順序カテゴリー反応の
探索的データ解析

第 **23** 回 高橋セミナー

2006 年 **1** 月 **28** 日

高橋 行雄

1/30/2006 9:44 AM

表紙裏

2006 年 1 月 23 日，新規作成，
ファイル名：E:¥Q8¥Q8_ANZ_S23_共変量¥S23_共変量.doc

順序カテゴリー反応の探索的解析の実践

目　次

1. はじめに 1

2. JMPによる順序ロジスティック回帰の練習 3

2.1. 臨床試験のデータ 3
2.2. パラメータの推定値の意味付け 9
2.3. JMPのバージョン 5 とバージョン 6 の比較 10
2.4. JMPとSASの標準誤差が不一致となる原因 11
2.5. JMPによる順位和検定 14

3. JMPによる交互作用 16

3.1. 疾患の重症度と薬剤の交互作用 16
3.2. バージョン 6 での拡張 18
3.3. 個別症例のデータ 19

4. 治療群と共変量間の交互作用の探索 21

4.1. 交互作用の探索の必要性 21
4.2. 臨床試験データ 24
4.3. 治療コード開示後の解析 28
4.4. 施設と治療群の交互作用 32
4.5. 層ごとの有意差検定の是非 35
4.6. 旧統計ガイドラインでの共変量による調整の考え方 36
4.7. 共変量による調整の是非 41
4.8. 複数の共変量を含めた解析 43
4.9. CPMPが提示している留意点 46
4.10. 共分散分析の無作為化比較臨床試験における留意事項 48

5. SASコード 51

図 表 目 次

表示 2.1 臨床試験データ 3
表示 2.2 JMPテーブル clinical_2x3_v5 3
表示 2.3 「モデルのあてはめ」で薬剤を名義尺度で与えた場合 4
表示 2.4 薬剤を名義尺度で与えた順序ロジスティック回帰の標準出力 4
表示 2.5 順序ロジスティック回帰による推定 5
表示 2.6 推定値のグラフ 5
表示 2.7 「モデルのあてはめ」で薬剤を順序尺度で与えた場合 8
表示 2.8 薬剤を順序尺度で与えた順序ロジスティック回帰の標準出力 9
表示 2.9 累積対数オッズに対する推定値の計算 10
表示 2.10 対数オッズからオッズ比と 95%信頼区間の計算 10
表示 2.11 V5 のオプション 10
表示 2.12 V6 のオプション 11
表示 2.13 V6 の環境設定オプション 11
表示 3.1 重症度別・治療法別の反応 16
表示 3.2 JMPテーブル clinical_3x2x3_v5 16
表示 3.3 「モデルのあてはめ」で重症度と薬剤の交互作用を設定 17
表示 3.4 交互作用を含む順序ロジスティック回帰の標準出力 17
表示 3.5 ワルド検定による結果 18
表示 3.6 バージョン 6 の新しい交互作用プロファイル 18
表示 3.7 重症度と薬剤の交互作用プロファイル 19
表示 3.8 重症度と薬剤の交互作用プロファイル 19
表示 3.9 個別の 305 症例のリスト，clinical_305.jmp 20
表示 3.10 薬剤別の患者の背景因子の分布 20
表示 4.1 共変量の同定 25
表示 4.2 累積ロジットによる共変量の同定 26
表示 4.3 共変量への対応 28
表示 4.4 治療法のみの場合 28
表示 4.5 治療群と各共変量との累積反応率についての解析 29
表示 4.6 共変量が反応におよぼす影響 30
表示 4.7 治療群別の++反応率とオッズ比 30
表示 4.8 治療群別の +以上の反応率とオッズ比 31
表示 4.9 ICH の統計的原則における交互作用 31
表示 4.10 治療法と施設の主効果モデルとした解析 33
表示 4.11 施設と治療法の交互作用 33
表示 4.12 ICHの統計的原則における治療群と施設との交互作用 34
表示 4.13 質的交互作用・量的交互作用 35
表示 4.14 旧統計ガイドラインにおけるサブグループ解析 35
表示 4.15 旧統計ガイドラインでの調整手順の例示 37
表示 4.16 薬剤別の患者の背景因子の分布 38
表示 4.17 重症度別・治療法別の反応 39
表示 4.18 拡張Mantel法 39
表示 4.19 旧統計ガイドラインでの背景因子の偏りの調整 40
表示 4.20 浜田知久馬（2005），調整の是非 42
表示 4.21 順序ロジスティック回帰における変数減少法のSASプログラム 44
表示 4.22 追い出し基準を変えた場合の変数減少法 45
表示 4.23 追い出し基準が 0.15 の場合の線点図 45

1. はじめに

テーマは、「順序カテゴリー反応の探索的データ解析」である.

順序カテゴリー反応の探索的データ解析は，私（高橋）にとっても長年の課題であった.「田口の累積法」は，解決法の一つであったが，ロジスティック回帰の拡張である順序ロジスティック回帰によって新たな解析の道が開かれた. 順序ロジスティック回帰は，JMP のマニュアルで使われている用語で，一般的には累積オッズ比に比例性を仮定したロジスティック回帰である.

累積オッズ比の比例性を仮定するとは，どのようなことなのか．プラセボと実薬の 2 群の臨床試験で，薬効が 3 段階の順序カテゴリーの反応（++，+，－）の場合について示す．プラセボの反応が ++の率で 19.2%，+以上の率で 48.3%であり，実薬の有効率がそれぞれ，26.6%，59.4%とする．オッズ比は，

++： $\theta_1 = \dfrac{0.266/(1-0.266)}{0.192/(1-0.192)} = \dfrac{0.362}{0.238} = 1.52$

+以上： $\theta_2 = \dfrac{0.594/(1-0.594)}{0.483/(1-0.483)} = \dfrac{1.463}{0.934} = 1.57$

と計算できる．反応が++，反応が+以上のオッズ比が，共に 1.55 前後である．このことから，プラセボと実薬のオッズが，ほぼ比例的であることがわかる．このように反応率のオッズに比例性が仮定できるときに, 反応が 2 値の場合のロジスティック回帰を拡張したのが順序ロジスティック回帰である．オッズの対数を取って線形化した場合には，反応が++の場合と反応が+以上の場合には，プラゼボと実薬の対数オッズの差が等しいとの仮定となる.

JMP のロジスティック回帰は，反応が順序尺度の場合への拡張のみならず，説明変数が質的変数である場合に，名義尺度あるいは順序尺度として扱う機能を持ち，さらに質的変数間の交互作用も変数の積の形で解析モデルに組み込むことができる. このことは，分散分析，共分散分析スタイルの探索的なデータ解析も手軽に行えることになる.

臨床試験の統計解析で，「症例の背景因子の偏りの調整」の問題がある．反応が順序カテゴリーの場合に，従来は，順序カテゴリーを 2 値反応に切りなおして，この問題に対応してきた．また，従来のロジスティック回帰は，説明変数が量的変数に限られ，3 水準以上の質的変数を取り扱う場合，さらに交互作用を検討したいときには，自らダ

ミー変数をデータとして作成する必要があった. 豊富な機能を持つ JMP のロジスティック回帰は，順序カテゴリー反応である「症例の背景因子の偏りの調整」の問題をスマートに取り扱える.

探索的な順序カテゴリー反応の臨床試験データの解析につて，2 つのガイドラインに準拠して例示する．一つは，1998 年に厚生労働省が提示した ICH の統計的原則に示されている「共変量」考え方であり，他方は，1992 年に提示された統計ガイドラインでの「背景因子の調整」の考え方である.

「背景因子の調整」については，積極的に調整を行うべきとの意見と，ランダム化に基づいた基本的な解析を優先すべきとの調整に否定的な意見もある. これらににについて, EMEA（The European Agency for the Evaluation of Medicinal Products）の CPMP（Committee for Proprietary Medicinal Products）が 2003 年に提示した Points to Consider on Adjustment for Baseline Covariates は，この問題に対するヨーロッパの規制当局および製薬団体の共同の意見として尊重すべきである．浜田（2005）も，この問題について日本科学技術連盟の第 15 回の BioS コース講義資料で取り上げているので紹介する．清見，西田，西島（2005）は,「共分散分析の無作為化比較臨床試験における留意事項」と題して，吉祥寺東急イン第 4 回東京吉祥寺地区月例研究会で，Points to Consider on Adjustment for Baseline Covariates にたいする周辺のコメント・考察、留意事項の講演資料もこの問題の理解に役に立つ.

2. JMP による順序ロジスティック回帰の練習

2.1. 臨床試験のデータ

プラセボとA薬の 2 用量の薬効を，ダブルブラインド法でヒトに投与して 表示 2.1 の結果を得た．反応変数は（++：治癒，+：改善，–：無効）であり，順序に意味のある順序カテゴリー反応である．

表示 2.1 臨床試験データ

薬剤	++	+	–	合計
P 薬	29	44	78	151
A 薬	41	51	62	154

表示 2.2 JMP テーブル clinical_2x3_v5

clinical_2x3_v5

clinical_2x3_v5
Source
列(9/0)
薬剤_名義
薬剤_順序
x(1,-1)

	薬剤_名義	薬剤_順序	x(1,-1)	x(0,1)	効果	y	症例数	累積症例数	累積有効率
1	1:P	1	1	0	1:++	1	29	29	19.2
2	1:P	1	1	0	2: +	2	44	73	48.3
3	1:P	1	1	0	3: –	3	78	151	100.0
4	2:A	2	-1	1	1:++	1	41	41	26.6
5	2:A	2	-1	1	2: +	2	51	92	59.7
6	2:A	2	-1	1	3: –	3	62	154	100.0

薬剤：名義尺度， 効果：順序尺度， 症例数：連続尺度

順序カテゴリー反応の統計解析は，順位和検定が広く使われている．反応が 2 値の場合には，反応率を対数オッズ（ロジット）に変換し線形化したロジスティック回帰が広く適用されている．ロジスティック回帰は，反応が 3 値以上の場合にも拡張され統計ソフトにも組み込まれてきたが，その活用事例は乏しい．

そこで，JMPで簡単な事例による導入を行う．さらに，複数の共変量があり，反応が順序尺度である臨床試験の統計解析の事例を取り上げる. 表示 2.2 のJMPデータを使って，薬剤群間の“有意差検定”を，「順序ロジスティック回帰」を使ってみよう．

表示 2.3　「モデルのあてはめ」で薬剤を名義尺度で与えた場合

手法が「順序ロジスティック」に自動的に設定されていることを確認する.

表示 2.4　薬剤を名義尺度で与えた順序ロジスティック回帰の標準出力

順序ロジスティックのあてはめ 効果

度数: 症例数

モデル全体の検定

モデル	(−1)*対数尤度	自由度	カイ2乗	p値(Prob>ChiSq)
差	2.1824	1	4.364789	0.0367 *
完全	320.6698			
縮小	322.8522			

R2乗(U)	0.0068
オブザベーション(または重みの合計)	305

勾配による収束

あてはまりの悪さ(LOF)

要因	自由度	(−1)*対数尤度	カイ2乗
あてはまりの悪さ(LOF)	1	0.0109	0.021872
飽和モデル	2	320.6589	p値(Prob>ChiSq)
あてはめたモデル	1	320.6698	0.8824

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片[1:++]	−1.2275	0.1371	80.1697	<.0001 *
切片[2: +]	0.1635	0.1155	2.0030	0.1570
薬剤[1]	−0.2242	0.1076	4.3382	0.0373 *

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
薬剤	1	1	4.3648	0.0367 *

薬剤のデザイン行列：対比型（P 薬：1，A 薬：-1）

表示 2.2 が，JMPのバージョン 6 の標準出力である．薬剤についてのパラメータ推定値：-0.2242，カイ 2 乗：4.3382，P値：0.0373 から統計的に有意な差があると判断できる．その下の効果の尤度比検定では，尤度比カイ 2 乗：4.3648，P値：0.0367 となりパラメータ推定値のP値と若干異なるが，統計的には有意な結果である．パラメータ推定値のカイ 2 乗は分散共分散行列から導出されたWald検定となっている．尤度比検定は，表示 2.4 の「モデル全体の検定」の欄で出力されている(-1)＊対数尤度の差 2.1824 の 2 倍に一致する．

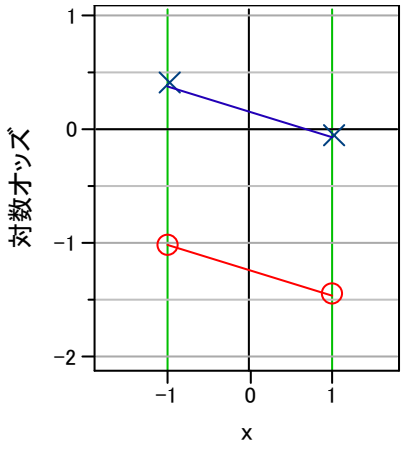
2 つの切片がパラメータ推定値にある．最初の切片は，反応が++の場合に対応し，第 2 の切片は，反応が+以上の場合に対応する．薬剤のパラメータの推定値が 1 つしかないのは，2 つの反応の対数オッズ変換値に対して共通の傾きを持つ回帰直線をあてはめていることによる．推定されたパラメータを用いて薬剤ごとの 2 つの反応率を推定した結果を 表示 2.5 に示す．

表示 2.5　順序ロジスティック回帰による推定

clinical_2x3_推定グラフ_v5

0/0列 / 2/0行	薬剤_名義	x	1:++	2:+以上	対数オッズ++	対数オッズ+以上	++推定値	+以上推定値
1	1:P	1	19.2	48.3	-1.437	-0.068	-1.4517	-0.0607
2	2:A	-1	26.6	59.7	-1.015	0.393	-1.0033	0.3877

++推定値：$-1.2275 + (-0.2242) * x$，+以上推定値：$0.1635 + (-0.2242) * x$



表示 2.6　推定値のグラフ

○：++実測値，—：++推定値，×：+以上実測値，—：+以上推定値

順序ロジスティック回帰の統計モデルについて,「JMP 統計およびグラフ機能ガイド（SGG）解説書」の統計的詳細を次に示す.

