

臨床第 1 相試験の計画と解析

第 7 回 高橋セミナー

高橋 行雄

表紙裏

改訂の記録

2001年3月14日，高橋 [BioS テキスト 1 相の試験計画と解析 2000] を元に全面改訂.

2001年4月7日，安全研，高橋セミナーのためにタイトルを一部変更，タイプミスの修正.

試験統計家のための 臨床第 1 相の計画と解析

目 次

1. はじめに	1
2. 初期の安全性試験のデザイン	5
2.1. 単回投与・群間比較	5
逐次増量・群間比較	5
試験デザインの良さ	7
試験デザインの部分的変更	8
2.2. 逐次増量・乱塊法	8
ラテン方格の良さ	9
他民族での初期用量の設定	10
部分的ラテン方格	10
同一被験者内・逐次増量 タイプ 1	11
同一被験者内・逐次増量 タイプ 2	12
交互パネル・同一被験者内・逐次増量	12
実験モデルの良さ	13
変量効果と固定効果	14
経時データの扱い	14
2.3. 連続投与試験	15
試験デザイン, 投与日数	15
バイタルサインの測定	15
トラフ値の変化	16
投与直後の変化	17
3. 生物学的同等性試験	19
3.1. 生物学的同等性試験の変遷	19
3.2. BE試験における信頼区間方式	19
3.3. 検出力方式と信頼区間方式の比較	21
3.4. クロスオーバー法の拡張	27

2×2 のラテン方格	27
2×2 のラテン方格の反復および追加.....	28
3×3 のラテン方格の応用.....	28
3×3 のラテン方格の変形.....	29
3.5. 新たな同等性の基準, Population, およびIndividual BE.....	30
3.6. 持越し効果の除去, 2×3 クロスオーバー試験	32
4. 用量反応	37
4.1. 用量反応関係の考え方.....	37
4.2. ヒトの薬物動態パラメータにおける用量反応.....	37
4.3. 1 被験者に複数の用量を投与.....	41
4.4. 冪モデルの性質.....	43
4.5. 信頼区間の活用.....	44
4.6. 平行性検定法の応用.....	45
5. 高齢者, 小児などにおける臨床薬理試験	49
5.1. 高齢者を対象にした臨床薬理試験.....	49
5.2. 小児に対する臨床薬理試験.....	49
5.3. 高齢者を対象にした薬物動態試験の事例.....	50
6. 食事の影響	53
6.1. ガイドラインでの記述.....	53
用量-反応関係の指針での記述.....	53
民族的要因ガイドラインでの記述.....	54
一般指針, 高齢者, 総括報告書ガイドラインでの記述	55
FDAのガイダンス.....	56
6.2. 2 元配置による実験.....	56
実験デザイン.....	56
効果の分離.....	57
生物学的判定.....	58
要因配置実験における用量比例性.....	60

7. 民族的要因	61
7.1. ICHガイドライン	61
7.2. 試験デザインの作成の例示	62
7.3. 民族的要因・演習 1	65
AUCの比較	65
体重によるAUCの調整	67
7.4. 民族的要因・演習 2	68
データの概要	68
AUC/doseを用いたプロット	69
用量反応	70
年齢と体重を共変量とした解析	71
7.5. データリスト	73
8. 交互パネルによるPKパラメータの解析事例	75
8.1. 実験デザイン	75
8.2. C _{max} についての用量反応	75
混合線形モデルによる用量比例性の検討	76
信頼区間方式での表示	77
9. 交互パネルによるタッピング検査の解析事例	79
9.1. 検証すべき仮説は何か，学習効果の抑制，促進？	80
9.2. タッピング数に対する薬物の影響	82

図 表 目 次

表 2.1	逐次増量・群間比較.....	5
表 2.2	欠測セルがある繰り返しが不揃いな 2 元配置.....	6
表 2.3	逐次増量・群間比較での脈拍数の経時変化（単位：回/分）.....	6
表 2.4	実施日に交絡する偏りの除去（単位：回/分）.....	7
表 2.5	実薬のみの低用量群.....	8
表 2.6	4×4 のラテン方格.....	9
表 2.7	部分的なラテン方格.....	10
表 2.8	同一被験者内・逐次増量（1）.....	11
表 2.9	同一被験者内・逐次増量（2）.....	12
表 2.10	交互パネル・同一被験者内・逐次増量.....	12
表 2.11	頑健性のチェックのためのデータ.....	13
表 2.12	連続投与試験の試験デザイン.....	15
表 2.13	血圧，および脈拍の測定ポイント.....	16
表 2.14	連投後のトラフ値の検討のための要約統計量.....	17
表 2.15	投与直後の短時間的の変化の要約統計量.....	18
表 3.1	シミュレーションデータの設定.....	21
表 3.2	シミュレーションデータの発生.....	22
表 3.3	作成された 1 組目のデータ.....	22
表 3.4	信頼区間方式 混合モデル.....	22
表 3.5	信頼区間方式 判定.....	23
表 3.6	信頼区間方式 不合格率.....	23
表 3.7	検出力方式.....	23
表 3.8	検出力方式 判定.....	23
表 3.9	検出力方式 不合格率.....	24
表 3.10	信頼区間方式と検出力方式の判定の比較.....	24
表 3.11	信頼区間と検出力方式の食違の例示.....	24
表 3.12	2×2 のラテン方格，繰返 6 の例.....	27
表 3.13	2×2 のラテン方格の反復.....	28
表 3.14	完全バランス型+3×3 ラテン方格.....	29
表 3.15	釣合い型不完備ブロック計画.....	30
表 3.16	2×4 のラテン方格法.....	31
表 3.17	順序効果.....	31
表 3.18	2×2 のラテン方格法，単純繰り返し.....	32
表 3.19	単純繰り返しの場合の順序効果.....	32

表 3.20	2×3 のクロスオーバー	32
表 3.21	2×3 の BE試験データ	33
表 3.22	2×3 のBE試験の解析プログラム	34
表 3.23	2×3 のクロスオーバー解析結果	35
表 4.1	一被験者ごとのCmaxの用量反応	41
表 4.2	AUCが 100ng・hr/mLとなるlog(dose)	47
表 5.1	トラフ値のデータ	51
表 5.2	時点間の相関構造	52
表 6.1	2 要因 2×2 クロスオーバー	56
表 6.2	AUCの結果,	57
表 6.3	投与量を共変量とした食事の影響解析プログラム	58
表 6.4	AUCでの食事の効果および差の 90%信頼区間	59
表 6.5	未変化体CについてのSASのアウトプット	59
表 7.1	人種的要因の影響を受けにくいことの根拠	61
表 7.2	民族的要因の影響を受けやすい根拠	62
表 7.3	シミュレーションDataの作成	63
表 7.4	信頼区間の計算と同等性の判定	64
表 7.5	シミュレーション 結果	64
表 7.6	健康成人における薬物動態	65
表 7.7	回帰係数の推定値	67
表 7.8	試験の概要	68
表 7.9	用量別の年齢, 体重, およびAUCの平均	68
表 7.10a	ETHNIC=EC, STUDY=1, SEX=male	73
表 8.1	交互パネルによる投与スケジュール	75
表 8.2	投与順, 投与量で整理したCmax	76
表 9.1	パネル 1 の投与後 1.5 時間目の学習効果	81
表 9.2	時点ごとのランダム係数モデルの傾きの推定値と 95%信頼区間	82
表 9.3	タッピング数の平均値データ	83
表 9.4	時点ごとの 10mg当たりのタッピング数の変化	84
表 9.5	タッピング数にたいする用量反応直線	85
表 9.6	タッピング試験・データリスト	87
図 2.1	最小二乗平均 (Lsmean) による偏りの調整	7
図 3.1	幾何平均値の差の 90%信頼区間	20
図 3.2	信頼区間方式と検出力方式との比較	21
図 4.1	用量反応が飽和している場合の回帰分析	38

図 4.2	回帰直線の切片が原点を通る	39
図 4.3	両対数での回帰直線	40
図 4.4	AUC/dose を用いた分散分析	40
図 4.5	被験者ごとの用量反応	42
図 4.6	係数 b を変えた場合の冪モデル	43
図 4.7	係数 a を変えた場合の冪モデル	43
図 4.8	AUC/Doseの 90%信頼区間の例示	44
図 4.9	生物学的同等性の判定基準の準用	45
図 4.10	冪モデルの当てはめ	45
図 4.11	複数の試験の用量反応の表示	46
図 5.1	年齢区分別の薬物濃度のトラフ値の変化	51
図 5.2	投与後 1~12 週のトラフ値の平均値 (対数) と年齢との関係	52
図 6.1	用量別の食事の効果	57
図 6.2	食事の差の 90%信頼区間の表示	58
図 6.3	変化体Cの用量反応	60
図 7.1	AUCの比較	65
図 7.2	体重とAUCの散布図および 50%確率楕円	66
図 7.3	平行な回帰直線の当てはめ	67
図 7.4	試験別用量別のAUC/doseの散布図	69
図 7.5	年齢および体重とAUC/doseの関連	70
図 7.6	両対数での回帰直線の当てはめ	70
図 7.7	試験別用量別調整済み \log (AUC/dose) の散布図	72
図 8.1	個人ごとの用量反応	76
図 8.2	単位用量当りの C_{max} , 個別データ	77
図 9.1	0mg, 40mg 投与時のタッピングの学習効果	79
図 9.2	学習効果の計量, 40mg群の 1.5 時間目	80
図 9.3	用量の増加による学習効果の変化	81
図 9.4	投与前日のタッピング回数	83
図 9.5	投与 1.5 時間後のタッピング数	84
図 9.6	10mg当たりのタッピング量の変化	86

1. はじめに

一連の ICH のガイドラインが整備された結果として、二重盲検比較臨床試験のみならず、すべての臨床試験に対して、その計画、実施、データの集積、データの解析、探索的解析、報告書の作成、規制当局との質疑応答などに、統計的な裏付けが必要とされるようになってきた。そのために、一つ一つの臨床試験に「試験統計家」の参画が必要となってきた。10 年ぐらい前迄は、実薬対照の臨床試験の症例数の計算と、有意差検定ができれば、「試験統計家」が勤めることができたのであろう。しかし、現在「試験統計家」に求められる役割は、広汎であり、また専門的でもある。

1977 年に設立されたイギリスの製薬企業統計協会 (Statisticians in the Pharmaceutical Industry) が公表している「GUIDELINES FOR STANDARD OPERATING PROCEDURES for Good Statistical Practice in Clinical Research, <http://www.psiweb.org/pdf/gsop.pdf>」に、試験統計家の果すべき役割が広範囲に詳細に示されている。また、「Careers for Statisticians within the Pharmaceutical Industry, <http://www.psiweb.org/pdf/careers-layout.pdf>」は、大学の統計学科の学生向けのパンフレットであるが、前臨床、臨床 (第 1~4 相)、製造、申請時などでの試験統計家の役割を述べている。これらを通読することにより、「試験統計家」の果すべき役割を再認識することができよう。

このテキストは、医薬品の開発のために行われている少数の健康な被験者、または少数の患者を対象にした薬物動態・薬力学的検討を行う臨床第 1 相試験に焦点を当てている。これらの試験のための実験 (試験) デザインは、治療効果を証明するための臨床試験のデザインとは異なる多様な側面を持っている。もちろん、それに伴い、様々な統計的アプローチもバラエティに富んでいる。限られた紙面では、語り尽くせないほど多岐にわたっている。さらに学習するための成書を章ごとに取り上げるので、参考にしてもらいたい。

さて、本テキストは、第 1 相試験に関連する統計的諸問題について概観する。1992 年に制定された「新医薬品の臨床評価に関する一般指針」によれば、「第 1 相試験は、治験薬を初めてヒトに適用する試験で、原則として少数の健康男子志願者において、治験薬について臨床安全用量の範囲ないし最大安全量を推定することを目的とし、あわせて吸収・排泄などの薬物動態学的検討を行ない、第 2 相試験に進み得るかの判断資料を得るための試験である。」としている。この定義は、第 1 相試験をかなり限定していた。

1997年10月には、ICHで合意された「臨床試験の一般指針、<http://www.nihs.go.jp/dig/ich/efficacy/e8/e8step4j.pdf>」では、臨床薬理試験を含む次の定義が変わった。

- a) 初期の安全性及び忍容性の推測
- b) 薬物動態
- c) 薬力学な評価
- d) 初期の薬効評価

Roodaら(1988)は、これに「生物学的利用性、あるいは生物学的同等性の試験」を含めている。Roodaらは、第1相試験での様々な試験デザインを示しているが、それらの試験デザインの統計解析については、これらの試験の目的は限定されており、観察が中心であるため、統計解析が行なわれないことが多いと述べている。ところが、現実には、臨床効果と結び付くようなデータを副次的に得る場合もある。このようなデータは、第2相以後の臨床試験と異なり、かなり細な測定ポイントが設定されている場合が多く、これらの実験デザイン画に適合した統計解析が必要である。

ヒトにおける薬理作用を検証するための、いわゆる臨床薬理試験では、薬物の特性、実験のランダム化の方式、個人内および個人間の誤差のバランスなどにより、様々な実験デザインが必要となる。これらの試験データの解析には、実験のランダム割付を前提にした統計的有意差検定の枠組みよりも、統計モデルをベースにした統計的アプローチの方が、実験全体の評価を行う上で実用的である。試験デザインと統計解析は車の両輪であり、設計上の基本的な考え方と統計解析時における注意点などについて合わせて概説する。

第1相ならびに臨床薬理試験のデータの特徴は、経時的な観察をきめ細く行うことであろう。その結果として、例えば、投与前値と各時点間に対応のある場合の t 検定、あるいはDunnettの多重比較が行われたり、各時点ごとに薬剤間で多重比較を繰り返している公表論文を目にすることが希ではない。これらは、検定の多重性の問題を回避しようとはしているが、本質的には検定の多重性の問題を回避していない。これらの経時データの解析については本テキストでは詳しく触れないので、高橋(2000)のBioSセミナーのテキストを参照してもらいたい。

生物学的同等性の試験を除いては、これらの試験の試験デザインと例数設計、あるいは統計解析については、ほとんど論じられていない。ICH - E3「治験の総括報告書の構成と内容に関するガイドライン（1995），<http://www.nihs.go.jp/dig/ich/efficacy/e6/e6.html>」が公表されて以来、生物学的同等性の試験も含め、第1相試験における統計的諸問題を取り扱う機会が増えてきた。さらに、ICH - E5「外国臨床データを受け入れる際に考慮すべき民族的要因（1998），<http://www.nihs.go.jp/dig/ich/efficacy/e5/e5-ka.pdf>」が公表・実施されたことにより、薬物動態パラメータについて民族的要因の探索、これらを踏まえたブリッジング試験のデザインなどについて、ますます試験統計家の活動が期待される場面が増えてきた。

薬物動態は、第1相での試験のみならず、対象患者集団においても評価することも規制当局から求められるようになった。これは、ICH - E7「高齢者に使用される医薬品の臨床評価に関するガイドライン（1993），<http://www.nihs.go.jp/dig/ich/efficacy/e7/e7.html>」に示されている明確な要求である。これは、老化に伴い様々な機能低下が起きるのなかで、薬物濃度が非高齢者に比べて異なるか否か、異なるならば医薬品の適性使用を促す情報として必要なためである。

ICH では、さらに「E11：小児集団における医薬品の臨床試験に関するガイダンス（2000），<http://www.nihs.go.jp/dig/ich/efficacy/e11/iyakusin1334-2.pdf>」を示した。このガイダンスでは、小児を、早産児、正期産新生児、児童、および青少年に分け、それぞれの年齢階層ごとに、肝と腎のクリアランスが異なること注意した薬物動態試験の実施を強調している。

このような臨床試験においても試験デザイン、および必要症例の算出など試験統計家の果すべき役割が存在する。

すでに述べたように、ヒトを対象とした第1相試験では、多くの制約条件があり、一般的な実験計画のデザインが採用できないことがしばしば発生する。しかしながら、その制約条件の範囲内であっても、各種の実験デザインの選択が可能である。本章では、それらのデザインの利点、および欠点について述べ、適切な試験デザインの選択法について述べる。制約条件として、

- ・ヒトに最初に投与する場合は、薬物は低い用量から高い用量へ段階的にゆかねばならない。
- ・ヒトに対する安全性が確保できている用量以下であれば、乱塊法、ラテン方格

法（クロスオーバー法）などの，釣合い型の実験モデルが適応できる．

- ・ 実験例数が小規模である．一投与量あたり 6 例程度，全体でも 30 例ぐらいが上限に近い．

これらの様々な試験デザインと統計解析に対する標準的なテキストの必要性を強く感ずるようになってきた．これまで BioS で何回か第 1 相試験の計画と解析について講義をしてきた．新しいテーマについてはテキストの参考資料として添付してきたが，テキストとしてはまとまりに欠けてきたので，再度，全面的に書き換えることにした．

文献

Rodda, B.E.他 (1988). *Biopharmaceutical Statistics for Drug Developement*, 日本語訳, 医薬統計学 (1992). サイエンティスト社, 253-303.

高橋 行雄 (2000), 試験統計家のための経時データの解析, 日本科学技術連盟 BioS テキスト.

2. 初期の安全性試験のデザイン

主に健常者を対象としている安全性と忍容性の検討を目的とした臨床試験では、多くの制約条件があり、一般的な実験計画のデザインが採用できないことがしばしば発生する。

第1は、実験計画法の基本であるランダム化が制約されることである。これは、ヒトに対する安全性および倫理上の面から、薬物の投与は低い用量から高い用量へ段階的にゆかねばならないためである。もちろんヒトに対する安全性が確保できている用量の範囲内での薬力学的試験などの場合には、乱塊法、ラテン方格法（クロスオーバー法）、あるいは釣合い型不完備ブロック計画法などの実験モデルを適応できる場合もある。

第2は、経時測定データが結果の大部分を占める。薬物投与後の経時変化を、日および時間単位で計測される血圧などのバイタルサイン、日単位で測定される臨床検査データなどの計量値が多いことである。

第3は、実験例数が小規模である。様々な制約条件、あるいは慣例により、1投与量あたり6例程度となることが多い。実験全体として30例ぐらいが上限となる。

2.1. 単回投与・群間比較

逐次増量・群間比較

ヒトへの初めて投与する薬物は、まず安全性と忍容性を確認するために、表 2.1 に示すように逐次的に投与量を増加する臨床試験が行われる。パネル1では、8人の被験者をプラセボ群に2例、5mg群に6例となるようにランダムに割り付け、実施日1で単回投与を行い、24時間ないし48時間にわたって様々な検査を一定時間ごとに行なう。

表 2.1 逐次増量・群間比較

パネル	被験者番号	実施日 1	実施日 2	実施日 3
1	1, 2	Placebo 5mg		
	3, 4, ...8			
2	9, 10		Placebo 10mg	
	11, 12, ...16			
3	17, 18			Placebo 20mg
	19, 20, ...24			

次のステップとして、異なる 8 人の被験者に対しプラセボ群に 2 例、10mg 群に 6 例となるようにランダムに割り付け、同様な手順で用量を増やして行く。この例では、最終的にプラセボ群、5mg 群、10mg 群、および 20mg 群がそれぞれ 6 例となるように計画されている。

この試験デザインで実施日を無視し、投与量で整理すると、1 元配置モデルとみなすことができる。投与量と実施日で整理すると表 2.2 に示すように、欠測セルがある繰り返しが不揃いな 2 元配置とまみなすこともできる。実際の解析では、1 元配置と見做すことができるのならば、それにこしたことはない。一般的な試験であっても、このように、これまでの古典的な実験計画法の教科書には含まれないような欠測セルがある試験デザインとなる。

表 2.2 欠測セルがある繰り返しが不揃いな 2 元配置

投与量	実施日 1	実施日 2	実施日 3	全体
Placebo 群	2 例	2 例	2 例	6 例
5mg 群	6 例			6 例
10mg 群		6 例		6 例
20mg 群			6 例	6 例
全体	8 例	8 例	8 例	24 例

この試験デザインの理解のために、ある降圧剤投与による脈拍数の変化を検討した試験から一部を抜き出したデータ表 2.3 に示す。脈拍数は、第 1 相試験の場合では、午後になると増えるという日内変動をしばしば経験する。そのためにプラセボ群との比較で、どのくらいの用量から脈拍数に変化が現れるかが検討の課題となる。これまでの動物試験から、ヒトでは、薬物濃度が 2 時間目で最大 (Tmax) になることが予測されたので、午前 8 時に薬剤の投与を開始し、脈を投与直前、2 時間後、8 時間後に測定（実際には、1 日に 10 時点前後）した結果である。

表 2.3 逐次増量・群間比較での脈拍数の経時変化（単位：回/分）

投与量	実施日 1 / パネル 1				実施日 2 / パネル 2			
	ID	0 h	2 h	8 h	ID	0 h	2 h	8 h
Placebo 群	1	65	66	74	5	60	59	68
	2	62	60	69	6	64	62	75
5mg 群	3	61	64	67	—			
	4	64	67	75	—			
10mg 群	—				7	65	78	82
	—				8	58	69	74

試験デザインの良さ

逐次増量する試験では、実施日に起因する測定値の系統的な偏りに対して頑健であることが求められる。Placebo群がな実薬のみの場合には、系統的な偏りを直接受けてしまい、安全域あるいは忍容域の判断ミスを誘発しかねない。表 2.4 に示すように実施日 2 に系統誤差 α が全ての測定値にあったとしよう。10mg群の平均値は 73.5 回/分であるが、系統誤差 α が負であれば、過少評価を呼び、系統誤差が正であれば過大評価となる。これをPlacebo群との差で評価をすれば、系統誤差 α の大小によらず差の大きさは不変である。これが、このデザインの利点となる。

表 2.4 実施日に交絡する偏りの除去 (単位: 回/分)

試験日 投与量	実施日 1 / パネル 1			実施日 2 / パネル 2			α を 5.0 とした場合		
	2 h	平均	差	2 h	平均	差	単純 平均	最小二 乗平均	差
Placebo 群	66			$59 \pm \alpha$					
	60	63.0	0.0	$62 \pm \alpha$	$60.5 \pm \alpha$	0.0	64.25	64.25	0.0
5mg 群	64								
	67	65.5	+2.5				65.50	66.75	+2.5
10mg 群				$78 \pm \alpha$					
				$69 \pm \alpha$	$73.5 \pm \alpha$	$+13.0 \pm 0$	78.50	77.25	+13.0

偏り α を 5.0 とした。SAS/GLM プロシジャで [model y = dose panel] としたときの最小二乗平均 (Lsmean) の Placebo 群との差は、パネル 2 のデータに一律に加えた α の影響を受けない。

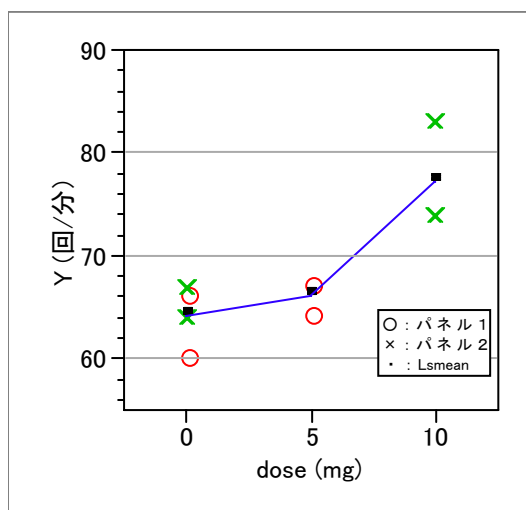


図 2.1 最小二乗平均 (Lsmean) による偏りの調整

Lsmean は、0mg のときに 4 個のデータの平均 64.25 であるが、5mg および 10mg のときは 2 個データの平均ではない。これは、0mg のときのパネル 1 とパネル 2 の 2 個の平均値からのずれ 1.25 を加減している。

ただし、この実験デザインは、Placebo 群の症例数が実薬群に比べて相対的に小さく、被験者の誤差変動の影響を受けやすいので、Placebo 群のみで実施日間に無視できない差が生じたときにのみ使う解析方法である。通常は、単純な 1 元配置モデルでの解析を基本とすべきである。

試験デザインの部分的変更

表 2.1 の試験デザインで投与量が最小の投与量 5mg 群であるが、この群では Placebo 群を除き、表 2.5 のように実薬の症例数を 3 例程度にすることもある。これは、もしも有害事象が出たとすれば、用量設定が高めであり、低用量からやり直すか、あるいは次の 10mg 投与を変更して 5mg の投与に切り替えるかの判断をすればよいだろうとの考え方である。いわば、本試験前の予備試験としての性格をもつ。

表 2.5 実薬のみの低用量群

パネル	被験者番号	実施日 1	実施日 2	実施日 3
1	1, 2, 3	5mg	Placebo 10mg	Placebo 20mg
2	4, 5 6, 7, ..., 11			
3	12, 13 14, 15, ..., 19			

逐次増量・群間比較の利点は、一人あたりの負荷が少ないことであるが、欠点は、多くの被験者数が必要なことである。統計的な欠点は、試験日に交絡する偏りの除去に制約がきついことである。また、薬力学的な特性値を副次的に得ようとする場合には、個人間誤差で用量群間の比較を行わなければならないので、1 群 6 例程度では、薬効の検出力が低く薦められない。もちろん副次的変数については予備実験として考えているのならば問題はない。

2.2. 逐次増量・乱塊法

半減期が短い医薬品の場合には、1 人に 1 週間おきに逐次増量して行くことも考えられる。これは、被験者の総数を少なくできる利点と共に、投与量群間の比較に、個人間誤差でなく、相対的に小さくなる個人内誤差を用いることによる統計的な検出力の向上に魅力がある。

ラテン方格の良さ

生物を対象にした様々な実験で、ラテン方格が持ちいられることがしばしばある。これは、実験時期に由来する偏りを統計的に除去できるためである。低用量から高用量へ逐次増量しなければならないとの制約が全くなければ、Placeboを含めて3用量の試験デザインは、表 2.6 に示す4×4のラテン方格が基本デザインである。それぞれの用量群の症例は8例と等しくなる。

表 2.6 4×4 のラテン方格

投与順	被験者番号	実施日 1	実施日 2	実施日 3	実施日 4
1	1, 2	Placebo	10mg	20mg	40mg
2	3, 4	10mg	Placebo	40mg	20mg
3	5, 6	20mg	40mg	Placebo	10mg
4	7, 8	(40mg)	20mg	10mg	Placebo

試験デザインの良さは、若干のデザインの変更によっても保てれる。例えば (40mg) の投与がためらわれるのであれば、投与しなくてもよいし、10mg でも 20mg でもよい。

ラテン方格は、実施日に由来する実験上偏りを統計的に除去できる。さらに被験者間の誤差と被験者内の誤差を統計的に分離し、投与量間の比較において被験者内の誤差を用いることができる。このことが、基本デザインであるゆえんである。

