

# 投与前値を含むクロスオーバー法での経時データの解析

高橋 行雄  
BioStat 研究所(株)

Analysis of Longitudinal Data on Crossover Designs with Baseline  
Yukio Takahashi  
BioStat Research Co.,Ltd.

**要旨：** 投与前値を含む経時データの群間比較において、投与前値を共変量とする線形混合モデルによる解析、あるいは、投与前値を各群で共通とした解析が、平均への回帰現象を受けにくく、投与前値からの差による群間比較より統計的に優れていることが知られている。クロスオーバー法による臨床薬理試験では、投与前値を含む経時データがしばしば得られる。同一症例内の各実験時期の投与前値の変動が少ないとみなせるのであれば、投与後のデータのみ、あるいは、前後差での解析を行うことの妥当性がある。しかしながら、同一症例内で投与前値に無視できない変動があり、また投与後値との間でなんらかの相関関係があった場合には、投与前値を共変量とした解析を行う必要がある。各種の統計モデルを適用した際に、薬剤群間の差の信頼区間がどのように変化するかを、SAS for Mixed Model 2ed. (Little (2006)) 5.1 節で例示されているクロスオーバー法による 1 秒あたりの呼吸量の経時データを用いて比較する。さらに、それぞれの統計モデルによる誤差構造についての比較も行う。

**キーワード：** QT/QTc 試験, 経時データ, MIXED プロシジャ, 共分散分析, クロスオーバー法

## 1. はじめに

ICH-S7b 「ヒト用医薬品の心室再分極遅延 (QT 間隔延長) の潜在的可能性に関する非臨床的評価」が、2009 年 10 月 23 日にステップ 5 となり、大動物を使ったクロスオーバー法によるテレメトリーQT/QTc 試験が多くの研究施設で行われるようになってきた。この試験から複数の投与前値を含む超多時点の経時データが得られる。このようなデータに関して、どのような経時データの解析を行うかについては適当な文献・成書がない。

ICH-S7b には、統計解析についての具体的な記載がないので、ICH-E14 「非抗不整脈薬におけるQT/QTc 間隔の延長と催不整脈作用の潜在的可能性に関する臨床的評価 (2009)」を参考にする。ICH-E14 の 2.2 節 QT/QTc 評価試験に「QT/QTc 評価試験の目的は、被験薬に心室再分極に対する一定の大きさ以上の薬理作用があるか否かを決定することであり、その値はQT/QTc 間隔の延長として検出される。規制当局が関心をもつ基準値レベルについては後述するが、QTc 間隔への作用の平均値としておよそ 5ms であり、95%信頼区間の上限を 10ms とするものである。」と信頼区間方式による判定基準が示されている。

2.2.4節の QT/QTc 評価試験の解釈には、「同様の考え方にに基づき、QT/QTc 評価試験が陰性とは、その薬剤の QTc 間隔への時間を一致させた平均効果の最大値に対する 95%片側信頼区間の上限が 10ms を下回る場合を指す。この定義は、被験薬の QT/QTc 間隔への作用の平均がおよそ 5ms を超えないことを合理的に保証するために選択されている。時間を一致させた差の最大値がこの基準値を超える場合、試験結果は陽性とされる。」と判定基準を定めている。

ベースライン値については、3.2 節 QT/QTc 間隔データの解析に「QT/QTc 間隔のベースラインに比しての延長は注意すべき徴候であるが、それらは平均値への回帰や極端な値を選択したためなど薬物療法に無関係な要因による変化である可能性があるので、QT/QTc 間隔のベースラインとの差の解釈は複雑である。」と注意している。さらに、「QT/QTc 間隔データは、中心傾向 (central tendency) の解析 (例えば、平均値、中央値) 及びカテゴリカル解析の両方の形で示すべきである。どちらの解析も、臨床上的リスクを評価する際の適切な情報となり得る。」と中心傾向の解析の必要性が強調されている。

中心傾向とは、質問紙を用いた評価では「どちらともいえない」といったようなほぼ中心に回答が集まる現象として知られている。なお、3.2.1 節で「被験薬が QT/QTc 間隔へ与える作用の解析は、最も一般的には、時間を一致させた被験薬群とプラセボ群の平均値の差 (ベースライン値による調整後) の、収集の全期間を通じた最大値を用いて行われる。」と述べられ、いくつかの例示もあるが、漠然としていてどのような解析なのかが不明瞭である。「平均値への回帰」については、丁寧な説明が別にされているので、一般的に用いられている投与前値を共変量とした解析でもないようである。

ICH-E14 の Q&A 問 6 に「ベースライン値の必要性について説明して下さい。また、ベースライン値が必要な場合、QT/QTc 評価試験がクロスオーバー試験と並行群間比較試験のデザインで実施されるそれぞれの場合について、ベースライン値の測定方法を説明して下さい。」とあり、回答では、被験薬の投与に先立って同時刻に測定されたベースライン、投与前のベースラインの 2 つがあり、クロスオーバー法では投与前のベースラインを用いることが適切と述べている。

これらの文脈から、「ベースライン値による調整」とは、クロスオーバー法では、投与前値からの差であることが推測され、投与前値を共変量とした調整ではないようである。

そこで、Littell ら (2006) が SAS for Mixed Model 2ed. 5.1 節で例示しているクロスオーバー法による 1 秒あたりの呼吸量の経時データを用いて、各種の統計モデルを適用した際の、主要評価時点における薬剤群間の差の信頼区間について比較検討し、テレメトリー QT/QTc 試験の経時データの解析法の参考とすることにした。

## 2. データの構造と判定基準

### 2.1 データの概要

Littell らの呼吸機能の経時データを表 1 に示す。これは、呼吸機能の改善を目的にした薬剤 T について、標準薬 S とプラセボ P を対象にした 24 症例のクロスオーバー試験の結果である。各群について、1 秒あたりの呼吸量 FEV1 (Forced Expiratory Volume in 1<sup>st</sup> second) が単回投与前から 8 時間後まで 1 時間ごとに測定されている。なお、文献では 24 症例に対して 3 種の薬剤をランダムに割り付けたと述べられているだけで、実験順序・時期などのデータは含まれていない。