統計的詳細

JMPのモデル

この付録では、いろいろな種類の応答モデルとその因子、計画行列のコード変換、およびパラメータ化について説明します。また、マニュアルの各章で説明されている分析方法についても、さらに詳しく解説します。

JMPシステムでは、連続、順序、名義の3種類の応答モデルに線形モデルがあてはめられます。因子に関する詳細のほとんどは、どの応答モデルでも同じですが、グラフィックとプロファイルは、連続量の応答変数だけにあって、順序尺度と名義尺度の応答変数にはありません。

効果をあてはめ、仮説検定を組み立てるためには、計画行列がコード変換され、パラメータ化されますが、その方法はコンピュータのプログラムによって異なります。JMPでは、SASシステムのGLMプロシジャとは違うコード変換を採用していますが、ほとんどの場合、結果はどちらのシステムでも同じになります。以下の節では、JMPでのコード変換を詳しく紹介し、業界標準として認められているSASのGLMプロシジャと異なる点を、特にていねいに説明します。

順序尺度の応答

順序尺度の応答は、一連の平行なロジスティック曲線を累積確率にあてはめてモデル化されます。どの曲線も、計画パラメータは同じですが、切片は異なります。これは、次のように表されます。

$$j = 1, \ldots, r-1 \text{ に関して } \mathrm{P}(y \le j) = F(\alpha_j + \mathbf{X}\beta)$$

ここで、rは応答の水準数、$F(x)$ は次のような標準ロジスティック累積分布関数です。

$$F(x) = \frac{1}{1+e^{-x}} = \frac{e^{x}}{1+e^{x}}$$

これを別の式で表現するには、連続量の潜在変数であるzを使います。zの値によって、名義尺度の応答が決まります。

$$y = \begin{cases} r & \text{if } \alpha_{r-1} \le z \\ j & \text{if } \alpha_{j-1} \le z < \alpha_j \\ 1 & \text{if } z \le \alpha_1 \end{cases}$$

zは、線形モデルと誤差の潜在関数で、次のように表されます。

$$z = \mathbf{X}\beta + \varepsilon$$

εは、ロジスティック分布に従います。

これらのモデルは、応答の持つ順序尺度的な性格が認識され、名義モデルよりパラメータがずっと少なくなります。最尤法での計算が反復して行われるにもかかわらず、計算が速いので、使いやすいモデルです。

このモデルを適用する場合に，反応が++の場合の A 薬の P 薬に対するオッズ比と，反応が+以上のオッズ比が異質でないことが，順序ロジスティックモデルを適用する必要条件である．

説明変数が名義尺度の場合，順序尺度の場合に JMP ではダミー変数を自動的に内部で生成する．

名義尺度の因子

名義尺度の因子は、変換されて、計画行列のインジケータ変数になります。SAS GLMでは名義水準ごとに1つのインジケータ列が作成されますが、JMPでは、最後の水準を除くすべての水準に、同じインジケータ列が作成されます。最後の名義水準が起こったときは、他の列からそれぞれ1が引かれます。たとえば、名義尺度の因子であるAに3つの水準があり、それぞれGLMとJMPで次のようにコード変換されているとします。

図 A.1

	GLM			JMP	
A	**A1**	**A2**	**A3**	**A13**	**A23**
A1	1	0	0	1	0
A2	0	1	0	0	1
A3	0	0	1	-1	-1

GLMの線形モデルでは、計画行列の列の間に線形の従属性があり、最小2乗法による解は一般逆行列を取ります。A3パラメータがゼロになるような解が選択されます。

JMPでは、線形モデルの計画行列はコード変換され、欠損セルや偶然による共線性がない限り、フルランクになります。最後の水準におけるA効果のパラメータは、その他の水準を合計してマイナスをつけたものなので、すべての水準にわたってパラメータを合計すると、ゼロになります。

パラメータの解釈

注: 名義水準のパラメータは、その水準の予測応答と全水準にわたる予測応答の平均の差であると解釈できます。

因子水準の計画列は、その水準の0／1インジケータから最後の水準のインジケータを引いて作成されます。パラメータが上記のように解釈されるのは、このコード変換のためです。

図 A.2

JMPパラメータレポート	解釈	計画列のコード変換
切片	全水準の平均	**1**′
A[1]	$\alpha_1 - 1/3(\alpha_1 + \alpha_2 + \alpha_3)$	(A==1) - (A==3)
A[2]	$\alpha_2 - 1/3(\alpha_1 + \alpha_2 + \alpha_3)$	(A==2) - (A==3)

順序尺度の因子

順序尺度の因子は、名義尺度の因子とは別の方法でコード変換されます。パラメータ推定値の解釈、検定、最小2乗平均も異なります。

順序尺度の因子では、最初の水準は対照水準（ベースライン水準）として扱われ、パラメータは応答に与える効果の測度になります。コード変換は、因子の水準が異なる用量を表している場合に適しています。その際、最初の用量はゼロとみなされます。

図 A.10

項	コード変換した列		
A	A2	A3	
A1	0	0	コントロール水準、用量はゼロ
A2	1	0	少ない用量
A3	1	1	多い用量

説明変数が順序尺度と定義されていると，名義尺度とは異なるダミー変数が生成される．パラメータの推定値は，第 1 水準と第 2 水準の差，第 2 水準と第 3 水準の差, . . . となる．説明変数が 2 水準の場合に順序尺度と設定すると，0，1 型のダミー変数となるので，名義尺度の変数でも，意図的に順序尺度として，0，1 型にしたほうがパラメータ推定の解釈がしやすい場合がある．

表示 2.7　「モデルのあてはめ」で薬剤を順序尺度で与えた場合

表示 2.8　薬剤を順序尺度で与えた順序ロジスティック回帰の標準出力

順序ロジスティックのあてはめ 効果

度数: 症例数

モデル全体の検定

モデル	(−1)*対数尤度	自由度	カイ2乗	p値(Prob>ChiSq)
差	2.1824	1	4.3648	0.0367 *
完全	320.6698			
縮小	322.8522			

R2乗(U)	0.0068
オブザベーション(または重みの合計)	305

勾配による収束

あてはまりの悪さ(LOF)

要因	自由度	(−1)*対数尤度	カイ2乗
あてはまりの悪さ(LOF)	1	0.0109	0.021871
飽和モデル	2	320.6589	p値(Prob>ChiSq)
あてはめたモデル	1	320.6698	0.8824

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片[1:++]	−1.4517	0.1808	64.4887	<.0001 *
切片[2: +]	−0.0607	0.1584	0.1467	0.7017
薬剤_順序[2−1]	0.4484	0.2153	4.3389	0.0373 *

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
薬剤_順序	1	1	4.3648	0.0367 *

薬剤のデザイン行列：参照型（P 薬：0，A 薬：1）

2.2.　パラメータの推定値の意味付け

表示 2.9 に，薬剤を 0，1 型で与えた場合のパラメータ推定値を用いて薬剤の水準ごとに 2 つの反応の対数オッズの推定値の計算過程を示す. 薬剤を名義尺度で与えた場合に 1，-1 型のダミー変数となりパラメータ推定値は異なるが，互いに換算可能である.

JMPの出力（パラメータ推定値）

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片[1:++]	-1.4517	0.1808	64.4887	<.0001
切片[2: +]	-0.0607	0.1584	0.1467	0.7017
薬剤_順序[2-1]	0.4484	0.2153	4.3389	0.0373

表示 2.9　累積対数オッズに対する推定値の計算

薬剤	x(0, 1)	効果	症例数	累積	累積有効率	対数オッズ実測値	切片	共通傾き	対数オッズ推定値
							切片1		
1	0	1:++	29	29	19.2	−1.4371	−1.4517	0	−1.4517
2	1	1:++	41	41	26.6	−1.0150	−1.4517	0.4484	−1.0033
							切片2		
1	0	2: +	44	73	48.3	−0.0680	−0.0607	0	−0.0607
2	1	2: +	51	92	59.7	0.3930	−0.0607	0.4484	0.3877
1	0	3: −	78	151	100				
2	1	3: −	62	154	100				

この表は，Excel のオブジェクトで貼り付けてあるので，クリックすると計算法を確認できる．

表示 2.10　対数オッズからオッズ比と 95%信頼区間の計算

	対数オッズ
推定値	0.4484
SE	0.2153

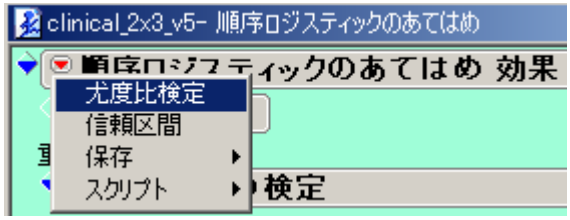
	対数オッズ	指数
推定値	0.4484	1.5658
L95%	0.0265	1.0268
U95%	0.8703	2.3877

この表は，Excel のオブジェクトで貼り付けてあるので，クリックすると計算法を確認できる．指数の欄が，オッズ比と 95%信頼区間の推定値である．

2.3.　JMP のバージョン 5 とバージョン 6 の比較

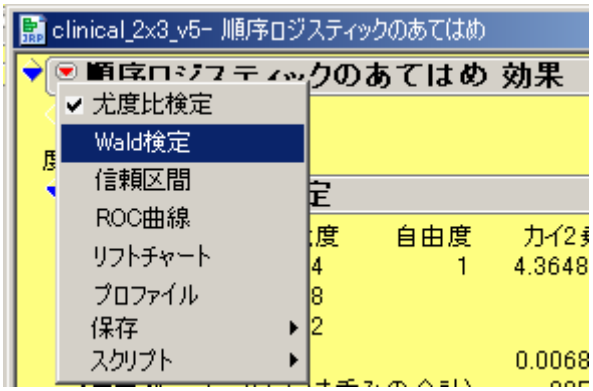
バージョン 6 では，解析機能が拡充されている．

表示 2.11　V5 のオプション



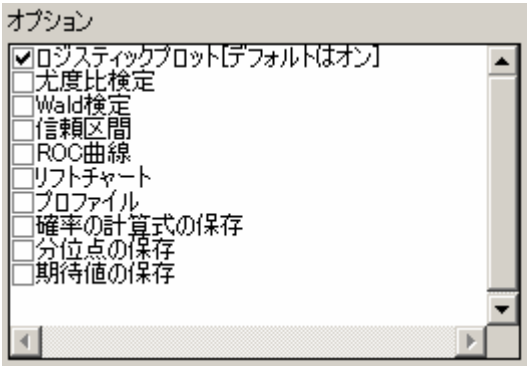
Wald 検定が標準出力，尤度比検定は追加指定

表示 2.12　V6 のオプション



尤度比検定が標準出力，Wald 検定は追加指定

表示 2.13　V6 の環境設定オプション



2.4.　JMP と SAS の標準誤差が不一致となる原因

どちらも最尤法によりパラメータの推定値を求めているのであるが，JMP では，フィシャーのスコアリング法でパラメータ推定値を求めている．SAS の GENMOD プロシジャではニュートン・ラフソン法でパラメータ推定値を求めている．両者のパラメータの推定値は一致するが，求められた分散共分散行列は若干異なるために，推定値の標準誤差が結果として一致しない．

SAS の LOGISTIC 回帰では，2 つの解法を選択できるようになっている．そのために，どちらの方法を使ったのかは，報告書に示す必要がある．どちらも近似解であり，どちらか一方を推薦する確固たる根拠がない．質的変数を複数のダミー変数に展開した上で，それらの変数（因子）全体についての包括的な検定も，尤度比検定とワルド（Wald）検定の 2 通りの出力があり，これもなのかも明示する必要がある．以下に，SAS のマニュ

アルの記述を示す.

GENMOD プロシジャ

Maximum Likelihood Fitting

The GENMOD procedure uses a ridge-stabilized Newton-Raphson algorithm to maximize the log-likelihood function $L(\mathbf{y}, \boldsymbol{\mu}, \phi)$ with respect to the regression parameters. By default, the procedure also produces maximum likelihood estimates of the scale parameter as defined in the "Response Probability Distributions" section (page 1402) for the normal, inverse Gaussian, negative binomial, and gamma distributions.

LOGISTIC プロシジャ

Iterative Algorithms for Model-Fitting

Two iterative maximum likelihood algorithms are available in PROC LOGISTIC. The default is the Fisher-scoring method, which is equivalent to fitting by iteratively reweighted least squares. The alternative algorithm is the Newton-Raphson method. Both algorithms give the same parameter estimates; however, the estimated covariance matrix of the parameter estimators may differ slightly. This is due to the fact that the Fisher-scoring method is based on the expected information matrix while the Newton-Raphson method is based on the observed information matrix. In the case of a binary logit model, the observed and expected information matrices are identical, resulting in identical estimated covariance matrices for both algorithms. You can use the TECHNIQUE= option to select a fitting algorithm.

JMP の結果と対比できるように，SAS の GENMOD プロシジャと LOGISTIC プロシジャによるプログラムと結果の出力を次に示す．LOGISTIC プロシジャは，2 つの最尤法の解法を示す.