初めてヒトに投与する第1相の安全性・忍容性の試験では使えないが、民族的要因の探索のために行なう薬物動態・薬力学試験などでは、このデザインを積極的に使うべきである。実施日1での最高用量40mgの投与が、新たな民族での安全性の観点からためらわれるのであれば、デザインは一部崩れるが、実施せずに欠測としても試験全体のバランスは若干崩れてはいるが、部分的なバランスの程度のほうが多いために、試験全体の良さが保たれている。試験管理の観点から被験者を別扱いするのが望ましくないとの理由で10mgあるいは20mgを投与しても、同様に試験全体の良さは失われない。

1群当たり8人でなく10人にしたい、逆に6人に減らしたいなどの場合にも、4×4のラテン方格が使えないので被験者ごとに投与順をランダムに割付けようと思う人がいるかもしれない。これは、乱解法といわれている実験デザインであるが、このような場合でも4人単位でラテン方格に割付け、残りの被験者に4×4のラテン方格の行をランダムに選択することが望ましい。

一般線形モデルをベースにした現代の統計ソフト SAS/GLM プロシジャなどは、実験

デザインが若干崩れたとしても、試験デザインの部分的にでもバランス（直交性）を考慮した統計解析が自動的に行えるようになっている。また、実験計画的にバランスがとれているデザインを実験目的に合わせて積極的に変更した場合に、どのように一部変更したら実験期計法的には最適なのかとの研究が行なわれており、その成果が SAS/OPTEX プロシジャ、SAS/JMP の DOE などで活用できるようになっている。これらのことが、バランスが取れた実験計画を、実際の実験の制約条件に合わせて変更すること積極的に薦めている根拠である。

他民族での初期用量の設定

しばしば耳にするのは、「この医薬品は日本人では、初めての投与なので、計画した最小用量から安全性を確認しつつ逐次増量すべきである。」との主張である。場合によっては、すでに海外で市販されている場合にも、「日本では始めてなので」という主張にも遭遇する。

先行する海外の臨床試験で、80mg まで単回投与が行われ、40mg の 7 日間連投試験が行われていたとしよう。日本で民族的要因の探索のために、何 mg が単回投与の最小用量であろうか。10mg であろうか、20mg なのか、40mg でもよいのか、その科学的根拠は何か。これに答えるためには、先行する試験の AUC などの PK パラメータと有害事象の発現の関係をロジスティック回帰分析によって解析し、さらに各用量における AUC の分布状態などを把握する必要がある。その上で、民族的要因の大きさを、データに基づいて推量した上で安全係数、2 分の 1 なのか、3 分の 1 なのか、あるいは 4 分の 1 なのかを、前もって定めた上で試験デザインを考えるとよい。

部分的ラテン方格

4×4 のラテン方格は、いきなり高用量の投与を受ける被験者がいて、逐次増量の制約が無視されてしまい、第 1 相の安全性と忍容性の検討のための試験デザインとしては使えない。そこで、4×4 のラテン方格を変形して、表 2.7 に示すように、同一被験者に対して逐次投与量を増加してゆく実験モデルが考えられる。

表 2.7 部分的なラテン方格

投与順	被験者番号	実施日 1	実施日 2	実施日 3	実施日 4
1	1, 2	Placebo	10mg	20mg	40mg
2	3, 4	10mg	Placebo	20mg	40mg
3	5, 6	10mg	20mg	Placebo	40mg
4	7, 8	10mg	20mg	40mg	Placebo

この実験モデルは、一見ラテン方格のように見えるものの、実験順序のバランスが崩れているが、部分的には2×2のラテン方格が構成されている。Placebo と 10mg, Placebo と 20mg, Placebo と 40mg において部分的に 2×2 のラテン方格が互いにオーバーラップして存在している。このデザインの利点欠点は次のようである。

利点：被験者数が比較的少なくてすむ。

実施日に交絡する偏りを除去できる。

投与量間の比較に個人内誤差が使える、検出力が高くなる。

欠点：一人あたりの負荷が多い。

実施日が、実薬の用量群の数に比べ1回増える。

この例では、3用量にたいして、4実施日となる。

同一被験者内・逐次増量 タイプ 1

部分的なラテン方格スケジュールは、時期に交絡する実験の場の偏りに頑健であるが、それを若干犠牲にすることにより、同じ時期の範囲内で投与量群を 表 2.8 に示すように増やすことができる。

表 2.8 同一被験者内・逐次増量 (1)

投与順	被験者番号	実施日 1	実施日 2	実施日 3
1	1, 2, 3	Placebo	20mg	40mg
2	4, 5, 6	10mg	Placebo	40mg
3	6, 8, 9	10mg	20mg	Placebo

表 2.8 のスケジュールの難点は、Placeboの投与例数が、実薬の投与例数よりも多いことである。このデザインでPlaceboの役割は、時期に交絡する実験の場の偏りが万が一あったときのリスク回避の為でもあり、できるだけ減らしたいとの気持もある。

利点：被験者数が比較的少なくてすむ。

実施日に交絡する大きな偏りを除去できる。

欠点：一人あたりの負荷ややが多い。

個体間の誤差が、用量間の比較に含まれ、比較精度が落ちる。

プラセボ投与の例数が、実薬群より多くなる。

この例では、Placebo 群 9 例、実薬群各 6 例である。

同一被験者内・逐次増量 タイプ 2

表 2.9 のようにプラセボの投与がない被験者 7, および 8 を増加すれば例数のバランスは保てる。極論すれば、各時期にプラセボのように共通な “投与量” が入っていれば、どの試験デザインでも時期に交絡する偏りを除去できることに変わりはないのである。この点については、本節の「実験モデルの良さ」で触れる。

表 2.9 同一被験者内・逐次増量 (2)

投与順	被験者番号	実施日 1	実施日 2	実施日 3
1	1, 2	Placebo	20mg	40mg
2	3, 4	10mg	Placeob	40mg
3	5, 6	10mg	20mg	Placebo
4	7, 8	10mg	20mg	40mg

同一被験者内での増量試験の問題点は、被験者を長期間にわたって拘束しなければならない点である。安全性の確保の面から、最低 1 週間毎に行うことを考えると、最長でも時期 4 が現実的と思われる。

利点：被験者数が比較的少なくてすむ。

実薬投与の人数の設定が自由にできる。

欠点：一人あたりの負荷ややが多い。

実施日に交絡する偏りを除去しにくい。

用量間の比較に個体間の誤差が入り込み比較精度がやや落ちる。

交互パネル・同一被験者内・逐次増量

同一被験者の負担を減らしたい場合には、表 2.10 に示すように複数のパネルを構成することにより、一被験者あたりの負担を軽減できる。この方法は、2 つのパネルに交互に投与することから交互パネル (An alternative panel) といわれている。

表 2.10 交互パネル・同一被験者内・逐次増量

パネル	投与順	被験者番号	実施日 1	実施日 2	実施日 3	実施日 4
1	1	1, 2	Placbo		40mg	
	2	3, 4, 5, 6			40mg	
	3	7, 8			Placebo	
2	4	9, 10		Placebo		80mg
	5	11, 12, 13, 14				80mg
	6	15, 16				Placebo

第 1 のパネルに 10mg と投与し、安全性を確認し、次に第 2 のパネルに 20mg を投与し、次の第 1 のパネルに戻って 40mg 投与する、といった手順で逐次増量して行く。実例を、第 8 章 及び 第 9 章 で示す。

実験モデルの良さ

実験には様々な誤差が存在する。目的とする薬剤効果にそれらの影響を受けにくいことが、実験デザインの良さである。また、系統的な偏り、ある実験日に一律に測定値が一定の方向に偏ってしまうようなことがあっても、薬剤効果にそれらの影響を受けにくいことが求められる。表 2.6 のラテン方格は、これらの条件を満たす基本的な試験デザインとして知られている。

さて、表 2.9 同一被験者内・逐次増量 (2) の場合は大丈夫であろうか。実験の制約条件より、試験モデルを変形する場合には、これらの偏りの影響の程度を計測しておくといよい。実例を示そう。Placebo の母集団の反応を 105mg/dL、薬剤 10mg では、110mg/dL、薬剤 20mg では、120mg/dL、薬剤 40mg では、140mg/dL とし、4 人の被験者全てが母集団の反応になるとしよう。ただし、実施日 2 では、全ての被験者に一律に 10mg/dL 増加する偏りが発生し、さらに、投与順 2 の被験者にすべての実験日で一律に 10mg/dL 増加する偏りが発生していたとしよう。この事により、被験者 2 の実施日 2 は Placebo 投与にもかかわらずデータは 125mg/dL と高めになってしまう。

この結果として、Placebo 群の平均値は $105+20/3=111.7$ 、10mg 群は $110+10/3=113.3$ 、20mg 群は $120+30/3=130.0$ mg/dL、40mg 群は $140+10/3=143.3$ mg/dL と反応が 20mg 群から起き、40mg に増量しても反応があまり伸びないような結果となる。

表 2.11 頑健性のチェックのためのデータ

投与順	被験者	実施日 1		実施日 2 (+10)		実施日 3	
1	1	Placebo	105	20mg	120+10	40mg	140
2	2 (+10)	10mg	110+10	Placebo	105+10+10	40mg	140+10
3	3	10mg	110	20mg	120+10	Placebo	105
4	4	10mg	110	20mg	120+10	40mg	140

このデータに用量群だけでなく、投与順、および実施日を因子として含めた SAS の GLM プロシジャによる用量群の Lsmeans を計算することにより Placebo 群の Lsmean は 110.8、10mg 群は 115.8、20mg 群は 125.8、40mg 群は 145.8mg/dL と Placebo の平均より

大きくなるが、用量群間の差は母集団の平均値の差と同じになっている。これが試験デザインが、よきしない系統誤差の影響を避けられることの確認である。

試験デザインの良さを簡便的に計る統計量は、デザイン行列についての相関係数行列である。ラテン方格は、因子が互いに直交しているので、対角要素がすべて1、それ以外が0となる。ラテン方格を崩して行くと因子相互に相関が出始める。相関が小さい方が、よい試験デザインとなる。

変量効果と固定効果

用量間の比較を精度よく行うためには、一人の被験者の中で全ての用量を投与することである。これにより、用量間の比較に個人間の誤差と分離された個人内の誤差を用いることができるからである。ただし、実験結果を一般化して表現したい場合には、個人間誤差と個人内誤差を加えた誤差を用いなければならない。

一被験者に複数回の用量を投与した場合は、一般的に被験者を変量 (Random) 効果とし、投与量などを固定 (Fixed) 効果として、SAS の Mixed プロシジャなど線形混合モデルによって計算することになる。用量間の差の推定の SE は、個人内の誤差のみが用いられ、各用量の平均値の推定値の SE には、2つの誤差が合成される。

経時データの扱い

これまでBioSでの「第1相試験計画と解析」、および「経時データの解析」では、得られたデータをそのまま経時データとして取り扱うことを基本にしていた。これは、SASのMixedプロシジャの出現により様々な経時データの統計解析が可能となったことに筆者自身が興奮していたことが原因であった。しかし、時点を固定効果とした経時データの解析は、いずれにしても探索的解析としての側面を色濃く持っており、何らかの統計的な結果を得ようとするならば、投与前値を共変量とし、実験目的に添った経時データの要約統計量に対して共分散分析を行うことが望ましいと思うようになった。投与前値を共変量とする場合も、SASのMixedプロシジャなどの線形混合モデルにより簡単に扱える。これについては、第9章で事例を示す。

2.3. 連続投与試験

試験デザイン，投与日数

単回投与試験に引き続き連続投与試験が行われる。投与期間については、「旧：毒性試験法ガイドライン，医薬品製造指針 1997 年度版，薬業時報社. 172-175」で 1 ヶ月の反復投与毒性試験があれば単回または 1 週間以内の連続投与ができるとされており，7 日間の連続投与が一般的に行われている。

ただし，ここにも ICH のガイドラインにより変化が起きている。ICH-M3：医薬品の臨床試験実施のための非臨床安全性試験の実施時期についてのガイドライン（2000 年 12 月），<http://www.nihs.go.jp/dig/ich/multi/m3/jstep5.pdf>」によれば，日本では，4 週間のげっ歯類，2 週間の非げっ歯類の反復投与試験により，第 1 相～第 3 相試験で 2 週間まで投与を認めるとの変化があった。

この経過については，「2 週間の反復毒性が試験による雄性制生殖器への影響の可否についての研究（70-79），医薬品規制ハーモナイゼーション推進国際共同研究平成十年年度研究業績報告書，<http://www.nihs.go.jp/dig/ich/houkoku/index.html>」で報告されている。今後の連続投与試験の計画において，注意を払ってもらいたい。

試験デザインとしては，被験者を拘束する期間が長いこともあり，プラセボをおいた単一用量試験，あるいは表 2.12 に示すような，逐次増量試験スケジュールとなることが多い。投与の前日も含め，投与期間中に様々なバイタルサイン，臨床検査，薬物濃度などが経時的に測定される。

表 2.12 連続投与試験の試験デザイン

パネル	被験者番号	実施日 1	実施日 2
1	1, 2, 3 4, ... , 9	Placebo 10mg	
2	10, 11, 12 13, ... , 18		Placebo 20mg

バイタルサインの測定

バイタルサインを例にすると，一日一回投与の場合であれば，表 2.13 に示すように頻回な測定がなされる。これらの測定値は，薬物の種類によっては，効果を表わす場合もあれば，安全性の確認のために頻回な測定が行なわれる場合もある。1 人分の測定データにおいて試験日それぞれについて経時変化を観察していることになる。

表 2.13 血圧, および脈拍の測定ポイント

投与日	測定 時間										
	0 h	投与	1 h	2 h	3 h	4 h	5 h	6 h	8 h	10 h	12 h
前日	$x_{i,0,0}$			$x_{i,0,2}$		$x_{i,0,4}$		$x_{i,0,6}$	$x_{i,0,8}$		
1 日目	$x_{i,1,0}$	↑	$x_{i,1,1}$	$x_{i,1,2}$	$x_{i,1,3}$	$x_{i,1,4}$	$x_{i,1,5}$	$x_{i,1,6}$	$x_{i,1,8}$	$x_{i,1,10}$	$x_{i,1,12}$
2, 3 日目	$x_{i,j,0}$	↑		$x_{i,j,2}$		$x_{i,j,4}$		$x_{i,j,6}$	$x_{i,j,8}$		
4 日目	$x_{i,4,0}$	↑	$x_{i,4,1}$	$x_{i,4,2}$	$x_{i,4,3}$	$x_{i,4,4}$	$x_{i,4,5}$	$x_{i,4,6}$	$x_{i,4,8}$	$x_{i,4,10}$	$x_{i,4,12}$
5, 6 日目	$x_{i,j,0}$	↑		$x_{i,j,2}$		$x_{i,j,4}$		$x_{i,j,6}$	$x_{i,j,8}$		
7 日目	$x_{i,7,0}$	↑	$x_{i,7,1}$	$x_{i,7,2}$	$x_{i,7,3}$	$x_{i,7,4}$	$x_{i,7,5}$	$x_{i,7,6}$	$x_{i,7,8}$	$x_{i,7,10}$	$x_{i,7,12}$
8 日目	$x_{i,8,0}$										

i 番目の被験者について示した. それぞれの変数について n 人分の表となる.

このデータを経時データの解析の問題としてとらえると, 試験時期, および投与量については, 欠測セルのある 2 元配置となっている. 試験全体としては, 1 次単位が欠測セルのある 2 元配置, 2 次単位が投与日, 3 次単位が測定時間の 2 段分割実験とみなせる. ただし, このような経時データとしての解析は探索的にならざるを得ない.

これらのデータについて, 一般的なまとめ方としては, 2 パネル分のプラセボ投与群をプールしたうえで, 試験日・経過時間別に投与量ごとの平均, 標準偏差のプロット図を作成し, 第一日目の投与直前値を基準とし, 投与量ごとに対応のある t 検定をする. さらに, 試験日・経過時間ごとに, 投与量群間で対応のない t 検定, あるいは Tukey の多重比較をする, などが論文を散見する. しかし, 比較的少数のデータに対なので, 効果の見逃しがおきる. また, 数多くの検定を繰り返して行なうために, 検定の多重性, 多重比較を行なっているにもかかわらず, 試験日・経過時間ごとに繰り返すためにおきる有意水準が低下による見かけの差などに惑わされることになる.

この問題を回避するためには, このデータから検証したい主要な統計的な命題を前もって定め, 複数の副次的な命題も明確にすることである. 幾つか例示しよう.

トラフ値の変化

第 1 の命題は, いわゆるトラフ値 (trough, 地溝, 波の谷, 気圧の谷) が投与量によって増加していないことの検証である. そのために, 表 2.14 の計算表のように, 投与前日の 0h と第 1 日目の 0h の平均を, 改めて投与前 X とし, 投与第 2 日目から第 8 日目 (あるいは, 後半の 5~8 日目) の平均を改めて投与後 Y とし, 投与前値 X を共変量とし, Placebo と 10mg の Lsmean の差, Placebo と 20mg の Lsmean の差とその 95% 信頼区間によって検証するとよい. 乱暴などと思われるかもしれないが, 薬剤の投与により何らかのトラフ値の

上昇があるとするならば、それを見つけるためには統計的には精度のよい方法である。

表 2.14 連投後のトラフ値の検討のための要約統計量

		投与前値			投与後			
		前日 0h	第1日目 0h	平均： x 0h	第2日目 0h	...	第8日目 0h	平均： y 0h
Placebo	1	$x_{1,0,0}$	$x_{1,1,0}$	$x_{1,..,0}$	$y_{1,2,0}$...	$y_{1,8,0}$	$y_{1,..,0}$
	2	$x_{2,0,0}$	$x_{2,1,0}$	$x_{2,..,0}$	$y_{2,2,0}$		$y_{2,8,0}$	$y_{2,..,0}$
	:							
10mg	7	$x_{7,0,0}$	$x_{7,1,0}$	$x_{7,..,0}$	$y_{7,2,0}$		$y_{7,8,0}$	$y_{7,..,0}$
	8	$x_{8,0,0}$	$x_{8,1,0}$	$x_{8,..,0}$	$y_{8,2,0}$		$y_{8,8,0}$	$y_{8,..,0}$
	:							
20mg	13	$x_{13,0,0}$	$x_{13,1,0}$	$x_{13,..,0}$	$y_{13,2,0}$		$y_{13,8,0}$	$y_{13,..,0}$
	14	$x_{14,0,0}$	$x_{14,1,0}$	$x_{14,..,0}$	$y_{14,2,0}$		$y_{14,8,0}$	$y_{14,..,0}$
	:							

表 2.13では、第*i*番目の被験者のデータをすべて x として表したが、ここでは投与前値を x とし、投与後を y と表した。

投与直後の変化

第2命題は、投与後のバイタルサインの上昇の比較である。短時間の血圧の上昇の検討は、単回増量投与試験での命題であり、連投試験では、より低い用量で反復した場合に、統計的に検出できることになるかが課題であろう。

このためには、表 2.15 に示した計算表のように、被験者ごとに第2日目から7日目（あるいは、後半の5～8日目）までの投与直前0hの平均値を改めて投与前 x_i とし、各被験者の2h～4hの各投与日ないで平均をとり、さらに第2日目から7日目までの平均値を投与後 $y_{i..}$ として、投与前値を共変量とした解析が薦められる。統計的な帰無仮説は、単回投与で得られたバイタルサインの変化と等しいか、である。

ここに示した命題は、考え方の例示であって、どのような命題にするかは、それぞれの医薬品の性質によって異なる。

表 2.15 投与直後の短時間的の変化の要約統計量

		投与前：0h の平均			投与後：2h～4h の平均				
		第 2 日目 . . .	第 7 日目	平均 x	第 2 日目 . . .	第 7 日目	平均 y		
Placebo	1	$x_{1,2,0}$. . .	$x_{1,7,0}$	$x_{1,..,0}$	$y_{1,2,..}$. . .	$y_{1,7,..}$	$y_{1,..}$
	2	$x_{2,2,0}$		$x_{2,7,0}$	$x_{2,..,0}$	$y_{2,2,..}$		$y_{2,7,..}$	$y_{2,..}$
	:								
10mg	7	$x_{7,2,0}$		$x_{7,7,0}$	$x_{7,..,0}$	$y_{7,2,..}$		$y_{7,7,..}$	$y_{7,..}$
	8	$x_{8,2,0}$		$x_{8,7,0}$	$x_{8,..,0}$	$y_{8,2,..}$		$y_{8,7,..}$	$y_{8,..}$
20mg	:								
	13	$x_{13,2,0}$		$x_{13,7,0}$	$x_{13,..,0}$	$y_{13,2,..}$		$y_{13,7,..}$	$y_{13,..}$
	14	$x_{14,2,0}$		$x_{14,7,0}$	$x_{14,..,0}$	$y_{14,2,..}$		$y_{14,7,..}$	$y_{14,..}$
	:								

表 2.13では、第*i*番目の被験者のデータをすべて*x*として表したが、
ここでは投与前値を x とし、投与後を y と表した。

文献

Milliken, G. A. and Johnson, D.E.(1997). Analysis of Messy Data, Chapman & Hall.

Rodda, B.E.他 (1988). Biopharmaceutical Statistics for Drug Development, 日本語訳, 医薬統計学 (1992). サイエンティスト社, 253-303.

松山裕, 山口拓洋 (2001), 医学統計学のための線形混合モデル, サイエンティスト社.

高橋行雄 (1996)、各種の分割実験および経時測定データの解析、SAS ユーザ会 '96 論文集 (263-286).

高橋行雄 (1997) , 臨床第 1 相試験の計画と解析モデル, 日本計量生物学会・日本応用統計学会抄録集.

3. 生物学的同等性試験

3.1. 生物学的同等性試験の変遷

生物学的同等性 (BE) 試験に例数設計の概念が盛り込まれたのは、緒方、鮫島 編 (1992) によれば、1980年5月30日付けの薬審718号「医薬品の製造又は輸入の承認申請に際し添付すべき資料の取扱い等について」の中の「生物学的同等性に関する試験基準」であった。「旧：臨床試験の統計解析に関するガイドライン (1992), 旧：統計ガイドライン」の出る10年以上前であり、例数設計に関して先駆的な役割を果していたといえよ。1980年の2つの製剤を生物学的に同等というための基準は、検出力アプローチといわれているものである。

この基準は、1997年に「後発品の生物学的同等性試験ガイドライン」によって変更となった。この方法はUSAの「*Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design*, (1992)」で採用された2つの片側検定を用いた信頼区間方式である。USAでは、2001年1月に「*Statistical Approches to Establishing Bioequivalence*,」が発行され、これまでの規準、信頼区間方式 (Average BE) に加えて Population BE および Individual BE が新たに追加されている。

臨床試験の様々な場面において血中の薬物濃度測定が行われており、その統計的な評価を行う上で、この信頼区間方式は示唆に富む。旧：統計ガイドラインの影響により、臨床試験の主要評価変数に対する設定根拠のみならず、その中で薬物濃度測定を行う症例数についても、規制当局から「後づけもでもよいから症例数の根拠」が求められる経験をしてきた。最近では、設定根拠を試験計画書に含めるようにしている。

薬物濃度をめぐる統計解析は、差を検出するよりも同等であることを示す場合が比較的多く、BE試験で用いられている信頼区間方式が、例数設計、あるいは結果の表示法として応用範囲が広い。

3.2. BE試験における信頼区間方式

BE試験ではPKパラメータ、主にAUCおよびCmaxの標準製剤と試験薬の平均値 (幾何平均) の差の90%信頼区間が、図3.1に示すように標準薬の平均 (幾何平均) を100%としたときに、80%から125%の範囲に入っているときに、生物学的に同等と判断する

ことになった。

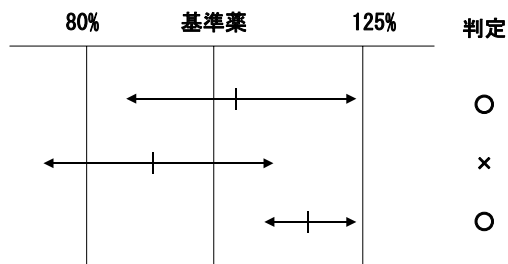


図 3.1 幾何平均値の差の 90%信頼区間

従来は、標準製剤の PK パラメータの推定値、個人内の分散の推定値を用いて、試験薬 R が標準薬 T の推定値の±20%離れたときの検出力が 80%となる症例数としていたので、その症例数は、理論的な計算により求めることが出来る。

検出力方式による判定基準

基準 PK パラメータ : AUC, Cmax

最小検出差 : 20%

有意水準 : 0.10 ~ 0.05

検出力 : 0.80 以上

現在、日本でも 1997 年にガイドラインの改訂が行われたが、これまでの基準で用いられている、「検出力+有意差検定+信頼区間」を組合せた方式に比べ、信頼区間方式は、すっかりとした合理的な判定方式となっている。下図に両方式のによる同等性の判定例の相違について示す。

信頼区間方式の場合の例数設定に、検出力方式を準用した場合の問題点を示す。検出力方式では、標準薬 R と試験薬 T の (平均値の差/差の SE) の有意差検定の結果により検出力を計算しているが、信頼区間方式では、試験薬 T の平均値±差の 90%信頼幅が、標準薬 R の平均値 (幾何平均) の 80%から 125%の範囲に入っていることを基準にしている。

信頼区間方式による判定基準

PK パラメータ : AUC, Cmax

標準製剤と試験薬の幾何平均値の差の 90%信頼区間

標準製剤の 80%と 125%の範囲

そのため、図 3.2 に示すように例 4 の場合、差の 90%信頼幅が大きく、検出力方式では有意差が出ないために同等と判定がされ場合でも、信頼区間方式では（試験薬平均値±差の 90%信頼幅）が基準範囲からはみ出てしまい同等ではないと判定がされる場合がある。また、例 3 の様に上限が 120%を越え、判定が逆となる場合も有り、検出力方式を準用することは問題である。このため、90%信頼区間が 80%から 125%に含まれる確率を高めるように、目標例数を設定するのが適切と考えられる。

図 3.2 信頼区間方式と検出力方式との比較

	80%	100%	125%	信頼区間方式	検出力方式
例 1				○	○
例 2				×	×
例 3				○	×
例 4				×	○
例 5				○	○
例 6				×	×

3.3. 検出力方式と信頼区間方式の比較

信頼区間方式と検出力方式の判定基準の違いを詳しく調べるために、シミュレーションを行い、その結果より、両方式の性能評価を行う。参照薬をRとし、試験薬をTとする。それぞれのAUCは対数正規分布に従うものとし、表 3.1 に示すように常用対数で平均値がそれぞれ 2.0、個人間のSDは 0.0791、個人内のSDは 0.0607 とした。