表 1 FEV1 の経時変化

drug	patient	0	1	2	3	4	5	6	7	8
T(a)	201	2.46	2.68	2.76	2.50	2.30	2.14	2.40	2.33	2.20
T(a)	202	3.50	3.95	3.65	2.93	2.53	3.04	3.37	3.14	2.62
	⋮									
T(a)	232	2.49	3.73	3.51	3.16	3.26	3.07	2.77	2.92	3.00
S(c)	201	2.30	3.41	3.48	3.41	3.49	3.33	3.20	3.07	3.15
S(c)	202	2.91	3.92	4.02	4.04	3.64	3.29	3.10	2.70	2.69
	⋮									
S(c)	232	2.79	4.10	3.85	4.27	4.01	3.78	3.14	3.94	3.69
P(p)	201	2.14	2.36	2.36	2.28	2.35	2.31	2.62	2.12	2.42
P(p)	202	3.37	3.03	3.02	3.19	2.98	3.01	2.75	2.70	2.84
	⋮									
P(p)	232	2.88	3.04	3.00	3.24	3.37	2.69	2.89	2.89	2.76

<http://support.sas.com/publishing/bbu/59882/59882.zip> からダウンロードし整形.

表 1 に示した 3 例についての 図 1 に示す線グラフで経時変化の特徴を概観する. 症例により投与前値が異なり, 同じ症例の中でも各薬剤の投与前値に症例間ほどではないが差があり, プラセボ投与以外は最初の 1 時間目から反応があり, 8 時間目まで継続していることが観察される.

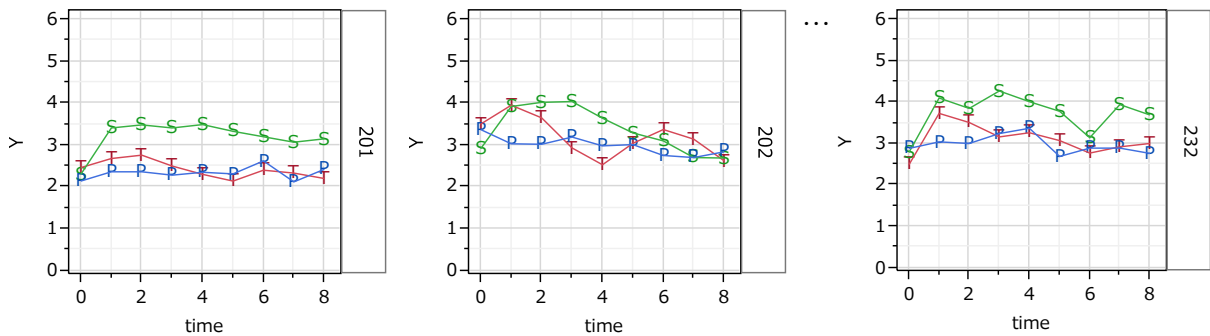


図 1 症例ごとの FEV1 の経時変化

全症例の経時変化の平均値 ( $y$ ) の推移および投与前からの差 ( $d$ ) について 図 2 に示す. S 薬および T 薬の投与後 1 時間目で反応がピークとなり, その後 8 時間目まで緩やかに減少している. P 群では大きな変動は見られない.

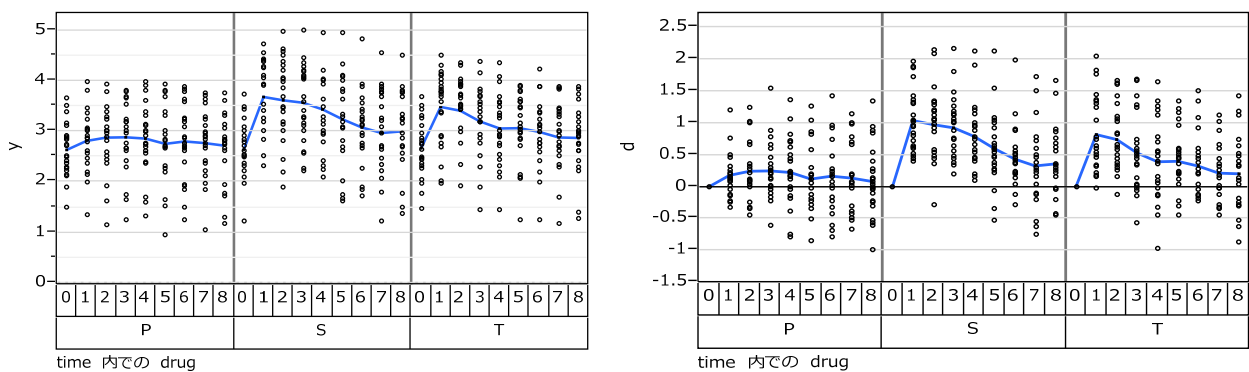


図 2 症例別 FEV1 の群ごとの経時変化

## 2.2 信頼区間方式による薬効の判定基準

この試験は薬効評価のための試験であり、有効性判定基準がないため、QT/QTc 評価試験に準じた信頼区間方式による判定基準を次のように別途定める。

主要な評価時期に対して、試験薬 T の片側 95%信頼区間の下限が、プラセボ P の同時期の点推定に対して 15%増の反応であれば、薬効がある用量とみなす。また、陽性対照である標準薬 S に対しても同様の基準を適用し、これを満たした場合はこの試験が適切に行われたと判断する。なお、試験薬 T を複数用量設定すれば、標準薬 S の投与量に対する等価用量の推定が可能となる。

並行群間試験の場合には、投与前値からの評価時点までの差  $d$  を主要変数とする場合には、投与前値を共変量として解析モデルに含めること。これは、主要評価時点の反応を  $y$  とし投与前値を共変量とした場合の薬剤間に関する解析結果と一致することが知られているためである。

なお、クロスオーバー法による実験の場合には、各症例のそれぞれの薬剤群の投与前値を共変量としなくとも薬剤群間差の推定が適切に行える可能性もあるので、投与前値を共変量として含める場合には、含めない場合の結果を示し、総合的な判断をすることが必要である。