SAS データセットの作成

```
Title 'clinical2x3.sas  2006-1-26  Y.Takahashi ' ;

data c2x3 ;
    input drug y r ;
datalines ;
1   1       29
1   2       44
1   3       78
2   1       41
2   2       51
2   3       62
;
```

GENMOD プロシジャ，ニュートン・ラフソン法（尤度関数の 2 回偏微分）

```
proc genmod data=c2x3 ;
    class drug / param=reference ref=first  ;
    freq  r ;
    model y = drug / dist=multinomial link=clogit ;
    run ;
```

The GENMOD Procedure

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept1	1	-1.4517	0.1810	-1.8064	-1.0971	64.36	<.0001
Intercept2	1	-0.0607	0.1587	-0.3717	0.2503	0.15	0.7021
drug 2	1	0.4484	0.2153	0.0265	0.8704	4.34	0.0373
Scale	0	1.0000	0.0000	1.0000	1.0000		

表示 2.8 に示した JMPの結果と比較する．切片 1 のStandard Error は，JMPでは，0.1808 であるが，GENMODプロシジャでは，0.1810 と 0.002 大きい．どちらが正しいということではなく，いずれも近似値と解釈すべきである．

LOGISTIC プロシジャ，フィッシャーのスコア法（尤度関数の 1 回偏微分）

```
proc logistic data=c2x3 ;
    class drug / param=reference ref=first  ;
    freq  r ;
    model y = drug / link=clogit technique=fisher expb ;
    run ;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-1.4517	0.1808	64.4841	<.0001	0.234
Intercept 2	1	-0.0607	0.1584	0.1467	0.7018	0.941
drug 2	1	0.4484	0.2153	4.3382	0.0373	1.566

表示 2.8 に示した JMPの結果と比較する．切片 1 のStandard Error は，JMPでは，0.1808 であり，LOGISTICプロシジャでも technique=fisher オプションでの結果は，0.1808 と一致する．Exp(Est) の欄のdrugの 1.566 は，P薬を 0，A薬を 1 としているので，P薬を基準としたオッズ比の推定値となっている．

LOGISTIC プロシジャ，ニュートン・ラフソン法（尤度関数の 2 回偏微分）

```
proc logistic data=c2x3 ;
    class drug / param=reference ref=first  ;
    freq  r ;
    model y = drug / link=clogit technique=newton expb ;
    run ;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-1.4517	0.1810	64.3563	<.0001	0.234
Intercept 2	1	-0.0607	0.1587	0.1463	0.7021	0.941
drug 2	1	0.4484	0.2153	4.3382	0.0373	1.566

LOGISTIC プロシジャで，technique=newton オプションを使うと GENMOD プロシジャの切片 1 の Standard Error の結果は 0.1810 と一致する．

2.5. JMP による順位和検定

JMP による順位和検定は，「2 変量のあてはめ」で反応を連続量とした場合の分散分析のオプションで実施できる．検定方法は，次の 3 種類である．

- **Wilcoxonの順位スコア**は、データの順位そのものです。Wilcoxonの検定は、ロジスティック分布の誤差に対して最も効果的な順位検定です。
- **メディアン順位スコア**は、順位がメディアン（中央値）の順位の上か下かによって1または0の値を取ります。メディアン検定は、二重指数分布の誤差に対して最も効果的な順位検定です。
- **Van der Waerdenの順位スコア**は、データの順位をオブザベーション数+1で割り、正規分布関数の逆関数を使って正規スコアに変換したものです。Van der Waerden の検定は、正規分布の誤差に対して最も効果的な順位検定です。

X変数に水準が2つしかないときは、該当する2標本検定の正規近似と一元配置検定のカイ2乗近似が計算されます。表 6.1「ノンパラメトリックな検定の選択の目安」(126 ページ) は、Xの水準数とそれに適した検定の対応表です。

表 6.1 ノンパラメトリックな検定の選択の目安

コマンド	2水準	2水準以上	適している分布形態
ノンパラメトリック-Wilcoxon	Wilcoxonの順位和	Kruskal-Wallis	ロジスティック
ノンパラメトリック-メディアン	2標本メディアン	*k*-標本メディアン (Brown-Mood)	両側指数
ノンパラメトリック-Van der Waerden	Van der Waerden	*k*-標本 Van der Waerden	正規

反応を 1，2，3 と連続量で与えた場合に JMP での順位和検定の結果は，次のとおりであり，3 種類の検定の P 値は，すべて有意水準 5%より小さい.

Wilcoxon/Kruskal-Wallisの検定(順位和)

水準	度数	スコア和	スコア平均	(平均-平均0)/標準偏差0
1	151	24590.5	162.851	2.081
2	154	22074.5	143.341	-2.081

2標本検定(正規近似)

S	Z	p値(Prob>|Z|)
24590.5	2.08110	0.0374＊

一元配置検定(カイ2乗近似)

カイ2乗	自由度	p値(Prob>ChiSq)
4.3339	1	0.0374＊

標本サイズが小さいので、この近似は精度がよくありません。統計表を使って検定してください。

メディアン検定(メディアンより上に位置する点の数)

水準	度数	スコア和	スコア平均	(平均-平均0)/標準偏差0
1	151	83.558	0.553364	2.045
2	154	68.442	0.444429	-2.045

2標本検定(正規近似)

S	Z	p値(Prob>|Z|)
83.557895	2.04505	0.0409＊

一元配置検定(カイ2乗近似)

カイ2乗	自由度	p値(Prob>ChiSq)
4.1822	1	0.0409＊

標本サイズが小さいので、この近似は精度がよくありません。統計表を使って検定してください。

Van der Waerdenの検定(正規分位点)

水準	度数	スコア和	スコア平均	(平均-平均0)/標準偏差0
1	151	15.725	0.10414	2.069
2	154	-15.725	-0.10211	-2.069

2標本検定(正規近似)

S	Z	p値(Prob>|Z|)
15.724676	2.06914	0.0385＊

一元配置検定(カイ2乗近似)

カイ2乗	自由度	p値(Prob>ChiSq)
4.2814	1	0.0385＊

標本サイズが小さいので、この近似は精度がよくありません。統計表を使って検定してください。

3. JMP による交互作用

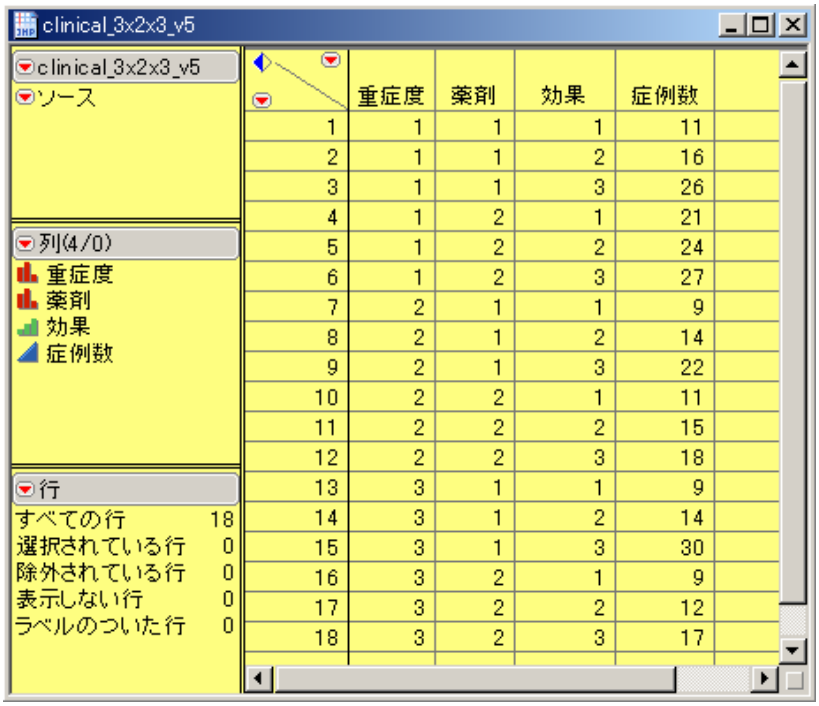
3.1. 疾患の重症度と薬剤の交互作用

順序ロジスティック回帰で交互作用の求め方，結果の見方，結果の解釈について例示する．疾患の重症度別・薬剤別の反応を，表示 3.1 に示す．

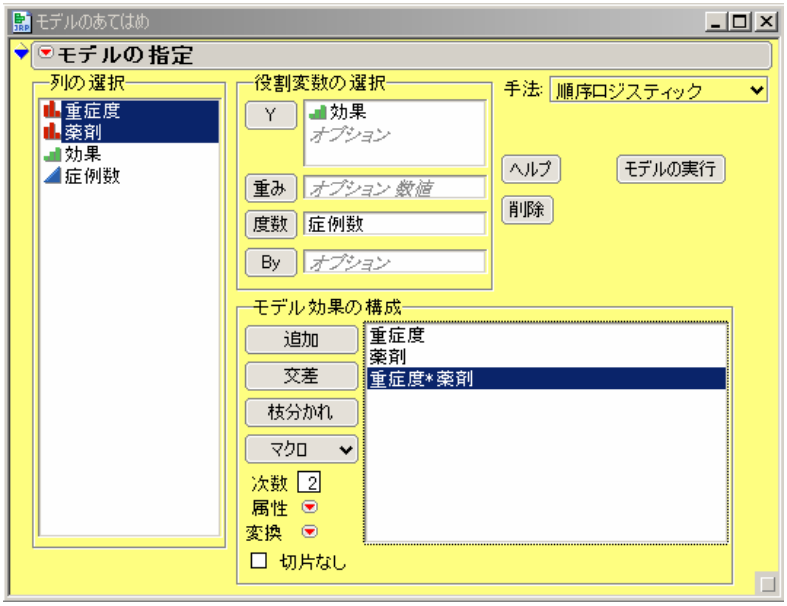
表示 3.1 重症度別・治療法別の反応

	P 薬				A 薬			
	++	+	−	計	++	+	−	計
1:軽度	11	16	26	53	21	24	27	72
2:中等度	9	14	22	45	11	15	18	44
3:高度	9	14	30	53	9	12	17	38

表示 3.2 JMP テーブル clinical_3x2x3_v5



表示 3.3　「モデルのあてはめ」で重症度と薬剤の交互作用を設定



表示 3.4　交互作用を含む順序ロジスティック回帰の標準出力

順序ロジスティックのあてはめ 効果

度数: 症例数

モデル全体の検定

モデル	(−1)*対数尤度	自由度	カイ2乗	p値(Prob>ChiSq)
差	2.90298	5	5.805968	0.3256
完全	319.94920			
縮小	322.85218			

R2乗(U)	0.0090
オブザベーション(または重みの合計)	305

勾配による収束

あてはまりの悪さ(LOF)

要因	自由度	(−1)*対数尤度	カイ2乗
あてはまりの悪さ(LOF)	5	0.05582	0.111647
飽和モデル	10	319.89337	p値(Prob>ChiSq)
あてはめたモデル	5	319.94920	0.9998

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片[1]	−1.2468	0.1395	79.89	<.0001*
切片[2]	0.1494	0.1176	1.61	0.2039
重症度[1]	0.1272	0.1468	0.75	0.3861
重症度[2]	0.0383	0.1584	0.06	0.8090
薬剤[1]	−0.2064	0.1099	3.53	0.0604
重症度[1]*薬剤[1]	−0.0269	0.1467	0.03	0.8543
重症度[2]*薬剤[1]	0.0531	0.1584	0.11	0.7375

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
重症度	2	2	1.2584	0.5330
薬剤	1	1	3.5403	0.0599
重症度*薬剤	2	2	0.1141	0.9445

表示 3.5　ワルド検定による結果

効果に対するWald検定

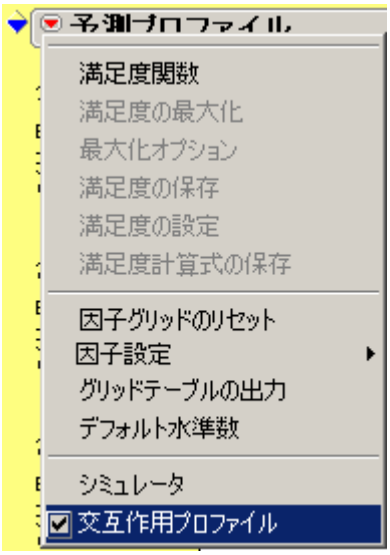
要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
重症度	2	2	1.2575	0.5333
薬剤	1	1	3.5258	0.0604
重症度*薬剤	2	2	0.1136	0.9448

重症度と薬剤の交互作用は，表示 3.4 および 表示 3.5 に示すように有意ではない．交互作用があった場合には，重症度と薬剤を組み合わせた場合の効果の推定値のグラフを作成して考察する必要がある．バージョン 5 までのJMPでは，反応が計量値の場合に付いて「交互作用プロファイル」によってグラフ化が手軽にできたのであるが，ロジスティック回帰では，サポートされていなかった．

