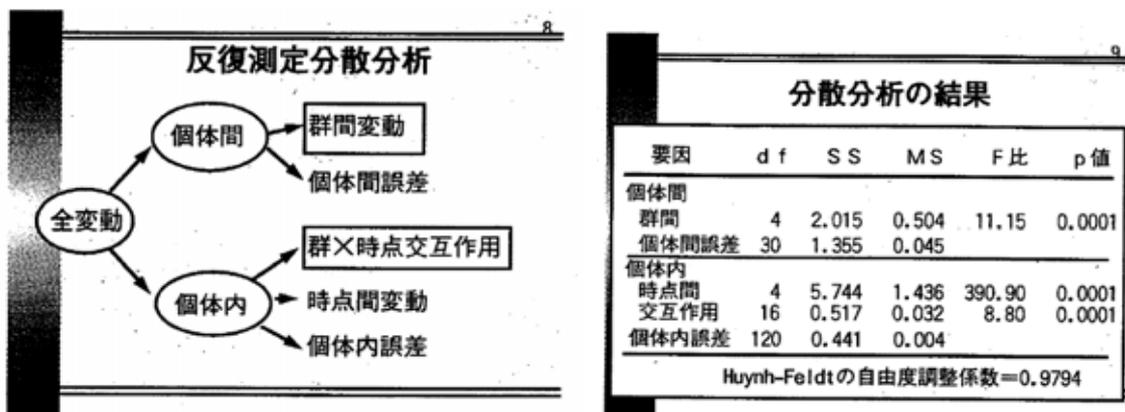
1番目の試験は、抗炎症剤の抗浮腫作用に関するスクリーニング試験である。ここでは、被験物質の4用量、Vehicle および陽性対照の6群が設定され、起炎剤によるラットの足の容積の膨張が5時点で測定された。化合物の抗浮腫作用は膨張に対する抑制効果として評価された。足の容積の経時測定データには、経時的推移がなめらかで、しかもすべてのラットで類似している、時間経過に伴ってラット間のバラツキが拡大する、等の特徴があった。これらのデータの特徴を踏まえ、浮腫容積(起炎剤の投与後の足の容積-投与前の足の容積)の経時測定データに反復測定分散分析(repeated measures

ANOVA) を適用した。この分散分析の通常の役目は、群×時点交互作用により各群の平均的推移が平行か否かを調べることである。ここではさらに、この交互作用の中身を追跡した。具体的に「時間の1次傾向」と「時間の2次傾向」を表す2個の対比をそれぞれ要約指標として用量反応関係を検討した。また、群間変動をもとにして「5時点を通じた平均」を要約指標として用量反応関係を評価した。その結果、浮腫容積の経時的推移の「傾き」、および5時点を通じた浮腫容積の平均について、ともに1次および2次の有意な用量反応関係が認められた。そして、これらの両指標のいずれでも、化合物の最高用量 (10mg/kg) は Vehicle に比して有意に低値であった。

図1：抗炎症剤のスクリーニング試験
足の容積の経時測定データ



反復測定分散分析の適用



曲線のあてはめ

2. 反応値の経時的推移がなめらかで、しかもすべての個体で類似している、その曲線を記述する関数が存在する、その関数が現象的意味づけのできるパラメータをもつ、等の条件が揃えば時間上の曲線のあてはめも有効である。実際に、抗浮腫作用の薬理試験の経時測定データには、負の指数成長曲線のあてはめが可能であった。この関数は、浮腫容積の最大限界値および膨張速度を表す2個のパラメータをもつ。これらの2個のパラメータの推定値（非線型混合効果モデルのあてはめによる）を経時的推移の要約指標として、用量反応関係を検討することもできる。

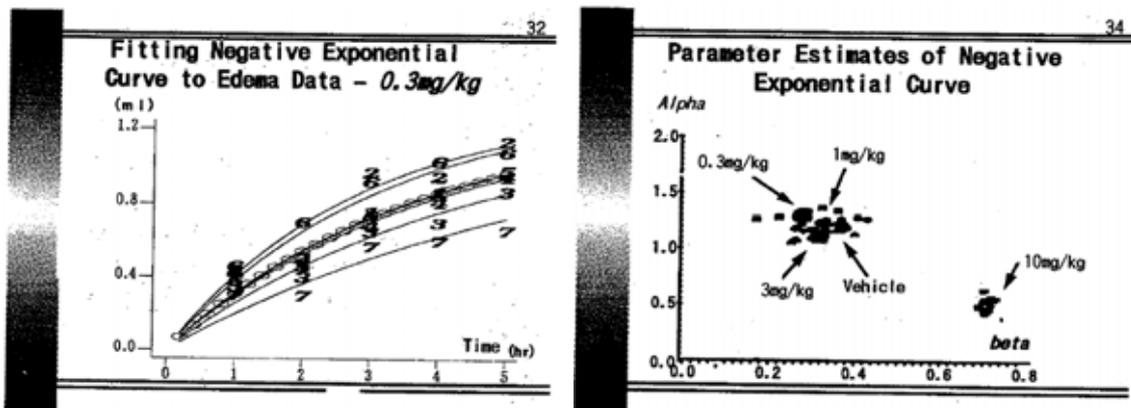


表 2.1 JMP のデータ (浮腫の容積_浦狩 JPAM.JMP)

浮腫の容積_浦狩 JPAM									
9/0列	化合物	NO.	体重	0	1	2	3	4	5
42/0行	1 A0_Vehicle	1	165	1.42	1.8	2.1	2.26	2.34	2.39
	2 A0_Vehicle	2	170	1.44	1.74	1.8	2.1	2.29	2.45
	3 A0_Vehicle	3	170	1.42	1.85	1.93	2.15	2.23	2.29
	4 A0_Vehicle	4	170	1.42	1.77	1.93	2.1	2.23	2.31
	5 A0_Vehicle	5	170	1.44	1.8	1.9	2.2	2.26	2.29
	6 A0_Vehicle	6	170	1.44	1.88	2.12	2.42	2.42	2.61
	7 A0_Vehicle	7	160	1.39	1.74	2.23	2.37	2.42	2.48

中略

29	A4-10mg/kg	1	170	1.47	1.71	1.8	1.88	1.9	2.04
30	A4-10mg/kg	2	175	1.44	1.8	1.88	1.93	1.96	1.99
31	A4-10mg/kg	3	170	1.42	1.74	1.74	1.8	1.8	1.9
32	A4-10mg/kg	4	165	1.42	1.71	1.71	1.77	1.82	1.85
33	A4-10mg/kg	5	180	1.44	1.77	1.85	1.99	2.12	2.12
34	A4-10mg/kg	6	175	1.42	1.71	1.71	1.77	1.85	1.88
35	A4-10mg/kg	7	170	1.44	1.77	1.9	1.88	1.99	2.04

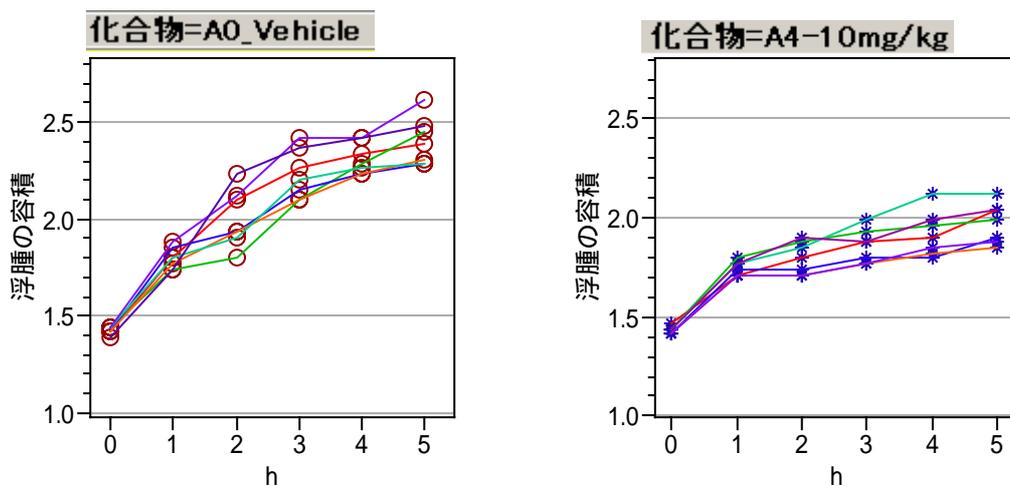


図 2.1 JMP による生データのプロット (浮腫の容積_浦狩 JPAM_T.JMP)