また、測定時点を含めた解析を追加する場合には、主要評価時点での薬剤群間差に関する各種の推定を行い、総合的な薬効の判定に加えても差し支えない。

## 3. 主要変数についての解析

### 3.1 主要変数についての解析

クロスオーバー法では、症例を変量効果、薬剤を固定効果とした解析が適切とも思われるが、生物学的同等 (BE) 試験では、症例を固定効果とした解析が定式化されている。そこで、両者での解析結果を比較し、結果の判定にどのような影響をあたえるのかを検討した。

反応のピークとなる 1 時間目のデータを対象として、症例、および薬剤を固定効果とした 2 元配置繰り返しなしの分散分析表を 表 2 示す。「平均平方の構造」は、GLM プロシジャの RANDOM ステートメントで症例を指定することによって得られる。この構造から症例についての分散成分を計算した結果を付け加えてある。

症例を変量効果とした MIXED プロシジャの REML 法 (制限付き最尤法) による解析では、変量効果に関する分散成分の出力と固定効果とした薬剤の自由度および  $F$  値 のみが出力されるが伝統的な 表 2 の様式の分散分析表の出力はない。

表 2 1 時間目における 2 元配置とした分散分析表と分散成分

要因	自由度	平方和	平均平方	$F$ 値	$p$ 値	平均平方の構造	分散成分
patient	23	27.1613	1.1809	13.02	<.0001	$\sigma_e^2 + 3\sigma_{patient}^2$	0.3634
drug	2	9.9948	4.9974	55.08	<.0001	$\sigma_e^2 + 24\sigma_{drug}^2$	
誤差	46	4.1734	0.0907			$\sigma_e^2$	0.0907
全体	71	41.33					

症例を固定効果とするか、変量効果とするかで、表 3 に示すように各薬剤の信頼区間に異なる結果を与える。プラセボ群の平均値は 2.8150 の 15% 増は 3.2373 であり、であり、症例を固定効果とした場合の S 薬の信頼区間の 90%下限は 3.5818, T 薬は 3.3855 と大きく離れている。他方、症例を変量効果とみなした場合に T 薬の信頼区間の 90%下限は 3.2553 と下方に広がり、T 薬の場合には

プラセボ群の平均値の 15%増である 3.2488 をわずかに上回る結果である。

表 3 1 時間目における 2 種類の SE および信頼区間

薬剤	平均	症例を固定効果			症例を変量効果		P群の15%増	
		SE	L 90%	SE	L 90%			
P	2.8150	0.0615	2.7118	0.1376	2.5816			
S	3.6850	0.0615	3.5818	0.1376	3.4516	3.2373		
T	3.4888	0.0615	3.3855	0.1376	3.2553	3.2373		
薬剤	薬剤	差	SE	L 90%	SE	L 90%	p 値	P群の15%
S	P	0.8700	0.0870	0.7240	0.0870	0.7240	<.0001	0.4223
T	P	0.6738	0.0870	0.5278	0.0870	0.5278	<.0001	0.4223
S	T	0.1963	0.0870	0.0503	0.0870	0.0503	0.0288	

症例を固定効果とした場合、変量効果とした場合の薬剤群の信頼区間の 90%下限が異なるのは、以下に示すように症例に関する分散成分を加味するかしないかによって説明される。

症例を変量効果とした場合の分散成分は、平均平方の構造から症例の分散  $\hat{\sigma}_{patient}^2$  は、

$$\hat{\sigma}_{patient}^2 = (V_{patient} - V_e) / 3 = (1.1809 - 0.0907) / 3 = 0.3634$$

と推定できる。症例を固定効果とした場合に、各薬剤の固定  $SE_{drug}$  は、24 症例の平均値に対するものなので、

$$\text{固定 } SE_{drug} = \sqrt{V_e / 24} = \sqrt{0.0907 / 24} = 0.0615$$

である。これに対し、症例を変量とした場合には、

$$\text{変量 } SE_{drug} = \sqrt{(V_e + \hat{\sigma}_{patient}^2) / 24} = \sqrt{(0.0907 + 0.3634) / 24} = 0.1376$$

誤差分散（誤差の平均平方と同じ）に症例に関する分散成分を加えた結果となる。症例を固定効果とした場合の SE は、同一の症例を対象として実験を繰り返した場合の各薬剤の母平均値に関するものであり、症例を変量効果とした場合の SE は、別の症例を対象にした場合となっていて、一般化可能性の観点からは、症例を変量効果とした場合の SE から信頼区間を計算することが望ましい。

薬剤間の差  $SE_{drug}$  は、症例を固定効果とした場合でも、変量効果とした場合でも、

$$\text{群間差 } SE_{drug} = \sqrt{2V_e / 24} = \sqrt{2 \times 0.0907 / 24} = 0.0870$$

と同じである。これは、異なる症例に対する実験であっても、各々の症例の反応の大きさに違いがあったとしても、同じ症例内での薬剤間の比較なので、症例に関する分散成分が入り込まないからである。プラセボ群の平均値の 15%は 0.4223 であり、T 薬の差の信頼区間の 90%下限は 0.5278 であるので、ゆとりをもって T 薬の薬効が証明されたことになる。

プラセボの平均値の 15%増増による評価と、プラセボと T 薬の差がプラセボの平均値の 15%増を用いた判定と 2 通りが考えられる。どちらが適切なのだろうか。

### 3.2 投与前値からの差での解析

投与前値からの差（変化量）による解析は、元データに比べて変化の大きさが明確で、相対的な薬効の比較がしやすいとの利点もあり、また投与前値の症例間の変動を除去できるために元データ

での解析よりも望ましいのではないかとされている。クロスオーバー法の場合は、固定効果、あるいは変量効果として症例をモデルに組み込んでいるので、SE がどのように変化するか検討する。

図 3 に示すように投与前と投与後の 1 時間目の FEV1 の間の相関は、プラセボ群で相関係数 0.82, S 薬 0.74, T 薬 0.61, といずれも 0.5 以上であるので、投与前からの差での解析が、元データでの解析よりも望ましいとも思われる。これは、並行群間試験の場合であり、クロスオーバー試験でも成り立つかを検討する。