表 3.1 シミュレーションデータの設定

```

data W01 ;
  NEX    = 100 ;          /* Number of Iteration */
  seed   = 970903 ;      /* change the seed like 0 */
  NSAMP  = 10 ;          /* Total subject number */
  LMEAN  = log10(100) ;  /* mean of AUC */
  LMEAN_R= log10(100) ;  /* mean of AUC */
  LS1    = log10(120) - LMEAN ; /* Inter Subject */
  LS2    = log10(115) - LMEAN ; /* Intra Subject */
    
```

100 組, n1=n2=10, 平均値 : log10 (100) = 2.00

個人間 SD =log10 (120) -2.0=0.0792, 個人内 SD = 0.0607

シミュレーションデータは、判定方式の違いの理解を主体としているので、1 群 10 症例、100 組のデータを発生させた。追試の便宜を考慮してして、乱数の seed を定数で与えた。

表 3.2 シミュレーションデータの発生

```

NDRUG = 2 ;
do EX = 1 to NEX ;
  do SUBJ = 1 to NSAMP ;
    LSD1 = LS1*rannor(SEED) ;
    do DRUG = 1 to NDRUG ;
      if DRUG=1 then LM = LMEAN ;
        else LM = LMEAN_R ;
      TIME = DRUG ;
      if (SUBJ>NSAMP/2) then
        if DRUG=1 then TIME=2 ; else TIME=1 ;
      LSD2 = LS2*rannor(SEED) ;
      LAUC = LM + LSD1 + LSD2 ;
      AUC = 10**LAUC ; output ;
    end ; end ; end ;

```

信頼区間方式のために LAUC を検出力方式のために、 $AUC=10**LAUC$ を計算している。

表 3.3 作成された 1 組目のデータ

OBS	EX	SUBJ	LSD1	DRUG	LM	TIME	LSD2	LAUC	AUC
1	1	1	0.062	1	2	1	0.052	2.115	130.37
2	1	1	0.062	2	2	2	-0.011	2.050	112.45
3	1	2	0.141	1	2	1	-0.058	2.083	121.13
4	1	2	0.141	2	2	2	-0.032	2.109	128.56
⋮									
19	1	10	-0.003	1	2	2	-0.104	1.892	78.00
20	1	10	-0.003	2	2	1	-0.006	1.989	97.70

LSD1 は個人間の SD で、同じ SUBJ 内では同じ値になっている。

OBS=1 の $LAUC=2.000+0.062+0.052=2.115$

OBS=2 の $LAUC=2.000+0.062 - 0.011=2.050$

表 3.4 信頼区間方式 混合モデル

```

proc mixed data=W01 ;
  class DRUG TIME SUBJ ;
  model LAUC = DRUG TIME ;
  random SUBJ ;
  lsmeans DRUG / cl alpha=0.10
            diff=control('1') ;
  make 'lsmeans' out=LSMEAN1 ;
  make 'diffs' out=DIFF1 ;
  by EX ;

```

model ステートメントの y は $\log_{10}(AUC)$

lsmean ステートメントにより、差の 90%信頼区間の出力を指定している

表 3.5 信頼区間方式 判定

EX	対数平均 _LSMEAN_	差 _DIFF_	下限% P-LOW1	上限% P-UPP1	判定 JUDGE1
1	1.959	0.043	99.7	122.4	0
2	1.947	-0.003	87.3	113.2	0
3	1.970	0.052	98.9	128.5	X
4	1.995	0.027	97.4	116.4	0
5	1.962	0.019	94.2	116.0	0
:					

EX=3 は、上限の 90%信頼区間が 128.5%と基準 125%を越えている。
プログラムはここでは省略しているが、判定結果が 0 と X で示されている。

表 3.6 信頼区間方式 不合格率

JUDGE1	Frequency	Percent	Cumulative Percent
0	81	81.0%	81.0%
x	19	19.0%	100.0%

100 組のデータ中、19 組が不合格となった。これは、個人内の SD が大きいためである。
個人内の SD が小さければ、合格数が増え、個人内の SD が大きければ、不合格数が増える。

表 3.7 検出力方式

```
proc mixed data=W01 ;
  class DRUG TIME SUBJ ;
  model AUC = DRUG TIME ;
  random SUBJ ;
  lsmeans DRUG / cl alpha=0.10
              diff=control('1') ;
  make 'lsmeans' out=LSMEAN2 ;
  make 'diffs' out=DIFF2 ;
  by EX ;
```

model ステートメントの y は AUC となっている。
検出力の計算は、Data ステップで別途計算している。

表 3.8 検出力方式 判定

EX	LSMEAN	DIFF	20%以下 PERCENT	有意差 _PT_	検出力 POWER	範囲 J-RANGE	判定 JUDGE2
1	93.2	9.6	10.3	0.093 *	0.958	*	X
2	90.6	-1.1	-1.3	0.865	0.819	0	0
3	95.7	15.9	16.6	0.095 *	0.667 *	*	X
7	108.2	-9.5	-8.8	0.194	0.901	*	0
8	110.4	-13.3	-12.1	0.008 *	1.000	0	0
14	112.3	-17.2	-15.3	0.002 *	1.000	*	X
19	106.2	23.2	21.9 *	0.001 *	0.997	*	X

*印がそれぞれの判定基準で不合格。EX=19 は T 薬が 23.2%で *、有意水準は $\alpha=0.10$ 、
EX=8 は有意な差であるが 90%信頼区間が $\pm 20\%$ 内なので総合判定 JUDGE2 が 0 と合格となっている。

表 3.9 検出力方式 不合格率

JUDGE2	Frequency	Percent	Cumulative Percent
0	72	72.0%	72.0%
X	28	28.0%	100.0%

100組のデータ中、28組が不合格となった。信頼区間方式と同じデータを用いているので、判定基準として厳しい方式だった。いいかえれば、このシミュレーションの場合は、信頼区間方式は判定基準が弱まったといえる。

表 3.10 信頼区間方式と検出力方式の判定の比較

		検出力方式 JUDGE2		Total
		0	X	
信頼区間方式 JUDGE1	0	69	12	81
	X	3	16	19
Total		72	28	100

2つの判定を比べた場合に、検出力方式で不合格で、信頼区間方式で合格となるは100組中12組あり、主な原因は、検出力不足による不合格例を救うことであった。

表 3.11 信頼区間と検出力方式の食違の例示

EX	信頼区間	検出力方式	食違	80	100	125%
1	0	X	X			
2	0	0			<----->	
3	X	X			<----->	----->
54	X	0	X		<----->	----->
74	X	0	X	<----->		
80	0	X	X		<----->	
90	0	X	X	<----->		
95	X	0	X	<----->		

信頼区間でX、検出力で合格なのは EX=54, EX=74, EX=95 の3組である。

信頼区間で合格、検出力方式でXは12組あるが、代表的なEX=80とEX=90を示した。

EX=1の場合は検出力は足りているが、

有意差があり、かつ90%信頼区間が±20%内にないので不合格である。

簡単なシミュレーションではあるが、信頼区間方式は、従来の検出力方式に比べて、BEの基準を満たす場合が多くなる。いいかえれば、基準が緩やかになったのである。

BE 試験シミュレーションプログラムリスト

```

Title1 'BES02G2.sas 1997-9-3, 2001-2-12 Y.Takahashi ' ;
Title2 '<<< BE / Power Method vs Confidence Limit >>>' ;

Title3 '==== Data Set =====' ;
data W01 ;
  NEX      = 100 ;          /* Number of Iteration */
  seed     = 970903 ;      /* change the seed like 0 */
  NSAMP    = 10 ;          /* Total subject number */
  LMEAN    = log10(100) ;   /* mean of AUC */
  LMEAN_R  = log10(100) ;   /* mean of AUC */
  LS1      = log10(120) - LMEAN ; /* Inter Subject */
  LS2      = log10(115) - LMEAN ; /* Intra Subject */
  NDRUG    = 2 ;
  do EX = 1 to NEX ;
    do SUBJ = 1 to NSAMP ;
      LSD1 = LS1*rannor(SEED) ;
      do DRUG = 1 to NDRUG ;
        if DRUG=1 then LM = LMEAN ;
          else LM = LMEAN_R ;
        TIME = DRUG ;
        if (SUBJ>NSAMP/2) then if DRUG=1 then TIME=2 ;
          else TIME=1 ;
        LSD2 = LS2*rannor(SEED) ;
        LAUC = LM + LSD1 + LSD2 ;
        AUC = 10**LAUC ;
        output ;
      end ;
    end ;
  end ;

Title3 '==== Confidence Limit =====' ;
proc mixed data=W01 noitprint noprofile noclprint ;
  class DRUG TIME SUBJ ;
  model LAUC = DRUG TIME ;
  random SUBJ ;
  lsmeans DRUG / cl diff=control('1') alpha=0.10 ;
  make 'lsmeans' out=LSMEAN1 ;
  make 'diffs' out=DIFF1 ;
  by EX ;

data M1 ;
  set LSMEAN1 ;
  if DRUG = '1' then output ;
  keep EX _LSMEAN_ ;

data D1 ; set DIFF1 ; drop DRUG _DRUG ;

data J1 ;
  merge M1 D1 ;
  P_LOW1 = 10**((_LSMEAN_ + _LOWER_) / 10**_LSMEAN_ * 100.0 ;
  P_UPP1 = 10**((_LSMEAN_ + _UPPER_) / 10**_LSMEAN_ * 100.0 ;
  by EX ;
  if P_LOW1 <= 80 or P_UPP1 >= 125 then JUDGE1 = 1 ;
  else JUDGE1 = 0 ;

proc print data=J1 ;
  where EX <= 40 ;
  var EX _LSMEAN_ _DIFF_ _SE_ _ALPHA_
    _PT_ _LOWER_ _UPPER_ P_LOW1 P_UPP1 JUDGE1 ;
  format _LSMEAN_ _DIFF_ _SE_ _LOWER_ _UPPER_ 6.3 ;
  format _T_ _PT_ P_LOW1 P_UPP1 6.3 ;
  format P_LOW1 P_UPP1 6.1 ;

proc freq data=J1 ; table JUDGE1 ; run ;

Title3 '==== Power Method =====' ;
proc mixed data=W01 noclprint noitprint noprofile ;
  class DRUG TIME SUBJ ;
  model AUC =DRUG TIME ;
  random SUBJ ;
  lsmeans DRUG / cl diff=control('1') alpha=0.10 ;
  make 'lsmeans' out=LSMEAN2 ;

```

```

make      'diffs'      out=DIFF2      ;
by        EX ;

data M2 ;
set      LSMEAN2 ;
if      DRUG = '1' then output ;
keep    EX _LSMEAN_ ;

data D2 ; set DIFF2 ; drop DRUG _DRUG ;

data J2 ;
merge M2 D2 ;
PERCENT= _DIFF_ / _LSMEAN_ * 100.0 ;
DELTA = _LSMEAN_ * 0.20 ;
LAMDA = DELTA / _SE_ ;
T0 = tinv(1-_ALPHA_/2, _DF_, 0) ;
Q1 = probt(-T0, _DF_, LAMDA) ;
Q2 = 1 - probt( T0, _DF_, LAMDA) ;
POWER = Q1 + Q2 ;
P_LOW2 = (_LSMEAN_ + _LOWER_) / _LSMEAN_ * 100.0 ;
P_UPP2 = (_LSMEAN_ + _UPPER_) / _LSMEAN_ * 100.0 ;
if abs(PERCENT) < 20.0 then J_PER = 0 ; else J_PER = 1 ;
if _PT_ > 0.10 then J_P = 0 ; else J_P = 1 ;
if POWER > 0.80 then J_POWER = 0 ; else J_POWER = 1 ;
if P_LOW2 >= 80 and P_UPP2 <= 120
then J_RANGE = 0 ; else J_RANGE = 1 ;
JUDGE2 = J_PER + J_P + J_POWER ;
if JUDGE2 >= 2 then JUDGE2 = 1 ;
if J_PER = 0 and
   J_POWER = 0 and
   J_P = 1
then JUDGE2 = J_RANGE ;
by EX ;

proc print data=J2 ;
where EX <= 40 ;
var EX _LSMEAN_ _DIFF_ _SE_ _ALPHA_
PERCENT _PT_ POWER DELTA P_LOW2 P_UPP2
J_PER J_P J_POWER J_RANGE JUDGE2 ;
format _LSMEAN_ DELTA _DIFF_ _SE_ PERCENT _LOWER_ _UPPER_ 6.2 ;
format LAMDA 6.2 ;
format _T_ _PT_ POWER 6.3 ;
format P_LOW2 P_UPP2 6.1 ;

proc tabulate data=J2 format=5.0 ;
class J_PER J_P J_POWER J_RANGE JUDGE2 ;
table J_PER*J_P*J_POWER all, J_RANGE JUDGE2 all / rts=25 ;
run ;

Title3 '==== Power Method vs Confidence Limit =====' ;
data J3 ;
merge J1 J2 ;
if JUDGE1 = JUDGE2 then JUDEG=' ' ; else JUDGE='X' ;
by EX ;

proc freq data=J3 ; table JUDGE1*JUDGE2 ;run ;

proc print data=J3 ;
where JUDGE='X' ;
var EX P_LOW1 P_UPP1 JUDGE1
P_LOW2 P_UPP2 JUDGE2 JUDGE ;
format P_LOW1 P_UPP1 6.1 ;
format P_LOW2 P_UPP2 6.1 ;
run ;

options ps=50 ls=66 ;
proc timeplot data=J3 ;
id EX JUDGE1 JUDGE2 JUDGE ;
plot P_LOW1='<' P_UPP1='>'
/ npp overlay hiloc ref=80 100 120 125
axis=70 to 140 by 2 ;
run ;

```

3.4. クロスオーバー法の拡張

ラテン方格の列方向に投与時期を、行方向に被験者を割付けた場合に、同一の被験者の薬剤が交叉（交差と同義）する様子から交叉（クロスオーバー）試験といわれている。生物学的同等性試験（BE 試験）のように、薬物の投与順序に制約が無い場合には、ラテン方格が多用される。代表的なのは、2 薬剤の場合の 2×2 のラテン方格である。

従来から用いている R 製剤を、賦型剤などを変更して新しい T 製剤を開発した時に、R 製剤と T 製剤を実際に被験者に投与し、実際の血中薬物濃度から計算される薬物動態パラメータについて R 製剤と T 製剤を比較するのが BE 試験である。

2×2 のラテン方格

2×2 のラテン方格の行方向に被験者を列方向に投与時期を割付け、2 つの製剤をクロスして割付ける。行方向に 6 人を割付けたのが表 3.12 である。この実験モデルは、個人間の誤差を精度よく抽出する方法として優れている。

表 3.12 2×2 のラテン方格，繰返 6 の例

投与順	被験者	実施日 1	実施日 2
1	1, 2, ..., 6	T	R
2	7, 8, ..., 12	R	T

2×2 のラテン方格は、一般論として順序効果（持ち越し効果ともいう）が存在したときに脆弱である。BE 試験では、R 製剤も T 製剤も有効成分は同じなので順序効果が出たために失敗することは考えにくいですが、患者を対象にした臨床試験では、注意を要する。

極端な例だが、実薬の後の Placebo の効果は、実薬の効果がよかった場合にその効果が Placebo に交叉した後も持続するような場合である。Placebo の投与が最初の場合の場合と比較して、このように投与の順による影響が出ること順序効果といわれている。順序効果の可能性が疑われるある場合には、クロスオーバー試験は避けるべきである。

これは、順序効果を統計的に分離できないことと、もしも、存在が実験結果から疑われた場合は、クロスオーバー法を前提とした例数設計が崩れ、統計的な検出力が低下することになり、実験そのものが失敗する危険性が大きいからである。

2×2 のラテン方格の反復および追加

実験の場の制約により 1 度に多人数の被験者を捌ききれない場合も生じる。主に採血に関連するのである。試験計画で、30 分目に採血する場合に、5 分間隔で薬剤を投与したとすると、採血者が 1 人の場合には、30 分目の最後の被験者が 1 時間目の最初の被験者が重ならないようにするためには、 $30/5 = 15$ 人が上限である。

このような場合に、表 3.13 のように実施日 1 と実施日 3 において 2×2 のラテン方格で実験をし、実施日 2 と実施日 4 で別の 2×2 のラテン方格を組むとよい。解析に際しては、「反復」を主効果とし、解析モデルに含め、実施日 1 と実施日 2 は投与 1 回目、実施日 3 と実施日 4 は、投与 2 回目としてもよい。また、反復を無視して解析をしても、製剤間の差の比較のための個人内誤差は同じである。

表 3.13 2×2 のラテン方格の反復

パネル	投与順	被験者	反復 1	反復 2	反復 1	反復 2
			実施日 1	実施日 2	実施日 3	実施日 4
1	1	1, 2, 3	T		R	
	2	4, 5, 6	R		T	
2	1	7, 8, 9		T		R
	2	10, 11, 12		R		T

分散の過小評価などで、2×2 のラテン方格の BE 試験で例数不足が起きたときに、例数を追加して再度 2×2 のラテン方格による試験を行い、2 つの試験を併合することが、日本のガイドラインでは認められている。例数設計時に、分散を推定するために十分なデータがない場合には、あらかじめ 2 回目の追加実験の計画を含めておくるとよい。ただし、追加実験の例数は、試験計画書で、1 回目の試験で得られて分散も用いて行うようにすべきである。

3×3 のラテン方格の応用

異なる用量を含む製剤を新たに製剤開発したときに、従来の R 製剤との比較を行いたい。さらに、異なる用量間でも全体の用量を揃えて比較する必要があったとしよう。このような場合に、1 人の被験者に 3 種の薬剤を順番に投与する 3×3 のラテン方格が使える。

表 3.14 のパネル 1b が基本的な 3×3 のラテン方格の割付である。順序効果の観点で見ると、R 製剤の後は T1 製剤しかなく、T1 製剤の後は T2、T2 製剤の後は R 製剤となり、

その逆がない。BE試験では、同等であるとの生物学的な判定をするために、個人内誤差が小さい場合であっても、経験的に被験者が10人ぐらいにない。9人と例数設計でなった場合には3製剤の全ての異なる投与順、 $3! = 6$ 通りをパネル1のように6人に割付け、残りの3人を 3×3 のラテン方格に割付けるとよい。

これにより、R製剤の後のT1製剤は、ID-1, ID-5, T1製剤の後のR製剤はID-3, ID-6と6人の部分でバランスが取れている。他の全ての投与順にも隣り合う2実施日間でバランスが取れている。

完全バランス型の試験デザインは、順序効果があったとしても統計的にその効果を取り出して、製剤間の比較可能性を保つことができる。これについては、本節の「持越し効果」を参照のこと。

表 3.14 完全バランス型+ 3×3 ラテン方格

パネル	投与順	被験者	実施日 1	実施日 2	実施日 3
1a	1	1	R	T1	T2
	2	2	R	T2	T1
	3	3	T1	R	T2
	4	5	T1	T2	R
	5	5	T2	R	T1
	6	6	T2	T1	R
1b	1	7	R	T1	T2
	5	8	T2	R	T1
	4	9	T1	T2	R

3×3 のラテン方格の変形

比較したい製剤は3種類であるが、被験者の拘束時間を短くしたい場合には、ラテン方格の最後の実施日を落としても、部分的なバランスが保てられている試験デザインが考えられる。表 3.14 の実施日3を落としたのが表 3.15 である。R製剤、T1製剤、およびT2製剤の症例数は、全て6例と一致し、実施日後とにみても、R製剤、T1製剤、およびT2製剤がそれぞれ3例とバランスが取れている。一見すると延べ実験例数が減らせるような気がするが、実験計画上、それぞれの製剤で9症例を確保する必要があるとした場合には、1bのパネルを3例から6例に増やす必要性がある。

表 3.15 釣合い型不完備ブロック計画

パネル	投与順	被験者	実施日 1	実施日 2
1a	1	1	R	T1
	2	2	R	T2
	3	3	T1	R
	4	4	T1	T2
	5	5	T2	R
	6	6	T2	T1
1b	1	7	R	T1
	5	8	T2	R
	4	9	T1	T2

3.5. 新たな同等性の基準, Population, および Individual BE

生物学的同等であるとの基準は、時代と共に変化してきた。ある基準が制定されると、不都合な事例が提起され、それを解決しようと新たな基準作りが行われてきた。日本では、1997年に「後発品の生物学的同等性試験ガイドライン」によって定められた信頼区間方式が、最新の基準である。この規準は、USAでは1992年より *Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design*, を踏襲したものであるが、USAでは、2001年1月に「*Statistical Approches to Establishing Bioequivalence*, <http://www.fda.gov/cder/guidance/3616fnl.htm>」が発行され、これまでの規準、Average BEに加えて Population BE および Individual BE が新たに追加された。

Population BE は、処方適合性 (Prescribability) を保証する基準である。これは、初めて医薬品が処方される患者に対して期待される AUC の平均値と分散が、製剤 R と製剤 T が同程度であることを保証しようとする基準である。この基準は、本質的に大きい個人内の分散のために、従来の基準 Average BE では、製剤 T の BE を保証するために例数設計上、実施不可能なような例数となるような場合の新たな基準である。Population BE は、製剤 R の個人内分散と個人間分散を合わせた分散が CV で約 20% 以上の場合に、動的に「80%と 125%の範囲内」の基準を広げる基準ともいえる。

Individual BE は、互換性 (Swichability) を保証する基準である。これは、すでに製剤 R が投与されている患者に、製剤 T に処方変更できることを保証しようとする基準である。この基準は、製剤 R の個人内分散と被験者間の交互作用を考慮して、「80%と 125%の範囲内」の基準を広める方向と狭める方向を併せ持った基準ともいえる。

このための試験デザインとして、2×2 のクロスオーバーの拡張が提言されている。

なお、Average BE は、同一製剤内による変動を考慮していない、さらに、製剤と被験者の交互作用の変動を考慮していない。

これは、本質的に変動の大きい製剤 R については、例数設計上 Average BE では、現実的でない被験者数、例えば 100 例以上など、が必要となり、新たな製剤 T の開発を断念するようなことが起きる。Population BE および Individual BE は、このような場合での BE 試験の新たな規準となっている。これらの詳細については、Chow S.C. and Liu J(2000), *Design and Analysis of Bioavailability and Bioequivalence Study. 2nd edition* の 14～16 章を参考にしてもらいたい。

2×2 のクロスオーバーの拡張

同じ製剤内での PK パラメータの変動が大きい場合には、多くの被験者を必要とする場合がある。その場合に、被験者を増やすよりも同じ被験者に繰り返し投与を行うことも考えられる。それでも同様の結果を得ることができる。投与順によっては、持越し効果があったとしても、表 3.16 のように“TT”および“RR”の投与順が含まれるのでそれを統計的に取り除くことができる。

表 3.16 2×4 のラテン方格法

投与順	反復 1		反復 2	
	実施日 1	実施日 2	実施日 3	実施日 4
1	T	R	R	T
2	R	T	T	R

表 3.17 順序効果

投与順	実施日	前薬	後薬
1	1, 2	T	R
2	3, 4	T	R
2	2, 3	T	T
1	2, 3	R	R
1	3, 4	R	T
2	1, 2	R	T

2×2 のラテン方格法を単純繰りに繰り返した場合は、投与順が“TR”および“RT”のりような 2 通り場合のみになり、持越し効果を統計的に取り除くことができない。

表 3.18 2×2 のラテン方格法, 単純繰り返し

投与順	反復 1		反復 2	
	実施日 1	実施日 2	実施日 3	実施日 4
1	T	R	T	R
2	R	T	R	T

表 3.19 単純繰り返しの場合の順序効果

投与順	実施日	前薬	後薬
1	1, 2	T	R
1	3, 4	T	R
2	2, 3	T	R
1	2, 3	R	T
2	1, 2	R	T
2	3, 4	R	T

表 3.20 のように 2×4 のデザインの最後の実施日 4 を除いた 2×3 の場合でも持越し効果があつたとしても, 統計的に取り除くことができる.

表 3.20 2×3 のクロスオーバー

実施順	反復 1		反復 2
	実施日 1	実施日 2	実施日 3
1	T	R	R
2	R	T	T

3.6. 持越し効果の除去, 2×3 クロスオーバー試験

クロスオーバー試験は, 臨床薬理試験で頻繁に用いられる試験デザインである. また, 生物学的同等性試験 (BE 試験) においては, 標準的に用いられている. これは, 個人間のバラツキよりも個人内のバラツキが, 経験的に小さい事が知られており, 同等性の基準をクリアーするためにも, 同一被験者に比較したい薬剤をすべて投与し, 薬剤間の差の比較に個人内誤差を使用したいためである.

クロスオーバー試験も, 基本的には, 時間経過を含む解析と同様に, 個人間誤差, 個人内誤差など複数の誤差をもつ, 解析モデルである. BE 試験では, AUC または, C_{max} などの代表特性値を用いるが, 同一人内で何らかの異なる因子, ここでは製剤の種類, が割り付けられていることに変わりがない.