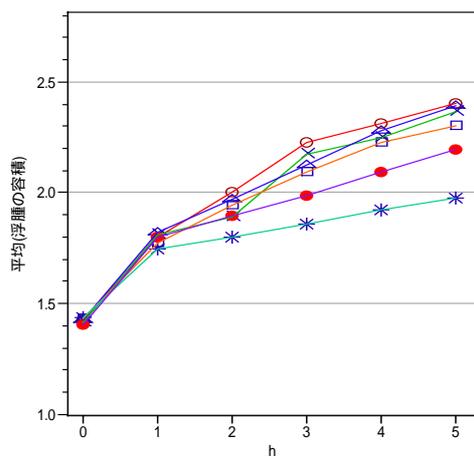


図 2.2 JMP による平均値のプロット (浮腫の容積_浦狩 JPAM_T の要約.JMP)

: Vehicle , × : 0.3mg/kg , □ : 1mg/kg , △ : 3mg/kg , * : 10mg/kg , ○ : 陽性対照

2.2. HPD 投与前後の DA 遊離量のデータ

2 番目の試験は、中枢神経薬が脳内移行性をもつことを裏付けるための試験であった。

そこでは、被験物質の 3 用量と Control の 4 群が設定され、ラットへのハロペリドール (HPD) の投与により誘発されるドーパミン (DA) の遊離量が、HPD 投与前の 3 時点と投与後の 12 時点において 20 分間隔で測定された。化合物の脳内移行は DA の遊離量に対する抑制効果として評価された。解析には、投与前の 3 時点の平均に対する投与後の各時点での比 (DA 遊離比率) を用いた。DA 遊離比率の経時測定データには、経時的推移がラットごとに不規則である、ラット間のバラツキが群間で相違する (共分散行列の不均一性) という特徴があった。とくに、共分散行列の不均一性は反復測定分散分析の前提を破るものであり、さらに複雑な統計モデルに基づく解析が必要かもしれない。ここでは、DA 遊離比率の経時的推移の曲線下面積 (AUC) を要約指標として用量反応関係を評価した。

その結果、AUC 値について 1 次の有意な用量反応関係が認められ、化合物の中用量および高用量での AUC 値は Control 群に比して有意に低値であった。ただし、AUC 値には経時的推移の差を識別できないという弱点のあることに注意が要る。1 番目の試験の場合と同様の手順に沿って反復測定分散分析により、経時的推移の傾きの用量反応関係を評価するには、たとえば隣合う 3 時点ずつの DA 遊離比率を平均化し、12 時点データを 4 時点データに変換すればよい。このとき、共分散行列の不均一さが補正される。12 時点データの 4 時点データへの変換は、恣意的な要素が強いため、その場合の解析結果が即座に受け入れられるか否かには問題が残る。

図 2：中枢神経薬の効力を裏付ける試験
DA 遊離比率の経時測定データ

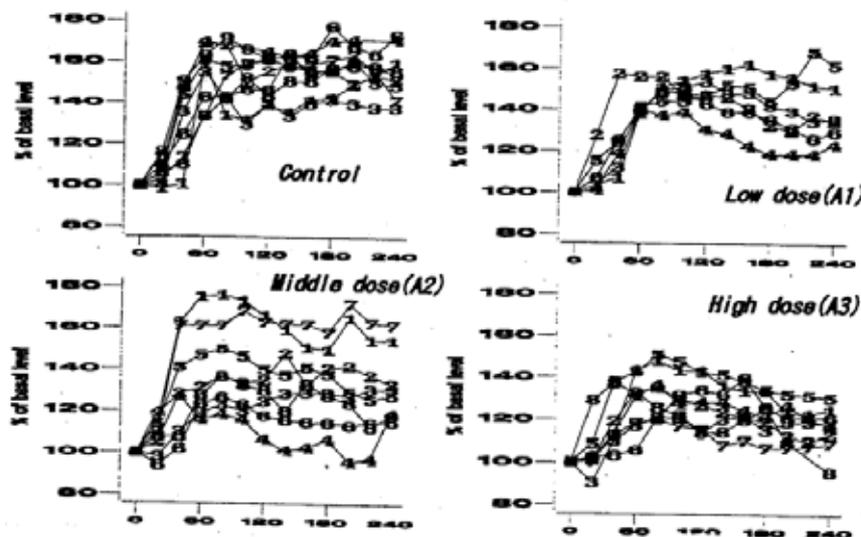


表 2.2 JMP のデータ (ドパミン遊離量_浦狩 JPMA.JMP)

群	NO	-40	-20	0	20	40	60	80	100	120	140	160	180	200	220	240
○ 1 A0_Control	1	17.2	16.8	16.6	16.6	17	26.8	22.6	22.2	23.5	22.9	23.9	24.1	25.2	26	27
○ 2 A0_Control	2	10	10.7	11.3	12.4	15.9	18	17.9	16	16.6	17	16.7	16.6	16	16.4	16.2
○ 3 A0_Control	3	11.3	11.8	11.2	11.5	15.5	17.7	16.3	14.8	16	15.3	16	16.2	16	15.8	15.8
○ 4 A0_Control	4	15	14.4	14.4	16	21.5	24.8	26.5	26.5	24.1	23.7	24	24.9	25	26.8	25
○ 5 A0_Control	5	37.2	37	36.5	39	53	59.4	58.2	58.5	59.5	58.3	57.6	57.5	59.3	55.5	55
○ 6 A0_Control	6	10.2	10.2	10.7	11.4	12.9	14.8	14.8	15.2	15.2	17	16.8	18.4	17.4	17	18
○ 7 A0_Control	7	38.5	38.5	38.5	39.5	43.8	51.5	59	61.3	62.3	60.8	61.2	62.1	63	60.2	54.7
○ 8 A0_Control	8	20.2	20.6	20.2	21.6	22.4	27.2	28.8	31	29.2	30.7	30.7	32	32.2	32.2	31.8
○ 9 A0_Control	9	32.6	33.7	35.5	38.8	51.1	55.3	58.3	56.3	54.7	55.5	52.5	53.2	54.3	53.6	52.4
× 10 A1	1	20.5	19.9	19.2	20.5	21.2	27.4	30	30.6	31.3	31.8	32.2	31.3	31	30.2	30
× 11 A1	2	14.9	14.9	15	19	23.5	23.3	23.3	22	22	21.8	21	19.8	19.4	20.5	20.1
× 12 A1	3	30.6	31.9	31.4	32.3	35	43.8	46.7	47.3	48.1	46.3	46.4	44.5	43.7	42.1	42.5
× 13 A1	4	26	26	26.3	26.3	31	36.6	35.8	36.2	34	33.5	32	31	31	31	32.2
× 14 A1	5	17.5	19.2	18.5	21.2	22.8	26	26.5	26.5	27	28	27.8	26.5	28.2	31	29.8
× 15 A1	6	24.8	25.4	25.8	27	31.7	35.2	37.7	37.5	36.2	35.2	35.2	34.3	33.2	32	32.8
△ 16 A2	1	20	20	21.3	23.6	33.3	35.8	36	35.4	34	32.6	30.8	30.6	33.8	31.6	31.6
△ 17 A2	2	16.8	17.6	17.6	18.2	22.2	22.7	23.6	23.1	23.3	25.6	23	24.4	24.6	23.8	23.1
△ 18 A2	3	22	21.4	21.8	21.3	23.8	25.8	26.7	26	27	27.9	28	28	27.6	28.4	28.4
△ 19 A2	4	7.6	7.4	8.4	9.3	9.9	9	9.2	9	8.3	7.9	8	8.3	7.5	7.6	9.3
△ 20 A2	5	35	35.2	34	39.3	48.9	51	52	50.8	48.5	47.6	49	47.8	48	43.7	44.3
△ 21 A2	6	41.7	39.5	38.3	37.4	40.5	49	50.3	48.9	46.8	46	45.6	45.5	45	45.6	45.8
△ 22 A2	7	9	11.4	11.8	11.8	17.3	17.3	17.3	18.1	17.5	17.4	17.4	17	18.5	17.5	17.4
△ 23 A2	8	11.6	12.7	12	13	13	15.4	16.5	16	15.7	14.5	16.1	15.4	15	13.6	14.3
■ 24 A3	1	27.5	28.5	28.2	28.5	31.6	40.2	41.5	40.2	40.2	38.2	37.5	37.5	34.6	34.5	35
■ 25 A3	2	25.4	25.8	25	25.6	30.2	33.7	34.3	33	32.2	31.4	30.8	30.2	28.6	28	28.8
■ 26 A3	3	34.8	35.5	32.7	31	37.7	40	41.5	41.5	39	41.2	41.8	39.5	40.8	40.8	40.4
■ 27 A3	4	38.1	37	36.8	38	42.2	48.8	50.8	47.5	47.8	47.8	46.4	45.4	47.3	45	43.6
■ 28 A3	5	48.7	48	48.7	52.5	66.7	69	73	71.5	69	68.4	67	65	64	63.7	63.2
■ 29 A3	6	23.2	24	24.2	24.4	24.5	25.1	28.5	31.3	31.9	31.4	33.3	30	28.8	28.8	28.8
■ 30 A3	7	34.3	32.5	32.2	33	37	39	39.5	38.5	38	35.7	36.4	35.3	35.3	35.5	36
■ 31 A3	8	8.5	7.8	10	11.3	12	11.6	11	10.7	10	10	10.5	11	9.6	6.4	8.4