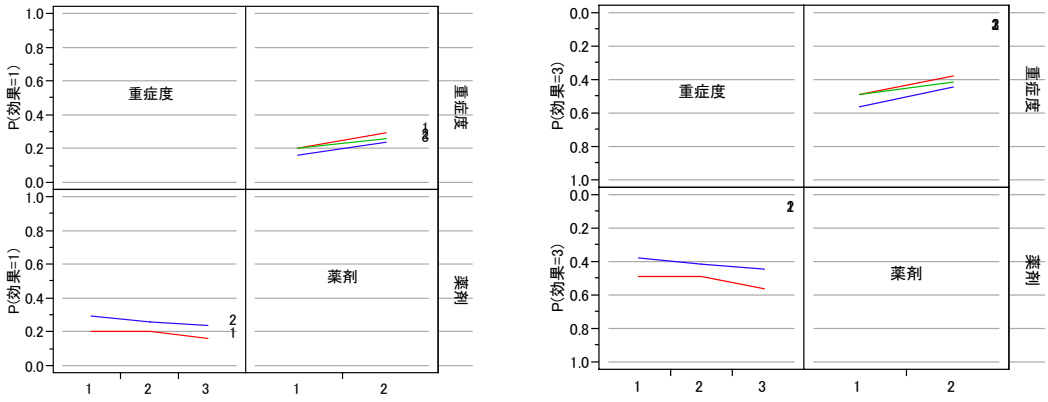
3.2.　バージョン 6 での拡張

バージョン 6 では，表示 2.12 に示すオプションの中に「プロファイル」が追加されている．「プロファイル」の出力は，1 変量ごとのグラフである．この中のオプションに「交互作用プロファイル」と「グリッドテーブルの出力」により，交互作用の検討に必要な推定値とグラフが得られる．

表示 3.6　バージョン 6 の新しい交互作用プロファイル



表示 3.7　重症度と薬剤の交互作用プロファイル



右の図は，反応が++の場合の推定値である．左は反応が－の場合で，目盛りを反転して，反応が＋以上のグラフとして使えるようにした．これは，グラフが反応が＋以上ではなく，反応が＋の場合の推定値が出力されているために，考察しにくいためである．

表示 3.8　重症度と薬剤の交互作用プロファイル

clinical_3x2x3_推定値v6

clinical_3x2x3_推定

列(6/0)
重症度
薬剤

行
すべての行 6
選択されている行 0

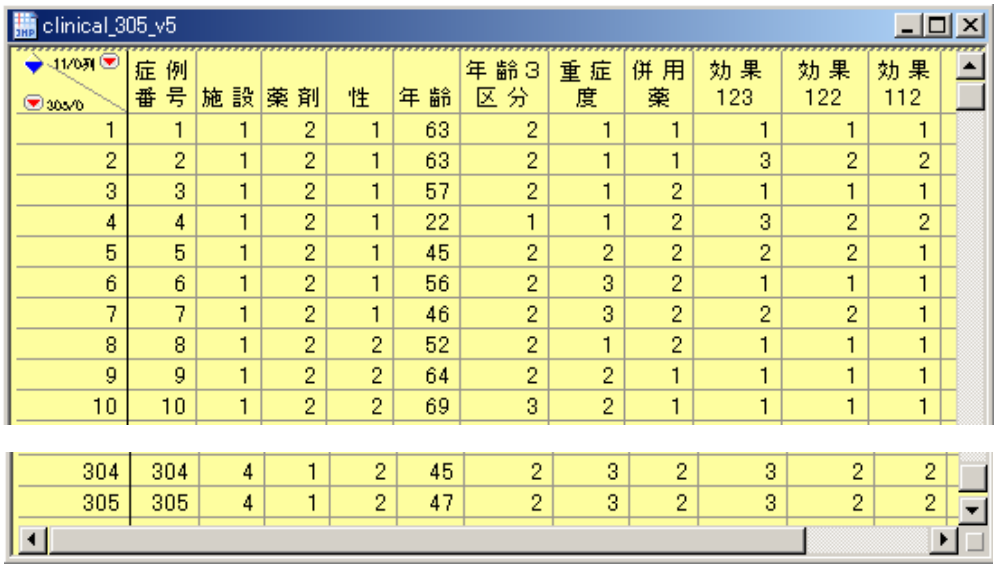
	重症度	薬剤	P(効果=1)	P(効果=2)	P(効果=3)	計
1	1	1	0.205	0.305	0.489	1
2	1	2	0.292	0.333	0.375	1
3	2	1	0.204	0.305	0.491	1
4	2	2	0.258	0.326	0.416	1
5	3	1	0.162	0.276	0.562	1
6	3	2	0.235	0.319	0.446	1

3.3.　個別症例のデータ

これまで示してきた表形式で集計したデータから順序ロジスティック回帰の事例を示してきたのが，これらは個別データから要約したものである．全体で 305 症例のデータの一部を次に示す．

第 4 章 で，この 305 症例の生データを用いて，治療前（投与前）の共変量を考慮した探索的データ解析を行う．解析は主にSASのLOGISTICプロシジャ（最尤解をニュートン・ラフソン法）を用いたが，JMPでも同様の解析を行うことができる．薬剤別の共変量の分布を 表示 3.10 に示す．

表示 3.9　個別の 305 症例のリスト，clinical_305.jmp


clinical_305_v5

症例番号	施設	薬剤	性	年齢	年齢3区分	重症度	併用薬	効果123	効果122	効果112
1	1	1	2	1	63	2	1	1	1	1	1
2	2	1	2	1	63	2	1	1	3	2	2
3	3	1	2	1	57	2	1	2	1	1	1
4	4	1	2	1	22	1	1	2	3	2	2
5	5	1	2	1	45	2	2	2	2	2	1
6	6	1	2	1	56	2	3	2	1	1	1
7	7	1	2	1	46	2	3	2	2	2	1
8	8	1	2	2	52	2	1	2	1	1	1
9	9	1	2	2	64	2	2	1	1	1	1
10	10	1	2	2	69	3	2	1	1	1	1
304	304	4	1	2	45	2	3	2	3	2	2
305	305	4	1	2	47	2	3	2	3	2	2

表示 3.10　薬剤別の患者の背景因子の分布

		1:P 薬		2:A 薬	
全体		151	100.0%	154	100.0%
施設	1:A 病院	17	11.3%	12	7.8%
	2:B 病院	45	29.8%	50	32.5%
	3:C 病院	31	20.5%	27	17.5%
	4:D 病院	58	38.4%	65	42.2%
性	1:男	61	40.4%	80	51.9%
	2:女	90	59.6%	74	48.1%
年齢	45 歳未満	37	24.5%	35	22.7%
	45 歳以上	73	48.3%	92	59.7%
	65 歳以上	41	27.2%	27	17.5%
重症度	1:軽度	53	35.1%	72	46.8%
	2:中等度	45	29.8%	44	28.6%
	3:高度	53	35.1%	38	24.7%
合併症	1:無	64	42.4%	86	55.8%
	2:有	87	57.6%	68	44.2%

4. 治療群と共変量間の交互作用の探索

ヒトを対象とした臨床試験における順序ロジスティック回帰を用いた共変量の探索の考え方と事例を示そう．2 値または順序尺度の反応に対して，共変量がまったくなければ，治療群の水準を s，反応の水準を r としたときに $s \times r$ の分割表に対する何らかの有意差検定手法を適用することで済む．しかしながら，反応と関連を持つ共変量が存在したり，片方の治療群に関連を持つ共変量があったりする場合もある．過去の経験から明らかな共変量もあれば，未知の共変量もあるかもしれない．共変量の分布が治療群間でアンバランスになってしまうこともあり，この場合には，共変量を考慮した解析も必要となる．

厚生省は 1992 年に，新医薬品の承認審査に用いる臨床試験（治験）に対して，薬効評価に際して留意する事項を「臨床試験の統計解析に関するガイドライン」で示した．その後，1998 年には，欧州，米国，日本の 3 極の規制当局と製薬団体の 6 者が合意した新しい「臨床試験のための統計的原則」を示し（以後，ICH の統計的原則という），1992 年のガイドライン（旧統計ガイドラインという）を廃止した．この 2 つのガイドラインで，共変量を考慮した探索的な解析の考え方が大幅に変更された．

旧統計ガイドラインでは，「共変量」ではなく「背景因子」および「予後因子」が用語として使われているが，ICH の統計的原則では，「背景因子」はまったく使われておらず，ランダム化に際して考慮すべき事柄の記述で「予後因子」が使われているのみである．ICH の統計的原則では，「背景因子」ではなく「共変量」を用いた記述になっている．本章では，旧統計ガイドラインでの背景因子に関連する統計解析の考え方について時代的背景を含めて解説しつつ，ICH の統計的原則での共変量に関連する統計解析の考え方に順じて，反応が順序データの場合に対する解析の事例を示す．

4.1. 交互作用の探索の必要性

治療群と共変量間の交互作用の探索は，新しい治療薬が市販後に様々な状態の患者様に使われたときの，効果がどの程度期待できるのか，有害事象・副作用がどのような場合に発生しやすいのかの情報の提供に主眼がある．高齢者で腎機能障害がある患者様の効果・副作用はどうなのか，肝機能障害がある患者様の有害事象の発現率はどうか，疾患が重症な場合は効果が減弱するのか，有害事象の発現頻度が高くなる患者特性はなにか，どのような注意を喚起したら副作用の発現を抑えられるのか，など多様な問いに試験統計家は答える準備をしておく必要がある．もちろん，限られたデータからは，はっ

きりしたことはわからなくても，注意を喚起しておくべきことを見逃さないように探索的な統計解析を行わなければならない．

臨床試験においては，どのような手順で探索的な統計解析を行うのかを，結果が出る前に手順を定めておく必要がある．本章では，このような状況を想定した探索的な統計解析の手順を示すと共に，ある臨床試験の生データに若干の修正を行ったものを用いて，解析の進め方を例示しつつ留意点を示す．

交互作用を探索する解析は，臨床試験では必要に応じて行うものではなく，計画的に行わなければならない．計画的な探索的な解析というのは，実験で検証したい仮説をあらかじめ設定できないためである．探索した範囲を明示し，解析を尽くしたが，「交互作用を示唆するような結果は見出さなかった」のであれば，新しい治療薬の効果の普遍性を示唆している．これは，さらなる臨床試験によって追試されなければならないことは，もちろんである．

共変量が治療群間の反応の差に及ぼす影響

治療群に対するランダム化に加えて，明らかな共変量に対して治療群間の分布をできるだけ均一に保つような最小化法[1]などのランダム化の方法も適用される場合もあるが，適用しない場合もある．最小化法を適用した場合であっても，治療群間で別の共変量の分布が不幸にもアンバランスとなってしまう場合もある．

共変量を無視して治療群間の効果の差について有意差検定を行うと，共変量の分布の状態で幸運にも有意な差になったり，逆に不幸なことに有意な差がなくなったりする場合が時として起きる．どちらの場合になるかは，治療群の共変量の分布に依存する．特に，治療群間の反応に対する有意差検定の結果が $p = 0.05$ 内外の場合に深刻な状況となる．

これは，長い年月をかけてようやく実現にこぎつけた新しい治療薬の臨床試験の結果であればあるほど,深刻なにである．それは，効果があるか否かの基準として両側 5%の有意水準が使われているためである．そのために，何らかの統計的方法で，共変量の影響が治療群間の比較に影響を与えなくするような統計的な方法が必要となる．

あらかじめ予想される共変量がある場合には，臨床試験の開始前に，共変量を含んだ統計手法を明文化しておくことが必須である．臨床試験の結果が出た後の探索的な統計解析の結果は，後知恵とみなさざるを得ない．新しい治療薬が優れるかを示すためには，新たに認識されたな共変量を考慮した臨床試験を再度実施し，この共変量を考慮した統

[1] 最小化法の説明を入れる．

計解析で，治療群間の P 値が 0.05 以下となることを示す必要がある．

事後的な共変量の同定とその限界

共変量は，実験前に反応に影響を与えることは確実であるが，実験前にはそれを制御できない変数である．ダブルブラインド法で行われている臨床試験では治療群のコードが関係者に明らかにされる前に，あらかじめ共変量と想定した変数が，実際に集められたデータで共変量となっているのか，実際のデータに未知の共変量が存在するのかを治療群をすべて込みにして検討する必要がある．

新たに共変量が同定された場合に，これを考慮した解析を効果の証明のために用いる場合は，実験計画書に記載した解析方法を，規定の手続きを踏んで変更する必要がある．この手続きなしに行った解析は，いずれにしても事後的な解析結果であり，効果の証明には用いることができない．

治療群のコードが開示される前の共変量の同定に際しては，あらかじめ治療開始前の症例の人口統計学的データ，疾患に関連するデータ，臨床検査の治療前データなどの中から，広めにその候補を選択し，それらを 1 変数ごとに説明変数としたロジスティック回帰を行う．あらかじめ定めた有意水準αよりも P 値が小さい変数をすべて説明変数としたロジスティック回帰を行い，共変量の同定をする．念のために，共変量と同定された変数とそれ以外の変数の 2 因子交互作用がないことを確認しておく．交互作用が出た場合には，それらを組み合わせて新しい変数とすることも考えられる．

治療群をマスクしたままでは，各治療法に共通な共変量の同定は可能であるが，特定の治療法にのみ共変量をして作用する場合，あるいは，治療法によっては逆の方向になるような場合，これらは交互作用のある場合は，同定できにくくなる．

交絡因子と予後因子の区別

共変量は主に統計学で用いられている用語で，疫学研究での予後因子と同じ意味として使われている．疫学研究では，ある疾患の発生の有無を反応としたときに，1 因子ごとの解析では反応と関連する因子であるが，反応と関連ある 2 つの因子を組み合わせると片方の因子の反応に対する影響が見出せなくなるような因子を交絡因子と言う．2 つの因子の分布がアンバランスになり相関があるような場合に，交絡が起きる原因となる．

治療群を 1 つの因子，ある共変量を 2 つめの因子としたときに，治療群間で，共変量の分布がアンバランスとなった場合に，互いに治療群間の反応の差を増強したり弱めたりする．このような場合に，2 つの因子が互いに交絡した状態となる．共変量の分布が，治療群間で統計的に有意な差となることもあるが，若干のアンバランスは常に起きる．