MIXED プロシジャを用いることにより, 様々な試験デザイン, 欠測セルが出てしまっ

た場合など、多彩な問題に対して手軽に解析できる。ここに示すのは、個人内の AUC にバラツキが比較的大きい場合、持ち越し効果を確認したい場合などに用いられる 2x2 クロスオーバーを拡張した試験デザインである。

同じ人に同じ製剤を続けて実施日 3 に投与しているのが無駄のように思われるかもしれない。しかし、これは、持ち越し効果を分離するための優れたデザインなのである。表示 2.8 に例を示し、解析手順と結果の解釈をしてみよう。

表 3.21 2×3 の BE 試験データ

投与順	ID	実施日 1		実施日 2		実施日 3	
1:TRR	2	T	32.14	R	42.55	R	36.08
	5	T	51.85	R	47.37	R	56.50
	8	T	27.48	R	27.87	R	33.54
	10	T	17.32	R	18.93	R	19.05
	11	T	31.61	R	24.93	R	28.21
	13	T	39.26	R	37.36	R	45.09
	17	T	47.55	R	40.00	R	36.00
	18	T	36.15	R	35.20	R	33.26
	2:RTT	1	R	26.79	T	27.23	T
3		R	26.71	T	30.81	T	26.39
4		R	28.70	T	40.74	T	30.49
7		R	35.01	T	45.85	T	43.86
9		R	30.98	T	34.57	T	30.18
12		R	57.70	T	54.39	T	44.35
14		R	27.87	T	30.02	T	27.00
15		R	31.03	T	25.77	T	29.65
16		R	27.67	T	21.61	T	25.64

Chow & Liu (2000) Table 9.3.3

前節と同じように、MIXEDプロシジャの出力にそって結果を見てゆこう。DATAセットを作成する際に表示 2.8 に含まれない持ち越し効果の抽出のために表 3.21 に示すように、CX, CT, CRといったデータを含ませている。持ち越し効果というのは、前に投与されている薬剤の影響が、当該の薬剤の効果に何らかの影響、あるいは効果を与えていることを言う。BE試験の場合は、あまり考えられないのであるが、異なる薬剤の場合には、希に確認されていて、無視はできないのである。実施日 2 のCTは、実施日 1 のT薬のCarry Over効果という意味で、実施日 3 のCRは、実施日 2 のR薬のCarry Over効果という意味である。実施日 1 のCXは、前に投与した薬剤がないので、SQN=1、でもSQN=2でもCXとしている。MIXEDプロシジャの使い方は、表 3.22 に示すごとくである。

表 3.22 2×3 の BE 試験の解析プログラム

```

Title1 'BES04C.sas 1996-7-5 Y.Takahashi ' ;
Title2 '<<< Chow & Liu(2000) Table 9.3.3 >>>' ;

data D1 ;
  input SQN $ SBJ @ ;
  do EXP = 1 to 3 ;
    input DRUG $ CARRY $ Y @ ;
    YL = log10(Y) ;
    output ;
  end ;
  cards ;
1:TRR 2 T CX 32.14 R CT 42.55 R CR 36.08
1:TRR 5 T CX 51.85 R CT 47.37 R CR 56.50
:
2:RTT 1 R CX 26.79 T CR 27.23 T CT 31.60
2:RTT 3 R CX 26.71 T CR 30.81 T CT 26.39
:
;

proc mixed data=d1 order=data ;
  class SQN EXP DRUG CARRY SBJ ;
  model YL = SQN DRUG CARRY / ddfm=satterth ;
  random SBJ ;
  lsmeans SQN CARRY / cl alpha=0.10 ;
  lsmeans DRUG / cl alpha=0.10 diff ;
run ;

```

・class ステートメントで、すべての因子を定義する。・model ステートメントで、EXP: 実施日は、モデルに含めない。これは、解析に用いるデザイン行列の中で、EXP と CARRE のダミー変数の一部が完全に一致し、それらの LSMEAN が推定不能となるからである。EXP: 実施日の効果を求める場合には、CARRY: 持ち越し効果をモデルから除くこと。実際に指定して、結果を見てみるとよい。・random ステートメントでは、SBJ: 被験者を指定する。・lsmeans ステートメントでは、つぎのオプションを指定する。cl: 信頼区間の出力の指定 alpha=0.10: 90%の信頼区間の表示 diff: 水準間の差の出力。

解析結果を表 3.23 に示す。持越し効果は、 $P=0.3820$ と有意ではない。信頼区間法により判定を試みてみよう。

標準製剤の幾何平均は、 $10^{**}1.5126=32.55$ であり、製剤間の差の 90%信頼区間は $(-0.0158, 0.0311)$ なので、

$$\text{下限\%} = 10^{**}(1.5126 - 0.0158) / 32.55 * 100 = 96.54\%$$

$$\text{上限\%} = 10^{**}(1.5126 + 0.0311) / 32.55 * 100 = 107.4\%$$

と 80%~125%の範囲に入っている。

表 3.23 2×3 のクロスオーバー解析結果

Cov Parm		Estimate	
SBJ		0.01221677	
Residual		0.00230658	

Tests of Fixed Effects						
Source	NDF	DDF	Type III F	Pr > F		
SQN	1	16.2	0.18	0.6754		
DRUG	1	33	0.30	0.5861		
CARRY	2	33	0.99	0.3820		

Least Squares Means											
Effect	SQN	DRUG	CARRY	LSMEAN	Std Error	DF	t	Pr > t	Alpha	Lower	Upper
SQN	1:TRR			1.52789335	0.03805506	16.1	40.15	0.0001	0.1	1.4615	1.5943
SQN	2:RTT			1.50490259	0.03805506	16.1	39.55	0.0001	0.1	1.4385	1.5713
CARRY			CX	1.51396330	0.02840516	19.9	53.30	0.0001	0.1	1.4650	1.5630
CARRY			CT	1.50654625	0.02840516	19.9	53.04	0.0001	0.1	1.4575	1.5556
CARRY			CR	1.52868435	0.02840516	19.9	53.82	0.0001	0.1	1.4797	1.5777
DRUG		T		1.52020970	0.02773946	18.2	54.80	0.0001	0.1	1.4721	1.5683
DRUG		R		1.51258624	0.02773946	18.2	54.53	0.0001	0.1	1.4645	1.5607

Differences of Least Squares Means											
Effect	SQN	DRUG	_DRUG	Difference	Std Error	DF	t	Pr > t	Alpha	Lower	Upper
DRUG		T	R	0.00762346	0.01386417	33	0.55	0.5861	0.1	-0.0158	0.0311

文献

緒方宏泰, 鮫島政義 編 (1992), 医薬品のバイオアベイラビリティと生物学的同等性試験, 第2版, 薬業時報社.

高橋行雄, 大橋靖雄, 芳賀敏郎(1990): SASによる実験データの解析, 東大出版会

高橋行雄 (1997), 薬物濃度を用いた同等性試験の例数設計, 第16回日本SASユーザー会論文集, 149-152.

Chow S.C. and Liu J(2000), *Design and Analysis of Bioavailability and Bioequivalence Study*. 2nd edition, Dekker.

4. 用量反応

4.1. 用量反応関係の考え方

医薬品には、何らかの用量反応関係を期待しているが、それらについて一律に論じてはならない。平滑筋などを用いた *in-vitro* の試験なのか、急性毒性試験のような *in-vivo* 試験なのか、ヒトでの薬物動態試験なのか、治療効果に対する用量反応試験なのかによって、設定できる用量の範囲が制約され、観察する反応の（シグナル / ノイズ）が大きく異なる。

臨床試験の場合であれば、ICH の E4 の「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（1994）<http://www.nihs.go.jp/dig/ich/efficacy/e4/e4.html>」が、示唆に富む。この指針の 1.2 節に次のような一節がある。

1.2 用量選択における用量－反応情報の利用

医薬品の開始用量を選択する際に最も役立つことは、望ましい効果および望ましくない効果の双方について母集団の（群としての）平均的な用量－反応曲線の形状および位置を知ることである。この情報に基づき、さらに望ましい効果と望ましくない効果の相対的重要性に関する判断を加味して用量を選択することが最も望ましい。

他方、*in-vitro* での薬理実験では、用量－反応関係が、統計的に有意差がなければ、実験そのものが不適切であるといわれるぐらい、明瞭な用量－反応関係が実験結果として、もちろん統計的にも有意な反応であることが求められる。

本章では、平均的な用量－反応曲線の形状および位置を知るための統計解析と、明瞭な用量－反応関係を知るための統計解析の両方の立場からこの問題を取り上げる。

4.2. ヒトの薬物動態パラメータにおける用量反応

ヒトにおける単回投与後の PK（薬物動態）パラメータ、AUC および Cmax は、投与量の増加に従って増加する。この用量反応関係に対する統計モデルを取り上げる。簡単な問題と思われるかもしれないが、統計的には奥が深い問題であり、いくつかの統計的な方法、および生物学的な判定法の考え方について示す。

薬物濃度の測定は、一般的には安全性と忍容性試験で同時に行われることが多い。この場合に逐次的に投与量が増やされ、血中の薬物濃度も同時に計測される。試験デザインとしては、1 被験者 1 用量の場合、1 被験者にいくつかの用量が投与される場合がある。これらの 2 つの場合について示す。

統計モデルを用いた用量比例性の検討

用量比例とは、おもに AUC と Cmax について、その大きさが投与量に比例して増大することである。これを統計的にはどのように示したらよいのであろうか。単純に直線回帰の当てはめ、相関係数を示せばよいのであろうか。あるいは、AUC を投与量で割って、用量群間で有意差がないことを示せばよいのであろうか。どちらの方法も用量比例性を間接的に取り扱っており、判定の基準があいまいである。

比例の変わりに線形を用いて、薬物動態的には線形であるとの言い方もあり、気楽に使われることにより、人によって理解が異なりコミュニケーションが阻害される場合もある。単回帰分析を行い、傾きが統計的に有意な差があるので、用量線形性がある、との言い方が相互理解を妨げる代表である。これは、図 4.1 に示すように用量反応曲線が飽和しているような場合であっても有意な差となってしまうからである。そのために、PKパラメータについて用量反応を論ずるときには、別のアプローチが必要である。

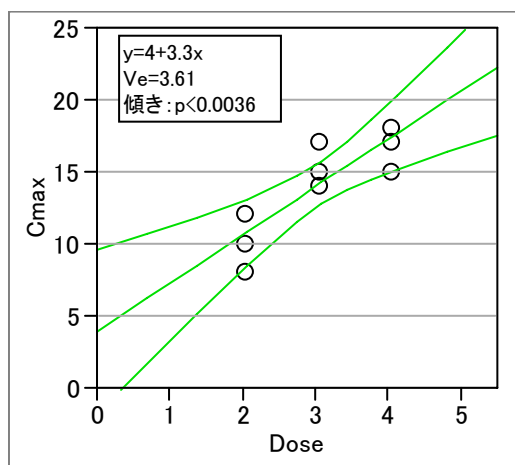


図 4.1 用量反応が飽和している場合の回帰分析

単回帰分析による方法

$Y = \alpha + \beta X$ を当てはめて、 $\alpha = 0$ ならば（回帰直線が原点を通る）ことを用いる。統

計的には、 α の 95%信頼区間が 0 を含む場合に用量比例性ありと判定する。しかし、用量が増加するに伴い AUC, C_{max} の分散も平均値に比例して増加することが多く、通常の単回帰分析ではなく重み付き最小二乗法を適用する必要がある。さらに、切片 α の推定は外挿となり、統計的に問題が残る。図 4.2 の例では α の 95%信頼区間が 0 を含むため用量比例性ありと判定する。

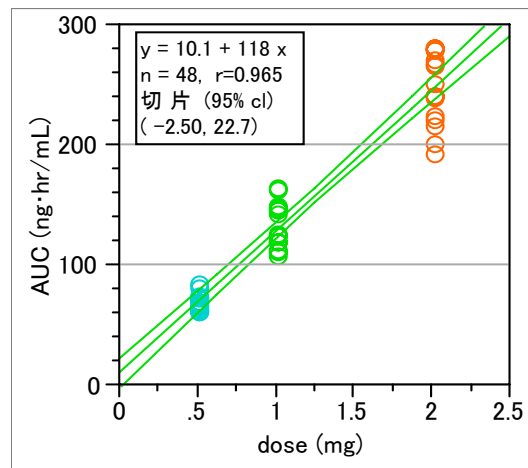


図 4.2 回帰直線の切片が原点を通る

このデータは、表 7.10 の AUC を用いている。

3 用量間の分散は統計的にも明らかに異なる (Bartlett, $\chi^2 = 12.62$, $df=2$, $P<.0001$)

対数変換後の直線回帰分析

投与量および PK パラメータについて対数を取り、回帰モデルを当てはめる。これは、投与量 x が 0 のときに PK パラメータも 0 となるような曲線を当てはめる冪モデルであり、 $b=1$ の場合に原点を通る直線となる。

$$y = a x^b$$

両辺に対数をとると、

$$\log(y) = \log(a) + \beta \log(x)$$

のように、単回帰分析となる。 β の 95%信頼区間が 1 を含む場合に用量比例性ありと判定する。図 4.3 の例では β の 95%信頼区間が 1 を含まないため用量比例性なしと判定する。

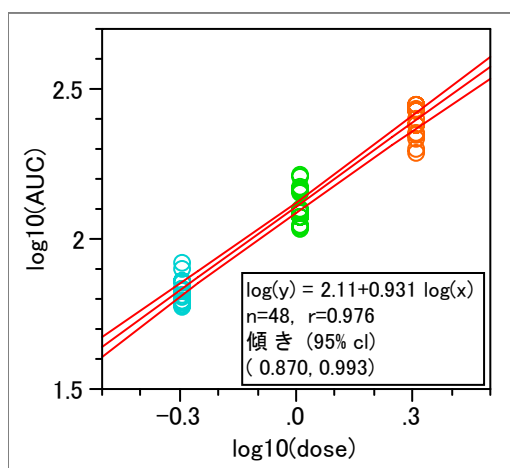


図 4.3 両対数での回帰直線

単位用量当りの AUC, Cmax における分散分析

投与量で割ったAUCまたはCmaxについて、投与量を名義尺度として一元配置分散分析を行い、有意差なしの場合に用量比例性ありと判定する。図 4.4 の場合に有意差がないので、用量比例性ありと判定する。AUC/投与量のグラフ表示は、生物学的同等性の基準を使うことにより、統計的判断と薬物動態学的判断がしやすい。

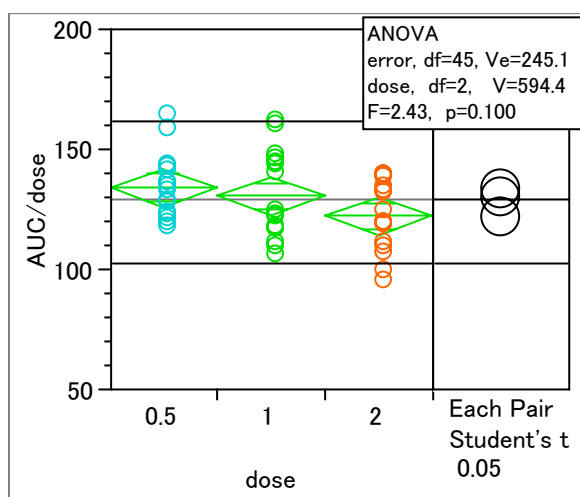


図 4.4 AUC/dose を用いた分散分析

各用量の菱形の中央が平均値，上下が 95%信頼限界
点線は総平均の 80%，125%の位置を示す。

この範囲に平均値の 95%信頼区間が入っていれば、生物学的に同等と判断する。

笠井，橋本ら（2000）は、シミュレーションにより、単回帰分析，対数変換後の回帰

分析，単位用量に変換した場合の3種類の方法について比較検討をした．その結果として，直線回帰分析 ($Y = \alpha + \beta X$) の $\alpha \neq 0$ の検定は有意水準5%を維持できず，非線形の検出力も低く薦められない．AUC/投与量に関する一元配置は有意水準を保持していたが，非線形の検出力は $\beta \neq 1$ の検定が優れるため対数変換後の直線回帰分析で代替可能である．対数変換後の直線回帰分析 ($\log(Y) = \alpha + \beta \log(X)$) を行い， β の点推定値と95%信頼区間を表示することを推奨する．と報告している．

4.3. 1 被験者に複数の用量を投与

表 4.1 に示すように，一被験者に複数の用量を投与する場合もある．この場合には，それぞれの被験者ごとに冪モデルにより係数 b を求め，全体として用量比例性を判断することが考えられる．そして，係数 b の平均値の95%信頼区間が1を含むことにより用量比例性を判断する．この2段階でのステップを，まとめて一度に行う方法もある．ランダム係数モデルといわれている方法で，IDを変量効果とした線形混合モデルを用いて，回帰係数 b の95%信頼区間を求める方法である．

表 4.1 一被験者ごとの Cmax の用量反応

ID	5mg	15mg	30mg	50mg	冪モデル b
9	.	31.8	87.66	134.00	1.210
10	12.5	.	71.53	121.03	0.983
11	24.2	36.6	80.99	.	0.647
12	15.9	52.2	.	104.37	0.813
13	12.6	.	78.01	124.42	1.001
14	.	41.5	87.51	98.98	0.742
15	16.3	37.4	91.52	.	0.944
16	21.6	45.1	.	104.7	0.686

このデータは，表 8.2 のパネル2のデータである．

それぞれの被験者に付いて冪モデルで係数 b を求めた結果を表 4.1 に示す．平均値は0.878，95%信頼区間は，(0.719~1.037) で，1を含んでいるので，用量比例的と判定される．

次に示すSASのMIXEDプロシジャがランダム係数モデルである．各被験者に切片が，

```
proc mixed data=tapping.pk ;
  class ID ;
  model CMAX_L = DOSE_L / solution clm ;
  random ID /solution ;
```

異なるが, 共通の傾きを持つ回帰直線を当てはめ, その 95%信頼区間を求めるのである. 傾き b の推定値は 0.8547 となり, その 95%信頼区間は, (0.7675~0.9419) と 2 段階に比べ範囲が小さくなり, 1 を含まないので統計的には用量比例的ではないとの結果となる.

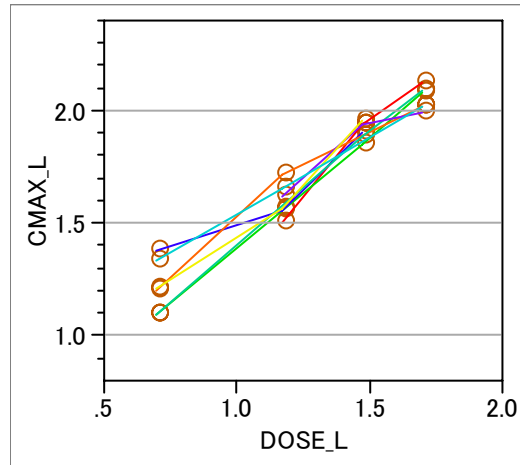


図 4.5 被験者ごとの用量反応

冪モデルの係数 b の統計量,

$n=8$, Mean=0.88, U 95%=1.04, L 95%=0.72 と上限が 1 を越えているので用量比例的と判断する

4.4. 冪モデルの性質

冪モデルは、図 4.6に示すように原点を通る曲線であり、係数 b が 1.0 のときは直線となり、係数 b が 1.0 より小さい場合は、用量 x の増加に対して y の増加が相対的に小さくなり、係数 b が 1.0 より大きい場合は、逆に用量の増加に対して、 y の増加が相対的に大きい。両対数グラフ上では、当然のことながら直線となる。

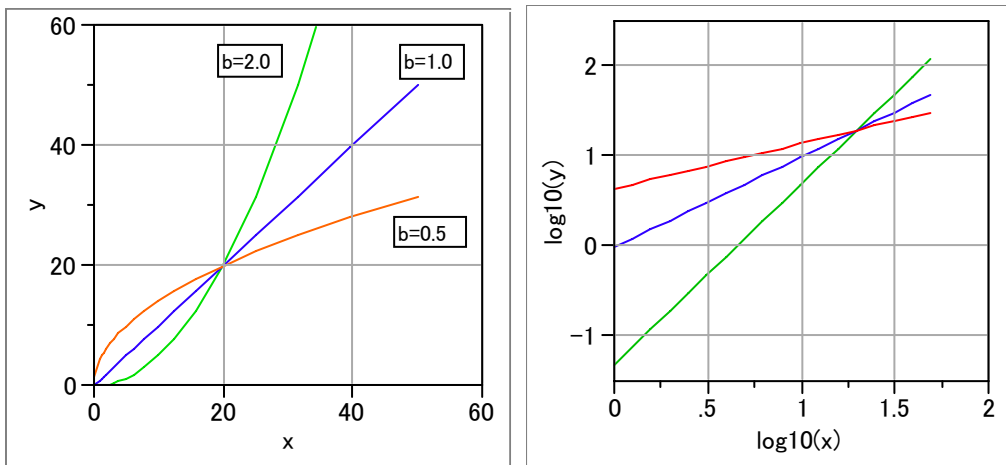


図 4.6 係数 b を変えた場合の冪モデル

$y = 20^{-1} x^{2.0}$, $y = 20^0 x^{1.0}$, $y = 20^{0.5} x^{0.5}$ の場合
(20, 20) を通る曲線となるように係数 a も変えている

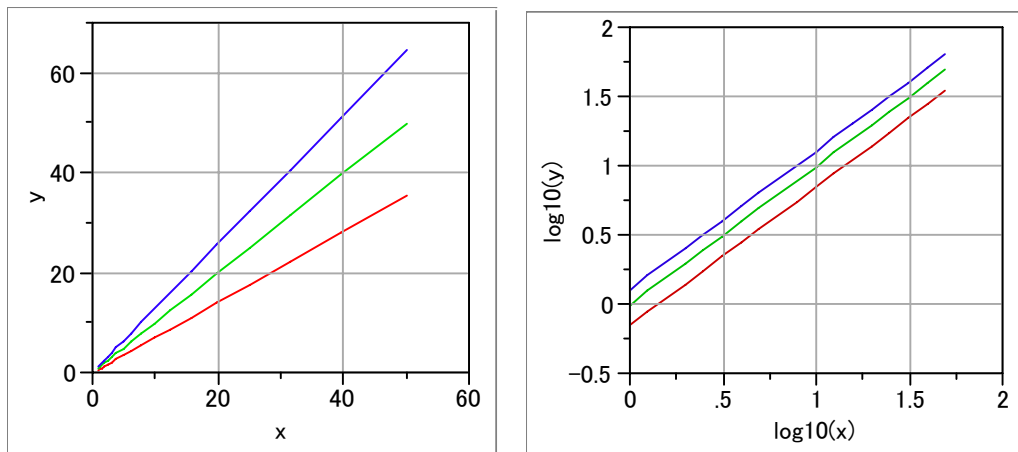


図 4.7 係数 a を変えた場合の冪モデル

$y = 1.3 x^{1.0}$, $y = 1.0 x^{1.0}$, $y = 0.65 x^{1.0}$ の場合

4.5. 信頼区間の活用

用量比例性の検討の場合にも BE 試験での信頼区間方式が役に立つ。しかしながら、基準となる投与量は一意に定めることが出来ないため、用量当りの AUC などであれば、用量群をプールした総平均（幾何平均）を用いて、同等性の基準として、総平均の 80% から 125% の範囲に、各用量の 90% または信頼区間が入っていれば、用量比例性が、成り立っているとみなす。この基準を使うのは、統計的には用量比例的でないと判定された場合でも、何らかの生物学的な許容範囲内であれば、用量比例的であるとの判定ができるからである。

BE 試験での基準である (80%, 125%) を使うのは、基準値を総平均としているので、甘い基準であるかのように思われるかもしれないが、用量比例性の検討の場合に、平均値の信頼区間には、個人間と個人内を合わせた誤差を用いることになるので、個人内の誤差を用いる BE 試験に比べて、甘い基準というわけではない。

いずれにしても、このような判定基準をも用いる場合は、事前に基準を設定しておくことは、当然の事である。交互パネル（1 被験者に複数用量の場合、1 被験者に複数人が人によって異なる用量のセットの場合）による例を示す。

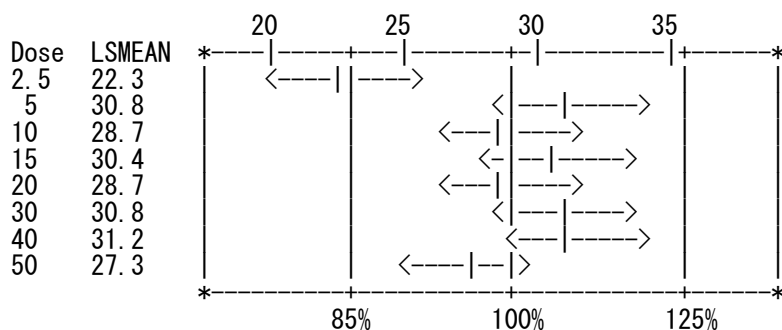


図 4.8 AUC/Dose の 90% 信頼区間の例示

この AUC は、8.1 節で実験から得られたものである。

5mg から 50mg の範囲で、用量比例性が成り立っていると見える。

信頼区間方式は、用量で PK パラメータを割っているのが、図 4.4 は AUC/Dose の例であるが、元のスケールおよび両対数の場合を 図 4.9 に示す。元のスケールでは扇形の信頼区間となり、両対数の場合は、平行な範囲となる。

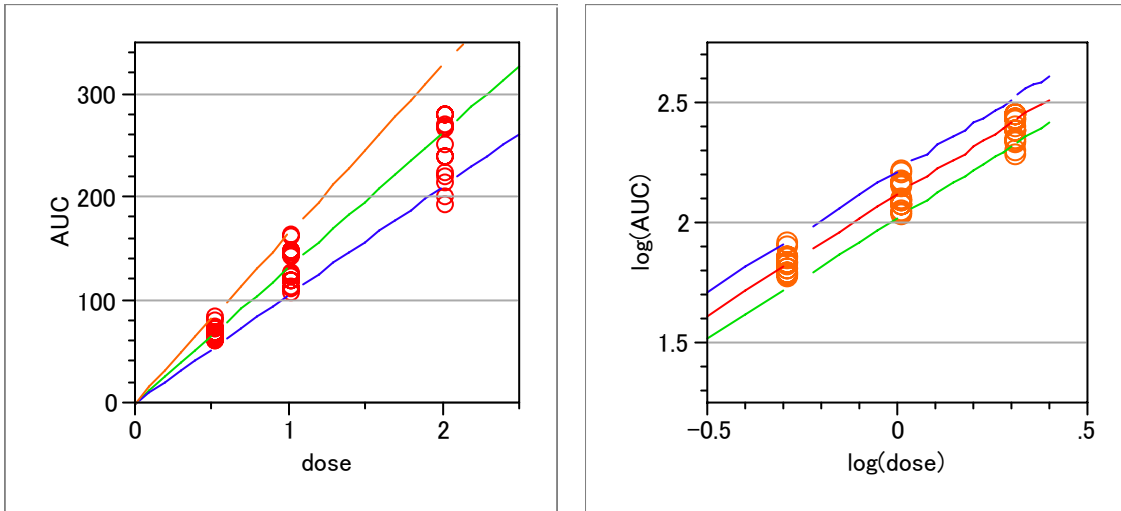


図 4.9 生物学的同等性の判定基準の準用

80%~125%の範囲は、 $y = 0.80ax$ 、 $y = 1.25ax$ で与えられる。両対数の場合、切片が $\log(0.80a) = -0.0969 + \log(a)$ 、 $\log(1.25a) = 0.0969 + \log(a)$ と上下同じ幅の平行線となる。