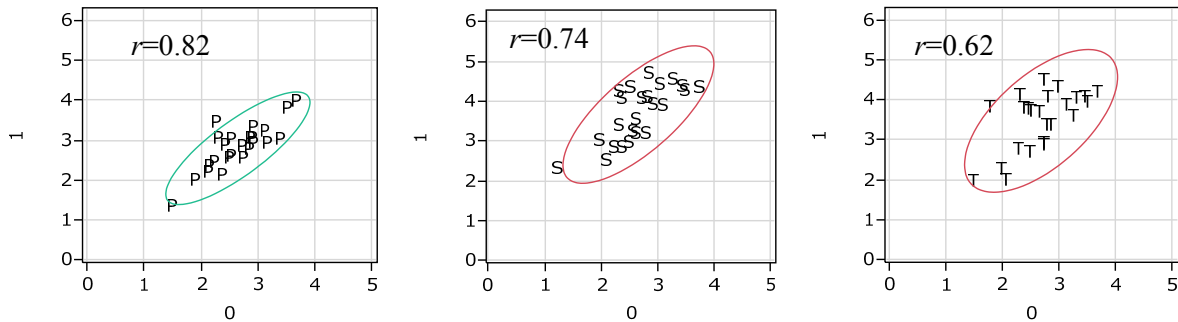


図 3 薬剤群ごとの投与前値と 1 時間目の相関関係

表 4 にクロスオーバー法ではなく並行群間試験とみなした結果を示す。元データでの T 薬の平均は 3.4888, SE は 0.1376 であるのに対し、投与前からの差のデータでは、T 薬の差の平均は 0.8204, SE は 0.0973 と元データに比べて SE が小さくなっている。プラセボと T 薬との差の SE は、元データで 0.1945 であり、信頼区間の 90% 下限は 0.3494 と P 群の 15% よりも小さいので、薬効があるかは判定保留となる。差の差のデータの SE は 0.1376 で、信頼区間の 90% 下限は 0.4119 と P 群の 15% よりも小さいので、薬効があるかは判定保留となる。

並行群間試験とみなした解析は、症例間の変動が信頼区間の 90% 下限の計算に含まれるために、厳しい判定となってしまいが、投与前からの差をとることによって SE が小さくなり、信頼区間の 90% 下限が狭まることが確認された。

表 4 並行群間試験とみなした解析結果

薬剤	元データでの群間比較			差のデータでの群間比較			P 群の 15% 増	
	平均	SE	L 90%	差の平均	SE			
P	2.8150	0.1376	2.5857	0.1792	0.0973			
S	3.6850	0.1376	3.4557	1.0413	0.0973		3.2373	
T	3.4888	0.1376	3.2594	0.8204	0.0973		3.2373	
薬剤	薬剤	差	SE	L 90%	差の差	SE	L 90%	P 群の 15%
S	P	0.8700	0.1945	0.5457	0.8621	0.1376	0.6327	0.4223
T	P	0.6738	0.1945	0.3494	0.6413	0.1376	0.4119	0.4223
S	T	0.1963	0.1945	-0.1281	0.2208	0.1376	-0.0086	

表 4 差の SE 0.0973 よりかなり小さい。しかしながら、症例を変量効果とした場合の SE は 0.0973 と同程度であるが、プラセボと T 薬の平均値の差の SE は固定効果でも変量効果でも同じ 0.0883 であり並行群間試験の場合の 0.1376 に比べて明らかに小さくなっている。

表 5 投与前値からの差についての信頼区間の 90%下限値

薬剤	差の平均	症例を固定効果		症例を変数効果		p 値	P 群の15%	
		SE	L 90%	SE	L 90%			
P	0.1792	0.0624	0.0744	0.0973	0.0154			
S	1.0413	0.0624	0.9365	0.0973	0.8775			
T	0.8204	0.0624	0.7156	0.0973	0.6567			
薬剤	薬剤	差の差	SE	L 90%	SE	L 90%	p 値	P 群の15%
S	P	0.8621	0.0883	0.7139	0.0883	0.7139	<.0001	0.4223
T	P	0.6413	0.0883	0.4930	0.0883	0.4930	<.0001	0.4223
S	T	0.2208	0.0883	0.0726	0.0883	0.0726	0.016	

投与前からの差での解析において、クロスオーバー法の良さは実感できたのであるが、表 3 に示した元データでの症例を変数効果とした場合の解析結果と比較してみよう。元データでの T 薬の SE は 0.1376 であるが、差のデータにした場合とした場合 0.0973 と小さくなっている。これは、差をとったことにより、症例間の変動が軽減されたことによる。

症例を固定効果としても変数効果としてもプラセボと T 薬の差の差の SE は 0.0883 と同じで、信頼区間の 90%下限も 0.4930 と同じであるが、表 3 の元データで場合の 0.5278 よりも信頼区間の 90%下限が小さくなり、投与前値からの差による解析のメリットが見いだせない。

### 3.3 元データでの解析か投与前値からの差のデータでの解析か

これらの結果を踏まえて、元データ、投与前値からの差のデータ、どちらの解析結果を用いたらよいのであろうか。あるいは併記するのがよいのだろうか。投与前データの総平均に比べ群平均が高めならば増加量が抑えられ、群平均が小さめならば増加量は多めとなり、元データと差のデータでの解析結果が異なってしまふ。これは、測定値がある範囲に限定されているような場合に、ある症例の投与前値が高目に出たとすれば、その次の測定ではそれ以上になる確率は低くなり、その症例の真の平均に近づくことになる。

「平均への回帰」現象は、ICH-E14 でも、3.2 節の 中心傾向 (central tendency) で説明されているが、投与前値を共変量とする共分散分析については言及されていない。投与前値の群平均が完全に一致していれば、「平均への回帰」現象の影響は受けないが、わずかでも異なると、「平均への回帰」現象の影響から逃れない。この結果として、元データでの結果と投与前からの差での結果が微妙にことなり、どちらか一方を使うと結果に対し、都合の良い方を使ったのではないかと疑われ、併記すれば、どちらの結果で判定することが望ましいのかと、詰問されることになりかねない。

### 3.4 投与前値を共変量とした解析

並行群間比較試験の場合には、投与前値を共変量とすることで、元データでも差のデータでも結果が一致することが知られているが、クロスオーバー法の場合でも一致するのであろうか。

図 3 に示したように投与前と 1 時間目の各薬剤での相関係数は、0.6 以上の相関となっている。クロスオーバー法の場合には、投与前値を共変量とすると、各群の相関よりも各症例内の相関が関与するようにも思われる。各症例内には 3 薬剤分のデータしかないが、1 時間目のデータについてそれぞれの平均値と SD で基準化したデータについて 図 4 に散布図を描き、50%の確率楕円を上

書きした結果を 図 4 に示す.