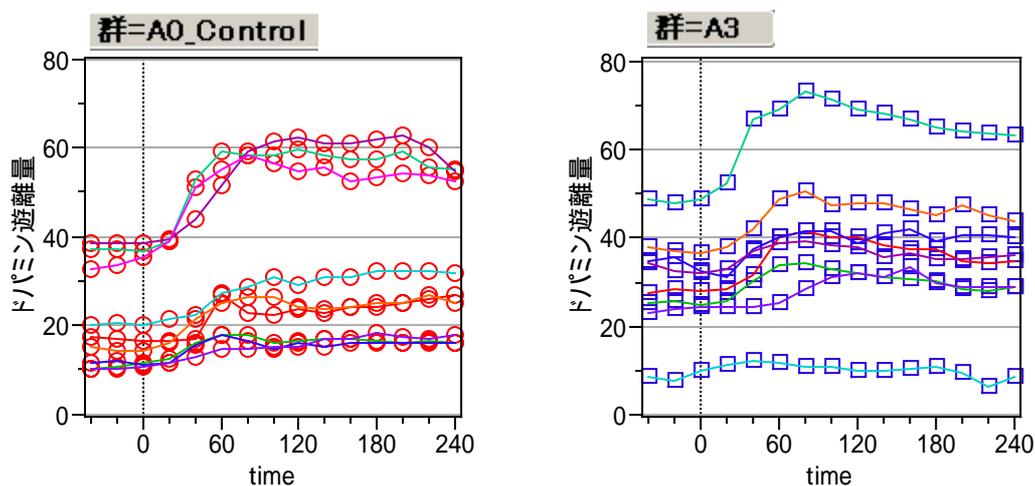


図 2.3 生データのプロット

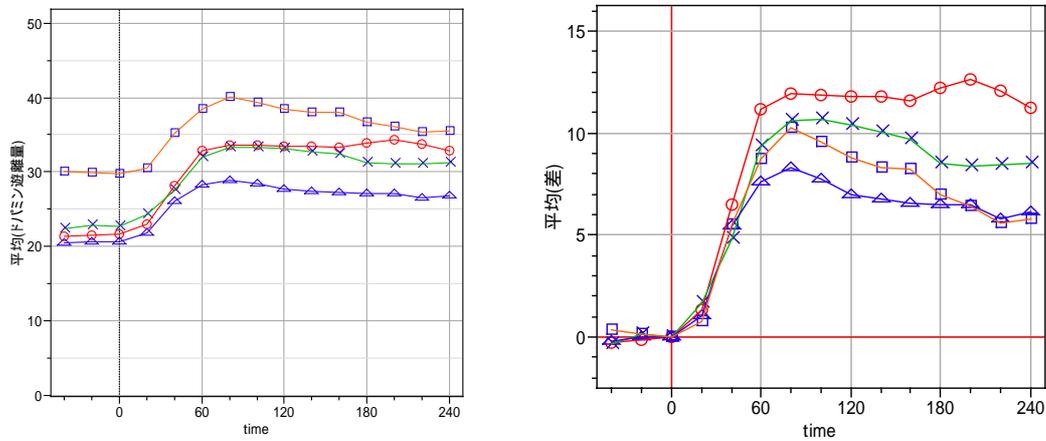
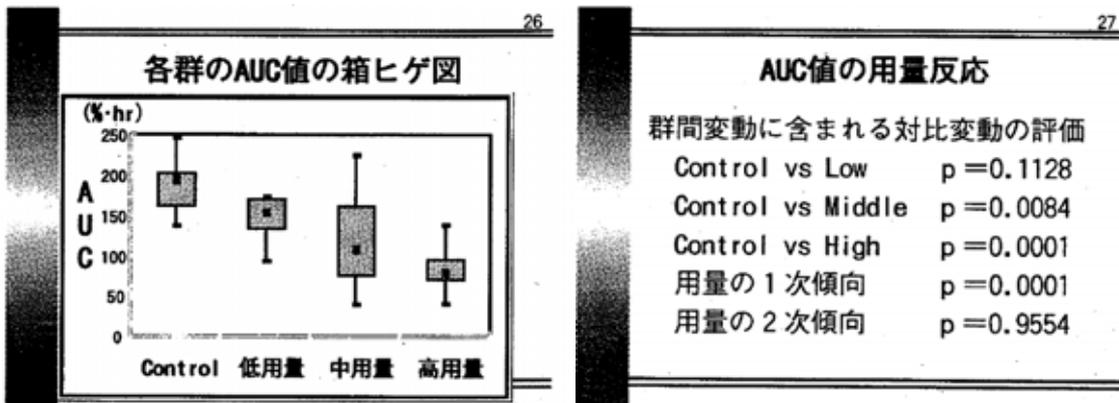


図 2.4 JMP による平均値のプロット (ドパミン遊離量_浦狩 JPMA_T の要約.JMP)
 □ : Control, × : A1, △ : A2, ○ : A3

要約統計量として AUC を使った場合 (ドパミン遊離量_浦狩 JPMA_T の要約.JMP)