図 4.3 の両対数グラフを元のスケールに直したのが、図 4.10 である。原点を通る直線のように見えるが、かすかに上に凸となっている曲線である。

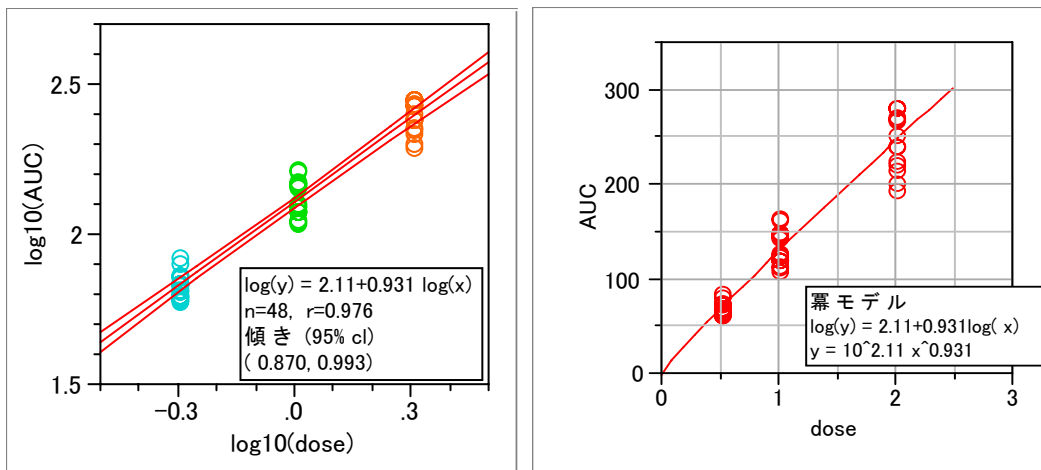


図 4.10 冪モデルの当てはめ

4.6. 平行性検定法の応用

複数の民族で用量反応試験が行なわれた場合の用量反応関係を解析するためには、生物検定法の 勾配比検定、平行性検定の考え方が役に立つ。図 4.11は、7 章の民族的要

因用いるデータの中から用量反応が得られる 3 試験の用量反応直線を引いたものである。3 試験間の交互作用は有意 ($p=0.0161$) であった。試験 1 の傾きは 0.931 ($0.870 \sim 0.992$) と試験 5 は 0.937 ($0.879 \sim 0.995$) と用量反応は平行であるが、試験 3 は 1.086 ($0.986 \sim 1.186$) と異なる。このような場合に、同じ反応を得るための投与量を求めると比較がしやすい。

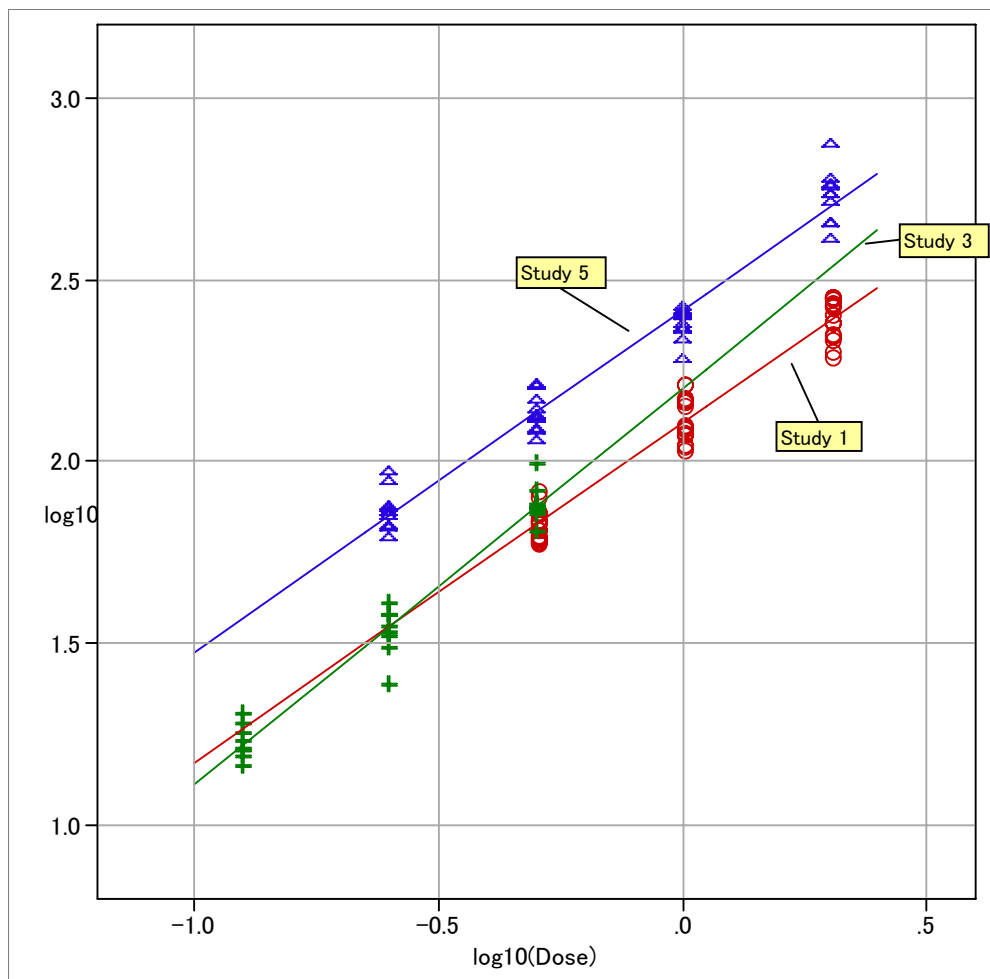


図 4.11 複数の試験の用量反応の表示

このような場合に、表 4.2 に示すように同じ反応を得るための投与量を求めると比較がしやすい。この例であれば、AUCが $100\text{ng}\cdot\text{hr}/\text{mL}$ となる $\log(\text{dose})$ は、試験 1 では、 0.76mg であり、試験 3 では、 0.65mg 、試験 5 では 0.36mg と約半分の量で、同じAUCが得られることがわかる。

表 4.2 AUC が 100ng・hr/mL となる log(dose)

試験番号	log(dose)	L95%t	U95%
1	-0.116	-0.135	-0.099
3	-0.187	-0.228	-0.138
5	-0.446	-0.475	-0.419

ある y となる x の逆推定, 95%信頼区間はフィラーの式による. 計算は SAS/JMP で行った.

文献

笠井英史, 橋本敏夫ら (2000), Phase I 試験における薬物動態値の線形性について, 医薬安全研
83 回定例会資料.

Kevin Gough et al : Assessment of Dose Proportionality : Report from the Statisticians in the
Pharmaceutical Industry / Pharmacokinetics UK Joint Working Party, Drug Information Journal
29,1039-1048(1995)

Little, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (1996) *SAS System for Mixed
Models*. SAS Institute.

高橋行雄 (1997), 薬物濃度を用いた同等性試験の例数設計, 第 16 回日本 SAS ユーザー会論文
集, 149-152.

5. 高齢者、小児などにおける臨床薬理試験

5.1. 高齢者を対象にした臨床薬理試験

高齢者を対象とした臨床試験の必要性が、ICHの有効性の分野の最初成果として1993年に「E7：高齢者に使用される医薬品の臨床評価法に関するガイドライン」によって示された。これ以前の臨床試験では高齢者は、薬効評価には適さないとの理由により多くの臨床試験の除外基準をして定めるのが常であった。

このガイドラインは、高齢者（65歳以上）に対して臨床薬理的アプローチを強調している。ガイドラインの緒言で「高齢者における薬物療法に関する多くの問題点は、年齢に関連する腎機能、肝機能、心機能などの変化および多剤併用などに基づく薬物感受性と薬物動態学的変化に由来する。したがって臨床試験も高齢者と非高齢者との薬物動態学的および薬力学的差異の有無およびその評価に重点を置いたものでなければならない。」と述べている。

5.2. 小児に対する臨床薬理試験

小児に対する臨床試験の必要性は、2000年末にICHの最新のガイダンス「E11：小児集団における医薬品の臨床試験に関するガイダンス」によって示された。ガイドラインでは、小児を、早産児、正期産新生児（0から27日）、乳幼児（28日から23ヶ月）、児童（2歳から11歳）、および青少年（12歳から16又は18歳）と5区分に分類している。そして、それぞれの年齢階層の特質を示し、慎重な臨床薬理的なアプローチを推奨している。例えば、肝と腎のクリアランスについて、次のように年齢階層による差異が述べられている。

早産児：腎及び肝クリアランス機構の未熟性

正期産新生児：腎及び肝クリアランス機構の未熟性熟であり、かつ急速に変化する

乳幼児：肝及び腎クリアランスが急速に成熟し続ける。1から2歳までに mg/kg で示される多くの医薬品のクリアランスは成人値を上回る可能性がある

児童：薬物クリアランスのほとんどの経路（肝臓及び腎臓）は成熟しており、クリアランスはしばしば成人値を上回る。薬物クリアランスの変動は、個々の代謝経路の成熟度に依存している可能性がある。

青少年：記載はないが、「青少年コンプライアンスの問題はあるにせよ、成人対象の

試験に比較的年齢が高い青少年を含むことは可能であろう」の記述により成人と同様に見なしている

これらの臨床試験において、医薬品の効果と安全性を調べるために、高齢者では 100 例程度とガイドラインには示されているが、臨床薬理的試験は、全ての患者に必要なのか、あるいは一部の患者なのか、これらは一律に決められない課題であり、成人における臨床薬理のデータを基に、それぞれの目的に合致した試験デザインおよび例数設計が必要であり、試験統計家の力量が示される場面でもある。

この問題に対しては、忍容性を主体に置いた第 1 相試験，PK/PD に主体をおいた第 1 相試験，および BE 試験などの経験を総合して対処すべき課題である。これらは、また民族的要因の探索にも通ずる試験統計家の新たな課題である。

5.3. 高齢者を対象にした薬物動態試験の事例

65 歳以上の高齢者における薬剤 R の安全性の評価および薬物動態を検討することになった。安全性については、症例数を増やして未知の有害事象の発見が主な目的である。薬物動態については、非高齢者に比べ投与前の薬物濃度のトラフ値が異常な上昇が起きていないか、もしも起きていた場合には、その原因の探索を目的とした。

統計的な例数設計は、非高齢者に比べて高齢者のトラフ値の平均的な上昇が 1.5 倍あった場合の検出力が 80%になるような例数とされた。非高齢者のエントリーは薬物濃度のトラフ値の測定を目的とし、高齢者の場合は薬物濃度の測定の症例例数に達した場合にはエントリーを打ち切るが、安全性の評価のためには別途定めた症例に達するまでエントリーを続けることとされた。

結果を表 5.1 に示す。トラフ値の測定は、1 週目、4 週目、8 週目、および 12 週目に行われ、高齢者に異常な上昇による中止例はなかったので、トラフ値の平均的な上昇について比較することにする。

図 5.1 に年齢区分ごとのトラフ値の経時変化を示す。この図からは高齢者におけるトラフ値の上昇はないと判断されが、どのような統計解析が、このような経時データに適しているのだろうか。本テキストでは、経時データをそのまま用いるのではなく、できる限り要約統計量を算出して対応することを薦めてきた。

表 5.1 トラフ値のデータ

ID	sex	age	1W	4W	8W	12W	ID	sex	age	1W	4W	8W	12W
1	2	68	11	17	12	9	14	2	52	5	7	.	.
2	2	67	7	6	4	4	15	2	44	5	10	8	7
3	2	72	4	5	11	8	16	1	64	14	10	9	8
4	2	70	9	7	12	9	17	2	41	8	6	7	7
5	2	67	6	7	7	8	18	2	52	8	8	8	.
6	2	70	5	6	6	12	19	2	47	11	14	.	.
7	2	65	15	10	14	18	20	2	48	5	6	4	7
8	2	65	10	7	4	8	21	2	45	11	6	15	10
9	1	72	4	7	4	7	22	2	43	6	7	4	5
10	2	64	7	6	21	13	23	2	64	4	6	5	5
11	1	68	5	6	6	.	24	1	39	5	8	13	7
12	1	72	4	9	7	5	25	1	44	10	11	20	.
13	1	89	5	4	8	5	26	1	49	21	14	16	17

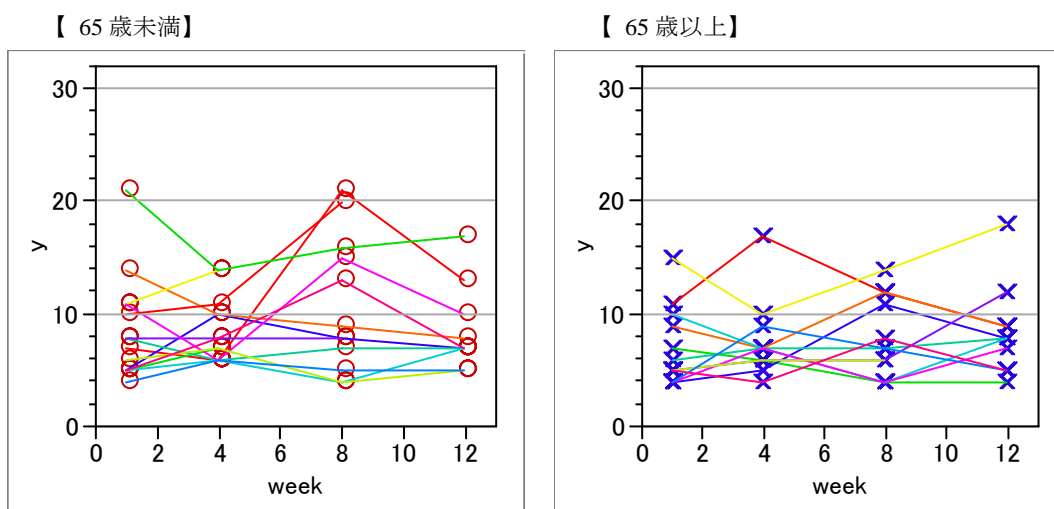


図 5.1 年齢区分別の薬物濃度のトラフ値の変化

薬物濃度のトラフ値の上昇は、薬物の代謝機能の衰退によって引き起こされ、徐々に上昇はするが、薬物濃度が増せば代謝量も増すという関係が成り立っているとすれば、ある濃度まで達すると定常状態になることが知られている。これは、ヒトにより異なることから、年齢とトラフ値の関連を見るための要約統計量としては、ある時点以後のトラフ値の平均を用いることが、個人内のトラフ値の変動を減少させるためには、妥当であろう。このことは、表 5.2 に示す時点間の相関構造が $r = 0.60$ 前後であることから裏付けられる。

表 5.2 時点間の相関構造

	1W	4W	8W	12W
1W	1	0.616	0.488	0.713
4W	0.616	1	0.306	0.398
8W	0.488	0.306	1	0.681
12W	0.713	0.398	0.681	1

当初の仮説は，非高齢者に比べて高齢者のトラフ値が 50%以上にならないことを検出しようとしていたので，図 5.2 (左) から高齢者は，非高齢者に比べトラフ値の平均的な上昇はないと考えられる．また 図 5.2 (右) の回帰分析からも加齢と共にトラフ値が上昇することはないと考えられる．

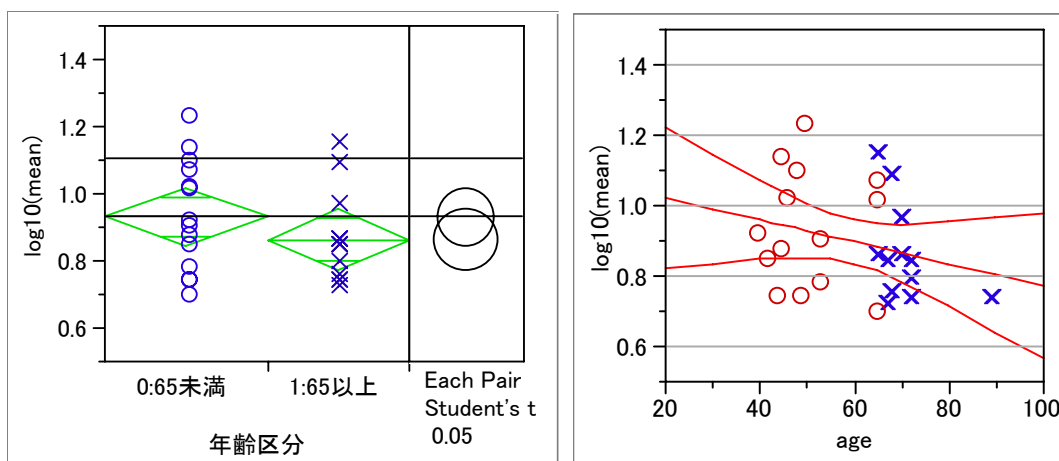


図 5.2 投与後 1～12 週のトラフ値の平均値（対数）と年齢との関係

左の図の菱形の上端と下端は平均値の 95%信頼区間を示している．高齢者の平均値は 95%の信頼区間の上限が，非高齢者の 50%増しのラインを超えていないので，高齢者のトラフ値の上昇はないと考えられる．回帰係数は負であり，加齢による薬物濃度の増加はない．回帰直線の 95%信頼区間の形状（回帰直線の傾きがプラスになる可能性がある）から傾きが有意な差でないことがわかる

6. 食事の影響

6.1. ガイドラインでの記述

医薬品製造指針(1997)によれば、医薬品の承認申請に必要な添付資料の(へ) 吸収、分布、代謝、排泄(ADME)の薬物動態試験中で、「3. ヒトにおける成績、食事の影響について試験成績と考察を必要に応じ食事の内容を含め記述する。」と記されていた。

1998年以降の医薬品製造指針には、ADMEのパートは、動物の薬物動態に限定して記され、ヒトの食事の問題は触れられていない。経口投与の薬物の場合、期待する薬物濃度を得るために食時の影響がどの程度あるかを調べる必要があると思われるのだが？

経口投与の薬物では通常、絶食下での投与した後の血中薬物濃度を測定している。食後に投与した場合に、健康男子を対象にした安全性と忍容性を検討する第1相臨床試験の中で、付け足し程度に行われることが多い。そのためにBE試験の場合とは比較にならないくらい少数の被験者の2x2のクロスオーバー試験が行われ、PKパラメータに対する対応のあるt検定の結果によって食事の影響について結論がだされがちである。一般的には、「食事の影響はない」とした方が簡単なので、少数例でのt検定は都合がよいと思惑しているかもしれない。

この場合にもBE試験の信頼区間方式により、各種のPKパラメータが、標準とみなす絶食下投与の平均値に対するパーセント表示が、対応のあるt検定よりも適切な医学的な判断を下すのには適している。

用量－反応関係の指針での記述

ICHのガイドラインE3「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」にしっかりと記載されている。このガイドラインのには、次に示すように「食事」が「第1章 はじめに」に2ヶ所に出てくる。

用量－反応情報を利用する際には、医薬品の薬物動態の個体差をもたらす因子が何であるかを可能な限り明らかにすることが重要である。その因子には、人口統計学的因子(例えば年齢、性別、人種)、他の疾患(例えば腎あるいは肝障害)、食事、併用療法、個々の患者の特徴(例えば体重、体型、併用薬、代謝の相違)が含まれる。

用量－反応試験においては、他の様々な面からも時間を考慮に入れるべきである。医薬品の効果発現の遅れが薬物動態学的、または薬力学的因子の結果である場合でも、所定の用量を用いた試験期間は、最大の効果が発現するのに十分な長さとするべきである。朝方の投与と夕方の投与とでも用量－反応関係が異なることもある。同様に、投与初期の用量－反応関係とそれ以後投与を継続した後のそれとが異なることもある。反応が一日用量ではなく、むしろ累積用量、投与期間（例えばタキフィラキシー、忍容性、または履歴現象）あるいは投与と食事との関係に関係していることもありうる。

民族的要因ガイドラインでの記述

ICH : E5 「外国臨床データを受け入れる際に考慮すべき民族的要因」についても「食事」のついて調べると、「外因性民族的要因 (Extrinsic Ethnic Factors)」の中の一つとして上げられている。

例えば医療習慣、食事、喫煙、飲酒、環境汚染や日光への暴露、社会経済的地位、処方された薬の服用遵守、並びに異なる地域の臨床試験の信頼性にとって特に重要なものとして、臨床試験の計画及び実施方法が挙げられる。

補遺C : 薬物動態、薬力学及び用量反応に関する考察

ICH 地域の3つの主要人種群（アジア人種、黒人種、白人種）についての薬物動態及び薬力学の評価とそれらの比較可能性は、ICH 地域での医薬品の承認において極めて重要である。基本的な薬物動態の評価として、吸収・分布・代謝・排泄 (ADME)、並びに適切な場合には、食事－薬物及び薬物－薬物相互作用を特徴づけなければならない。

補遺D : 医薬品の民族的要因による影響の受けやすさ

- ・ 生物学的利用率が高く、従って薬物吸収が食事の影響を受けにくいこと。
- ・ 薬物間、薬物－食事及び薬物－疾病間の相互作用の可能性が少ないこと。
- ・ 生物学的利用率が低く、従って薬物吸収が食事の影響を受けやすいこと。

質疑応答集

(民族的要因の影響の評価について)

Q15. 外因性民族的要因の主要なものと考えられる医療習慣や食事の影響はどのように評価すればよいのか。事例を重ねないと分からないと考えるべきなのか。

(答) 医療習慣や食事の内容が薬物の体内動態や効果の発現、薬効評価の方法などにどのように影響するかは、薬物によって異なるため、一概にその評価法を述べることはできない。最終的にはケースバイケースの判断を行う以外にはないが、類似した構造の薬物や同様の作用機序を有する薬物の経験から類推する方法は一つの方法と言えよう。

(臨床薬理試験について)

Q26. 外国在住の日本人（日本人の両親を持ち、日本で生まれ育ち現在外国で生活している）、絶食下投与と食後投与あるいは日系2世、3世（外国で生活している日本人の両親から生まれ、又は祖父母が外国に移住しさらに両親が日本人で外国で生活している）などを対象に実施した薬物動態試験データは、日本人の薬物動態に関するデータとして利用可能であるのか。

(答) 食事や環境などにより薬物動態に差が出ることが想定される場合を除いては、利用可能である。

一般指針、高齢者、総括報告書ガイドラインでの記述

ICH : E8「臨床試験の一般指針」には、バイオアベイラビリティに関係した1ヶ所のみである。「b) 薬物動態、経口投与される医薬品の多く、特に薬物放出が制御されるべく設計された製剤においては、バイオアベイラビリティに対する食事の影響を検討する試験は重要である。」

ICH : E7「高齢者に使用される医薬品の臨床評価法に関するガイドライン」には、2ヶ所に「食事」が出てくるが、「与薬時刻と採血時刻、食事との関係をも記録することが重要」、「薬物動態スクリーニングを実施する際も、血中濃度に影響を与える要因(食事、胃酸度、肝・腎・心機能等)をできるだけ押さえ」などのように、実験条件をそろえるとの注意になっている。ICH : E3「治験の総括報告書の構成と内容に関するガイドライン」では、「投与の時期(一日の中の時刻、投与間隔)及び食事と投与との関係を記述」などのように記録として記すことを求めている。

FDAのガイダンス

FDA では、「Food-Effect Bioavailability and Bioequivalence Studies (DRAFT), <http://www.fda.gov/cder/guidance/1719dft.pdf>」により，食事の影響についてのガイダンスを作成中である．このガイドラインでは，Food-effect bioavailability (BA) 試験と Food-effect bioequivalence (BE) 試験とを区別し，実施すべき時期について記している．

試験デザインは，2×2 のクロスオーバー法を，統計解析は AUC と Cmax の対数について 90%の信頼区間方式を推奨している．同等であるとの基準は，絶食下での投与を基準として，AUC では 80%–125%，Cmax では 70%–143%の範囲に差の 90%信頼区間が入ることを求めている．

6.2. 2元配置による実験

実験デザイン

食事の影響を検討するための試験は，FDA のガイダンスが示すように，開発の段階に合わせて逐次行うことが望ましい．次に示すのは，民族的要因の検討のために，PK パラメータの用量反応関係の探索も考慮しつつ，食事の影響を合わせて検討された．

投与量は3用量で，その中で表 6.1 に示すように2×2 のクロスオーバーとし，食前と食後を割付けた．例数設計は，BE試験に準じて計画した，用量反応関係は，おおよその形状を見るためとされていた．

表 6.1 2要因 2×2 クロスオーバー

用量	投与順	被験者	実施 時期	
			1	2
10mg	1	1, 2, 3	食前	食後
	2	4, 5, 6	食後	食前
20mg	3	7, 8, 9	食前	食後
	4	10, 11, 12	食後	食前
30mg	5	13, 14, 15	食前	食後
	6	16, 17, 18	食後	食前

実験結果を表 6.2 に，用量別に食前食後の線グラフを示す．個体間の変動が大きいですが，用量反応関係はうかがわれるもの，30mgでの伸びが少なく用量比例的とはいえない．食事の影響は，食前の方がややAUCが大きいようである．

食前・食後についての同等性の解析は，投与量を分類変数とみなし共変量としたクロスオーバー試験として解析する．BE とする判定は，常用対数変換後の AUC の平均値の差

の90%信頼区間が(-0.1~0.1)の範囲に入っていることをもって判定する。

表 6.2 AUCの結果,

dose	被験者	未変化体 C		代謝物 M	
		食前	食後	食前	食後
10mg	1	67.39	47.63	336.07	277.92
	2	19.48	9.58	117.31	94.98
	3	18.85	21.90	71.78	68.83
	4	35.38	22.79	125.79	119.90
	5	21.39	17.99	137.61	126.25
	6	23.44	20.18	121.60	103.14
20mg	7	31.30	44.41	205.01	247.45
	8	65.22	65.74	144.81	168.14
	9	43.81	37.02	279.75	260.18
	10	46.73	37.25	332.69	361.71
	11	27.71	20.62	123.82	110.29
	12	52.42	43.34	412.16	371.58
30mg	13	37.99	46.50	308.90	278.01
	14	—	53.58	—	418.88
	15	72.33	47.36	254.72	217.61
	16	77.89	58.46	317.00	310.27
	17	—	—	—	—
	18	61.27	68.70	305.17	311.61

No17 の被験者には投与されなかった。No14 は、食前での投与ミス。

用量別に食前と食後のAUCを図 6.1 に示す。未変化体Cでは、食前のほうがややAUCが大きいためである。代謝物Mもわずかであるが食前のほうが大きくなっている。

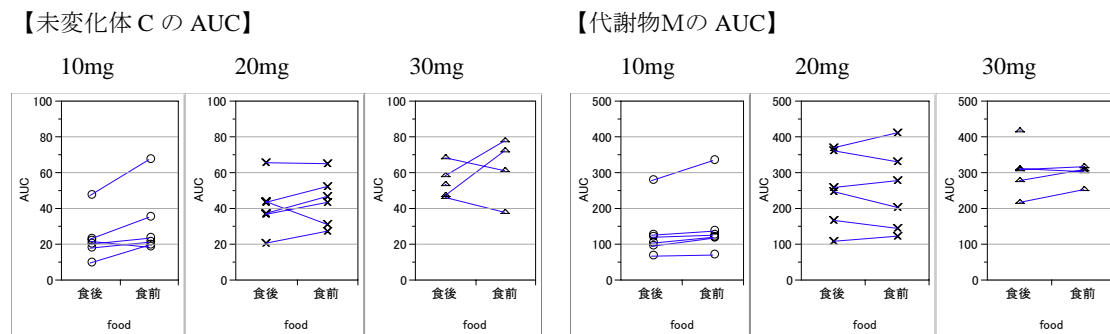


図 6.1 用量別の食事の効果

効果の分離

2因子要因配置クロスオーバにより、食事の効果の抽出は適切にできるのであろうか。これは、AUCの大きさが投与量によって異なっても、その違いは被験者の違いにとしててられ、食事による差は、それらの被験者内の誤差による比較であり、投与