したがって，共変量を含めた解析により，治療群間の差が増強されているのか，弱められているのかを，探索的な解析によって示す必要がある．

疫学のケース・コントロール研究法は，交絡がしばしば起きることが知られている．肺ガンの患者群をケースとし，それと対応するコントロール集団を設定する．反応を肺ガンの発生とし，肺ガンの発生率と関連するする喫煙歴と飲酒歴の 2 つの因子が 1 因子ごとの検討でピックアップされたとしよう．どちらが肺ガンの発生率が高くするような予後因子で，どちらが，交絡因子となるのであろうか．あるいは，共に予後因子なのであろうか．

飲酒の習慣がある人たちに喫煙者が多いという相関がある場合に，喫煙が予後因子であるとした場合に，1 因子ごとの解析では，飲酒は肺ガンの発生率を見かけ上が高めてしまう．喫煙と飲酒の 2 つの因子を考慮した解析で，飲酒は，肺ガンの発生率と無関係となれば，飲酒は交絡因子となる．

4.2. 臨床試験データ

ここで用いる臨床試験データは，標準治療に新しい治療薬を上乗せしたときの効果を検証することを目的とした臨床試験の結果である．標準治療に新しい治療薬のプラセボを投与する P 群と，標準治療に新しい治療薬を加えて投与する A 群の 2 群の試験で，ダブルブラインドにて行われた．治療群との交互作用を検討する変数として，試験実施施設，性，年齢，疾患の重症度，併用薬の有無の 5 変数を取り上げる．事前に知られている共変量はなく，治療群がマスクされている状態で，未知の共変量の探索を行う．症例数は全体で 305 例であり，反応は，++，+，－の 3 段階の順序変数である．

共変量の同定

治療開始前に得られた観測データで，事後的に共変量となりうる．症例の性，年齢などの人口統計学的データ，疾患の重症度などの疾患に関連するデータ，併用薬など治療法に関連するデータ，様々な治療開始前の臨床検査データなど，実際には 100 以上の変数があることはまれではない．これらの変数と治療効果の関係をすべて，治療コードが関係者に開示される前，治療効果と関連する共変量となる可能性のある変数をピックアップし，新たに発見されて共変量があれば，それを加味した検証的な解析法の是非を慎重に検討し，その結果を記録として残しておく．

年齢，治療前の臨床検査データなどで連続変数となるは，3 区分程度の順序変数に変換しておくことを勧める．切れ目の良い区分でも，大きさの順で自動的に 3 区分，ある

いは 4 区分としてもよい．区分化は，共変量と治療効果の関係が直線的であるのか否かの検討に役に立つ．年齢は，ここでは，45 歳未満，45 歳以上 65 歳未満，65 歳以上と 3 区分とする．疾患の重症度などの順序データで，ある水準の症例が極端に少ない場合などは，水準の併合も必要である．併用薬が多岐にわたる場合は，ある一定数以上の併用薬に限定することも必要である．

表示 4.1 に共変量の候補となる変数と治療効果のクロス集計表を示す．各変数の水準間の有効率の差が 10%前後の開きがあるのは，年齢の反応が++%の場合に 45 歳未満とそれ以上の間，重症度の場合に+以上%で軽度と高度の間で 9.2%の差が見受けられる．他の施設，性，および合併症では，治療効果に 10%以上の影響を与えていない．

表示 4.1　共変量の同定

		反応 1:++	++%	2:+	+以上%	3:-	全体
	全体	70	23.0%	95	54.1%	140	305
施設	1:A 病院	7	24.1%	8	51.7%	14	29
	2:B 病院	19	20.0%	34	55.8%	42	95
	3:C 病院	13	22.4%	19	55.2%	26	58
	4:D 病院	31	25.2%	34	52.8%	58	123
性	1:男	33	23.4%	44	54.6%	64	141
	2:女	37	22.6%	51	53.7%	76	164
年齢	1:45 歳未満	11	15.3%	29	55.6%	32	72
	2:45 歳以上	42	25.5%	46	53.3%	77	165
	3:65 歳以上	17	25.0%	20	54.4%	31	68
重症度	1:軽度	32	25.6%	40	57.6%	53	125
	2:中等度	20	22.5%	29	55.1%	40	89
	3:高度	18	19.8%	26	48.4%	47	91
併用薬	1:無	34	22.7%	42	50.7%	74	150
	2:有	36	23.2%	53	57.4%	66	155

比例オッズ性の検討

重症度と年齢の 2 因子について，統計的に有意な共変量となるのか，順序ロジスティック回帰を適用してみる．1 因子ごとで共変量となっても，2 因子の主効果モデルでは，どちらかが交絡因子となっていて，統計的に有意とならない場合もあるので，重症度と年齢を主効果とする順序ロジスティック回帰を行う．

年齢については，45 歳未満を参照水準とするダミー変数とし，重症度は 1，2，3 を連続量とみなすことにする．反応が順序データのなので，「++」vs「+以下」の 2 値とするのか，「+以上」vs「 - 」の 2 値とするのか，オッズ比がどの累積反応率にかかわらず等しいとみなせる場合は，順序データのままでの順序ロジスティック回帰が行えるが，オッズ比が比例的とみなせない場合には，別々に 2 値反応としたロジスティック回帰を

行うことになる．これは，治療コードの開示前に決める必要がある．

重症度が軽度と高度の場合 ++の区分でのオッズ比は 0.712 であり，+以上の区分では 0.786 とほぼ等しい．年齢の 45 歳未満と 65 歳以上のオッズ比は，++の区分では 1.843 であり，+以上の区分では 0.934 となり，年齢による差は，++の区分でだけで，+以上の区分では差がない．細かく見れば等しいとは言い切れないこともあるが，全体として等しいとみなしてよいかの判断をすることになる．ここでは，オッズ比が等しいとみなして，2 つの区分を生かした解析を行うことにする．

予後因子・交絡因子

表示 4.2 に示すように 45 歳以下との比較で，45 歳以上 65 歳未満のオッズ比と 95%信頼区間は，1.171（0.702～1.951），65 歳以上は 1.174（0.636～2.168）と 1 を含み，P 値も小さいので，交絡因子とみなすことにしよう．重症度が 1 増加する場合のオッズは，0.832（0.645～1.072）となり，P値は 0.1551 と有意な差にはならないが，予後因子である可能性がある．

統計的にも有意となる共変量が検出され，検証的な解析とすべきとおもわれた場合は，文献情報，あるいは先行する臨床試験で確認を取る必要がある．その上で，共変量を考慮した解析方法への変更を慎重に判断する．そして，倫理委員会での審議をへた上で，試験実施計画書の修正を行わなければならない．

表示 4.2　累積ロジットによる共変量の同定

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-0.9885	0.3220	9.4250	0.0021	0.372
Intercept 2	1	0.3951	0.3169	1.5539	0.2126	1.485
age3 2	1	0.1576	0.2606	0.3659	0.5452	1.171
age3 3	1	0.1605	0.3128	0.2632	0.6079	1.174
severity	1	-0.1844	0.1297	2.0216	0.1551	0.832

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
age3 2 vs 1	1.171	0.702	1.951
age3 3 vs 1	1.174	0.636	2.168
severity	0.832	0.645	1.072

SAS/LOGISTICプロシジャ（ニュウトン・ラフソン法）

共変量間の交互作用

さて，年齢と重症度のように，2 つの変数の交互作用の検討はどうしたらよいのであろうか．治療群を含めた解析が本命の解析であるので，2 つの共変量を組み合わせ場合には，3 因子交互作用を検討することになる．経験的には，3 因子交互作用が有意となった場合にそれを解釈し，説明することは困難が伴う．また，3 因子交互作用は，ある水準の組み合わせにはずれ値的な変動があった場合にも現れれ，その解釈はこじ付けになりかねない．

このような理由により，共変量の同定の過程では，2 因子交互作用の検討は，行わないとの選択肢ももある．しかしながら，2 つの併用薬が共に共変量となったような場合に，（共に有，片方が有，共に無）などのように組み合わせて一つの共変量とすることを視野に入れておくとよい．

治療コード開示前の準備

治療コードが開示される前は，共変量の分布が治療群によってアンバランスになるかどうかはわからないので，治療コードが開示された後は速やかに，治療群別の共変量の分布を検討すると共に，治療群と各共変量の 2 因子ごとの主効果モデルで治療群間の P 値が大きくなるか小さくなるかが第 1 の関心事となる．第 2 には，治療群と各共変量の間の 2 因子交互作用が示唆されるかどうかである．

治療群全体では共変量と言いがたい場合でも，片方の治療法にのみ共変量となる可能性が捨てきれない場合もある．そのような場合は，治療群と各共変量の間の 2 因子交互作用が示唆されるかどうかが第 1 の興味となる．予後因子の場合には，第 1 の関心事は，主効果モデルでの治療群の P 値となり，第 2 の関心事が交互作用となる．

治療コード開示後の解析の目的を，共変量と反応との関連に対応して，表示 4.3 に示すように，第 1 の関心事と第 2 の関心事を解析計画書にまとめておことにより，探索的ではあるが計画的な解析となる．

表示 4.3　共変量への対応

共変量	反応との関連	交絡因子か予後因子か	交互作用の可能性	第1の関心事	第2の関心事
施設	なし	-	なし	治療群のP値	交互作用
性	なし	-	なし	交互作用	-
年齢	少々あり	交絡因子	なし	交互作用	-
重症度	少々あり	予後因子	ややあり	治療群のP値	交互作用
合併症	なし	-	ややあり	交互作用	-

関心事が治療群のP値となっている場合は，主効果モデル
交互作用は，治療群と各共変量間の交互作用
共変量が施設の場合の治療群のP値は，ICH統計的原則の要求事項

4.3.　治療コード開示後の解析

表示 4.4 に順序ロジスティック回帰の結果を示す．治療群間のP値は 0.0373 であり，有意な差であった．オッズ比は 1.566，95%信頼区間は（1.027，2.3988）と 1 を含まないので，A薬の効果が統計的に検証された．さて，この結果は，共変量の分布の状況によって，P値が 0.05 より大きくなることが無いのであろうか．共変量との交互作用はどなのであろうか.

表示 4.4　治療法のみの場合

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-1.4517	0.1810	64.3563	<.0001	0.234
Intercept 2	1	-0.0607	0.1587	0.1463	0.7021	0.941
drug 2	1	0.4484	0.2153	4.3382	0.0373	1.566

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
drug 2 vs 1	1.566	1.027	2.388

SAS/LOGISTICプロシジャ（ニュウトン・ラフソン法）

次に,治療群と各共変量について交互作用のP値と主効果モデルでの治療群間のP値を示す．施設と治療群の交互作用はp=0.2755 と有意な差ではないので，主効果モデルでの治療群のP値がどうなるかは関心事ではなくなる．この理由については，4.4 節で詳しく述べる.治療群と性の交互作用はp=0.0278 なので,グラフを見て考察する必要がある．年齢との交互作用はp=0.3914 と有意な差ではない.予後因子である重症度と治療群の交

互作用はp=0.9445 と有意ではないので，主効果モデルでの治療群のP値を見ると，p=0.0542 と共変量を含まない治療群のみのP値に比べて大きくなっている．どのような原因で大きくなったかの考察が必要となる. 治療群と併用薬の交互作用はp=0.0195 なので，反応率のグラフを見て考察する必要がある.

表示 4.5　治療群と各共変量との累積反応率についての解析

共変量	区分数	交互作用P値	治療群の主効果P値	オッズ比	結果の解釈
治療群のみ	-	-	0.0373	1.566	治療群間に有意な差がある
施設	4	0.2755	0.0368	1.572	施設によらず治療効果の差は均一
性	2	0.0278	-	-	効果のグラフを描いて考察する
年齢	3	0.3514	-	-	年齢によらず効果の差は均一
重症度	3	0.9445	0.0542	1.519	治療群のP値の増加の原因の探求
併用薬	2	0.0195	-	-	効果のグラフを描いて考察する

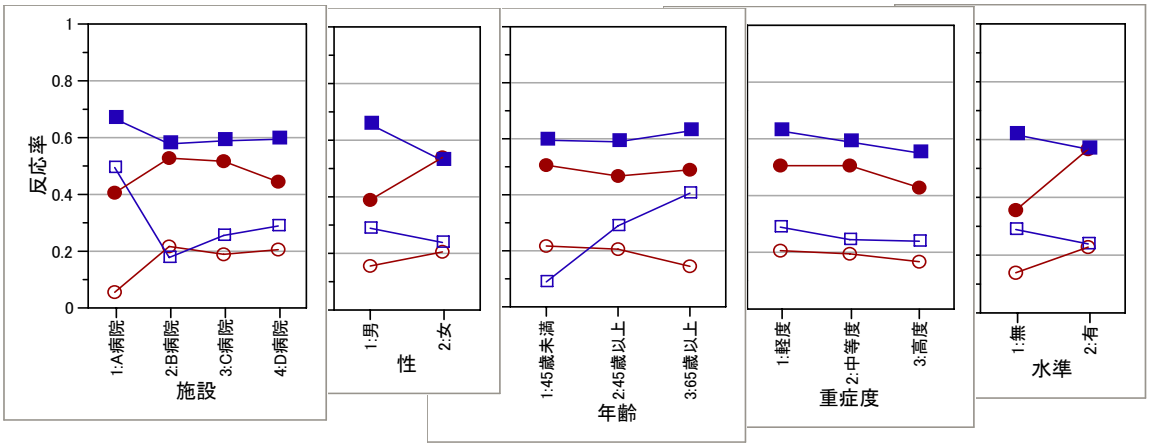
SAS/LOGISTICプロシジャ（ニュウトンラフソン法）

共変量ごとの反応率のグラフを
表示 4.6 に示す．施設はA病院で++の反応に40%以上の差がみられ，B病院ではP薬の++反応が高くなっていて，交互作用があるかのようであるが，p=0.2755 と有意な差ではない.