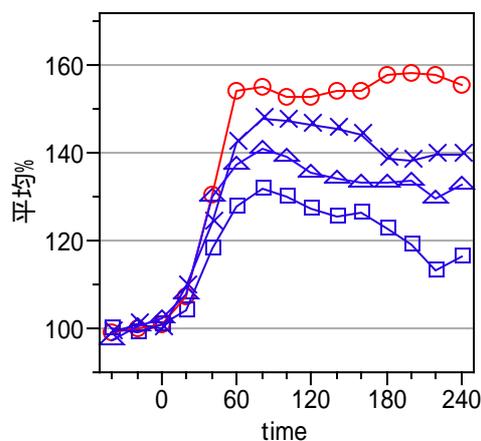


図 2.5 変化率の平均値の経時変化

○ : Control , × : A1 , △ : A2 , □ : A3

2.3. 解析の留意点

3. 経時測定値の解析において、反応値の時間的变化に対して視点が向けられつつある。本年に通知された ICH-E3 ガイドラインにも「基準値からの変化、傾きの解析、…といった複数の解析手法が妥当とみなせる場合には、計画された手法を明確にすること」、「可能な限り反応の時間的経過を記述すること」という記載がある。このガイドラインは臨床試験を対象としたものであるが、その考え方は非臨床の場合にもあてはまる。時点ごとの解析結果の移り変わりに一喜一憂するのではなく、**個体内の変化**をどう捉え、どのように評価するかを重視していくべきであろう。このとき、次の警句が注意を惹く。

We want the statistical significance of the change

not the change in the statistical significance. ~ M.G. Kenward

3. 経時データの解析について安全研での議論の集大成

BioS での講義資料からの抜粋

3.1. はじめに

新医薬品の開発過程において、様々な経時データが測定されている。ヒトを対象とした臨床試験のみならず、イヌなどを対象とした毒性試験でも、多くの経時データが測定されている。しかしながら、それらのデータに対して適切な統計解析が行われているとは言い難い。

共通する問題は、有意差検定の乱用である。測定時間ごとには、多重性を考慮した有意差検定をしているというのであるが、輪切りの多重性を考慮せずに繰返し行われている。さらに、時間方向にも多重性を考慮したと言いつつ、投与群ごと輪切りの有意差検定も行われている場合もある。多重性を考慮したと言いつつ、それを無視した有意差検定の洪水である。これは、多重性を考慮したと述べさえすれば、免罪符が得られたかのごとくである。

統計的な仮説検定の方法は広く世の中で用いられている統計手法の一つである。これは、多くの学問分野別に出版されている統計の入門書の主要なテーマの一つであることによって支えられている。しかし、経時データの統計解析について触れているのは極めて稀である。

そのため、多くの実験研究者は、実験結果をまとめる際に自ら学習した統計の教科書から得られた知識を用いようとするのは当然である。実際に統計計算を必要とするときには、多くの統計ソフトが2群間の有意差検定、あるいは多重比較をサポートしているために気楽に用いられているのである。

検定の多重性の問題でも、問題があると言うことは容易であるが、現実の問題に対し適切な解析手法を選択し、かつ適切にその結果を解釈し、適切な結論を導くことは容易でない。経時測定データを適切に統計解析しようとするとき、検定の多重性の問題をも包含する様々な問題が目の前に立ちはだかるのである。

統計的な仮説検定は、1元配置型の実験結果を前提にしているために、経時的なデータの解析に元々そぐわないのである。経時データの統計解析は、統計学の分野でも新しい課題であり、多くの統計ソフト、SASでさえも適切な解析がサポートされていなかった。SASのバージョン6、リリース6.08から少数例の場合の誤差の計算に機能が追加され実際に活用できるようになったばかりである。

経時データには、さまざまなタイプのデータがあり、一まとめにすることはできない。経済の分野で取り扱っている経済時系列データ、あるいは、脳波、筋電図、音声など生体から連続的に得られる電流の変化をデータとする時系列データの解析は、データを"波"として取り扱うフーリエ解析、ARMA(自己回帰移動平均)モデルなどを主な解析手段としている。

ここで取り上げる経時データの解析は、"時系列"というには短い、実験で予め測定点と定められた数時点から10時点程度のデータである。医薬品の開発過程でしばしば経験する経時的に測定されたデータの例をいくつか示してみよう。

短時間の変化

誘発電位の観察、反応が発現するまでの時間、変化パターンの分類
脳波の周波数帯域の変化
手術時のバイタルサインの変化

一日以内の変化

薬物の血中濃度の変化、半減時間、最大血中濃度
血圧の日内変動、日内リズム
睡眠深度の変化、レムの回数、各睡眠深度の割合
臨床薬理試験における生理活性物質の変化

長期の変化

臨床試験での臨床検査データの変化
反復毒性試験の臨床検査データの変化

これらのデータで経時測定データとしての解析の対象とするのは、一団体ごとに、一定の時間、週、月単位でデータが得られていることを前提とする。なお、これは、等間

隔という事ではなく、予め定められた時点である。

解析方法の分類は、吉村、大橋（1992）毒性試験データの統計解析、2.4 節 経時測定データの解析によれば、

- (A) 二元配置母数モデル
- (B) 枝分かれモデル
- (C) 長曲線モデル
- (D) 多変量モデル
- (E) 時系列モデル
- (F) クラス分けモデル

について解説されており、考慮すべき課題も次のように示されている。

- (1) 個体を母数と見做すか、変量と見做すのかの判断
- (2) 初期値を差し引くか否かの判断
- (3) 層の併合の問題
- (4) 時点に交絡する偏りの除去
- (5) 実験前の測定値の取り扱い

これらの、解析方法および課題だけでも広範な内容を含んでいるのであるが、この多くの課題に対して検討を重ね、その成果を SAS ユーザ会、医薬安全性研究会、計量生物学会などで発表してきた。このテキストは、これらの結果を整理したものである。

このテキストは、主に臨床試験に従事する試験統計家を目指す人達を前提に、経時データの統計解析で、ぜひ知っておいてもらいたい事柄について典型的な事例を織り交ぜながら述べたい。試験統計家の最大の責務は、試験計画である。それを全うするためには、統計解析の様々な事柄について幅広い知識が必要である。このテキストは、経時データを主要変数とする臨床および非臨床試験について適切な統計解析の計画および実施ができるようになることを目的としている。

3.2. 単回，反復，および逐次投与の区別

これまで述べてきたように述べたように，単回投与後の経時データと反復投与後の経時データでは，評価の視点が異なる．単回投与の場合には，生体内の薬物濃度に関連して生体反応が徐々に減少(増加)し，ある一定時間後には元に戻るようなバスタブ型(ベルシェイプ型)の反応を期待し，投与量よる形状，主に大きさの違いに興味があることが多い．これに対し，反復投与の場合には，

- 1) 初期の反応の早さに興味がある場合，
- 2) 維持レベルに興味がある場合，
- 3) 投与終了後の回復の形状に興味がある場合

などに区分される．

薬物を持続点滴し，逐次的にその量を増加させながら反応を観察するような薬理実験からも経時データが得られる．その用量反応は，ある投与量に曝露されたという条件下での反応であることから，単純な用量反応関係ではないことに注意しながら解析する必要がある．

3.3. 経時データの分類

経時データの解析は，これまで述べたように，投与方法，期待する反応の形状によって解析の視点が明らかに異なる．そこで，実験の型を前もって分類しておくことは，経時データの解析方針を考える上で有益である．

I 単回投与後の経時データ

- (a) 可逆的反応を前提とする場合
- (b) 非可逆的反応を前提とする場合
- (c) 日内変動を伴う反応の場合

II 反復投与の経時データ

- (d) ある反応に達するまでの変化
- (e) 下限値あるいは上限値に達した後の定常性
- (f) 反復投与後の可逆性

III 逐次増量時のデータ

- (g) 同一個体内での用量反応
- (h) 用量反応の比較を伴う場合

3.4. 検定の多重性の問題

そもそも、ダネット、あるいはチューキーの多重比較は、一元配置型の実験データを前提にした検定手法である。薬効を検証しようとする場合に、時点ごとに多重比較を繰り返して、ある時点で5%の水準で有意な差が出たからといっても、検定の多重性の問題により5%の有意水準が保たれてはいない。さらに、投与前と各時点の差の検定を追加した場合にも、さらなる検定の多重性の問題が生じる。