量の影響を受けないために、適切に評価ができる。統計解析は、被験者をランダム効果、投与量群、食事、投与時期を固定効果とし表 6.3 に示すようにSAS/MIXEDによる解析を行う。

表 6.3 投与量を共変量とした食事の影響解析プログラム

```

proc mixed data=d01 ;
  class DOSE ID PERIOD FOOD ;
  model LOG_AUC = DOSE PERIOD FOOD / ddfm=satterth solution ;
  random ID ;
  lsmeans DOSE PERIOD / alpha=0.10 cl ;
  lsmeans FOOD / diff alpha=0.10 cl ;
  by DRUG ;
run ;

```

生物学的判定

表 6.4 および図 6.2 に示すように、未変化体Cおよび代謝物M共に 90%信頼区間が 0 を超えていないので、有意差がある。生物学的な判定基準は、差の 90%信頼区間がプラスマイナス 0.1 を超えていなければよい。この基準を用いると、未変化体Cは、下限が -0.1230 となり生物学的に差がり、代謝物Mは生物学的には差がないと判定される。

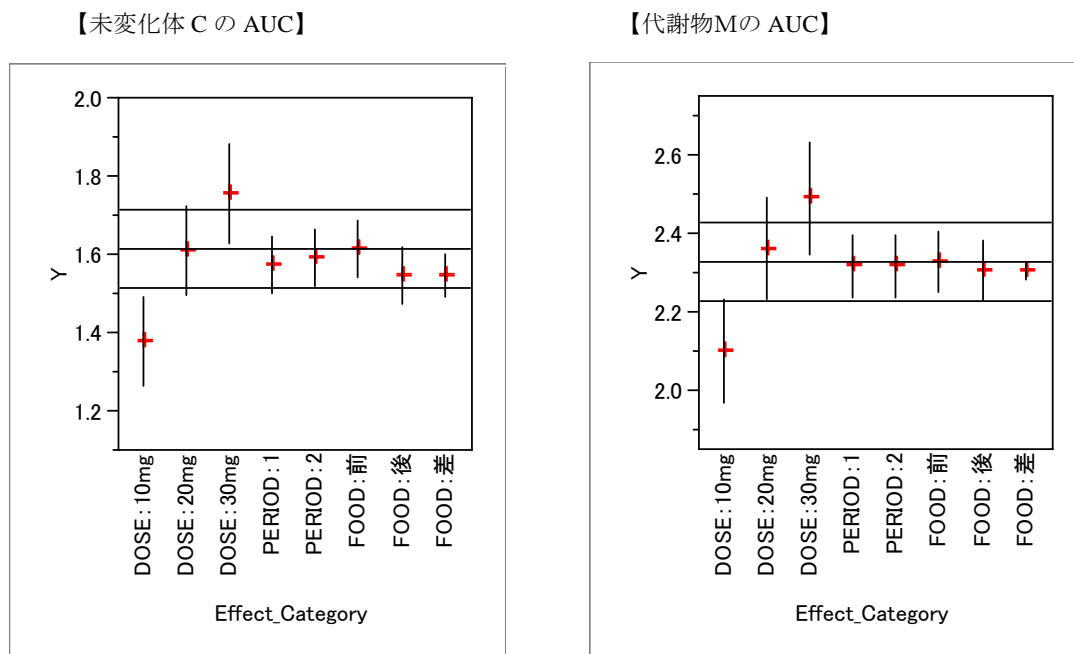


図 6.2 食事の差の 90%信頼区間の表示

表 6.4 AUC での食事の効果および差の 90%信頼区間

	category	LSMEAN	Lower	Upper
未変化体 C	食前	1.618	1.545	1.691
	食後	1.548	1.476	1.619
	差の 95%	1.548	1.495	1.601
代謝物 M	後-前	-0.0702	-0.1230	-0.0175
	食前	2.330	2.252	2.409
	食後	2.307	2.229	2.385
	差の 95%	2.307	2.284	2.330
	後-前	-0.0234	-0.0464	-0.0005

表 6.5 未変化体 C についての SAS のアウトプット

Covariance Parameter Estimates (REML)										
Cov Parm		Estimate								
ID		0.02141180								
Residual		0.00724484								
Solution for Fixed Effects										
Effect	DOSE	PERIOD	FOOD	Estimate	Std Error	DF	t	Pr > t		
INTERCEPT				1.80214157	0.07485903	17.2	24.07	0.0001		
DOSE	10			-0.37743443	0.09655797	14.4	-3.91	0.0015		
DOSE	20			-0.14581196	0.09655797	14.4	-1.51	0.1527		
DOSE	30			0.00000000		
PERIOD		1		-0.01919739	0.02999681	14.4	-0.64	0.5323		
PERIOD		2		0.00000000		
FOOD			食後	-0.07024874	0.02999681	14.4	-2.34	0.0341		
FOOD			食前	0.00000000		
Tests of Fixed Effects										
Source	NDF	DDF	Type III	F	Pr > F					
DOSE	2	14.2		7.94	0.0048					
PERIOD	1	14.4		0.41	0.5323					
FOOD	1	14.4		5.48	0.0341					
Least Squares Means										
Effect		LSMEAN	Std Error	DF	t	Pr > t	Alpha	Lower	Upper	
DOSE	10	1.37998407	0.06459389	14	21.36	0.0001	0.1	1.2662	1.4938	
DOSE	20	1.61160654	0.06459389	14	24.95	0.0001	0.1	1.4978	1.7254	
DOSE	30	1.75741850	0.07177097	14.7	24.49	0.0001	0.1	1.6314	1.8834	
PERIOD	1	1.57340434	0.04184729	18.7	37.60	0.0001	0.1	1.5010	1.6458	
PERIOD	2	1.59260173	0.04120411	17.9	38.65	0.0001	0.1	1.5211	1.6641	
FOOD	食後	1.54787866	0.04120411	17.9	37.57	0.0001	0.1	1.4764	1.6193	
FOOD	食前	1.61812741	0.04184729	18.7	38.67	0.0001	0.1	1.5457	1.6905	
Differences of Least Squares Means										
Effect	FOOD	_FOOD	Difference	Std Error	DF	t	Pr > t	Alpha	Lower	Upper
FOOD	食後	食前	-0.07024874	0.02999681	14.4	-2.34	0.0341	0.1	-0.1230	-0.0175

要因配置実験における用量比例性

食事の効果についての検討のためには、投与量を群として取り扱ったが、用量比例性の検討のためには、 $\log(\text{dose})$ で再解析を行う。未変化体 C の冪モデルでの係数 b とその 95%信頼区間は 0.785 (0.3810~1.1942)，未変化体 M は 0.822 (0.3574~1.2874) と広範囲である。用量比例性を精度よく求めるためには、被験者内で複数の投与量の設定が必要であるが、この実験の目的は、用量比例性の検討より、食事の影響が優先されたことによる当然の結果ともいえる。

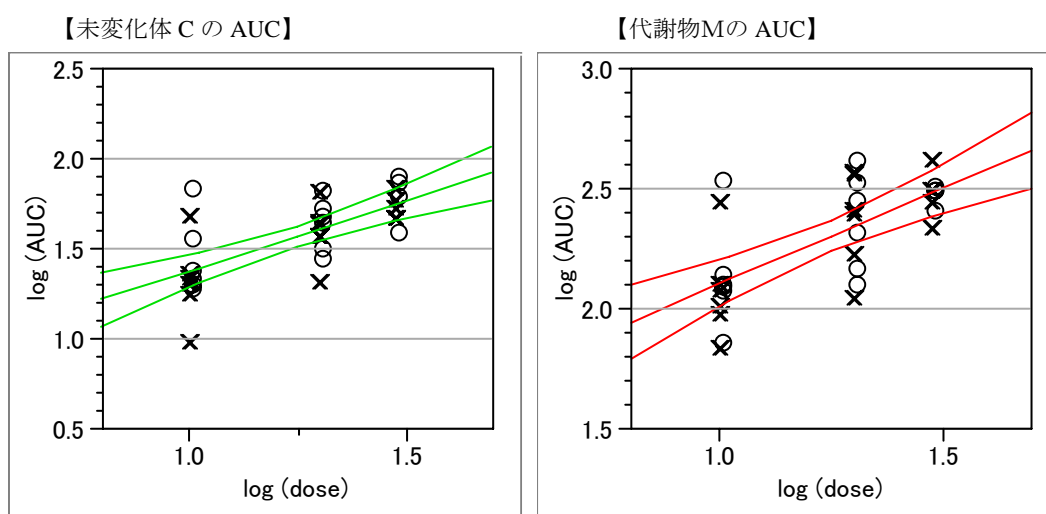


図 6.3 変化体 C の用量反応

Effect	PERIOD	FOOD	Estimate	Std Error	DF	t	Pr > t	Alpha	Lower	Upper
未変化体 C	LOG_DOSE		0.78762865	0.19110050	15.3	4.12	0.0009	0.05	0.3810	1.1942
代謝物M	LOG_DOSE		0.82238762	0.21819068	15	3.77	0.0019	0.05	0.3574	1.2874

SAS プログラム

```
proc mixed data=d01 ;
  class ID PERIOD FOOD ;
  model LOG_AUC = log_DOSE PERIOD FOOD / ddfm=satterth solution cl ;
  random ID ;
  by DRUG ;
run ;
```

7. 民族的要因

7.1. ICH ガイドライン

ICH のE5「外国臨床データを受け入れる際に考慮すべき民族的要因についての指針(1998)」は、これまで日本における臨床試験における海外との障壁を一気に開放した。日本の規制当局は、海外で行われた臨床試験を日本で行われた「ブリッジング試験」により申請資料として受け入れるとの宣言である。本章では、主に薬物動態における民族的要因について探索的な統計解析について焦点を当てる。

ガイドラインでは、医薬品の民族的要因による影響の受けやすさについて、下記のように述べている。また、人種的要因の影響を受けにくさと、受けやすさの原因を例示している。日本で、民族的要因の探索のために薬物動態試験を行う場合は、この例示に従って当概の医薬品の性質を推定した上で実験計画を立てる必要がある。

薬物動態、薬力学及び治療効果に民族的要因が与え得る影響に関して医薬品を特徴づけることは、新地域でどのようなブリッジング試験が必要となるかを決定する上で役立つであろう。医薬品の効果に対する民族的要因の影響は、その医薬品の薬理学上の性質、適応症、患者の年齢、性別によって異なる。民族的要因の影響を受けやすさを、医薬品のひとつの特性のみで予測することは出来ない。必要とされるブリッジング試験の種類は、結局のところ状況に応じて判断していく以外にないが、当該薬物の民族的要因による影響の受けやすさの評価が、その判断に役立つであろう。

表 7.1 人種的要因の影響を受けにくいことの根拠

要素	判定の例	根拠
薬物動態が線形性である	?	
推奨される用法・用量の範囲内では、有効性および安全性のいずれも薬力学的曲線が平坦である	Yes	
治療領域が広い	Yes	
代謝が最小限である、もしくは、代謝が多くの代謝経路で行われる	Yes	
生物学的利用率が高い	25%	
蛋白結合能が低い	No	
薬物-薬物、薬物-食事および薬物-疾病間の相互作用が少ない	Yes	
作用が全身的でない	Yes	
乱用の可能性が低い	Yes	

表 7.2 民族的要因の影響を受けやすい根拠

要因	判定の例	根拠
薬物動態が非線形である	?	
推奨される用法・用量の範囲内で、有効性及び安全性のいずれに関しても薬力学的曲線が急峻である（用量がわずかに変化するだけで、効果が大きく変化する。）	No	
治療量域が狭い。	No	
高度に代謝され、特に代謝経路が単一で、そのために薬物間相互作用の可能性が大きくなる	No	
遺伝多型を示すことが知られている酵素により代謝される	?	
民族により異なる酵素変換を受ける可能性があるプロドラッグとして投与される	No	
生物学的利用率の個体差が大きい	Yes	
生物学的利用率が低く、従って薬物吸収が食事の影響を受けやすい	Yes	
多くの併用薬とともに使用される可能性が大きい	Yes	
鎮痛剤や精神安定剤のように不適切な使用の可能性が高い	No	

7.2. 試験デザインの作成の例示

血中の薬物濃度の AUC において、民族的な差が有りや無しやを論ずる場合に、何らかの例数設計が必要である。基本的には同等性を証明することになる。そのために、生物学的同等性試験の基準が、この問題にも活用することができる。すなわち、先行する民族での試験の平均値の 80% から 125% の範囲に、後から行う民族、例えば日本人での平均値の 90% 信頼区間が入るような確率が 80% となるような症例数、などのように例数設定を行うことが、妥当と考える。

1) 試験の目的

- ・ R 薬については、海外において、ABC 医薬品として開発され、Phase2 まで臨床試験が終了している。
- ・そこで、臨床試験の重複を最小限にし、海外データを日本での新薬申請 (NDA) データとして用いるために、健常人での薬物動態について日本人と他の人種との差の程度を計るとともに、用量比例性および安全性についても検討する。

2) 試験デザイン設計書 海外の成績

海外での成績

海外での試験の結果より、R 薬の AUC は用量比例的であり、単位用量当たり

の \log_{10} (AUC) の平均および標準偏差は 2.58, 0.127 であった.

例数設計方式

海外の試験の AUC の幾何平均 $380.2\text{ng}\cdot\text{hr}/\text{mL}$ の 80% から 125% の範囲 (304.2 ~ 475.2) に, 本試験での幾何平均の 90% 信頼区間が入るような確率が 80% と なる症例数を目標とする.

3) プロトコールの記載例

- SAS を用いて, 正規乱数発生による 10000 回のモンテカルロシミュレーションにより症例数を算出した.
- その結果, 17 例の場合に 80% の確率で海外のデータと同程度との判定ができる.
- 用量比例性の検討を行うことも考慮し, 各用量で (2mg, 4mg, 8mg, 16mg) 5 例とする.

4) シミュレーション結果

シミュレーションは, 海外の試験を対象に, 日本で行う試験の AUC の平均と標準偏差が同じと仮定して, 生物学的に同等と判定できる例数を求める. 表 7.3 に SAS の data ステップによるシミュレーション・データ発生を示す. 実験に必要な症例数を 6~26 まで逐次的に増加させ, 各々の症例数で 10000 組のデータを発生させている.

表 7.3 シミュレーション Data の作成

```
data D01 ;                               /* PIS11B.sas */
  Seed   = 20010311 ;
  Lmean0 = 2.58 ;
  Lsd0   = 0.127 ;
  Nitr=10000; N1=6 ; N2=26 ; N3=2 ;
  do Nsamp=N1 to N2 by N3 ;
    do K = 1 to Nitr ;
      do i = 1 to Nsamp ;
        logY = Lmean0 + Lsd0*rannor(Seed) ; output ;
      end ; end ; end ;
  keep Lmean0 Nsamp K logY ;
```

信頼区間方式による例数設計, Lmean0 が参照する平均, Lsd0 が標準偏差, 対数正規分布に従うデータを発生させている.

表 7.4 では, 発生させたデータの 1 組ごとの平均値と SE を proc MEANS で計算し, その 90% 信頼区間を求めている. 再度 data ステップで, 90% 信頼区間が, 参照とする試験

の幾何平均 380.2ng・hr/mLの 80%から 125%の範囲 (304.2~475.2) に入れば合格とし、どちらかの片方でも範囲を超えれば不合格としている。

表 7.4 信頼区間の計算と同等性の判定

```
proc means data=D01 noprint ;
  var logY ;
  by Lmean0 Nsamp K ;
  output out=D02 mean=Lmean1 stderr=Lsel1 ;

data D03 ;
  alpha=0.10 ; set D02 ;
  C80 = 10**(Lmean0)*0.80 ;
  C125= 10**(Lmean0)*1.25 ;
  T = tinv(1-alpha/2, Nsamp-1) ;
  L1 = 10**(Lmean1 - T*Lsel1) ;
  U1 = 10**(Lmean1 + T*Lsel1) ;
  if (L1<C80) or (U1>C125) then JUDGE=' 1:No' ;
  else JUDGE=' 0:OK' ;

  keep Nsamp Judge ;

proc tabulate data=D03 noseps ;
  class Nsamp Judge ;
  table Nsamp, (Judge=' ' all)*f=10. / rts=10 ;
  keylabel n=' ' ;
run ;
```

表 7.5 シミュレーション 結果

	0:OK	1:No	All
Nsamp			
6	1648	8352	10000
8	2937	7063	10000
10	4573	5427	10000
12	6011	3989	10000
14	7171	2829	10000
16	7897	2103	10000
18	8575	1425	10000
20	9026	974	10000
22	9268	732	10000
24	9560	440	10000
26	9651	349	10000

「0:OK」が、90%の信頼区間が80%~125%の間に入っている数である。

Nsamp が 16 のとき 7897/1000=78.97%, Nsamp が 18 のとき 8575/1000=85.75%で同等と判定された。

従って、80%の確率で同等と判定するための症例数として 17 症例が必要となる。

7.3. 民族的要因・演習 1

AUCの比較

USAで行われた薬物動態試験と日本で行われた薬物動態試験から民族的な違いを検討し、今後の臨床試験に用いる投与量を検討することになった。表 7.6 に同一用量を投与して、代表的な薬物動態パラメータであるAUCを得た。以下の設問に答えよ。

表 7.6 健康成人における薬物動態

Ethnic	No	体重	AUC
JAPAN	1	55	130
	2	60	140
	3	60	125
	4	65	125
	5	70	130
	6	75	120
USA	7	65	135
	8	70	115
	9	80	130
	10	80	115
	11	85	125
	12	90	105

- 1) 実施場所に由来する民族に違いはあるか？ あるとすれば何が異なるのか？
- 2) AUC を縦軸，実施場所を横軸とした散布図を書き，実施場所による違いについて考察せよ。

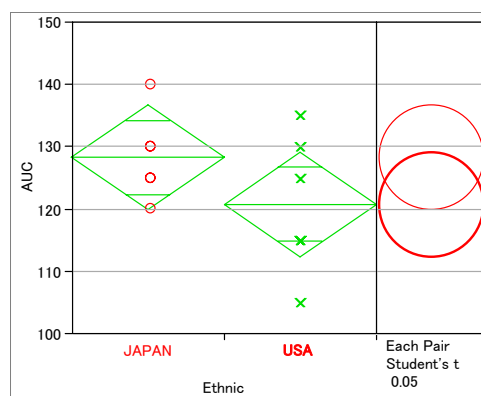


図 7.1 AUC の比較

- 3) AUC について，2 群間の差の t 検定による p 値を求め，その有意性について検討し，民族に由来する被験者の特性について考察せよ。

- 4) 体重について、AUC と同じように実施場所による違いについて考察せよ。
- 5) 次の表は、ある被験者にある薬物の経口投与後の血中の薬物濃度である。このデータを Excel または JMP にコピー&ペーストし、時間・薬物濃度曲線を描きなさい。

表 2 血中の薬物濃度の例

TIME, hr	Y ng/dL
0	.
1	42
2	267
3	399
4	385
6	291
8	195
12	72
24	10

- 6) AUC, Cmax, Tmax を計算せよ。

答は (3088.5, 399, 3) である。

- 7) 体重の重さは、AUC と相関があるか。実施場所ごとに検討せよ。50%の確率楕円を生データのプロット上に図示せよ。

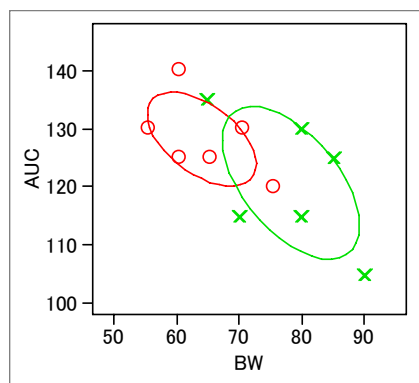


図 7.2 体重と AUC の散布図および 50%確率楕円

○ : 日本, × : USA

- 8) 体重と AUC 相関があるとしたら、因果関係はあるか。その原因は何か？
- 9) ICH : E3 ガイドライン, 11.4.2.1 節の「共変量」とは何か？
- 10) 予後因子と共変量の違いは何か？

- 11) 民族は AUC についての共変量か？
- 12) 体重は, AUC の予後因子か？
- 13) 共変量による調整とは何か.

体重によるAUCの調整

「体重による AUC の調整」は, 2 つの民族について, 切片は異なるが, 傾きは共通の回帰直線を当てはめ, 2 群の平均体重における AUC の推定値を比較することである. この方法は, 1 群の回帰分析を拡張したもので, 変数 J を, 被験者が JAPAN の場合は $J=1$, USA のばあいは $J=0$ とし, 体重も含めて回帰分析をすることによる計算することができる.

表 7.7 回帰係数の推定値

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	168.111	24.284	6.923	<.0001
J	-1.050	6.408	-0.164	0.8734
BW	-0.604	0.307	-1.965	0.0809

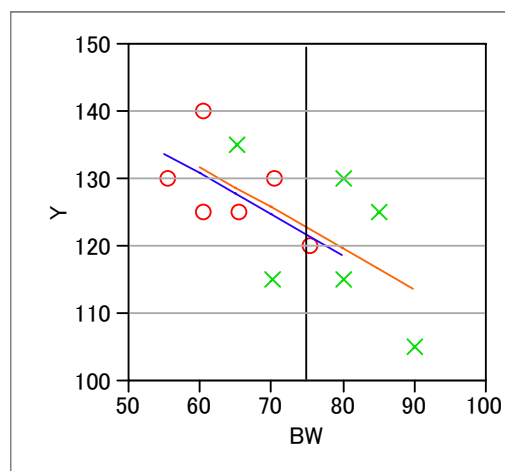


図 7.3 平行な回帰直線の当てはめ

○ : 日本, × : USA

- 14) 2 群別々の回帰式を参考にしながら, 傾きが 2 群で共通な回帰式をそれぞれ求めよ.
- 15) 2 群の平均体重は, 75kg である. 2 群の体重が 75kg のときの AUC の推定値を求めよ.
- 16) AUC の推定値の差とその 95%信頼区間を計算せよ.

7.4. 民族的要因・演習 2

データの概要

ECとJPで全体で5回の臨床試験が行われ、血中の薬物濃度が測定され、154人について表 7.10 に示すようにAUCが計算されている。これらの臨床試験から、民族的要因について検討せよ。それぞれの試験は、次のようである。データ・ファイル名は、Ethnic_EC_JPである。

表 7.8 試験の概要

試験番号	実施場所	対象	投与用量	人数
1	EC	健康人	3 用量	48 人
2	EC	患者	1 用量	34 人
3	日本	健康人	3 用量	24 人
4	日本	健康人	1 用量	8 人
5	日本	患者	4 用量	40 人

それぞれの試験の投与量ごとに人口統計学的なデータの要約統計量と AUC についての平均値を示す。基本的な解析結果を、前節と同じような考え方で、民族的な要因について考えよ。

表 7.9 用量別の年齢，体重，および AUC の平均

Study	Code	DOSE	n	男/女	Mean(AGE)	Mean(BW)	Mean(AUC)
1	EH	0.5	16	16/0	32.13	79.86	67.26
		1	16	16/0	30.19	78.93	130.91
		2	16	16/0	28.19	82.54	245.28
2	EP	0.5	34	0/34	51.09	68.20	95.14
3	JH	0.125	8	8/0	22.00	60.70	17.10
		0.25	8	8/0	23.13	66.82	34.38
		0.5	8	8/0	21.63	63.25	77.20
4	JH2	0.5	8	8/0	22.88	62.51	74.43
5	JP	0.25	10	0/10	62.90	54.03	74.41
		0.5	10	0/10	64.60	54.62	136.82
		1	10	0/10	63.70	56.84	239.87
		2	10	0/10	64.40	50.62	540.73

AUC/doseを用いたプロット

試験間に明らかな AUC の違いが観察される。年齢と AUC, 体重と AUC について散布図を示す。年齢が増加すると AUC が大きく、体重が増加すると AUC が減少することが観察される。

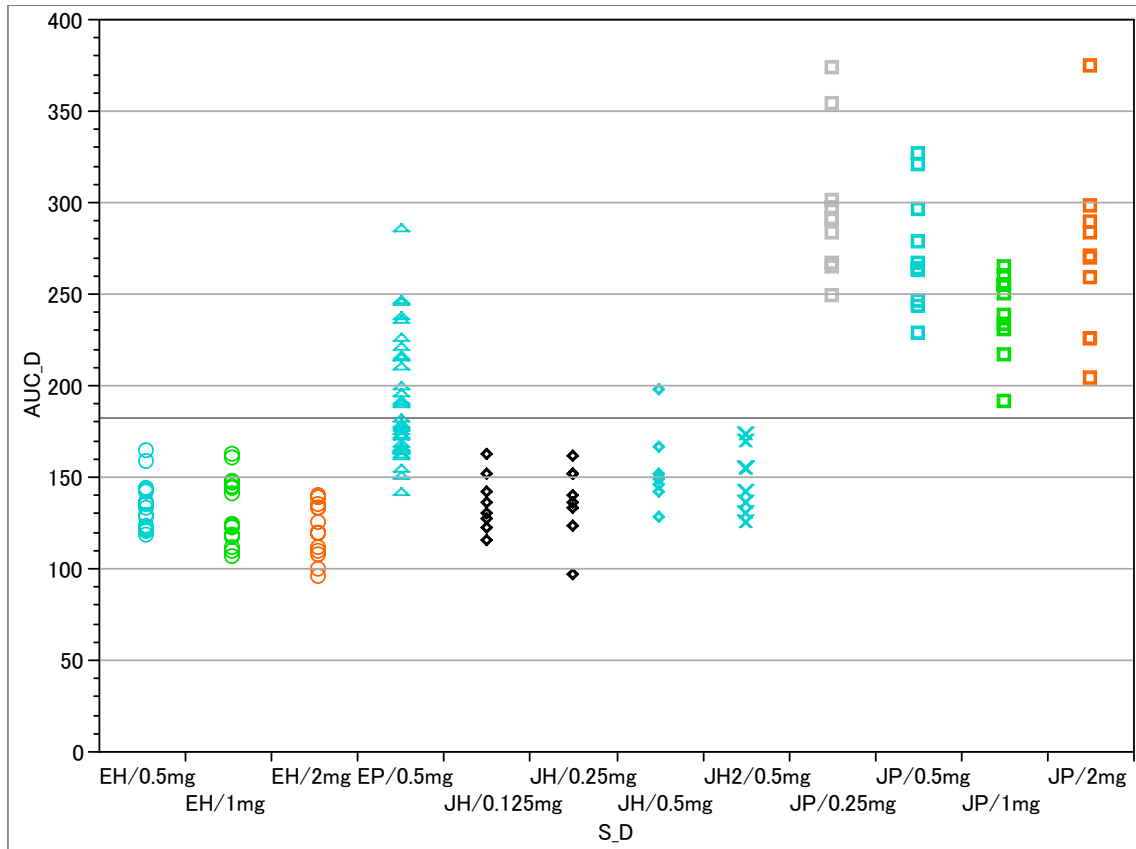


図 7.4 試験別用量別の AUC/dose の散布図

○ : EC,健康人,男, △ : EC,患者,女, ◇ : JP,健康人,男, × : JP,健康人,男, □ : JP,患者,女
 総平均は 182.56ng・hr/mL である。