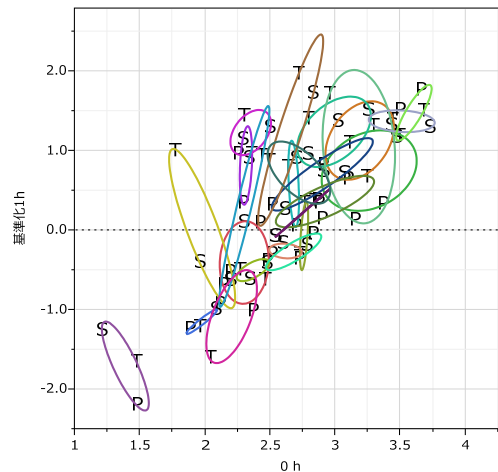


図 4 症例ごとの投与前値と 1 時間目の相関関係

投与前と基準化した 1 時間目の確率楕円から、症例内での 3 ポイントの相関は、正の場合もあるが、負となる場合もあり、また無相関の場合もあるが、全体的には正の相関構造が示唆される。これらの相関構造を念頭にしつつ、投与前値を共変量とした解析を行い、結果の解釈の参考とする。

症例を変量効果とし、1 時間目の元データ、および投与前からの差のデータについての結果を表 6 に示す。プラセボ群の調整平均 (LSMEAN) は、2.8252 で、その 15%増は 3.2490 であり、表 3 の症例を変量効果とみなした場合の *SE* が 0.1376 であるのに対し、投与前値を共変量とすることにより、0.0988 と大幅に減少し、S 薬の信頼区間の 90%下限は 3.5225 とかなり上回り、T 薬は 3.3076 とゆとりを持って上回るようになった。

投与前からの差について、もちろん調整平均は異なるが、*SE* は完全に一致し、クロスオーバー法であっても、元データと差のデータでも同じ結果が得られることが確認された。ただし、残念なことに薬剤間の差の *SE* は、0.0870 から 0.0850 へとわずかな減少にとどまっている。この原因は、

表 6 投与前値を共変量，症例を変量効果とした場合の信頼区間の 90%下限値

薬剤	前値を共変量：元データ			前値を共変量：差のデータ			P群の15%増		
	調整平均	<i>SE</i>	<i>L</i> 90%	差(調整)	<i>SE</i>				
P	2.8252	0.0988	2.6585	0.1759	0.0988				
S	3.6892	0.0987	3.5225	1.0399	0.0987		3.2490		
T	3.4743	0.0988	3.3076	0.8250	0.0988		3.2490		
薬剤	薬剤	差	<i>SE</i>	<i>L</i> 90%	差の差	<i>SE</i>	<i>L</i> 90%	<i>p</i> 値	P群の15%
S	P	0.8640	0.0849	0.7213	0.8640	0.0849	0.7213	<.0001	0.4238
T	P	0.6491	0.0850	0.5063	0.6491	0.0850	0.5063	<.0001	0.4238
S	T	0.2149	0.0850	0.0721	0.2149	0.0850	0.0721	0.0151	

クロスオーバー法なので「症例」がモデルに含まれており、薬剤群間の差とした場合に各症例内での投与前値の共変量としての調整の役割の寄与がほとんどなくなってしまったと解される。



#### 4. すべての測定時点を用いた解析

##### 4.1 分割実験とみなした解析法の応用

並行群間試験の場合には、すべての測定時点を用いて解析を行うことで症例内の変動を抑えることができ、1時間目についての薬剤群の比較における推定精度が向上することが期待される。クロスオーバー法の場合についても、同様に推定精度の向上があるのだろうか。

症例、薬剤、時点の3因子をランダム化の順序を考えた分割実験として考える。各症例は互いに独立していて、その中で3薬剤がランダムに割り付けられ、その中で1, 2, ..., 8時間目のデータがランダムに測定されたとみなす。表7に3因子交互作用まで平方和を分解した結果を示す。症例は変量効果、薬剤は固定効果、薬剤×症例は変量効果で1次誤差、時間は固定効果であり、時間×症例、時間×薬剤×症例は、症例が変量効果なので変量効果となりも、同じ2次要因内なので合わせて2次誤差とする。

表7 分割実験とみなした場合の平方和

	要因	自由度	平方和	平均平方	役割
ブロック	patient	23	223.97	9.7378	変量効果
1次要因	drug	2	25.78	12.8913	固定効果
	drug×patient	46	23.44	0.5096	変量：1次誤差
2次要因	time	7	17.17	2.4529	固定効果
	time×patient	161	12.21	0.0759	変量：2次誤差
	time×drug	14	6.28	0.4486	固定効果
	time×drug×patient	322	18.28	0.0568	変量：2次誤差
	全体	575	327.14		

表7を組替えて表8に分散分析表としてまとめ直し、平均平方の構造から、分散成分計算した結果を示す。ここに示した分散分析表は、1, 2, ..., 8時間の測定が完全にランダム化されたとみなした解析であり、そのために2次誤差の自由度が483とインフレーションを起こし、時間×薬剤のF検定が有意になりやすいとの批判があり、そのために自由度の補正が定式化されているが、ここでは言及しない。