多重性を回避するためには、これまで述べてきたように、試験目的に合わせた時点为主要変数として有意差検定を行うことにつぎる。不幸にも、主要変数に有意差を見出せない場合には、潔く実験計画の失敗と認め、新たな実験を行う覚悟を持つべきである。副変数には有意差が出たからといって、薬効が統計的に証明されたことにはならない。

主要変数は、ある1時点の経時データである必要はない。単回投与試験であれば、反応が最大となる前後の時点の平均値を主要変数とし、ばらつきを減少させることも考えられる。

3.5. 全ての時点を用いた検定

これまで述べてきたのは、経時データをそのまま用いるのことに伴う検定の多重性の発生を食い止め、従来型の仮説検定の世界で事を完結したいとの願があったからである。全ての時点を用いた解析法は、基本的には探索型の統計解析となる。

- 1) 時点を多変数とみなした多変数分散分析モデル
- 2) 時点を固定効果とみなした枝分かれモデル (自由度の調整法も含む)
- 3) 個体を変量効果、全ての時点を固定効果とした線形または非線形混合モデル
- 4) 各測定時点間の相関構造を考慮した GEE 法 (Generalized Estimating Equation Method)

5) その他：パターン分類，時系列モデル

などがある．

さらに，厄介なのは，測定された時点間の相関を無視することができないことである．全時点のデータをまとめて解析する場合には，時点間の相関構造を明示的に指定する必要がある．さらに付け加えると，時点間の分散が無視できないほど異なる場合があり，分散の安定化をするか，重み付き混合モデルを使わざるを得なくなる場合もある．これは単回投与後の数時間から 24 時間ぐらいまでの薬効を経時的に測定する時にしばしば経験する．図 3.1 には，剤形ごとに経時的な変化を示したものである．時点ごとの SD が薬物濃度に関連して明らかに異なる．このような経時データを扱う場合には，変数変化をして分散の安定化を計るか，測定値の逆数を重みとするような配慮が必要である．

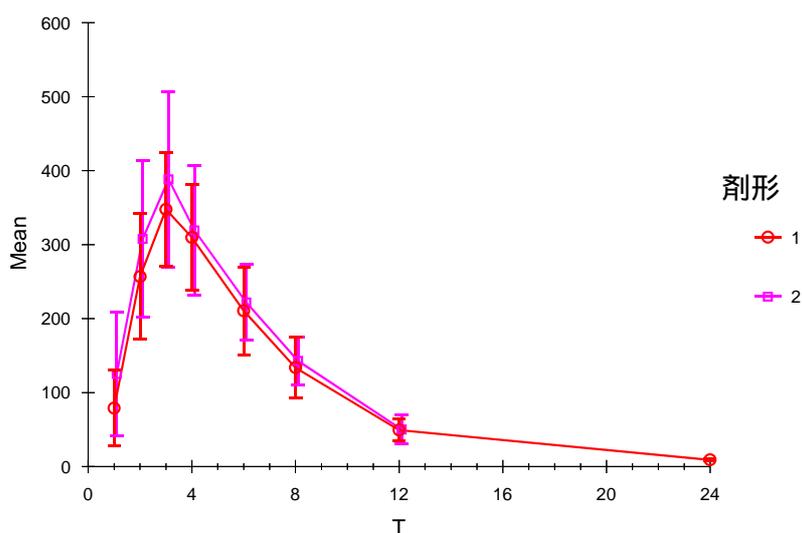


図 3.1 剤形別薬物濃度の経時推移

平均 \pm SD

3.6. 良くある質問

1 元配置の実験モデルが適用できる場合に，分散分析を多重比較に先立ち行う必要があるのか，との問題に対して，比較の型があらかじめ定められていれば，分散分析のような予備検定は必要なく多重比較を直接行って差し支えない．全ての時点を用いた解析手法も，これと同じ問題であり，主要変数についての検定がしっかり行ってあれば，

全ての時点を用いた解析手法をあえて行う必要ない。

とはいえ、検定の多重性の批判にもめげずに、経時データに対する分散分析（時点を固定効果とみなした枝分かれモデル、自由度の調整法も含む）に引き続き、輪切りの分散分析の是非についての質疑が絶えない。これは、幾つかの統計計算ソフトにこのような解析手順が組み込まれていて、その手順の是非について判断に迷っているからである。そして、この問題に対する標準的な解析の考え方が、一般的な統計の教科書で示されてこなかったためでもある。

これまでも述べてきたのであるが、断定的に言えば、「経時データに対する分散分析に引き続く輪切りの分散分析」は、探索的な方法であり、統計的な仮説検定には使えない。次に、医薬安全性研究会で A さんの質問に対する回答を示す。

(1) 1 回の試験から多くの情報を得ようとして何が悪い？

答：検証すべき命題を多重性、多項目性を考慮すれば OK

検証すべき命題が無い試験は科学的な試験とは認めない、許されない

(2) 1 次解析法として repeated measures ANOVA は、一般的になるか？

答：たいした答えが出ないので、はやらないであろう、薦められない

(3) 無相関、等分散を仮定してよいか？

答：No! 相関構造は常に注意しなければならない

(4) 投与前からの差および率への変換は問題ないか？

答：投与前と投与後のデータに、ほとんどの生物実験で相関が認められることから、統計的な検出力の低下が起こるので、薦められない、止めるべきだ

(5) 交互作用が量的に有意な場合に時点をおしなべて群間の検定をしてよいか？

交互作用が有意な場合、時点ごとに群間の検定をしてよいか？

答：探索的な解析としてならば、好きにすればよい

(6) Repeated measures ANOVA 後に時点での輪切りの多重比較を容認するか？

答：検証的な観点ならば容認しない、図を見て解釈できる以上のものは得られない

(7) 作用の発現時期、持続時間、可逆性を統計的検定で解決できないか？

答：薬理屋のプライドにより定義する問題であり、
検定で解決しようとするれば自己矛盾に陥る、
考えを変えなさい

(8) 非臨床薬理、毒性、臨床薬理で手法を使い分けるべきか？

答：試験分野ではなく、検証すべき命題で使い分けるべきである

(9) 用量反応性を知る手法はあるか？

答：経時データのままで手法は思い当たらない

(10) 順序分類データに対しての方法論、ツールがあるか？

答：あるかと問われれば、SAS の CATMOD、GENMOD プロシジャがあるが、
それらの方法を適用するための制約などがきつく実務に使ったことはない

3.7. 投与前値の調整法の比較

ヒトを対象とした臨床試験，大動物を対象にした毒性試験，あるいは薬理試験などから得られた経時データで，2種類の治療の比較を行うことを考えよう．それぞれの個体について，投与後のデータのみならず投与前値と投与後の差，あるいは比に変換したデータについての比較検討が，投与前における2群間の平均値のアンバランスを消去すると共に，個体間のばらつきを消去する方法としてしばしば用いられている．

それらの変換データに対する有意差検定の性能が，投与前値を共変量とした共分散分析による比較に比べ劣ることが，実際の臨床試験データの各種の臨床検査データに基づき，Frison, S. and Pocock, S.J. (1992) によって報告されている．

A群の投与前値を x_{Aj} , $j=1, \dots, n_A$, B群の投与前値を x_{Bj} , $j=1, \dots, n_B$ 投与後の結果をそれぞれ y_{Aj} , $j=1, \dots, n_A$, y_{Bj} , $j=1, \dots, n_B$ としよう．それぞれの群の投与後の分散を σ^2 としたときに，投与後の2群間の平均値の差 ($\bar{y}_{A.} - \bar{y}_{B.}$) の分散は，

$$\text{Var}(\bar{y}_{A.} - \bar{y}_{B.}) = (1/n_A + 1/n_B) \sigma^2$$

となる．それぞれの群の個体ごとに投与後の結果から投与前値を引いて平均値を求め，2群間の差 ($(\bar{y}_{A.} - \bar{x}_{A.}) - (\bar{y}_{B.} - \bar{x}_{B.})$) の分散は，それぞれの群内の投与前と投与後の相関を ρ としたときに

$$\text{Var}((\bar{y}_A - \bar{x}_A) - (\bar{y}_B - \bar{x}_B)) = (1/n_A + 1/n_B)(1 - \rho)\sigma^2$$