性別のグラフから，++反応率でも，+以上の反応率でも，P 薬では男性が女性よりも反応が低く，A 薬では逆になっている．この結果として，交互作用が p=0.0278 と有意な差になったのであるが，この結果の解釈には注意が必要である．これまでの臨床試験で性別が予後因子となることがあったのかどうか，A 薬は，男性で反応率が高くなっているが，性によらず反応が一定で，差があるような結果は偶然の変動なのかなどを考慮しつつ，短絡的な結果の解釈に慎重にならなければならない.

年齢は，45 歳未満の++反応率で A 薬より P 薬が高くなり交互作用がありそうであるが p=0.3914 と有意ではない．重症度は，軽度・中等度・高度と反応率は若干低下するが P 薬と A 薬の差はほぼ一定である．主効果モデルでの治療群間の P 値が，p=0.0542 と有意な差ではなくなるのは，重症度の分布がアンバランスとなっている影響と思われる.

併用薬の無い場合にP薬の反応が低く，有る場合に高くなっているので予後因子と考えられる．しかしながら，A 薬の場合は，併用薬の有と無しの反応の差は小さく，予後因子となっていない．併用薬の有る場合には，治療法間に差が無く，併用薬の無い場合に A 薬の治療効果が出るが，併用薬がある場合には，A 薬を加えても治療効果が増強されないという結果である.



表示 4.6　共変量が反応におよぼす影響

Y　P++　A++　P+>　A+>

共変量の水準ごとの治療法別の反応率を 表示 4.7 と 表示 4.8 に示す．治療法間の比較には，層ごとの有意差検定ではなく，オッズ比のみを示している．層ごとの有意差検定は，実験計画書に検証したい仮説として明記されている場合にのみ実施すべきである．その際には，適切な検出力を保てるような症例数設計がなされていなければならない．オッズ比を示したのは，交互作用が出た場合の解釈を容易にするためであり，有意差検定の代わりとなる 95%信頼区間が一人歩きしないように，あえて示さない．

表示 4.7　治療群別の++反応率とオッズ比

		1:P 薬				2:A 薬				オッズ比
全体		29	/	151	19.2%	41	/	154	26.6%	1.53
施設	1:A 病院	1	/	17	5.9%	6	/	12	50.0%	16.00
	2:B 病院	10	/	45	22.2%	9	/	50	18.0%	0.77
	3:C 病院	6	/	31	19.4%	7	/	27	25.9%	1.46
	4:D 病院	12	/	58	20.7%	19	/	65	29.2%	1.58
性	1:男	10	/	61	16.4%	23	/	80	28.8%	2.06
	2:女	19	/	90	21.1%	18	/	74	24.3%	1.20
年齢	1:45 歳未満	8	/	37	21.6%	3	/	35	8.6%	0.34
	2:45 歳以上	15	/	73	20.5%	27	/	92	29.3%	1.61
	3:65 歳以上	6	/	41	14.6%	11	/	27	40.7%	4.01
重症度	1:軽度	11	/	53	20.8%	21	/	72	29.2%	1.57
	2:中等度	9	/	45	20.0%	11	/	44	25.0%	1.33
	3:高度	9	/	53	17.0%	9	/	38	23.7%	1.52
合併症	1:無	9	/	64	14.1%	25	/	86	29.1%	2.50
	2:有	20	/	87	23.0%	16	/	68	23.5%	1.03

表示 4.8　治療群別の +以上の反応率とオッズ比

		1:P 薬			2:A 薬			オッズ比
全体		73	151	48.3%	92	154	59.7%	1.59
施設	1:A 病院	7	17	41.2%	8	12	66.7%	2.86
	2:B 病院	24	45	53.3%	29	50	58.0%	1.21
	3:C 病院	16	31	51.6%	16	27	59.3%	1.36
	4:D 病院	26	58	44.8%	39	65	60.0%	1.85
性	1:男	24	61	39.3%	53	80	66.3%	3.03
	2:女	49	90	54.4%	39	74	52.7%	0.93
年齢	1:45 歳未満	19	37	51.4%	21	35	60.0%	1.42
	2:45 歳以上	34	73	46.6%	54	92	58.7%	1.63
	3:65 歳以上	20	41	48.8%	17	27	63.0%	1.79
重症度	1:軽度	27	53	50.9%	45	72	62.5%	1.60
	2:中等度	23	45	51.1%	26	44	59.1%	1.38
	3:高度	23	53	43.4%	21	38	55.3%	1.61
合併症	1:無	23	64	35.9%	53	86	61.6%	2.86
	2:有	50	87	57.5%	39	68	57.4%	1.00

これまで示してきた解析の例示は，表示 4.9 に示すI CHの統計的原則の第 5.7 節「部分集団、交互作用及び共変量」の考え方に順じている．ここには，治験実施計画書に記述すべき内容，試験開始後（ランダム化後）に測定された共変量の扱いなどの原則が述べられいる．反応が順序データではなく 2 値の場合でも，あるいは計量値の場合でも，本章に示した解析の考え方は同じである．

表示 4.9　ICH の統計的原則における交互作用

5.7 **部分集団、交互作用及び共変量**

主要変数が試験治療によるものとは別の作用と、系統的に関連している例がよくみられる。例えば、年齢や性といった共変量と関連がある場合、又は多施設共同治験では異なる施設で試験治療を受けた被験者という部分集団間に差がみられる場合がある。共変量の影響又は部分集団の効果を調整することは予定した解析の重要部分となる場合があり、したがって治験実施計画書に記述すべきである。主要変数に重要な影響を及ぼすと予想される共変量と要因は、試験開始前に議論して確認しておくべきであり、精度を向上させ、試験治療グループ間のバランスの欠如を埋め合わせるため、それらを解析でどう取り扱うかを考慮すべきである。計画時に一つ以上の因子を層別因子として用いたならば、それらの因子を解析時に考慮することが適切である。調整することの潜在的な有益さが疑わしい場合にはしばしば、調整しない解析に主要な関心を払うと宣言し、調整した解析はそれを補うものとすることが薦められる。施設の効果及び基準となる時点での主要変数の測定値の役割には特別な注意を払うべきである。しかし、ランダム化後に測定された共変量を主要な解析で調整することは薦められない。なぜならばそのような共変量は試験治療の影響を受けている可能性があるからである。

試験治療の効果自身もまた部分集団や共変量により異なる可能性がある。例えば、効果は年齢とともに減少する場合があり、被験者の特別な診断分類では、より大きな効果があることもある。そのような交互作用は予想されたり、交互作用自体に重要な

関心がある場合（例えば高齢者医療など）があり、したがって部分集団別解析又は交互作用を含んだ統計モデルは予定した検証的解析に含まれる場合がある。しかし、多くの場合、部分集団別解析又は交互作用解析は探索的であるため、探索的であることを明確に確認しておくべきである。それらの解析では、全体でみられた試験治療の効果がすべて一様であるかどうかを探索すべきである。一般に、そのような解析は、まず問題となっている統計モデルに交互作用項を加えることから始めるべきであり、適切な被験者の部分集団内、又は共変量によって定められる層内を更に探索的に解析することによって補われる。探索的解析である場合、これらの解析結果は慎重に解釈すべきである。試験治療の有効性（若しくは有効性がないこと）、又は安全性に関する結論は、どのようなものであっても、探索的な部分集団別解析のみに基づいていては受け入れ難い。

4.4. 施設と治療群の交互作用

新しい治療法が従来の治療法に比べて優れていると検証できたとしよう. その結果を受けて，様々な条件で治療効果がどのくらい期待できるのか，できないのか，副作用が増強されるのはどのような場合なのか，などの探索的な解析を例示してきた．新しい治療法は，臨床試験に参画した施設以外で，その効果を再現できるのだろうか．臨床試験を実施できる施設は，ごく一部の施設に限定されていが，多くの施設でその効果をどのくらい期待できるのであろうか，一般化可能性があるのだろうか.

施設ごとの治療効果はどのくらい変動するのであろうか. どこの施設でも臨床試験で検証された新たな治療法の効果が出せるのであろうか. 施設により臨床効果はかなり変動することは経験的に知られている．ある施設では，治療法間で効果の差がまったく見られないのに，別の施設では大きな差があったりする．既存の治療法の効果でも施設間で大きな違いがあることをしばしば経験する．

施設を共変量とした解析についてI CHの統計的原則では，表示 4.12 に示すように第3.2 節で考え方を明示している．施設を共変量とした解析では，治療群との交互作用を含まない主効果モデルでの治療群間のP値を，交互作用を含めるモデルに優先することを求めている. その上で, 交互作用のP値を示し, 結果の考察を求めている. 表示 4.3 での共変量が施設の場合に第 1 の関心事に「治療群のP値」, 第 2 の関心事として「交互作用」としたのは，I CHの統計的原則の要求に答えるためである．

表示 4.10 に示すように治療法と施設の主効果モデルでの治療法の P 値は 0.0368 となり, 表示 4.4 に示したように施設を含まない治療法のみのモデルの P 値 0.373 に比べてわずかに大きくなっている．治療法と施設の交互作用は，

表示 4.6 を見ると施設Aでの反応が他の 3 施設とことなり交互作用が出そうであるがP値は 0.2755 と微妙である. 探索的な解析を, ++反応率でも, +以上の反応率に対して行っ

た結果を 表示 4.11 に合わせて示す．反応が+以上の場合に交互作用のP値は 0.1272 と小さくなっている．施設Aで，P薬の反応が他の 3 施設より低くなり，A薬の反応が高くなっている．施設Aは，他の 3 施設に比べて症例数が 1/3 程度なので，偶然による変動なのか，何か説明できる原因があるのか，さらなる検討が必要となる．

表示 4.10　治療法と施設の主効果モデルとした解析

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
DRUG	1	4.3600	0.0368
HOSP	3	0.0485	0.9972

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-1.4294	0.3717	14.7876	0.0001	0.239
Intercept 2	1	-0.0382	0.3619	0.0111	0.9160	0.963
DRUG 2	1	0.4524	0.2167	4.3600	0.0368	1.572
HOSP 2	1	-0.0507	0.4008	0.0160	0.8993	0.951
HOSP 3	1	0.0138	0.4283	0.0010	0.9744	1.014
HOSP 4	1	-0.0276	0.3923	0.0050	0.9439	0.973

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
DRUG 2 vs 1	1.572	1.028	2.404
HOSP 2 vs 1	0.951	0.433	2.085
HOSP 3 vs 1	1.014	0.438	2.347
HOSP 4 vs 1	0.973	0.451	2.098

SAS/LOGISTICプロシジャ（ニュウトンラフソン法）

表示 4.11　施設と治療法の交互作用

		++反応		+以上の反応		順序反応	
Effect	DF	Wald x2	p	Wald x2	p	Wald x2	p
DRUG	1	5.5073	0.0189	1.7835	0.1817	4.798	0.0285
HOSP	3	1.9726	0.5781	1.2135	0.7498	1.5235	0.6769
DRUG*HOSP	3	5.6984	0.1272	1.2453	0.7422	3.8733	0.2755

SAS/LOGISTICプロシジャ（ニュウトンラフソン法）

表示 4.12　ICH の統計的原則における治療群と施設との交互作用

3.2 多施設共同治験

前略

試験治療の効果の推定と検定に用いる統計モデルは治験実施計画書に記載すべきである。試験治療の主効果は、最初に施設と試験治療の交互作用を含まず施設間差を考慮に入れるモデルを用いて調べることができる。モデルに常に交互作用を含めると、試験治療の効果が施設間で均一な場合、主効果の検定の効率が低下する。試験治療効果の不均一性が真に存在する場合には、主効果の解釈には様々な議論がある。

後略

用語集

交互作用（質的及び量的）　Interaction (Qualitative & Quantitative)

試験治療の対比（被験薬と対照薬との差など）が、他の要因（施設など）により変わる状況。量的な交互作用とは、要因のレベルが異なるとそれに応じて対比の大きさが変わることであり、質的な交互作用とは、要因のレベルの少なくとも一つにおいて、対比の方向までもが変わることである。

Q&A

Q7. 3.2 では「試験治療の主効果は、最初に施設と試験治療の交互作用を含まず施設間差を考慮に入れるモデルを用いて調べることができる」とあるが、交互作用の有無を最初に検討しないのはなぜか。

（答）

本ガイドラインは、試験治療効果をまず調べるという立場で記載されており、このような考え方に立って、施設と試験治療の交互作用を含めず施設間差を考慮に入れるモデルを用いて調べる方法が記載されている。これは、交互作用を考慮しなくてよいという意味ではない。むしろ、このようなモデルで解析を行った場合には、仮に交互作用が存在する場合であっても、その影響を上回る差が試験治療間に存在することが必要である。