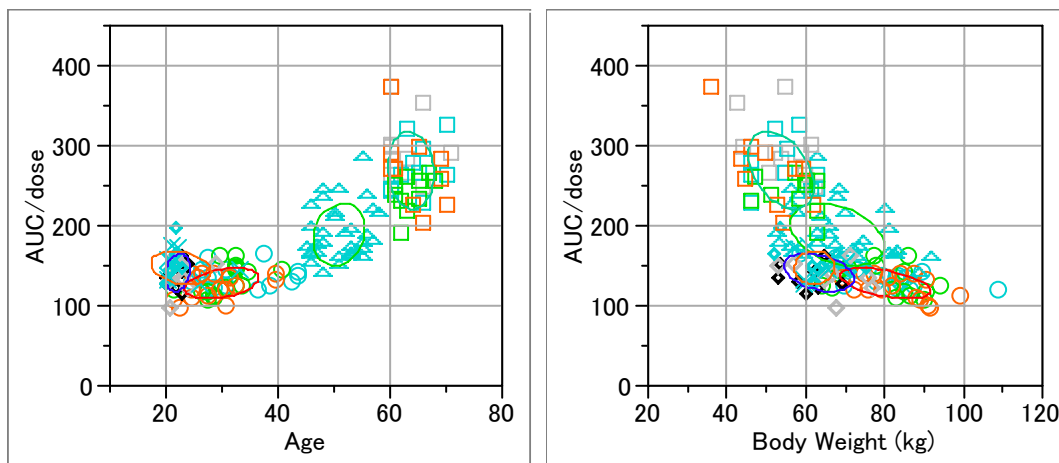


図 7.5 年齢および体重と AUC/dose の関連

○ : EC,健康人,男, △ : EC,患者,女, ◇ : JP,健康人,男, × : JP,健康人,男, □ : JP,患者,女
 各試験ごとに確率楕円 (50%) を示した

用量反応

用量反応は、3 試験ともほぼ用量比例的であるが、試験番号 5 は、試験番号 1 と試験番号 3 に比べて、明らかに低用量側にシフトしている。

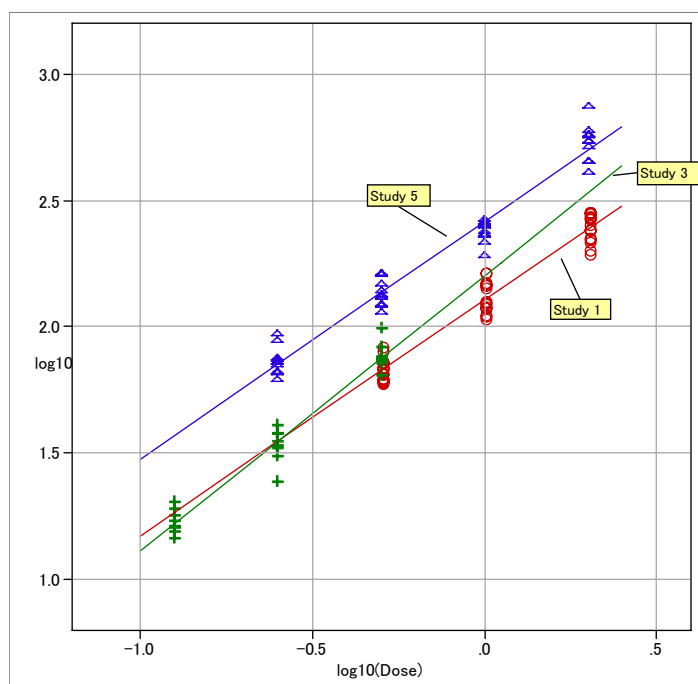


図 7.6 両対数での回帰直線の当てはめ

Term	傾き Estimate	Std Error	Lower 95%	Upper 95%	用量比例
STUDY=="1"	0.9312552	0.030654	0.8695512	0.9929591	×
STUDY=="3"	1.0858769	0.048396	0.9855091	1.1862446	○
STUDY=="5"	0.9369954	0.028792	0.8787091	0.9952817	×

年齢と体重を共変量とした解析

- 1) 試験を固定効果，年齢と体重を共変量とした共分散分析を行う。

```
/* AGE10=AGE/10, BW10=BW/10 */
proc mixed data=D02 ;
  class STUDY SEX ;
  model AUC_D_L = STUDY AGE10 BW10 / solution ;
run ;
```

Solution for Fixed Effects						
Effect	STUDY	Estimate	Standard Error	DF	t Value	Pr > t
Intercept		2.4248	0.07511	147	32.28	<.0001
STUDY	1	-0.1563	0.04116	147	-3.80	0.0002
STUDY	2	-0.08000	0.02058	147	-3.89	0.0002
STUDY	3	-0.1433	0.04778	147	-3.00	0.0032
STUDY	4	-0.1289	0.04978	147	-2.59	0.0106
STUDY	5	0
AGE10		0.02519	0.01085	147	2.32	0.0216
BW10		-0.02933	0.005292	147	-5.54	<.0001

- 2) 年齢と体重の回帰係数を用いて，すべて被験者を年齢が 50 歳，体重が 60kg となるように $\log(\text{AUC}/\text{dose})$ を調整する。

$$\text{AUC}_{\text{adj}} = \text{L_AUC_D} + (50 - \text{AGE}) * 0.002519 + (60 - \text{BW}) * -0.002933$$

- 3) 調整済み $\log(\text{AUC}/\text{dose})$ について生物学的同等性の基準を当てはめ民族的要因を検討する。

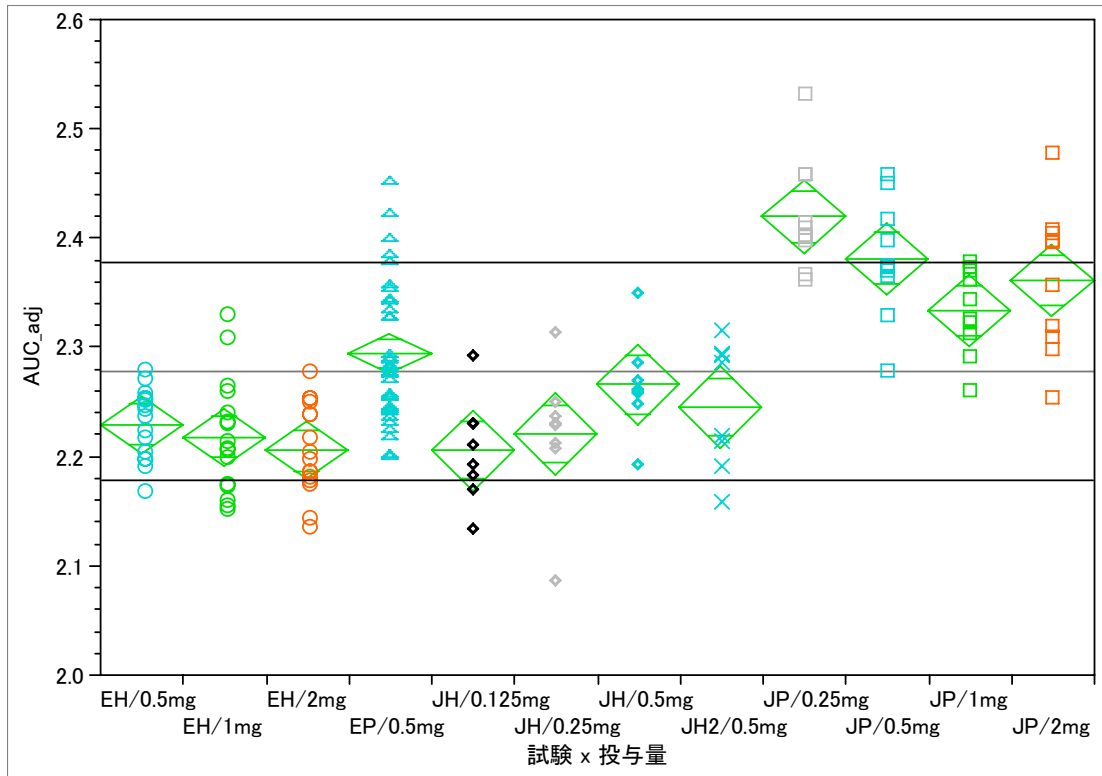


図 7.7 試験別用量別調整済み $\log(AUC/dose)$ の散布図

○ : EC,健康人,男, △ : EC,患者,女, ◇ : JP,健康人,男, × : JP,健康人,男, □ : JP,患者,女
総平均は 2.279, 同等性基準範囲 (2.179~2.279), 菱形の中央が平均, 上下が平均値の 95%信頼区間
(この図は JMP で作成したものであり, 90%信頼区間の表示機能がない)

図 7.4 に比べて試験 1~5 の被験者のプロットが総平均の方向にシフトしていることが読み取れる.

7.5. データリスト

表 7.10a ETHNIC=EC, STUDY=1, SEX=male

0.5mg				1mg				2mg			
ID	AGE	BW	AUC	ID	AGE	BW	AUC	ID	AGE	BW	AUC
2	26	72	71.46	1	32	77	147.73	3	39	90	264.83
5	27	108	58.97	7	21	81	118.22	6	24	74	278.27
11	36	85.8	60	9	32	67.8	144.89	10	28	86.5	278.39
15	24	77	64.02	13	34	84	140.64	16	39	83	278.15
23	43	82.5	68.02	21	32	70	146.15	17	24	71.5	239.11
24	27	84.5	66.45	22	30	66	121.84	18	24	87.5	218.81
27	38	87	61.26	28	27	89.5	106.58	26	27	98.5	222.39
29	42	83	64.27	32	28	82	109.62	30	21	88.5	213.87
34	25	82	67.37	38	29	76.6	160.71	36	30	78.5	237.92
37	31	78.5	68.02	40	27	90	116.89	39	26	76.5	278.47
42	27	67	79.21	47	33	93.5	123.17	43	32	81.4	249.27
45	30	69	61.72	48	32	85	162.22	46	29	75	238.62
49	37	58	82.18	50	40	67.5	143.71	51	31	68.5	266.57
53	28	82	60.61	55	32	63.5	124.5	52	22	91	191.4
62	30	72.5	71.99	58	26	84	116.84	57	30	90.3	198.34
64	43	89	70.68	60	28	85.5	110.84	63	25	80	270

表 7.10b ETHNIC=EC, STUDY=2, SEX=female

0.5mg											
ID	AGE	BW	AUC	ID	AGE	BW	AUC	ID	AGE	BW	AUC
1	48	88.6	71.015	13	47	67.5	88.445	25	55	92	80.924
2	52	52.5	96.073	14	53	68	82.05	26	53	73.5	86.489
3	51	73	80.452	15	56	58.5	119.1	27	56	66	94.765
4	55	63	142.83	16	47	81	82.711	28	57	80	110.55
5	48	68.5	123.39	17	54	83	83.863	29	46	66	99.667
6	48	68	117.74	18	57	74.5	90.707	31	53	81	97.664
7	48	53	89.424	19	50	58	107.92	32	46	56.5	113.16
8	53	82.5	82.403	20	53	57	82.713	34	55	63	88.812
9	46	68.3	82.889	21	58	70	107.54	35	51	68	75.482
10	45	56.2	87.493	22	55	72	85.912	36	46	59.5	82.104
11	48	62.8	95.27	23	46	67	77.082				
12	51	56.5	122.97	24	50	64	105.21				

表 7.10c ETHNIC=JP, STUDY=3, SEX=male

0.125 mg				0.25 mg				0.5 mg			
ID	AGE	BW	AUC	ID	AGE	BW	AUC	ID	AGE	BW	AUC
1	22	63.2	15.437	11	21	64.8	33.385	21	20	62.4	64.523
2	23	60.6	14.524	12	21	67.9	24.47	22	22	52.3	83.359
3	20	53.3	17.152	14	29	57.5	38.173	23	23	75.6	74.563
5	21	58.1	16.326	16	23	53.1	38.098	24	21	63.6	73.185
6	23	64.7	20.387	17	22	71.5	40.651	25	22	53.9	99.346
8	23	62.6	17.88	18	22	66.8	35.146	27	22	64.5	75.133
9	20	69.5	16.039	19	24	77	30.941	28	21	71.6	71.495
10	24	53.6	19.092	20	23	76	34.161	30	22	62.1	76

表 7.10d ETHNIC=JP, STUDY=4, SEX=male

0.5 mg			
ID	AGE	BW	AUC
52	22	56.9	85.101
53	20	66.4	77.745
54	21	70.1	77.492
55	26	59.5	62.869
56	20	59.3	65.518
57	20	62.1	68.547
59	21	60.6	86.777
60	33	65.2	71.415

表 7.10e ETHNIC=JP, STUDY=5, SEX=female

0.25 mg				0.5 mg				1 mg				2 mg			
ID	AGE	BW	AUC	ID	AGE	BW	AUC	ID	AGE	BW	AUC	ID	AGE	BW	AUC
1	60	44	74.445	22	60	61.6	121.72	42	63	63	217.61	61	70	62.1	451.38
2	71	58.9	73.02	23	70	58.1	163.43	43	63	47	260.43	63	66	54.5	408.9
3	60	61.3	75.452	24	66	55.5	148.09	44	62	62.8	191.53	64	64	52.9	451.86
4	66	42.4	88.694	25	65	54.9	133.59	41	65	62.7	255.74	65	60	59.4	541.07
7	63	53.7	71.059	26	60	62.7	123.33	45	62	45.9	230.55	66	60	36	749.55
8	66	62.3	62.282	27	63	52.2	160.62	47	61	51	238.32	67	61	57.4	542.4
10	60	52.2	72.615	29	66	46.3	114.45	49	61	60	250.03	68	65	46	596.63
11	63	50.7	66.72	30	70	46	131.98	50	68	60.4	255.09	69	60	49.9	580.1
12	60	60.1	66.223	32	64	46.1	139.52	51	65	58.4	233.87	70	69	44.4	517.67
13	60	54.7	93.579	33	62	62.8	131.52	52	67	57.2	265.52	71	69	43.6	567.72

文献

高橋行雄 (1997), 薬物濃度を用いた同等性試験の例数設計, 第 16 回日本 SAS ユーザー会論文集, 149-152.

8. 交互パネルによる PK パラメータの解析事例

8.1. 実験デザイン

この実験は、被験者を2つのパネルに分け、交互に単回投与量を逐次的に増やす実験モデルである。そして、Placebo が、各試験日に2人、最終的にすべての被験者に投与されるようにデザインされている。被験者 No.6 は、途中で別な人に交代があったので区別して取り扱うために No.6' とした。

表 8.1 交互パネルによる投与スケジュール

パネル	投与順	被験者	試験日							
			1	2	3	4	5	6	7	8
1	1	3, 5	0mg		10mg		20mg		40mg	
	2	2, 4	2.5mg		0mg		20mg		40mg	
	3	7, 8	2.5mg		10mg		0mg		40mg	
	4	1, 6, 6'	2.5mg		10mg		20mg		0mg	
2	5	9, 14		0mg		15mg		30mg		50mg
	6	10, 13		5mg		0mg		30mg		50mg
	7	12, 16		5mg		15mg		0mg		50mg
	8	11, 15		5mg		15mg		30mg		0mg

パネル (Panel) : 同じ実験スケジュールで管理される被験者、パネルには、羽目板などの意味の他に討論会の討論者団の意味でも用いられており、ここでは後者の意味で用いている。

投与順 (Sequence) : 同一の薬物が投与される被験者の集まり、A 群 (group), B 群のように表現する場合もできるが、10mg 群, 20mg 群, などの表現と紛らわしいので投与順とする。

被験者 (Subject) : 被験者番号またはコード、一人一人が、一意に識別できる番号とすること。

試験日 (Study Day, Period) : 薬剤の投与が行われる日で、試験日と定義する。生物学的同等性試験 (BE 試験) では、時期と表現されてことが多い。第 1 相試験の場合には複数の「時期」があるので試験日とした。

8.2. Cmax についての用量反応

経口投与後の薬物濃度曲線から、AUC, Cmax, Tmax, $t_{1/2}$ などの PK パラメータが求められているが、ここでは Cmax について用量で整理した結果を表 8.2 に示す。同一被験者ごとに用量反応関係を見たのが 図 8.1 である。どの被験者の用量反応も直線的である。投与量の増加に従い分散が大きくなっているため、用量および Cmax を対数変換した図も合わせて示す。

表 8.2 投与順, 投与量で整理した Cmax

パネル	投与順	被験者	2.5mg	5mg	10mg	15mg	20mg	30mg	40mg	50mg	
1	1	3	.	.	22.20	.	52.80	.	126.10	.	
		5	.	.	45.50	.	60.30	.	114.80	.	
	2	2	6.20	.	.	.	65.40	.	101.70	.	
		4	8.70	.	.	.	49.60	.	103.50	.	
	3	7	6.00	.	21.30	.	.	.	110.80	.	
		8	4.90	.	24.40	.	.	.	127.60	.	
	4	1	7.30	.	29.10	.	48.50	.	.	.	
		6	7.70	.	35.00	
		6'	42.50	.	.	.	
	2	5	9	.	.	.	31.80	.	87.66	.	134.00
14			.	.	.	41.50	.	87.51	.	98.98	
6		10	.	12.50	.	.	.	71.53	.	121.03	
		13	.	12.60	.	.	.	78.01	.	124.42	
7		12	.	15.90	.	52.20	.	.	.	104.37	
		16	.	21.60	.	45.10	.	.	.	104.70	
8		11	.	24.20	.	36.60	.	80.99	.	.	
		15	.	16.30	.	37.40	.	91.52	.	.	
平均値			6.80	17.18	29.58	40.77	53.18	82.87	114.08	114.58	
SD			1.36	4.78	9.30	7.20	8.35	7.42	10.99	13.86	
CV			20.0	27.8	31.5	17.7	15.7	8.9	9.6	12.1	

混合線形モデルによる用量比例性の検討

一人の被験者に複数の用量投与がなされているので、共分散分析の考え方に基づいて、それぞれの被験者に共通の傾きを持つ回帰直線を当てはめる必要がある。被験者ごとに直線を引くと 図 8.1 のようになる。対数変換後の用量反応を見ると、ほとんど全ての被験者の用量反応は平行的である。

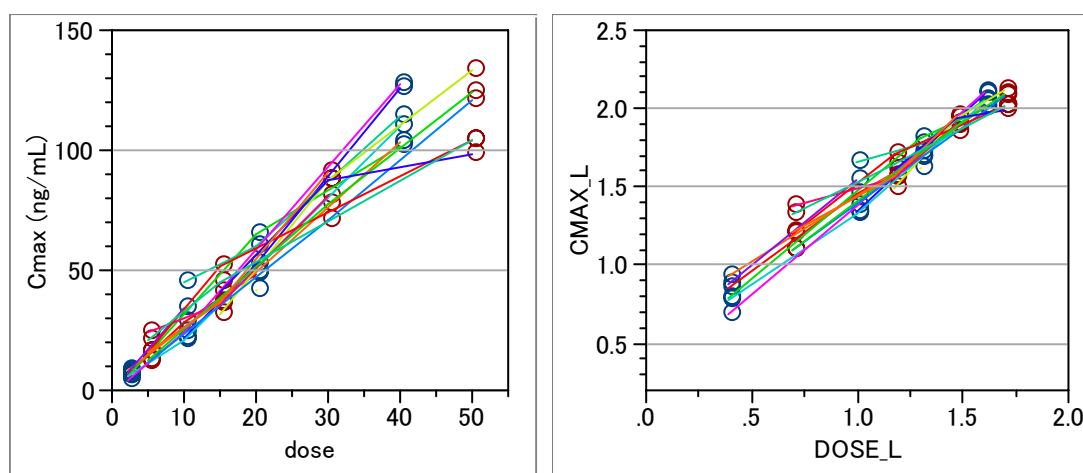


図 8.1 個人ごとの用量反応

対数変換後のデータに、被験者をランダム効果とした回帰分析により用量比例性を検討する。

```
proc mixed data=DD.pk ;
  class ID ;
  model Cmax_L = Dose_l / cl ;
  random ID ;
run ;
```

傾きは 0.9495, 95%信頼区間 (0.8895 ~ 1.0095) と 1.0 を含んでいるので, 用量比例性が成り立っている。

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
ID	0.00000000
Residual	0.00737237

Solution for Fixed Effects

Effect	Estimate	Std Error	DF	t	Pr > t	Alpha	Lower	Upper
INTERCEPT	0.49702713	0.03649790	16	13.62	0.0001	0.05	0.4197	0.5744
DOSE_L	0.94952037	0.02936587	30	32.33	0.0001	0.05	0.8895	1.0095

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
DOSE_L	1	30	1045.50	0.0001

信頼区間方式での表示

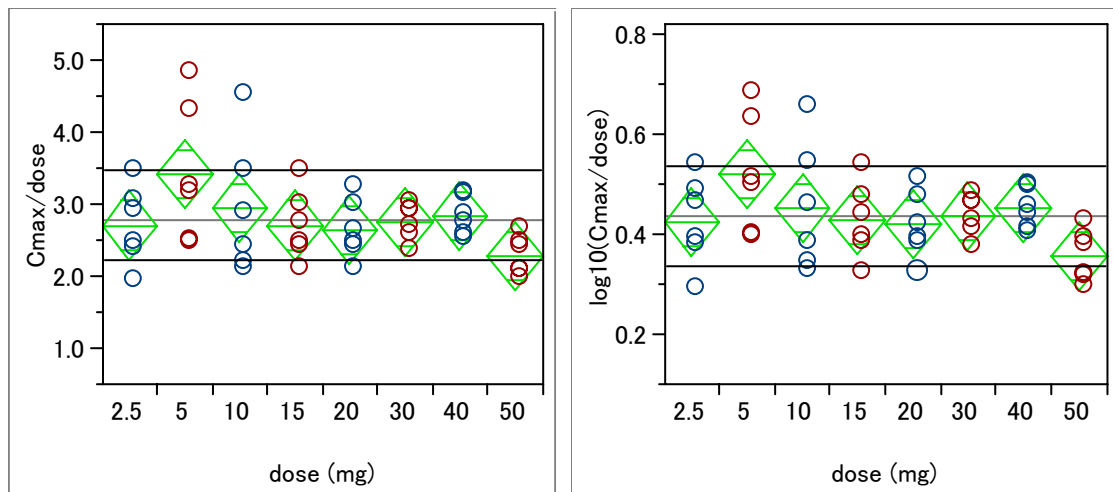


図 8.2 単位用量当りの Cmax, 個別データ

生物学的同等性領域は, 生データでは総平均の 80%~125%であり,
常用対数では総平均の -0.10 ~ +0.1 である。

文献

高橋行雄, 大島信之 (1997), 交互パネル法による臨床第 1 相試験, 医薬安全研 70 回定例会資料.

高橋行雄 (1997), 薬物濃度を用いた同等性試験の例数設計, 第 16 回日本 SAS ユーザー会論文
集, 149-152.