表8 分割実験とみなした場合の分散分析表

	要因	自由度	平方和	平均平方	F: 1次	F: 2次	平均平方の構造	分散成分
ブロック	patient	23	223.97	9.7378	19.11		$\sigma_e^2 + 24\sigma_{patient}^2 + 8\sigma_{drug \times patient}^2$	0.3845
1次要因	drug	2	25.78	12.8913	25.30		$\sigma_e^2 + 192\sigma_{drug}^2 + 24\sigma_{time \times drug}^2 + 8\sigma_{drug \times patient}^2$	
	drug×patient	46	23.44	0.5096	1	19.11	$\sigma_e^2 + 8\sigma_{drug \times patient}^2$	0.0558
2次要因	time	7	17.17	2.4529		38.86	$\sigma_e^2 + 72\sigma_{time}^2 + 24\sigma_{time \times drug}^2$	
	time×drug	14	6.28	0.4486		7.11	$\sigma_e^2 + 24\sigma_{time \times drug}^2$	
	誤差	483	30.49	0.0631		1	$\sigma_e^2$	0.0631
	全体	575	327.14					

批判にさらされている分散分析表をあえて持ち出したのは、平均平方の構造から分散成分の推定ができるからである。投与後の全時点を用いることにより、1時間目のデータのみで推定した分散成分よりも、1時間目の薬剤間の平均値を安定して推定することが可能となる。分散成分の推定においては、自由度のインフレーションは、平均平方の構造に示したように除去されることが確認され

る。

分割実験型の分散分析を行う統計ソフト（GLM プロシジャ，JMP/EMS タイプの適用）の致命的な欠陥は，分散分析表における自由度のインフレーションよりも，各種の水準間の比較にある．時間×薬剤の1時間目の推定値に対して薬剤群の推定平均の推定を行った時に起きる．表9に示すように，GLM プロシジャのRANDOM ステートメントで症例，薬剤×症例を変数と指定しても，平均平方の構造，分散分析表のF検定は適切に対応するが，薬剤群の推定平均および差の推定平均も，すべて2次誤差から次のようにSEが推定されているために，常に過大評価を招く．

$$1 \text{ 時間目の薬剤群の平均 } SE : SE_{1h,drug} = \sqrt{\hat{\sigma}_e^2 / 24} = \sqrt{0.0631 / 24} = 0.0513$$

$$1 \text{ 時間目の薬剤群間差の平均 } SE : SE_{1h,drug(diff)} = \sqrt{2\hat{\sigma}_e^2 / 24} = \sqrt{2 \times 0.0631 / 24} = 0.0725$$

症例，症例×薬剤を変数効果としたMIXED プロシジャ（JMP のREML 指定）の場合には，

1時間目の薬剤群の均SE：

$$SE_{1h,drug} = \sqrt{(\hat{\sigma}_{patient}^2 + \hat{\sigma}_{patient*drug}^2 + \hat{\sigma}_e^2) / 24} = \sqrt{(0.3845 + 0.0558 + 0.0631) / 24} = 0.1448$$

1時間目の薬剤群間差の平均SE：

$$SE_{1h,drug(diff)} = \sqrt{2(\hat{\sigma}_{patient*drug}^2 + \hat{\sigma}_e^2) / 24} = \sqrt{(0.0558 + 0.0631) / 24} = 0.0996$$

のように，変数効果の分散成分が含まれていて，それぞれのSEは大きく推定されている．MIXED プロシジャの場合には，推定された分散成分を合成したSEを算出し，適切な推定値となる．

表9 分割実験とみなした場合の1時間目の水準間の信頼区間

		1時間目	固定 (GLM)		変数 (MIXED)		
薬剤		平均	SE	L 90%	SE	L 90%	P群の15%増
P		2.8150	0.0513	2.7305	0.1448	2.5703	
S		3.6850	0.0513	3.6005	0.1448	3.4403	3.2373
T		3.4888	0.0513	3.4042	0.1448	3.2441	3.2373
薬剤	薬剤	差	SE	L 90%	SE	L 90%	P群の15%
S	P	0.8700	0.0725	0.7505	0.0996	0.7052	0.4223
T	P	0.6738	0.0725	0.5542	0.0996	0.5090	0.4223
S	T	0.1963	0.0725	0.0767	0.0996	0.0315	

#### 4.2 時点間の相関構造

MIXED プロシジャを用いた経時データの解析では，各種の時点間の相関構造を設定できるようになっていて，Littell らは，元データでの相関構造に対し，どのような相関構造が適合するか詳しく示し，図5に示すように投与後の各自時点間の相関は，時点が離れるにつれて大きくなるが，自己回帰型 (AR1 タイプ) ほどではないと考察している．また，AIC を用いた相関構造の選択では，相関構造を特定しない UN タイプが優れているが，BIC の観点からで貧弱な選択だと述べ，生物学的な観点からの検討が必要性を示唆している．

時点間の相関構造の選択は，多くの実験研究の裏付けが必要であり，また一定間隔で得られた経時データでなければ適用が困難であり，また解析しようとしている実験データそのものから特定で

きるものではない. ここでは, (0h, 1h, 2h, 4h, 8h) などのように測定間隔が異なる場合でも適用できる時点間の相関構造を平均的な相関とした CS タイプを用いることにする.