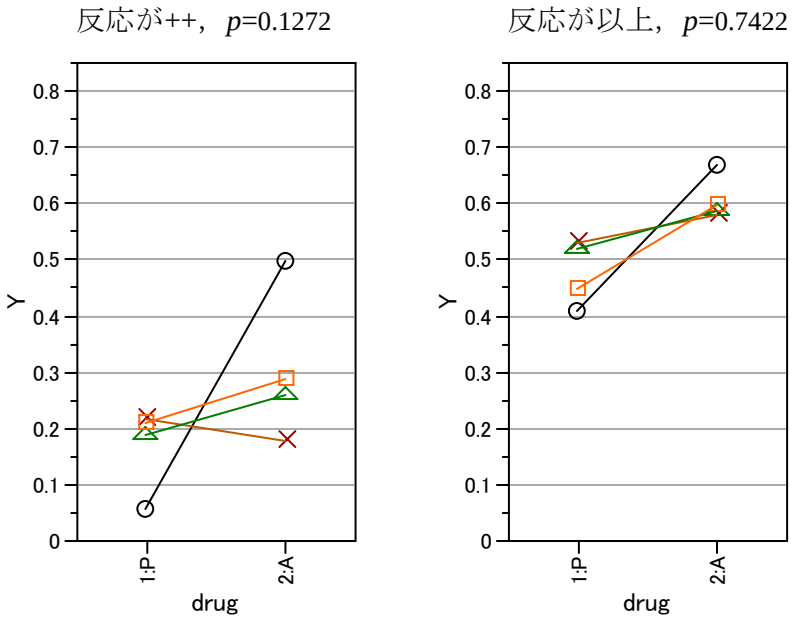
主効果の存在が確認された場合は、交互作用を調べ、交互作用が認められた場合には結果の安定性を確認することが必要である。特に質的な交互作用が認められた場合にはその理由を考察すべきである。理由が十分に説明できない場合は、追加の試験が必要となる。

また、主効果が認められないときには、仮に交互作用の検討の結果、試験治療がある条件の下に有効であることが示唆されるような場合であっても、新たにそのような条件を考慮した検証的試験を行うことが必要である。

質的交互作用か，量的交互作用か

施設と治療法の交互作用は，量的なのか，質的なのかを調べてみよう．表示 4.12 に示した ICH の統計的原則の抜粋の中の用語集に解説がある．これに準じて結果を考察してみる．

表示 4.13 に示すように治療法をX軸として，施設ごとにP薬とA薬の反応について線グラフで表す．反応が++%の場合にA施設は，CとD施設よりも反応の差が大きいが方向は同じであるが，B施設は差が負の方向となっており（対比の方向までもが変わっている），質的な交互作用が示唆されている．反応が+以上%の場合には，A施設の効果の差は大きいが，他の施設と傾きの方向は同じであり，P値がある程度小さければ量的交互作用が示唆されることになる．



表示 4.13　質的交互作用・量的交互作用

○：A 施設，×：B 施設，△：C 施設，□：D 施設

4.5.　層ごとの有意差検定の是非

「層ごとにどうなっているかを，単に見たいだけなのに，なぜ検定をしてくれないのか」という臨床試験の関係者の要求を試験統計家はそのまま受け入れてはならない．機械的に検定結果を示すことは，検定手法の誤用となるからである．有意差検定が試験統計家にとっても手軽に行えない時代では，気になる共変量が，予後因子なのか，交絡因子なのか，片方の治療法にのみ予後因子となるのか，それらの判断の初期の検討資料として層ごとに治療別の有効率と治療群間の有意差検定を機械的にしておくことは，それなりの価値があったと思われる．

しかしながら，コンピュータの普及と共に，さらに広汎に有意差検定が形式的に行われ，事後的な解析にもかかわらず，薬効の特徴づけのために層ごとの治療群間の有意差検定の乱用を誘発することになった．旧統計ガイドラインは，第 14 項で層ごと（サブグループ）の統計解析を，次のようにいさめていた．

表示 4.14　旧統計ガイドラインにおけるサブグループ解析

14. サブグループ解析 ある背景因子の特定の層（サブグループ）について、治験薬の効果を検証したい場合もある。しかし、サブグループごとに統計的推論を繰返すことに伴う多重性の

問題は、最も解決の困難なもののひとつである。複数の背景因子のそれぞれについてサブグループ解析を繰り返した場合には、多項目比較と同様に、背景因子間の相関構造の扱いに絡む未解決の問題があり、第 1 種の過誤の正確な制御が困難になる。複数の背景因子の組合わせによるサブグループ解析では、各サブグループでの被験者数の減少に伴う検出力の低下という第 2 種の過誤の確率の増大も問題になる。検証的な目的でのサブグループ解析では、サブグループごとの充分な被験者数の確保を計画段階から図る必要がある。

性や年齢階層などのサブグループによって治験薬と対照薬の効果の様式が異なるという交互作用は、治験薬の効果を評価する上で極めて重要な情報である。しかし、治験薬と同一系統の作用機序をもつ実対照薬との比較試験においては、治験薬の非臨床試験やそれ以前の臨床試験の成績の充分な検討なくしては、交互作用の予測は困難なことが多い。検証的データ解析では、それ以前の治験薬の効果に関する情報を総合し、薬剤と背景因子間の交互作用に注目して、事前に層別する背景因子を絞り込むことが不可欠である。

旧統計ガイドラインが制定された当時は，有意差検定は手軽につかえる統計手法であったが，本章で示したロジスティック回帰による交互作用の検討は統計ソフトを自由使える試験統計家にとってもたやすいことではなかった．そのために，旧統計ガイドラインは，次節で示すように，層をなす分割表の有意差検定を注意深く使うように求めていた．

すべての結果に対して，何か結果を述べる場合に有意差検定の結果を付ける必要があると思い込んでいる有意差検定症候症の人たちが存在する．これらの人たちに試験統計家は暖かく接しなければならない．「大切なのは，新しい治療法に対する予後因子の探索なのです．このためには，治療群間の比較ではなく，片方の治療だけに予後因子となる共変量の探索が必要なのです．層ごとの検定では，この問題の解決とはならないのです．」などのように誠意を尽くして説明する必要がある．

多くの有意差検定症候症の人たちと接し，試験統計家は「経験のある臨床医」になら無ければならない．試験統計家は，なにかを探し出したいとの関係者の気持ちを大切にし，探索的解析を通じて関係者と友好関係を築き上げなければならない．

4.6. 旧統計ガイドラインでの共変量による調整の考え方

共変量の探索に分割表に対する有意差検定を主体にしていた時代には，次のようなストラテージに基づいて解析がなされてきた．

1. 共変量の分布を治療法間で比較する．
2. 共変量の各層ごとに療法間で反応の比較を行う．
3. 分布に差があれば，共変量を考慮して治療法間で反応の比較を目視的に行う．

表示 4.15 に示すように旧統計解析ガイドラインでは，予後因子（共変量）が治療群間に偏りが生じているかどうかの有意差検定を行い，そのP値が 15%前後でも注意深く検討を行うように求めていた. 反応が 2 値の場合について交互作用の点検はBreslow-Dayの検定が例示され，各層に共通の薬効差があることを仮定できた場合に，第 3 章で示した層をなす 2×2 の分割表の解析で述べたMantel-Haenzel法の適用が薦められていた．これらの手順はICHの統計的原則にはもはや見出せないが，長年の慣行でもあり，適用されがちでもあるので，その統計的な問題点を指摘しつつ代替案を提示したい.

表示 4.15　旧統計ガイドラインでの調整手順の例示

1	背景因子の不均衡の検討
2	処置群間で背景因子の分布の検定 有意確率が15%前後も考慮する 複数の種類の統計手法で異なる種類の不均衡の点検
3	背景因子（予後因子）の分布 処置群間で異なっている場合
4	やむをえない場合は統計的調整 各層に共通の薬効差があることを仮定 2値の場合Mantel-Haenzel法 3分類以上，拡張Mantel法 2値の場合，交互作用の点検はBreslow-Dayの検定

処置群間で背景因子の分布の検定は必要なのか

臨床試験での治療法の割付は，実施計画で規定している登録基準に該当し，除外基準に抵触しない症例が選択さ，インフォームドコンセントが得られた場合に，行われる．あらかじめ定められた手順でランダム化・ブラインド化された治療法，ここでは P 薬または A 薬のどちらかが割り付けられる．割り付に際して症例の背景因子（共変量）は考慮しなっかたとしよう.

どのような背景因子（共変量）をもつ症例が結果として集積されたかは，臨床試験の結果を一般化する際に不可欠なデータなのである. 重症な症例が結果として多かったのか，少なかったのか，どのような年齢分布であったのかなどである.

表示 4.16 施設と治療法は, 4×2 の分割表になっていて, 分割表に対する適合度の χ^2 検定の適用は妥当である．重症度と治療法の場合はどうであろうか．3×2 の分割表に対する適合度の χ^2 検定の適用は妥当である．さて，治療法を群とみなし，重症度を順序データとしたWillcoxonの順位和検定の適応は妥当なのだろうか．これは，重症度を反応とみなした 2 群間の有意差検定をしたことになり, 統計手法の誤用と言わざるをえ

ない．筆者たちもこのような検定は，他に手ごろな手順がないとので，便宜的な方法としてやむをえないという立場であった．

分割表の検定も同じではないかとの疑問に対しては，分割表の検定は，2 つの質的変数がどちらも説明変数であっても反応（目的変数）であっても変数間のアンバランスを検討するための方法として適応できる．治療群と重症度は，反応に対して共に説明変数であるので，群間の差を問題にするような有意差検定を行うことは誤用である．

ICH の統計的原則には，患者の背景因子の不均衡を有意差検定により見出せとの記述はもはやないので，筆者たちは，検定はしないとの立場に立っている．参考までに検定結果を見たいとの要望には，「有意な差がなくても，有意な差が有ったとしても，ICH の統計的原則に準じた統計解析にはまったく関係しないので，心配しなくてもよくなったのです．」と丁寧に説明している．

規制当局から申請後にリクエストされたらどうするのか，との質問には，「リクエストの基盤となるガイドラインはもはや 1998 年に効力を失っているので，もしも，そのようなリクエストがでできそうになれば，規制当局の統計家がたしなめることになるでしょう．」などと輸すことに している．

それでも見てみたいと，食い下がる関係者もいないわけではないので，昔の規制当局からの指摘事項集で頻度が高かったことは承知しているが，現在の指摘事項集では事例としてまったく見かけないが，それでもという根拠は何ですか，追加解析をするのには理由が必要なのですよと，やんわりと避難することにしている．

表示 4.16　薬剤別の患者の背景因子の分布

		1:P 薬		2:A 薬		分割表
全体		151	100.0%	154	100.0%	χ^2 検定
施設	1:A 病院	17	11.3%	12	7.8%	df =3
	2:B 病院	45	29.8%	50	32.5%	χ^2 = 1.7701
	3:C 病院	31	20.5%	27	17.5%	p = 0.6215
	4:D 病院	58	38.4%	65	42.2%	
性	1:男	61	40.4%	80	51.9%	df =1, χ^2 = 4.0921
	2:女	90	59.6%	74	48.1%	p = 0.0431
年齢	45 歳未満	37	24.5%	35	22.7%	df =2
	45 歳以上	73	48.3%	92	59.7%	χ^2 = 5.0968
	65 歳以上	41	27.2%	27	17.5%	p = 0.0782
重症度	1:軽度	53	35.1%	72	46.8%	df =2
	2:中等度	45	29.8%	44	28.6%	χ^2 = 5.3428
	3:高度	53	35.1%	38	24.7%	p = 0.0692
合併症	1:無	64	42.4%	86	55.8%	df =1, χ^2 = 5.5267
	2:有	87	57.6%	68	44.2%	p = 0.0187

各層に共通の薬効差があることを仮定した検定

旧統計ガイドラインの要求であった 2 値の場合，交互作用の点検は Breslow-Day の検定を行い，Mantel-Haenzel 法の適用，3 分類以上の場合には拡張 Mantel 法の適用が示されている．SAS の FREQ プロシジャは，層をなす分割表に対する各種の検定法をカバーしており，詳細は，第@@章で示すので，ここでは，重症度を層とする場合について，ロジスティック回帰の結果と対比して示す．

拡張Mantel法は，表示 4.17 に示すように，軽度について治療法と反応の 2×3 の分割表とし，同様に中程度，高度についての分割表全体を 3 つの層をなす 2×3 の分割表として検定する方法である．

表示 4.17 重症度別・治療法別の反応

	P 薬				A 薬			
	++	+	−	計	++	+	−	計
1:軽度	11	16	26	53	21	24	27	72
2:中等度	9	14	22	45	11	15	18	44
3:高度	9	14	30	53	9	12	17	38

SASのFREQプロシジャは 2×2 のオッズ比を拡張したCMH検定（Cochran-Mantel-Haenzel）法は，反応を順序データとした場合にも，名義尺度とした場合にも対応している．反応を 1，2，3 ようなスコアとした場合にCMH検定統計量は，表示 4.17 に示すように χ^2=3.5865（df = 1），p= 0.0583 となる．ロジスティック回帰の結果は，表示 4.5 より， p= 0.0542 であるので，ほぼ同様の結果となる．

表示 4.18　拡張 Mantel 法

FREQ プロシジャ
drug と y123 の要約統計量
層別変数 : severity
Cochran-Mantel-Haenszel 統計量(テーブルスコアに基づく)

統計量	対立仮説	自由度	値	p 値
1	相関統計量	1	3.5865	0.0583
2	ANOVA 統計量	1	3.5865	0.0583
3	一般連関統計量	2	3.6953	0.1576

反応が 2 値の場合には, Breslow-Dayの検定も合わせて出力される. 反応を++で区切った 3×2×2 場合に，Breslow-Dayの検定統計量は， χ^2=0.0641（df = 2），p= 0.9684 となる．反応を+で区切った場合に，Breslow-Dayの検定統計量は， χ^2=0.0881（df = 2），p= 0.9569 が得られる．ロジスティック回帰の場合には，表示 4.5 より，p = 0.9445 であるので，2 つのBreslow-Dayの検定とほぼ同様の結果となる．

これらの対比から，分割表の検定を拡張した方法は，ロジスティック回帰を拡張した方法で置き換えられることを示したきた. 層をなす分割表の検定を主体にした探索的な解析の弱点は，層についての効果を計量できないこと，複数の因子を同時に扱えないこと，交互作用は Breslow-Day の検定を別途しなければならないことである.