となる。

投与前値を共変量とした共分散分析で求められた両群の共通の傾きを β としたときに A 群の各個体の投与前 n 総平均 $\bar{x}_{..}$ における調整値は $y_{Aj(COV)} = y_{Aj} - \beta(x_{Aj} - \bar{x}_{..})$, B 群の各個体の調整値は $y_{Bj(COV)} = y_{Bj} - \beta(x_{Bj} - \bar{x}_{..})$ となる。各群の調整値の平均値の差 $(\bar{y}_{A(COV)} - \bar{y}_{B(COV)})$ の分散は、近似的に

$$\text{Var}(\bar{y}_{A(COV)} - \bar{y}_{B(COV)}) = (1/n_A + 1/n_B)(1 - \rho^2)\sigma^2$$

となり、各群 10 例、投与前と投与後の相関 ρ を 0.6 とした場合に

$$\text{POST: 分散} = (1/n_A + 1/n_B)\sigma^2 = 0.200\sigma^2$$

$$\text{CHANGE: 分散} = 2(1/n_A + 1/n_B)(1 - \rho)\sigma^2 = 0.160\sigma^2$$

$$\text{ANCOVA: 分散} = (1/n_A + 1/n_B)(1 - \rho^2)\sigma^2 = 0.128\sigma^2$$

とる ρ を 0 から 1 まで変化させたときの POST, CHANGE, および ANCOVA の分散は、表 3.1 に示す。

この結果から、CHANGE が、POST 比べて検定の際の分母となる分散が必ずしも小さい訳ではないことが示されている。投与前と投与後の時点間の相関が 0.5 を境にして、相関が小さい場合には、投与後の値のみを用いることが望ましいことがわかる。それに対して、ANCOVA は、POST, CHANGE 比較した場合に時点間の相関の大小に係わらず分散が常に小さくなることが示されている。

表 3.1 時点間に相関がある場合の分散の比較

ρ	POST	CHANGE	ANCOVA
0.0	$0.20\sigma^2$	$0.40\sigma^2$	$0.200\sigma^2$
0.2	$0.20\sigma^2$	$0.32\sigma^2$	$0.192\sigma^2$
0.5	$0.20\sigma^2$	$0.20\sigma^2$	$0.150\sigma^2$
0.8	$0.20\sigma^2$	$0.08\sigma^2$	$0.072\sigma^2$
1.0	$0.20\sigma^2$	$0.00\sigma^2$	$0.000\sigma^2$

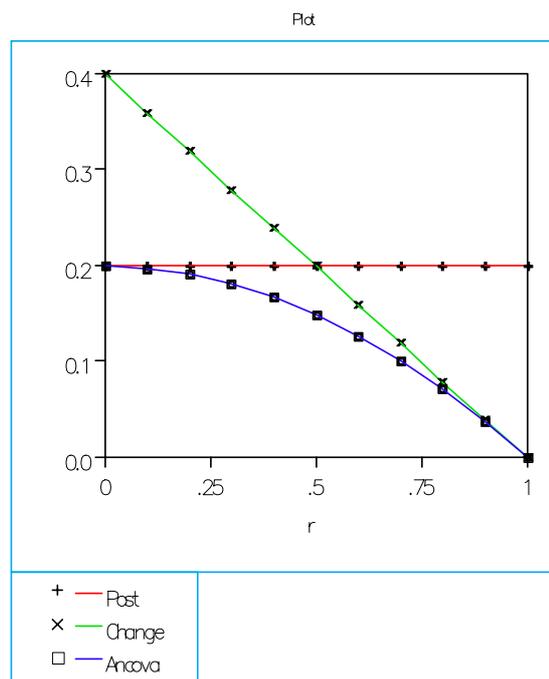


図 3.2 相関係数の変化による分散の比較

投与後の値を投与前の値で除して変化率 CAHNG%にすることも、差を取ることに同様に広く使われているが、投与前値にも投与後と同様の個体内変動がある場合には、必ずしも群間比較の性能が良い訳ではない。投与前値がたまたま低くなった場合に投与後の変化率の著しい増大が生じることもあり注意が必要である。このような場合に、対数変換した上で共分散分析をすることが望ましい。

統計的な性質として共分散分析が優れていることは確かであるが、常に必要な訳ではない。平均値の差の検定でP値が0.001となるような明らかな反応の場合にはどのような方法でも結論は同じであり、POSTのみでも良いし、CHANG、CHANG%で何でもよい。これに対し、実験例数の制約があるような大動物の毒性試験の場合などには、対照群との差の検定の感度を上げるためにANCOVAを積極的に使うべきである。

4. 解析の考え方，実験計画の再考

4.1. 差または比をとるべきか，投与前値を共変量とすべきなのかの判定

浦狩らの経時データはどのような性質を持っているのだろうか。実験の目的は検証的であるのだろうか。実験の規模から検証的な試験と考えられる。この試験の用量反応関係の形状から，この試験以後の用量設定をする際の参考されるのであろう。

浮腫の容積の実験で検証すべきことは，何であろうか。実験が成立する前提としては，Vehicle 群に対して，高用量 10mg/kg 群と陽性対象群が共に統計的に有意な差が同時に見出せることであろう。この場合には，検定の多重性の問題は発生しない。用量反応関係の形状を知ることが次の課題である。ICH-E4「新医薬品の承認に必要な用量 反応関係の検討のための指針」は，人での用量反応関係の統計的な考え方も示されているので，これに準じて実験動物の用量反応関係を取り扱うことを薦める。

表 4.1 経時データの性質

	経時データの性質	浮腫の容積	ドパミン遊離量
1	経時観察の長さ	1日以内	1日以内
2	薬剤投与回数	単回か？	単回か？
3	観察する反応	反応の立上り	可逆的反応
4	主要変数	最終時点	1~2時間目
5	副次変数	中間時点	最終時点
6	投与前値の CV	2%	45%
7	前値との相関（低用量）	正または負	0.98 以上
8	前値との相関（高用量）	正または負	0.97 以上
9	投与後の時点間	自己回帰型	0.98 前後で一定

浮腫の容積の場合は，変数変換は投与前値と各時点との正の一貫した相関が見られないことから，差あるいは比をとる必然性は無い。もちろん，投与前値を共変量とすることは避けなければならない。最終時点の反応で群間比較を行えばよい。中間時点の反応は副次的な解析とするのであるが，最終時点との関係から用量反応関係の性質を探索的に見るために，用量ごとに2つの時点間の確率楕円による検討が考えられる。

ドパミン遊離量は，投与前値の CV が 45% と大きく，かつ，投与前値と各時点の相関が 0.98 と高いので，投与前値からの差あるいは，投与前値を共変量とする共分散分析を行う必要がある。

浦狩らは，時点間の相関が CS タイプを前提としている分割実験タイプの分散分析をしているが，時点間の相関が自己回帰型であること，時間経過につれて変動が大きくなることから，適用はできない。

浮腫の容積

5 時間目の浮腫容積から，高用量 10mg/kg 群と陽性対象群が共に統計的に有意な差が同時に見出せているので，実験は成立したことになる．用量反応関係は，3mg/kg 群から出始めているようである．薬理作用は，相対的な問題であり，陽性対照と同程度の作用を示す本薬の用量を推定したい．反応が飽和していないので，シグモイド曲線をあてはめることはできないので，1mg/kg，3mg/kg，10mg/kg の対数用量にたいして便宜的に回帰直線をあてはめて，陽性対照の反応が平均値となる本薬の用量を推定することにする．