9. 交互パネルによるタッピング検査の解析事例

表 8.1 で示した交互パネルの実験では、同時に様々な検査が行われている。その中で、薬剤の単回投与した後に起きるタッピング能力の短時間的な作用を検討してみよう。タッピング検査は、図 9.1に示すように、投与前日、単回投与後の 1.5 時間、24.5 時間、48.5 時間に行なわれている。0mg は、パネル 1 とパネル 2 の 16 人分の結果であり、40mg はパネル 1 の 6 人分の結果が示されている。

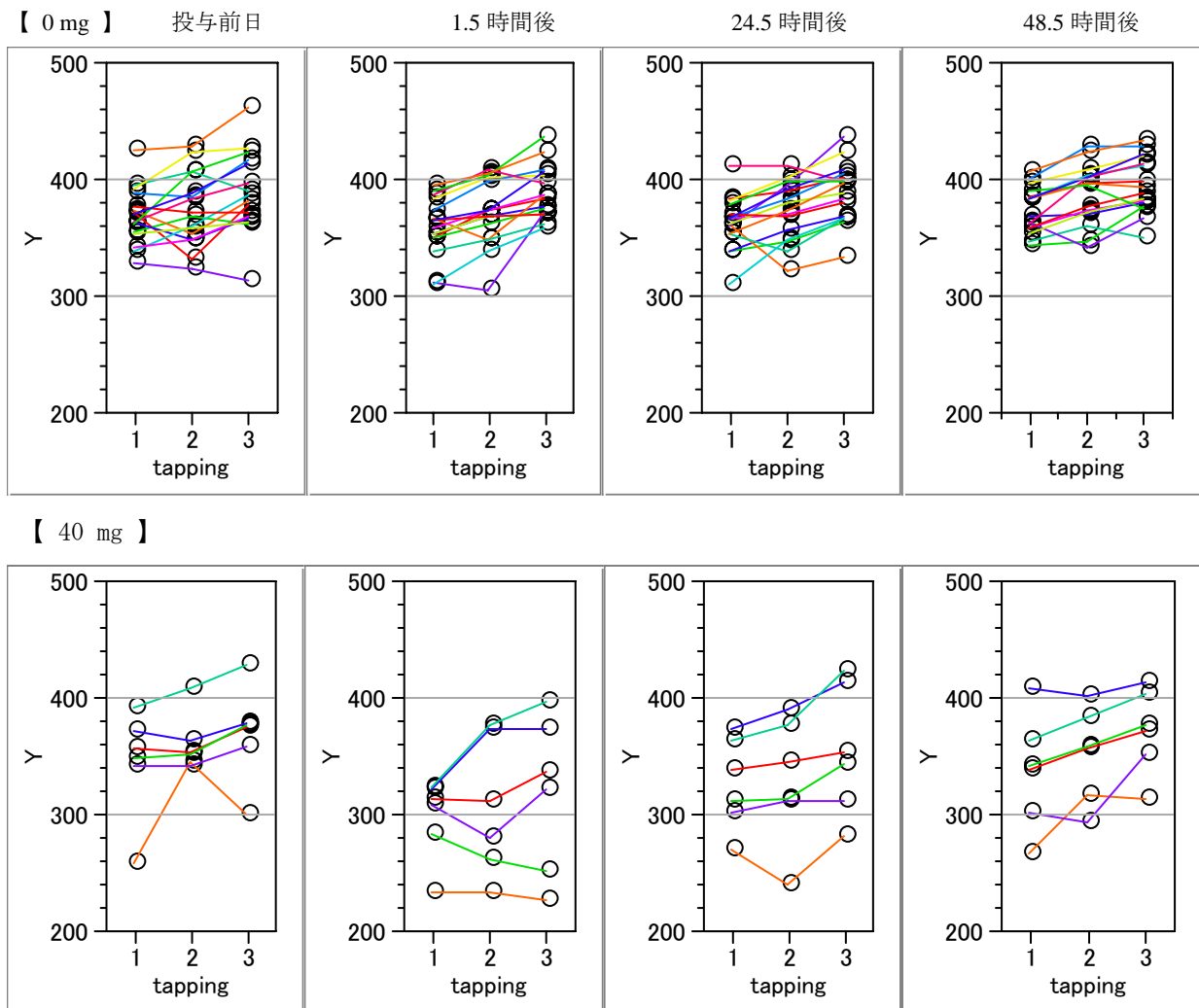


図 9.1 0mg , 40mg 投与時のタッピングの学習効果

タッピング検査は、一時点について 3 回の試行があり、学習効果により 1 回目より 2 回目、2 回目より 3 回目のほうがデータが一般的に大きくなることが知られている。40mg 群では、投与前日に 6 人中 1 人を除き学習効果が見られるが、1.5 時間目には、2 人に

は投与前日とは異なる学習効果の抑制が観察されている。その後、48.5 時間後には、もとに戻っている。

9.1. 検証すべき仮説は何か、学習効果の抑制、促進？

投与量の増加により、タッピングテストの学習効果が、抑制されていないことを示したいとしよう。0mg の結果から、投与前日、単回投与後 1.5 時間後、24.5 時間後、および 48.5 時間後それぞれの時点ではほとんど全ての被験者において傾きがプラスとなっており、学習効果は抑制されていない。

タッピング検査での“学習効果”を、3 回の試行の平均的な伸びと定義しよう。統計的には、タッピングの試行の 3 ポイントに回帰直線を当てはめ、その傾きにより求めることにする。

学習効果の抑制は、傾きが寝てくることであり、

学習効果の亢進は、傾きが立ち上がることであり、

により評価する。40mg の 1.5 時間目について回帰直線を当てはめたのが 図 9.2 である。

被験者 3 と 5 は、傾きがマイナスとなり学習効果の抑制が見られる。

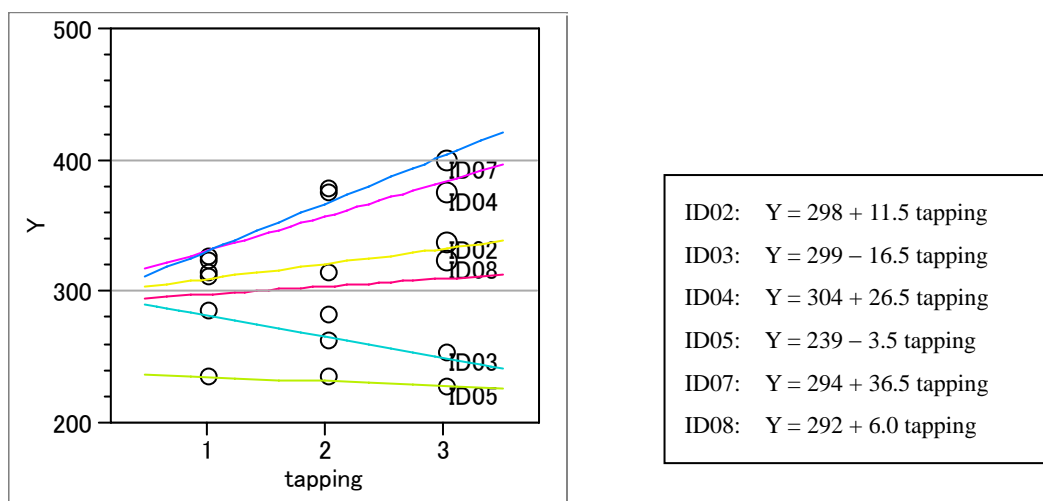


図 9.2 学習効果の計量, 40mg 群の 1.5 時間目

表 9.1 にパネル 1 の被験者について 1.5 時間目の“学習効果”を示す。このデータは、前章の表 8.2 で示した Cmax と同様の形式に学習効果が要約されたことになる。さて、検討したいのは、投与量の増加にしたがって、学習効果が抑制されるかであり、抑制さ

れるとすれば、その用量反応関係の形状が知りたいのである。統計的には、各被験者に投与量を X として学習効果を Y とし、被験者をランダム効果とした線形混合モデル（ランダム係数モデル）を考えよう。係数が負ならば抑制、正ならば亢進となる。

表 9.1 パネル1の投与後1.5時間目の学習効果

被験者	0mg	2.5mg	10mg	20mg	40mg
1	17.0	12.5	8.0	29.5	.
2	12.5	1.0	.	6.5	11.5
3	11.0	.	7.0	13.0	-16.5
4	10.0	24.5	.	41.0	26.5
5	11.0	.	4.0	-0.5	-3.5
6	.	3.5	8.5	.	.
6'	30.0	.	.	43.0	.
7	9.5	25.0	19.5	.	36.5
8	24.5	-0.5	25.0	.	6.0

図 9.3 に投与前日と投与後1.5時間の被験者ごとの用量反応関係を示す。投与1.5時間後には、投与前に比べ学習効果が被験者によりバラツキが大きくなっているようであるが、用量の増加に伴う傾きの低下は見られない。

この薬剤の経口投与後の血中の薬物濃度が最大になるのは、平均で0.9時間後、SDは0.5時間であることから、1.5時間目は、血中の薬物濃度が最大になる時間と一致しており、薬剤の短期的な作用を計量するためのポイントとして主要変数とみなせる。

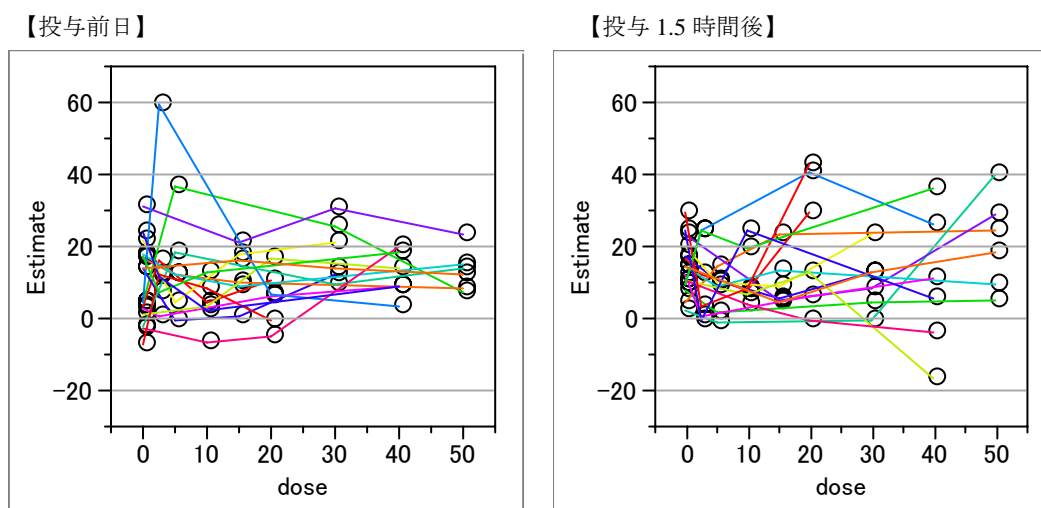


図 9.3 用量の増加による学習効果の変化

tapping 検査の3回の施行の傾きがY軸である。被験者ごとに複数の用量が与えられており、それを直線で結んだ。用量が増加しても水平方向の変化なので、学習効果の抑制も亢進も起きていないことがわかる。

表 9.2 にランダム係数モデルによる解析を、時点ごとに行った結果を示す。学習効果の用量反応（95%信頼区間）は、投与前日で 0.0368 (-0.1214~0.1951) 回/dose, 1.5 時間目で 0.1068 (-0.0473~0.2609) 回/dose, 24.5 時間目で 0.104 (-0.0392~0.2479) 3 回/dose, 48.5 時間目で 0.1787 (0.0390~0.3184) 回/doseと用量の増加に伴い抑制は見られず、逆に 48.5 時間目では亢進 ($p=0.0133$) している。但し、注意しなければならないのは、用量の増加に伴いタッピング検査の経験の回数が増えることと、実施時期が投与量と交絡しており、学習効果の亢進が「これでつらい試験ももうすぐ終わりだ、がんばろう」という効果が入っている可能性を否定できない。

表 9.2 時点ごとのランダム係数モデルの傾きの推定値と 95%信頼区間

時間	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
投与前日	0.0368	0.0786	0.47	0.6418	-0.1214	0.1951
1.5 時間後	0.1068	0.0766	1.39	0.1699	-0.0473	0.2609
24.5 時間後	0.1043	0.0713	1.46	0.1503	-0.0392	0.2479
48.5 時間後	0.1787	0.0694	2.57	0.0133	0.0390	0.3184

それぞれの時間ごとにランダム係数モデルに統計を行い、傾きの推定値、および、その 95%信頼区間を求めた。

実験がラテン方格で割り付けられていれば、実施時期に交絡する効果を統計的に除去できるが、この実験デザインは逐次増量となっているので、残念ながらこれを分離できない。何を実験の主要変数とするかによって、実験デザインを考えなければ、せっかくの実験結果に陰りを残すことになる。

9.2. タッピング数に対する薬物の影響

用量に依存する学習効果の抑制がなかったため、次に、タッピングの回数の抑制について用量反応関係を検討しよう。“反応 Y ”として、3回のタッピングの回数の平均値とする。投与前日と投与後 1.5 時間後の平均値を表 9.3 に示す。

図 9.4 から、被験者によってタッピング数が、かなり異なることがわかる。すなわち、タッピングの回数が多い人はいつも多く、少ない人はいつも少ないことが観察される。分散分析の結果でも、統計的にも被験者間により明らかな差がある。

図 9.5に投与 1.5 時間後の用量反応関係を示す。用量が増えるとタッピング数が減少する様子が観察される。ただし、ほとんど減少しない被験者、急激に減少する被験者が

表 9.3 タッピング数の平均値データ

パ ネ ル	投 与 順	被 験 者	0 mg		2.5 mg		10 mg		20 mg		40 mg	
			前日	1.5hr	前日	1.5hr	前日	1.5hr	前日	1.5hr	前日	1.5hr
1	1	3	361	367	.	.	348	357	355	324	360	266
		5	398	351	.	.	292	314	331	343	302	232
	2	2	363	363	354	357	.	.	362	360	363	321
		4	370	368	347	390	.	.	402	332	372	357
	3	7	414	397	404	409	400	402	.	.	411	367
		8	362	337	357	337	326	324	.	.	348	304
4	1	362	370	351	366	326	333	352	348	.	.	
	6	.	.	310	324	333	350	
	6'	323	331	310	334	.	.	
平均			369.1	360.5	353.8	363.8	337.5	346.7	352.0	340.2	359.3	307.8

			0 mg		5 mg		10 mg		30 mg		50 mg	
			前日	1.5hr	前日	1.5hr	前日	1.5hr	前日	1.5hr	前日	1.5hr
2	5	9	354	374	.	.	371	392	379	377	394	379
		14	398	411	.	.	422	385	418	393	391	373
	6	10	360	365	389	365	.	.	371	369	397	369
		13	374	369	404	365	.	.	388	364	404	354
	7	12	382	398	396	362	396	387	.	.	380	363
		16	439	409	430	396	401	395	.	.	458	410
	8	11	397	394	328	399	380	388	376	379	.	.
		15	392	382	401	367	371	363	385	359	.	.
平均			387.0	387.8	391.3	375.7	390.2	385.0	386.2	373.5	404.0	374.7

【投与前日】

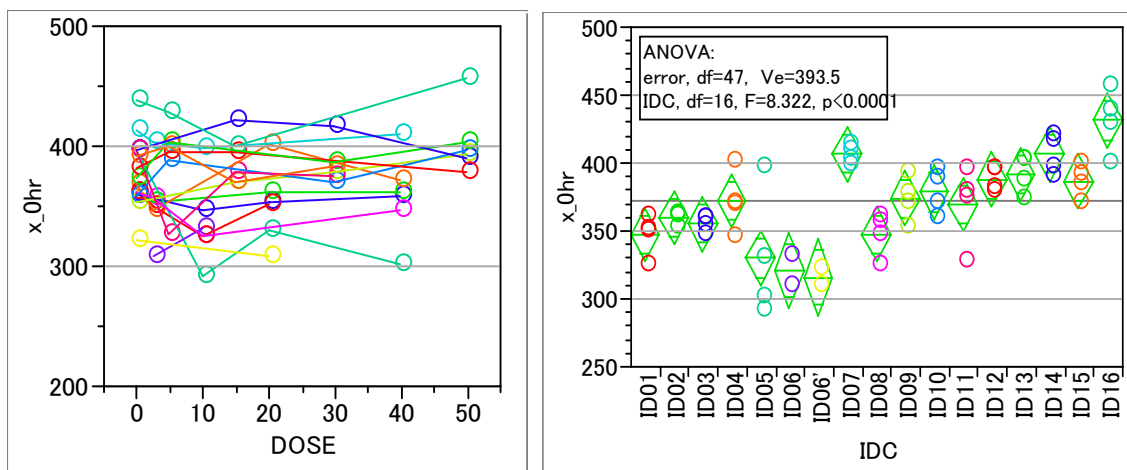


図 9.4 投与前日のタッピング回数

図左：17人それぞれについて薬剤の投与前日に“用量反応”が見られないことから、持ち越し効果のないことが確かめられた。逐次増量試験であるが、投与はランダムに行われたと見なせる。図右：被験者ごとの投与前日のタッピング数について1元配置分散分析を行った。

混在しており、一律にタッピング数が減少するわけではないようである。統計的には、被験者をランダム効果とし、共通の傾きを持つ回帰直線を被験者ごとに当てはめて、それらの回帰直線からの誤差により検定する。

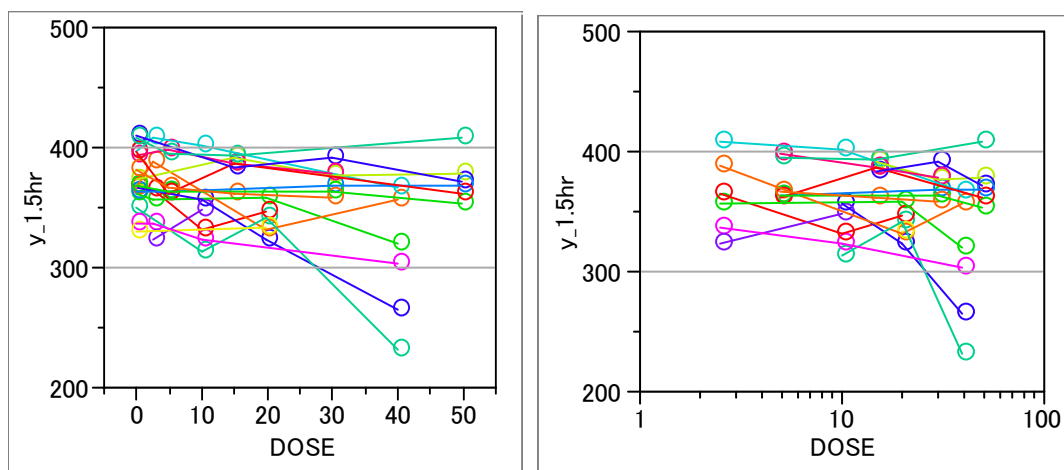


図 9.5 投与 1.5 時間後のタッピング数

17 人それぞれについて投与 1.5 時間目のタッピング数をプロットし、被験者ごとに線で結んだ。なお、投与量 0mg のデータも投与 1.5 時間後である。0mg での反応数が多い被験者に、用量を増やしてもタッピング数の減少が少ないようである。図右から、対数用量では、図左とまったく異なる用量反応関係が見出される。見出される。ただし、対数用量の図では、0mg でのタッピング数が含まれていないことに注意が必要である。

学習効果の場合と同様に、被験者をランダム効果、固定効果として投与量を連続量で与える。結果として、単位用量当たりのタッピング数が、ランダム係数モデルの傾きとして求められる。投与 1.5 時間後の変化量は、10mg 当たり - 6.34 回、95%信頼区間 (- 8.24 ~ - 4.41) が 0 を含まないので有意な差 ($p < 0.0001$) である。投与後 24.5 時間にも統計的な差が見られが、48.5 時間後には投与前日と同様に薬効は認められない。

表 9.4 時点ごとの 10mg 当たりのタッピング数の変化

時間	Estimate	SE	t	Pr> t	L95%	U95%
投与前日	0.8680	1.0056	0.86	0.3895	-1.1198	2.855
1.5 時間後	-6.3429	0.9751	-6.51	<.0001	-8.2709	-4.4149
24.5 時間後	-2.1472	0.8770	-2.45	0.0156	-3.8809	-0.4134
48.5 時間後	-0.0777	0.9149	-0.08	0.9324	-1.8863	1.7309

時間ごとに MIXED プロシジャを適用し、10mg 当たりの投与量の回帰係数 (傾き) の部分をまとめたものである。

統計的なまとめは、表 9.4 であるが、主要変数としての被験者ごとの 10mg 当たりのタッピング回数が示されていないので、表 9.5 のように時間ごと被験者ごとに回帰直線を当てはめた場合に、統計的に異なる結果を得ることになるので注意しなければならない。

表 9.5 タッピング数にたいする用量反応直線

被験者	投与前日	1.5 時間	24.5 時間	48.5 時間
1	-5.02	-12.65	-19.68	-2.09
2	1.09	-9.25	0.26	0.86
3	0.79	-26.36	-7.68	-0.94
4	4.68	-6.85	8.64	7.43
5	-17.49	-27.62	-20.17	-12.55
6	31.11	34.67	10.22	0.44
6'	-6.17	1.83	-23.67	6.50
7	0.59	-9.40	-3.21	-6.22
8	-2.12	-8.24	-6.76	-14.16
9	7.78	-0.02	0.89	5.04
10	4.24	0.88	-3.49	3.61
11	2.03	-5.95	-6.03	-5.11
12	-1.87	-4.22	2.51	0.52
13	3.04	-2.56	0.34	1.23
14	-1.99	-6.45	-1.26	4.19
15	-4.37	-6.67	-3.15	-4.11
16	5.30	1.48	3.10	2.38
平均	1.27	-5.14	-4.07	-0.76
SD	9.63	13.26	9.50	6.11
SE	2.34	3.21	2.30	1.48
U95%	6.22	1.67	0.82	2.38
L95%	-3.68	-11.96	-8.95	-3.90

被験者ごとの 10mg 当りのタッピング数の変化量を示している。

投与 1.5 時間後の 95%信頼区間が 0 を含んでいるので、統計的には有意な差ではない。

表 9.4 の p 値との食い違いが大きい。統計手法は、性能が最も良いものを実験前に選択しておかねばならない。ここで示したように事後的に行ってはならない。

表 9.5 の結果からは、投与 1.5 時間後では、10mg 当たりのタッピング数の変化は -5.14 回、95%信頼区間は (1.67 ~ -11.96) と 1 を含み統計的には差がない。ただし、データには、飛び離れ値が散見するので、箱ひげ図による評価を行う

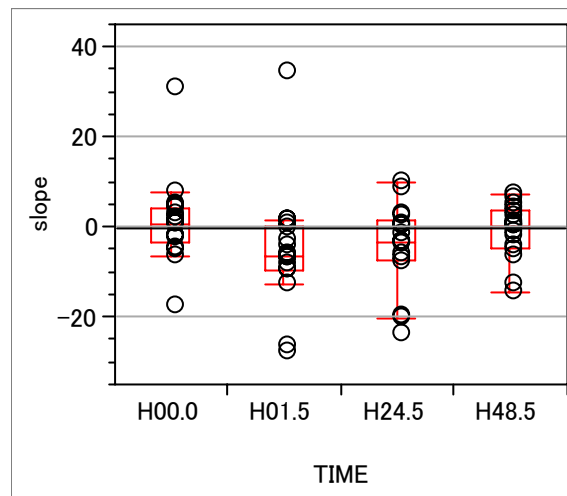


図 9.6 10mg 当たりのタッピング量の変化

表 9.6 タッピング試験・データリスト

PNL	DOSE	SQN	ID	EXP	投与24時間前			投与後1.5時間目			投与後24時間目			投与後48時間目		
					1	2	3	1	2	3	1	2	3	1	2	3
1	0	1	03	1	365	349	368	355	370	377	340	358	369	370	371	381
1	0	1	05	1	396	408	391	340	350	362	354	340	367	347	361	351
1	0	2	02	3	356	370	362	351	362	376	340	347	364	345	348	378
1	0	2	04	3	374	354	383	368	349	388	362	322	335	384	397	395
1	0	3	07	5	392	424	427	385	403	404	382	403	424	397	410	421
1	0	3	08	5	339	361	387	311	340	360	311	350	367	385	404	412
1	0	4	01	7	373	332	380	351	375	385	384	391	403	386	398	400
1	0	4	06'	7	329	324	315	313	306	373	361	396	437	365	342	368
1	2.5	4	01	1	333	354	366	351	371	376	342	370	370	363	371	389
1	2.5	2	02	1	356	348	358	355	360	357	343	335	361	356	338	357
1	2.5	2	04	1	285	351	404	362	396	411	360	393	390	355	358	404
1	2.5	4	06	1	293	318	318	321	322	328	330	333	332	347	338	341
1	2.5	3	07	1	395	408	410	384	410	434	377	406	413	399	406	420
1	2.5	3	08	1	345	357	369	336	340	335	318	328	344	335	342	347
1	10	4	01	3	323	323	332	326	332	342	307	341	347	339	349	365
1	10	1	03	3	344	348	351	352	354	366	332	342	350	334	344	368
1	10	1	05	3	297	295	284	308	319	316	334	323	303	349	320	395
1	10	4	06	3	326	330	343	343	346	360	346	330	342	359	337	331
1	10	3	07	3	388	399	414	382	402	421	376	403	418	379	402	420
1	10	3	08	3	324	326	329	295	333	345	311	319	327	330	333	355
1	20	4	01	5	340	354	362	322	342	381	331	348	362	372	385	406
1	20	2	02	5	355	364	368	352	364	365	352	364	373	347	359	377
1	20	1	03	5	341	350	375	310	325	336	319	327	350	373	387	413
1	20	2	04	5	404	384	418	285	345	367	356	357	391	409	415	421
1	20	1	05	5	334	335	325	340	350	339	337	311	386	370	396	412
1	20	4	06'	5	305	322	304	307	303	393	330	352	370	325	358	431
1	40	2	02	7	358	354	376	314	313	337	340	346	355	339	357	373
1	40	1	03	7	350	352	378	285	262	252	312	315	344	343	359	378
1	40	2	04	7	372	364	379	322	375	375	374	391	415	409	403	414
1	40	1	05	7	260	346	301	234	235	227	271	241	283	268	318	315
1	40	3	07	7	393	410	430	325	378	398	365	377	424	364	384	405
1	40	3	08	7	342	342	360	310	281	322	303	313	312	302	295	352
2	0	5	09	2	342	349	370	359	374	388	368	371	385	360	375	383
2	0	5	14	2	362	407	425	391	405	438	380	400	399	391	396	376
2	0	6	10	4	355	360	365	355	370	371	363	381	390	355	372	385
2	0	6	13	4	377	372	373	366	370	371	371	370	381	359	378	389
2	0	7	12	6	364	384	398	387	410	397	412	413	399	363	403	414
2	0	7	16	6	426	430	462	396	406	425	355	375	398	408	425	434
2	0	8	11	8	389	386	417	374	400	409	368	385	406	401	429	429
2	0	8	15	8	371	389	415	366	374	407	367	393	410	385	403	422
2	5	6	10	2	347	400	421	365	361	369	358	367	386	359	369	379
2	5	8	11	2	327	330	326	386	403	408	397	410	417	406	398	407
2	5	7	12	2	380	403	405	339	378	368	352	379	365	348	385	403
2	5	6	13	2	383	410	420	366	364	364	361	358	352	398	398	431
2	5	8	15	2	396	400	406	369	343	390	362	.	393	372	380	383
2	5	7	16	2	411	443	436	385	401	403	361	369	395	403	411	412
2	15	5	09	4	356	368	389	386	394	395	364	396	412	368	379	398
2	15	8	11	4	379	381	381	386	381	398	395	404	413	397	400	423
2	15	7	12	4	382	402	403	359	395	406	361	417	406	385	405	424
2	15	5	14	4	402	418	445	380	385	391	363	406	423	381	393	428
2	15	8	15	4	359	360	395	355	361	373	344	362	404	379	391	411
2	15	7	16	4	393	400	411	384	390	411	363	384	439	387	398	433
2	30	5	09	6	365	379	393	365	376	391	363	374	395	390	403	421
2	30	6	10	6	345	372	397	364	370	373	345	369	394	381	393	404
2	30	8	11	6	363	377	388	369	374	395	363	382	380	406	393	399
2	30	6	13	6	383	379	402	359	375	358	361	387	374	371	392	405
2	30	5	14	6	381	431	443	380	403	397	359	369	403	380	417	410
2	30	8	15	6	369	375	412	334	361	381	359	379	398	378	370	397
2	50	5	09	8	382	393	407	360	379	397	377	378	397	373	401	409
2	50	6	10	8	390	396	405	370	356	381	342	354	374	366	382	402
2	50	7	12	8	366	390	383	336	369	385	392	400	415	362	405	406
2	50	6	13	8	395	394	423	313	354	394	348	356	395	372	396	423
2	50	5	14	8	362	401	409	356	350	414	359	401	415	391	414	426
2	50	7	16	8	439	464	470	400	410	420	374	399	402	406	433	442

文献

Little, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (1996) *SAS System for Mixed Models*. SAS Institute.

Verbeke, G. and Molenberghs, G. (1997), *Linear Mixed Model in Practice*. springer. 日本語訳出版予定.

Verbeke, G. and Molenberghs, G. (2000), *Linear Mixed Models for Longitudinal Data*. Springer.

高橋行雄, 大橋靖雄, 芳賀敏郎 (1989) SAS による実験データの解析, 東大出版会.

高橋行雄, 大島信之 (1997) 交互パネル法による臨床第 1 相試験, 医薬安全研 70 回定例会資料.