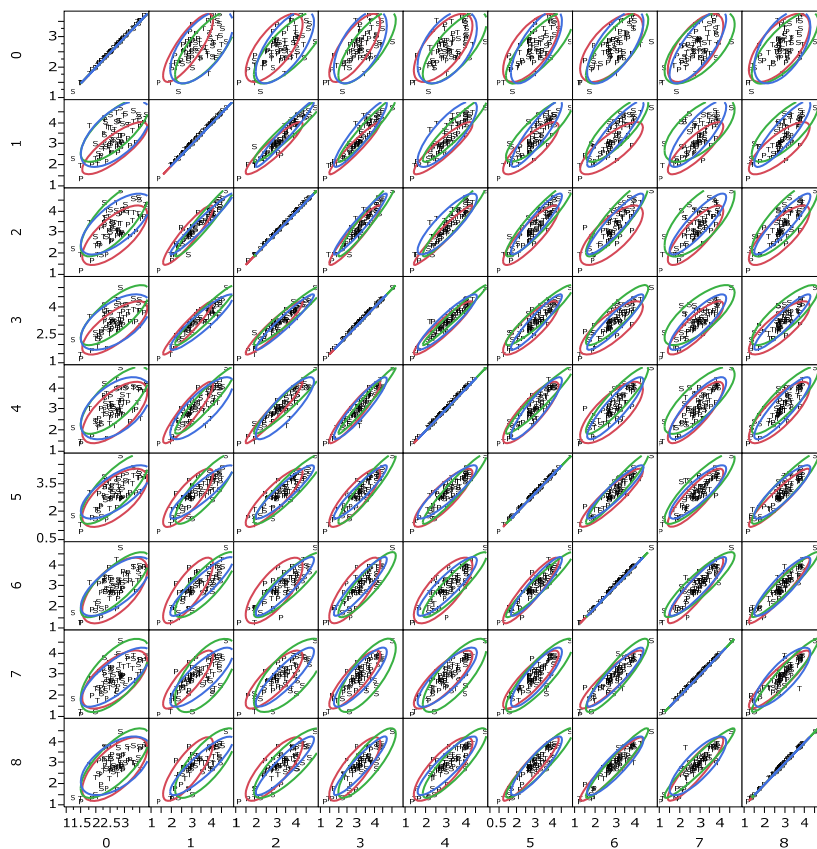


図 5 時点間の散布図と薬剤群ごとの 90%確率楕円

### 4.3 投与前値を共変量とした解析

投与前値を共変量として変量効果モデルに組み込むことにより, 1 時間目の分散成分が小さくなることが期待される. 表 10 に示すように,  $\hat{\sigma}_{patient}^2$  は, 0.3845 から 0.1669 と大幅に減少したが,  $\hat{\sigma}_{drug \times patient}^2$  および  $\hat{\sigma}_e^2$  の減少はほとんどなかった.

表 10 前値を共変量とした場合の分散成分

変量効果	前値を含まず 分散成分	前値を共変量 分散成分	変化
patient	0.3845	0.1669	大幅減少
drug×patient	0.0558	0.0557	ほとんど変わらず
残差	0.0631	0.0631	変わらず

表 11 に, 投与前値を共変量とした場合の 1 時間目における各薬剤の推定値, および薬剤群間の差の推定値と 90%の信頼区間を示す. 結果は, 表 6 で示した 1 時間目の時点を用いた共分散分析とほぼ同程度の結果であり, 労多くして功少なしであった.

表 11 前値を共変量とした場合の 1 時間目の推定値と信頼区間

		前値を含まず (表 9 再掲)			前値を共変量			
薬剤		平均	SE	L 90%	調整平均	SE	L 90%	P群の15%増
P		2.8150	0.1448	2.5703	2.8241	0.1091	2.6410	
S		3.6850	0.1448	3.4403	3.6888	0.1091	3.5056	3.2373
T		3.4888	0.1448	3.2441	3.4758	0.1091	3.2926	3.2373
薬剤	薬剤	差	差のSE	L 90%	差	差のSE	L 90%	P群の15%
S	P	0.8700	0.0996	0.7052	0.8646	0.0995	0.6999	
T	P	0.6738	0.0996	0.5090	0.6517	0.0996	0.4868	0.4223
S	T	0.1963	0.0996	0.0315	0.2129	0.0996	0.0481	0.4223

## 5. 考察

大動物を用いたテレメトリーTQ/TQ<sub>c</sub>試験ではクロスオーバー法での実験が定着しているため、信頼区間方式による判定を行うために、どのような統計解析が適切であるかについて、元データが公開され、また各種の解析結果についても提示されている Littell らの 1 秒あたりの呼吸量 FEV1 のデータを用いて検討した。

ICH-E14 で提示されている信頼区間方式では、時間を一致させた被験薬群とプラセボ群の平均の比較としている。並行群間比較の場合には、症例に関する変動が信頼区間に入り込み、信頼区間が広がる原因となる。さらにプラセボ群と被験薬群の平均の差の信頼区間は  $SE_{\text{群間差}} = \sqrt{2\hat{\sigma}_{e(\text{patient})}^2} / n$  となり、被験薬群の例数を 2 倍増やしても信頼区間の幅の  $\sqrt{2}$  分の 1 にしかならない。

並行群間試験の場合は、表 4 に示すように誤差分散に症例間の変動が含まれ SE が増大し、プラセボと T 薬の平均値の差の SE にも入り込んで、信頼区間の 90% 下限値を押し広げる原因となる。クロスオーバー法の場合は、表 3 に示したように症例間の変動を誤差変動から分離することにより誤差分散を大幅に減少することができる。

並行群間試験とクロスオーバー法による同時点の平均値の差の信頼区間の算出方法に明らかな違いがあり、圧倒的に並行群間比較が不利である。これは、平均値の差の信頼区間を前提にしているためであり、公平な判定を行うために判定基準の明確化が必要と思われる。次のような判定基準とすることにより、試験法による差異がなくなる。

判定 1. 並行群間試験でもクロスオーバー試験でも時間を一致させた被験薬群の片側 95% の信頼区間が、プラセボ群の平均値の 10 ms 増しの限界値を下回る場合に陰性とする (新たに処方される集団に対する評価)

表 3 に示した症例を変量効果とした場合に薬剤 T の信頼区間の下限は 3.2553、プラセボ P の 15% 増増しは 3.2373 と、0.0181 上回っており薬効が認められる。ここで計算されている信頼区間は、表 4 に示した並行群間試験とみなして解析した結果と同様であり、試験法による判定に差異が生じにくい。しかし、この判定基準だけでは、苦勞してクロスオーバー法で行った苦勞が報われない。表 3 に示したクロスオーバー法での両群間の差の信頼区間の下限は、0.5278、プラセボ P の 15% 増しは 0.4223 と、0.1055 とゆとりをもって上回っており、明確な薬効が認められる。しかしながら、この判定は同一症例の中で比較となり、新たに試験薬が処方される集団に対する平均的な増加を評価しているわけではない。したがって、2 つの判定を併記し、考察することを薦めたい。