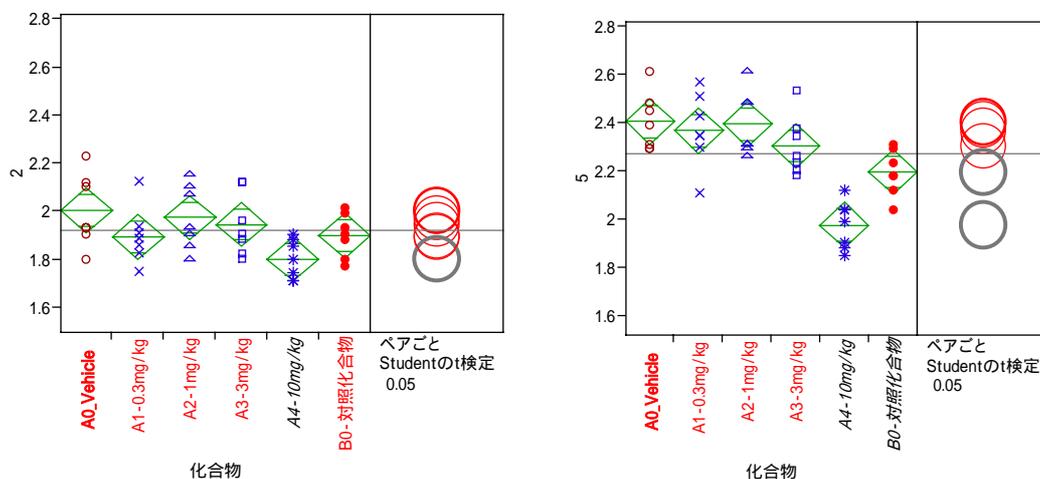


図 4.1 2 時間目（左）と 5 時間目（右）の浮腫容積についての用量反応関係

2 時間目では，薬効が出ているのであろうか．高用量 10mg/kg では有意な差とはなっていないが，陽性対照群に有意な差が出ていないので，2 時間目では，実験が成立しているとは言いがたい．この時点での用量反応関係を論ずることは差し控えるべきである．

浦狩らが逆指数関数をあてはめ，最大限界値と膨張速度を持つパラメータで散布図を示している．これに類似する検討を，2 時間目の浮腫容積を X 軸に，5 時間目を Y 軸とした散布図に用量ごとに 50% の確率楕円を描いてみよう．

浦狩らが行った個体ごとの 6 時点の経時データから推定された 2 つのパラメータほどのシャープさには欠けるが，5 時間目の浮腫容積は同じであるが，浮腫容積が速やかに膨張したのか，緩やかなのかの区別に少しは役にたつであろう．

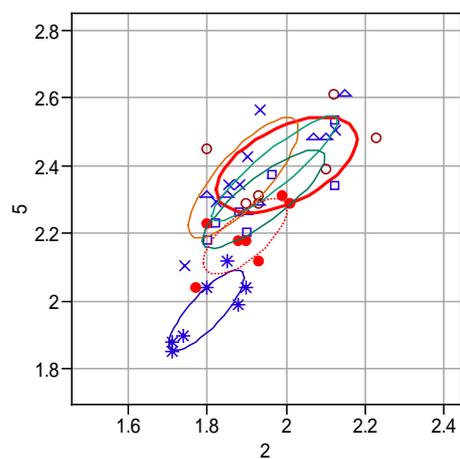


図 4.2 2時間目と5時間目の浮腫容積の用量

ドパミンの遊離量

生データの経時プロットから、投与前値が大きければ変化量も大きく、小さければ小さいとの関連が見受けられる。分布は平板な一様分布的であり、対数変換はなじまない。薬剤投与後の反応は個体ごとの投与前値の違いを凌駕するものではなく、薬効は投与前からの変化率が望ましい。

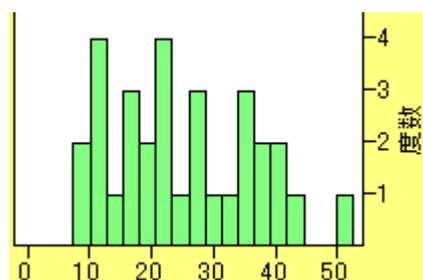


図 4.3 投与前3時点の平均値の分布

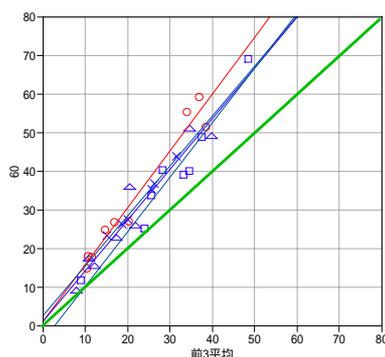
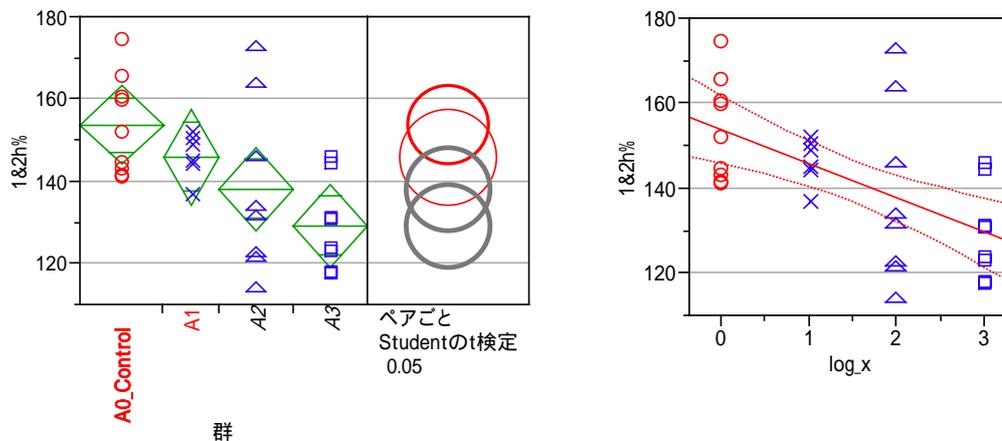


図 4.4 60 分値のドパミン遊離量と投与前 3 ポイントの平均値の関連

この実験で検証したいことは何であろうか。変化率の最大値とした場合に、個体によるピークの変動があるので、60分から120分間の4ポイントの平均を主要変数とし、用量反応関係を論ずることにする。投与量が示されていないので、Control群を対数で0、低用量を1、中用量を2、高用量を3とみなしてロジスティック曲線をあてはめ、形状パラメータの信頼区間で統計解析を行なうにしたい。ただし、この場合にロジスティック回帰は4パラメータが必要となり、Control群を含めて4用量しかないのでロジスティック曲線のあてはめはできない。便宜的にControl群と高用量群間で平均値の差が統計的に有意であることをもって実験が成立したとしよう。

用量反応関係について直線回帰を当てはめLOFを参考に用量反応関係を論ずることにする。副次的変数として、240分目での変化量について用量反応関係を論ずることにする。これらは、経時的プロファイルを解釈するための統計的なアプローチである。



あてはまりの悪さ(LOF)

要因	自由度	平方和	平均平方	F値
あてはまりの悪さ(LOF)	2	3.2453	1.623	0.0084
純粋誤差	27	5245.1077	194.263	p値(Prob>F)
合計誤差	29	5248.3530		0.9917
				最大R2乗
				0.3432

パラメータ推定値

項	推定値	標準誤差	t値	p値(Prob> t)
切片	153.81833	3.922308	39.22	<.0001 *
log_x	-8.098954	2.082218	-3.89	0.0005 *