判定 2. クロスオーバー法の場合、時間を一致させた被験薬とプラセボ群の差の片側 95% の信頼区間が、10 ms を下回る場合に陰性とする（個々の症例に対する評価）

クロスオーバー法の場合でも、ある時間の元データを使った場合、投与前値からの差のデータを使った場合で、投与前値を共変量とすることで、表 6 に示したように投与群の *SE*、および群間の *SE* が完全に一致することが確認された。これにより、元データを使った場合と投与前値からの差のデータを使った場合の結果の不一致がなくなり、両方で解析結果が異なった場合の解釈をどうするかという悩ましい問題が解決される。ICH-E14 では、中心傾向の解析のため投与前値からの差のデータについて時間を一致させたプラセボと被験薬の平均値の差に対して多面的に検討するように求めているが、投与前のベースライン値を共変量とした解析を標準的に使うべきである。

クロスオーバー法で、時期ごとの投与前値を共変量として用いた場合のいない場合に対する相対効率については、Yan (2011) によれば次の通りであった。すなわち、時期数が 2 の場合に相対効率は高いが、時期数が 3 で、薬剤数が 3 の場合には時点間の相関が高ければ相対効率は上がるが、高々数パーセントでしかない。今回の群間差の *SE* が元データの場合 0.5278、差のデータの場合 0.4930、共変量とした場合 0.5063 であり、Yan が示した結果と整合している。

クロスオーバー試験では症例を固定効果とするか変量効果とするかによって薬剤群の平均値に対する *SE* が大きく異なり、固定効果とした場合には、同じ被験者に再度実験をした場合の評価となり、一般化可能性を考えない判定になるので、推奨することができない。また症例を変量効果と指定しても REML 法での計算を行うことが必須である。

測定時点を固定効果としてモデルに取り込んだ、いわゆる反復測定共分散分析の適用については、クロスオーバー法での Littell らの実験データを用いた検討では、分析精度の向上を確認することはできなかった。並行群間試験の場合には、複数時点のデータを取り扱うことによって分散の安定化が図れ、分析精度の向上が期待できるが、クロスオーバー法の場合には、同じ症例に対する同時刻で繰り返し測定が解析モデルの中に含まれており、分析精度の向上を見いだせなかったと判断される。

今後、さらに検討を要するが、時点を含めた複雑な解析を行うよりも、時点ごとに、あるいは数時点を平均して、それぞれについて信頼区間による総合的な判定を行うことを推奨する。

## 文献

Yan Z. (1997), The impact of baseline covariates on the efficiency of statistical analyses of crossover designs. *Statistics in Medicine*. 32: 956-963.

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (2006) *SAS System for Mixed Models 2ed*. SAS Institute.

渡橋靖 (2009), QT 延長をいかに判定すべきか : ICH E14 に準拠した臨床試験デザインと統計解析法, 日薬理誌 (*Folia Pharmacol. Jpn.*) 133, 14-8.

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 ユーザ会論文集, 263-86.

高橋行雄 (2009), 薬理学研究における経時データ解析の考え方 — 血圧降下試験事例による解説 —, 日薬理誌 (*Folia Pharmacol. Jpn.*) 133, 325-31.

高橋行雄 (2010), 経時データに対する投与前値を考慮した解析モデルの比較検討, SAS ユーザ会論文集, 45-54.

付録：データリスト

patient	P(p)			S(c)			T(a)		
	0 h	1 h	d(1-0 h)	0 h	1 h	d(1-0 h)	0 h	1 h	d(1-0 h)
201	2.14	2.36	0.22	2.30	3.41	1.11	2.46	2.68	0.22
202	3.37	3.03	-0.34	2.91	3.92	1.01	3.50	3.95	0.45
203	1.88	1.99	0.11	2.08	2.52	0.44	1.96	2.28	0.32
204	3.10	3.24	0.14	3.02	4.43	1.41	3.44	4.08	0.64
205	2.91	3.35	0.44	3.26	4.55	1.29	2.80	4.09	1.29
206	2.29	3.04	0.75	2.29	4.25	1.96	2.36	3.79	1.43
207	2.20	2.46	0.26	1.96	3.00	1.04	1.77	3.82	2.05
208	2.70	2.85	0.15	2.70	4.06	1.36	2.64	3.67	1.03
209	2.25	3.45	1.20	2.50	4.37	1.87	2.30	4.12	1.82
210	2.48	2.56	0.08	2.35	2.83	0.48	2.27	2.77	0.50
211	2.12	2.19	0.07	2.34	4.06	1.72	2.44	3.77	1.33
212	2.37	2.14	-0.23	2.20	2.82	0.62	2.04	2.00	-0.04
214	2.73	2.57	-0.16	2.78	3.18	0.40	2.77	3.36	0.59
215	3.15	2.90	-0.25	3.43	4.39	0.96	2.96	4.31	1.35
216	2.52	3.02	0.50	3.07	3.90	0.83	3.11	3.88	0.77
217	1.48	1.35	-0.13	1.21	2.31	1.10	1.47	1.97	0.50
218	2.52	2.61	0.09	2.60	3.19	0.59	2.73	2.91	0.18
219	2.90	2.91	0.01	2.61	3.54	0.93	3.25	3.59	0.34
220	2.83	2.78	-0.05	2.48	2.99	0.51	2.73	2.88	0.15
221	3.50	3.81	0.31	3.73	4.37	0.64	3.30	4.04	0.74
222	2.86	3.06	0.20	2.54	3.26	0.72	2.85	3.38	0.53
223	2.42	2.87	0.45	2.83	4.72	1.89	2.72	4.49	1.77
224	3.66	3.98	0.32	3.47	4.27	0.80	3.68	4.17	0.49
232	2.88	3.04	0.16	2.79	4.10	1.31	2.49	3.73	1.24
平均	2.636	2.815	0.179	2.644	3.685	1.041	2.668	3.489	0.820
SD	0.518	0.577	0.337	0.546	0.707	0.477	0.554	0.728	0.584