図 4.5 60分から120分の平均ドパミン遊離量

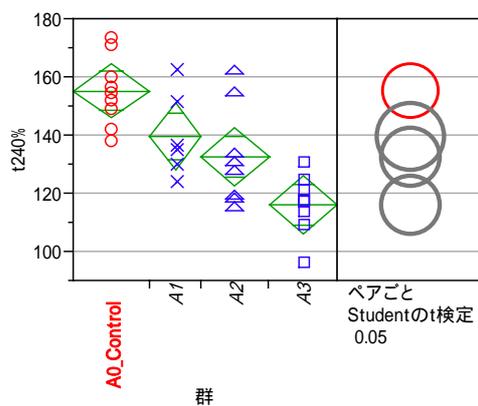


図 4.6 240分のドパミン遊離量の比較

4.2. 陽性対照と同程度の反応を起こす用量の推定

浮腫の容積

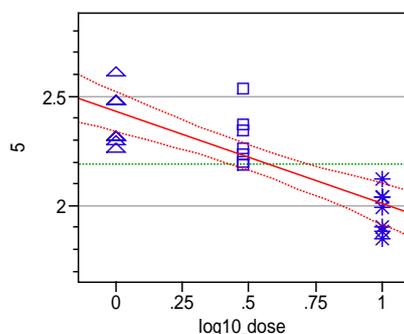
陽性対照と同程度の反応となる本薬の用量を推定値し、効力比を求めることが、この薬理試験の本来の目的であろう。しかしながら、陽性対照の用量が示されていないので、効力比は求められないが、本薬の陽性対照と同じ反応となる用量と95%信頼区間を推定しよう。

用量反応の観点から結果をみると、0.3mg/kg、1mg/kgは薬理反応が現れる用量以下であり、3mg/kg群から少し反応が現れ、10mg/kgでは明らかな反応が得られている。さらに用量を上げれば、反応は飽和すると思われる、反応全体としては、ロジスティック曲線などのシグモイド状の反応となるであろう。

直線のあてはめ

おおよその用量の推定ならば、3mg/kgと10mg/kgの対数用量の2点で反応の平均値間で直線を引いて、陽性対照の反応の平均値となる本薬の用量を逆推定すればよい。

1mg/kg群を加えた場合に、反応が直線的に減少しているとはいいがたい面もあるが、JMPの逆推定の機能を使って推定してみよう。



あてはまりの悪さ(LOF)

要因	自由度	平方和	平均平方	F値
あてはまりの悪さ(LOF)	1	0.05543419	0.055434	3.9200
純粹誤差	18	0.25454286	0.014141	p値(Prob>F)
合計誤差	19	0.30997704		0.0632
				最大R2乗
				0.7261

逆推定

5 予測値 log10_dose 下側限界 上側限界 1-Alpha
 2.190000 0.569385732 0.428977727 0.729891287 0.9500
 期待応答についての信頼区間

	ラベル	データ	antilog
1	予測値 log10_dose	0.5694	3.71
2	下側限界	0.4290	2.69
3	上側限界	0.7299	5.37

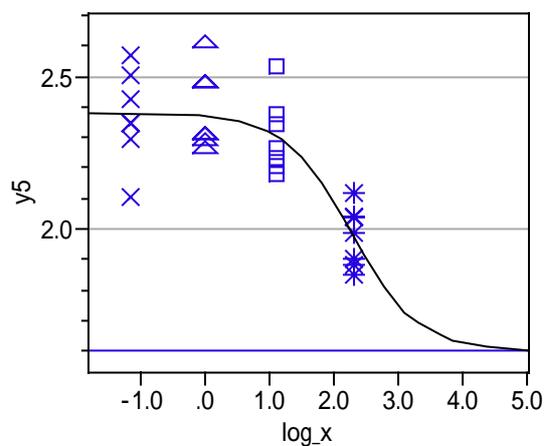
陽性対象群の反応の平均値は 2.19 であるので、それと同等の反応を得るための本薬の用量は、3.71mg/kg (2.69 ~ 5.37) と推定される。

ロジスティック曲線のあてはめ

反応が 1.6 で飽和することを仮定し、本薬の ED_{50} を推定しよう。JMP データセット (浮腫の容積_浦狩 JPMA_非線型.JMP) を推定してみる。

計算式

$$1.6 + \frac{\text{beta4} - 1.6}{1 + \text{Exp}(\text{beta1} * (\text{log_beta2} - \text{log_x}))}$$



パラメータ	推定値	近似標準誤差	下側信頼限界	上側信頼限界
beta1	-1.936437792	0.72257041	.	-0.9244555
log_beta2	2.2575463685	0.13460388	1.95505863	2.61008272
beta4	2.3811010005	0.03858149	2.30754366	2.47708271

	ラベル	log	antilog
1	推定値	2.26	9.56
2	下側信頼限界	1.96	7.06
3	上側信頼限界	2.61	13.60

ED50 は、9.56mg/kg (7.06 ~ 13.60) と推定される。

4.3. 実験計画の再考

浮腫の容積

薬理試験の実験計画が、平均値の差の有意差検定に汚染されている。浮腫の容積の場合、用量の設定範囲が狭いために反応がでる用量が 10mg/kg の高用量群だけである。1 用量当たり 7 匹を 3 匹にして用量設定の範囲を広げ、反応の全体像を把握すべきである。陽性対照も 1 用量に 7 匹としている。陰性対象と有意差検定をしたいためにこのような設定をしたと思われる。1 用量 2 匹にして 4 用量とすべきである。

それらの結果に、単純な平均値の比較のための有意差検定は、してはならない。ロジスティック曲線の傾きの 95% 信頼区間によって判定すべきである。

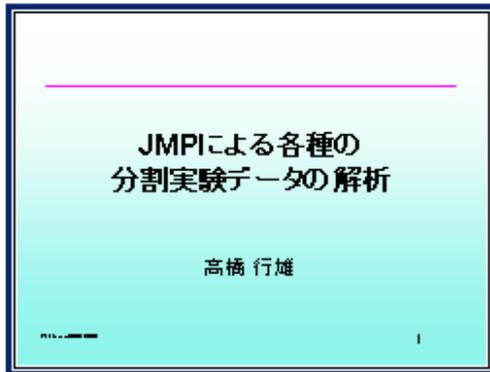
ドパミンの遊離量

この実験は、CV で 45% と大きな個体差があり、かつ用量反応も大きな個体差があるが、時点間の相関は 0.95 以上あることから、個体内では測定誤差も小さく経時的に安定した推移を示している。このことから、経時測定を 20 分起きを実施する必要はなく、薬物濃度の経時測定のような最初は密に、後は疎なサンプリングで十分である。

用量反応関係の検討も Control 群を含めて 4 用量では 4 パラメータのシグモイド曲線の当てはめは行なえない。各用量の症例数を半分にし、用量の設定範囲を広げ Control 群を含め 7 用量程度にすべきであろう。実験が成立しているかの検討には、Control 群と低用量をプールし、高用量 2 群をプールして、その平均値の差の検定をおこなうことでどうだろうか。用量反応関係は、4 パラメータのロジスティック曲線の当てはめができるようになる。

5. JMP による線形混合モデルの解析

JMP 関連の会合で紹介をしたパワーポイントを再掲します。



1

The slide has a light blue background with a white title box. The title is '目次' (Table of Contents). It lists four main items with their page numbers: '分割実験の考え方・JMPでの対応' (3), '1 度分割実験' (14), '欠測値がある分割実験の例' (42), and 'さらなる分割実験の活用に向けて' (55). A sub-item '実験配置での分割, 重欠近での分割' is listed under the last item. A small number '2' is in the bottom right corner.

目次	
◆ 分割実験の考え方・JMPでの対応	3
◆ 1 度分割実験	14
◆ 欠測値がある分割実験の例	42
◆ さらなる分割実験の活用に向けて	55
◆ 実験配置での分割, 重欠近での分割	

2

以下 PPT ファイル参照