表示 4.19　旧統計ガイドラインでの背景因子の偏りの調整

第２章．治験実施計画書に記載すべき統計解析関係事項
　略
９．解析方法
　略
(２)背景因子の扱い
　記録すべき背景因子とは、臨床的に治療予後と関連することが知られており、その処置群間での不均衡が比較試験結果に影響する可能性のある要因である。薬効に関する統計解析を行う前に、記録された背景因子に関する処置群間での比較可能性を点検する必要があり、その手順を規定しておく。
　背景因子の統計解析では、5%、1%の有意水準あるいは 95%、99%の信頼係数といった基準による評価ではなく、背景因子不均衡が有効性・安全性・有用性評価にどの程度影響するかという観点から検討されることが重要である。多くの場合、有意確率が１5%前後といった背景因子の不均衡についても注意深い検討が必要である。また、複数の種類の統計手法を用いて、異なる種類の不均衡を点検することも勧められる。
　背景因子の不均衡が生じた場合の対応についても検討しておく必要がある。統計的な調整を実施する可能性がある場合には、その解析手法について検討しておく。

（注）
20．背景因子の偏りの調整
　治療予後と強く関連する重要な背景因子（予後因子）の分布が処置群間で異なっている場合に、予後因子の影響を考慮せずに統計解析を進めると解析結果は歪んだものになる。臨床試験においては、既知及び未知の予後因子の分布が処置群間で均等になるように、無作為割付が可能な限り実施されている。無作為割付は処置群間の比較可能性を確保するための最良の手段であり、多くの場合、予後因子の影響を考慮せずに解析を行っても妥当な比較結果が得られる。しかし、特に症例数が少ない場合には、無作為割付を行ったとしても処置群間に予後因子の偏りが偶然に生じ、薬効差に関する単純な比較の妥当性が懸念される場合もある。予後との関連が強いことが明らかな予後因子の分布が処置群間で異なる場合には、やむを得ないときの手段として、適切な統計的調整を実施することも必要となる。検証的な臨床試験で予後因子の偏りに関する統計的調整を行う場合、処置群間での主要な予後因子の偏りの点検は薬効差の解析に先立って実施する必要がある。この際、予後因子自体の検定は参考的なものであり、予後因子の偏りが 薬効差の解析結果にどのように影響するかという観点から注意深く点検することが重要である。多くの場合、有意確率が 15%前後といった背景因子の不均衡についても注意深い検討が必要である。薬効差の統計解析が一旦なされてから、後知恵によって予後因子を選択して統計的調整を行った場合には、結果の公平性が保証できず、治験薬の薬効の検証とはならない。
　解析段階での予後因子の偏りの統計的調整としては種々の方法があるが、層別化による方法が最も基本的な方法といえる。層別化による方法は、偏りのあった予後因子について層別集計し、各層に共通の薬効差があると仮定してこの共通薬効差に関する統計的推論を行うものである。各層に共通の薬効差が存在するという統計モデルの妥当性については各層別集計において確認することが可能であり、臨床医学的な理解も直接的に得られる。統計的調整の具体的方法としては 2 値データについては Mantel-Haenszel 法な

どがあり、3 分類以上の順序分類データについては拡張 Mantel 法などが提案されている。共通薬効差という前提の点検（交互作用の点検）に、2 値データについては､例えば、Breslow-Day 検定などを用いることができる。計量データについても、種々の方法による統計的調整が可能であるが、交互作用の点検と共に、パラメトリックな手法を用いた場合にはその統計的方法の基礎にある諸仮定が満たされているかどうかの検討が必要である。

4.7. 共変量による調整の是非

治療法群間の P 値が 0.05 前後の時に，共変量を事後的な探索によって，それを含めた解析によって，P 値が 0.05 を境に前後することを時として経験する．その場合に臨床試験の結果をどのように結論づけたらよいか，判断に迷うことになる．その臨床試験の結果を基にして，薬の販売許可を申請したいという当事者の思いと，申請された規制当局は，販売許可を与えてよいかのじれんまに悩むことになる．

小規模な実験の結果ならば，P 値が 0.05 前後の効果は，たまたまよい結果が得られたぐらいに考え，スケールアップした実験で再確認を要する因子であろう．本章で示した臨床試験の実施は，数億円の費用と実社会の人たちを巻き込んだ研究であるために，再度，臨床試験を行なうことは，なかなか困難なことであり，P 値が 0.05 を前後する結果から適切な判断をしたいとの利害関係者のじれんまを引き起こすのである．

疾患の重症度を例にしてみよう．表示 4.1 の共変量の同定では，重症度が軽度・中等度・高度の順に，反応が+以上の場合に 57.6%，55.1%，48.4% と低くなって，弱い予後因子のようである． 表示 4.2 示すように 2 つの共変量の年齢と重症度を含んだロジスティック回帰から，年齢はP値が 0.4601 で弱い交絡因子，重症度はP値が 0.1722 で弱い予後因子と分類した．これらの結果から，治療法のコードが開示された後の共変量を含めた解析で，第 1 の関心事を治療群のP値とし,第 2 の関心事を交互作用としたのである．

表示 4.4 で, 治療群のみの解析結果からP値が 0.0373 と有意な差が得られたのであるが, 表示 4.16 に示すようにP薬とA薬の重症度の分布は, P薬で軽度の症例が 53 例 35.1%, A薬では 72 症例 46.8%となり，軽度の場合に全体の反応が高いので，A薬の反応がP薬に比べ若干高くなるような分布となっている．逆に，P薬で高度の症例が 53 例 35.1%, A薬では 38 症例 24.7%とP薬の反応がA薬に比べ若干低くなるような分布となっている．治療群ごとの重症度の分布から，重症度を考慮しない治療群のみの解析結果は，反応の差が開く方向となり, 重症度を共変量とした場合には反応の差が縮まる方向になることがわかる．

旧統計ガイドラインが要求していた治療群間の重症度のアンバランスを計量するた

めに，3×2 の適合度の χ^2 検定を行うと 表示 4.16 に示すようにP値が 0.0692 と無視できない状況である．表示 4.7 および 表示 4.8 に示すように，重症度の層ごとのP薬とA薬のオッズ比は，反応が++の場合に軽度・中等度・高度それぞれ，1.57，1.33，1.52であり，反応が+以上の場合に 1.60，1.38，1.61 とA薬の反応がP薬に比し高くなり，オッズ比の共通性に対するBreslow-Dayの検定結果は．反応が++の場合にp= 0.9684，反応が+の場合にp= 0.9569 が得られる．ロジスティック回帰では治療群と重症度の交互作用でも 表示 4.5 からP値は 0.9445 であるので，重症度を共変量とした主効果モデル，あるいは重症度を層とした拡張Mantel法による治療群のP値の統計的な調整が支持される．

重症度を共変量としたロジスティック回帰での治療群のP値は，表示 4.5 から 0.542 と有意な差でなくなっている．重症度を層とした拡張 Mantel 法では，P 値は 0.0583 と有意な差ではなくなる．調整をしない P 値がロジスティック回帰で 0.0373 となり，調整をすると P 値は 0.542 となり，5%の有意水準を結果が前後する．どちらの P 値を優先して判断したらよいのであろうか. 治療コードが明らかになる前に重症度を共変量とした解析を主たる解析法と計画書に明記した場合には，p = 0.0542 を受け入れなければならない．この場合に，症例数を増やした臨床試験を再度実施して，P 値が 0.05 以下になければ規制当局は A 薬の販売許可を与えないと判断するのであろうか．ならないのであろうか．あるいは，症例数は半分程度の臨床試験を実施し，2 つの臨床試験の結果をメタ・アナリシスの方法で併合し P 値が 0.05 以下となったら A 薬の販売許可をあたえるのであろうか.

重症度を共変量としない解析を主たる解析法と計画書に明記した場合で，P 値が 0.05 以上で，重症度を調整した場合に P 値が 0.050 以下となったような場合はどのような判断をしたらようのであろうか．再度，臨床試験を行なった後に販売許可を求める申請を行うと判断するのであろうか．調整した P 値が 0.05 以下となることに救いを求めて，規制当局に販売許可を申請した場合に，規制当局の判断はどうなるのであろうか．却下するのであろうか, 販売許可をあたえた後に追加の臨床試験の実施を条件に仮の販売許可にするのだろうか．いずれにしても両者に判断に際してじれんまを生じさせる．

浜田（2005）は，この問題について日本科学技術連盟の第 15 回のBioSコース講義資料で,「調整した解析を否定する意見」と「調整した解析を肯定する意見」を併記して対立点を整理しているので 表示 4.20 に示す.

表示 4.20　浜田知久馬（2005），調整の是非

背景因子の調整

調整した解析を否定する意見
1）無作為化割付を行っていれば、群間で大きなアンバランスは　生じるはずがなく、調整する必要はない.
2）無作為化割付を行っていれば、調整しない解析のα水準は保証されるので調整しなくても解析結果は妥当である.
3）調整する項目、調整する方法（層別、モデルによる方法）には様々なパターンが考えられ、これらの組み合わせによって、結果が整合しなくなったり．解釈しにくくなる場合がある.
4）1 つのデータに様々な方法を適用すると、検定の多重性の問題によって有意な結果が生じやすくなる．都合のよい解析結果だけを選んだのでは，後知恵解析のそしりを免れ待ない.
5）モデル等を用いた複雑な解析を行った場合、結論の妥当性はモデルの正しさに依存するため、モデルの妥当性について評価する必要性が生じるが，しばしばこの過程は困難である。
調整した解析を肯定する意見
1）無作為化割付によって、調整しない解析の妥当性は保証されるが、背景因子の影響が大きいときには、群間でバランスがとれていても．調整した解析の方が検出力が高くなる.
2）症例数が少なく、しかも背景因子が複数存在すると、大きなアンバランスが偶然によって生じる可能性がある．アンバランスが生じている場合には、調整しない解析では妥当な結論を導くことはできない.
3）p 値が 5%前後で有意だったとすると、偶然によって有意差が生じた確率は 5%は残ることになる．この意味で検定による判定は完全でない．重要な予後因子に偏りが生じ、調整した解析で有意差がなくなれば、第一種の過誤によって有意差が生じていた可能性が高くなるし、そうでなければ第一種の過誤の可能性は低くなる．調整した解析を行うことによって、結論の確信の度合いについての情報を得ることができる.
4）一連の解析を通じて、効果の修飾（交互作用）についての評価が可能になる.

4.8.　複数の共変量を含めた解析

これまでの示してきた解析は，治療群を主体とし共変量を一つごとに含めるという方法を示したきた．ロジスティック回帰を用いれば，複数の共変量を含めた解析を行うことが可能である．得られた観察データから統計的に有意となる因子を広い出したいというような探索的な場面であれば，躊躇なく気になる変数（因子）を含めたロジスティック回帰を実施し，探索的に無視できない因子および因子間の交互作用の抽出を行うであろう．

歴史的には，重回帰分析における変数選択の方法，変数増加法・変数減少法・変数増加減少法・変数減少増加法などの方法，これらのアルゴリズムにおける変数選択法に伴う抽出される変数の差異をなくすための総当り法などの影響を受けて，ロジスティック回帰でも受け継がれてきた．変数選択法は，モデルに新たに組みいれる基準，モデルから除く基準をあらかじめ設定しておくことにより，解析者の変数選択における恣意性を少しでも排除する方法として好まれている．

変数選択法の基本は，1 変数ごとに逐次的にモデルに加えるか除くかの判断を繰り返す．そのたに変数選択の方法により最終的に到達したモデルが異なることが生ずる．また，変数が 3 分類以上の分類データの場合に複数のダミー変数に展開すると別々の変数

として取り扱われてしまうといった問題点がある．さらに，交互作用をダミー変数として含めた場合に，主効果と交互作用をセットで取り扱わないと交互作用がモデルに含まれるが主効果が落ちてしまうことも起きる．このような問題が変数選択法には，付きまとっている．

最近では，これらの問題点が SAS の LOGSTIC プロシジャでは解消するようになったことは喜ばしい．とは言え，臨床試験の探索的な解析においては，共変量を一つごとに共変量と治療群の交互作用を調べることが基本である．ここで示した臨床試験では，制御しているのは治療群のみであり，他の複数の共変量については症例の登録に依存して変動する変数であり，それらの間の 2 因子交互作用はどうなっていいるか興味はあるが，治療群を含めた 3 因子交互作用の検討を誘発するような手順の定式化は困難を生ずる．

多くの因子を制御する直交表を用いた実験計画でも，3 因子交互作用は誤差的な変動として誤差とみし，有意な 2 因子交互作用から 3 因子の関与を示唆される場合でも，あえて 3 因子交互作用を統計的には求めず，2 因子交互作用のみの解釈を主体にしている．3因子交互作用は偶発的な結果の解釈はにつながるこじ付けになる可能性が増大するといった理由による．

とはいえ，全ての変数（共変量の候補）の 2 因子交互作用を含めた場合のモデルがどのようになるかを把握した上で 1 共変量ごとの解析結果を考察することは有意義である．次に 2 因子交互作用を全て含む場合の変数減少法のLOGISTICプロシジャを示す．MODELステートメントで変数の間の“｜”が交互作用をモデルに取り込めという表記であり，“@2”が 2 因子交互作用までをモデルに取り込めという指示となっている．そして，変数をモデルから追い出す基準はslstay=0.20 で変数のP 値が 0.20 以下である場合に設定している．さらに，slstay=0.15 の場合，slstay=0.05 の場合に付いても結果を 表示 4.22 に示す．

表示 4.21　順序ロジスティック回帰における変数減少法の SAS プログラム

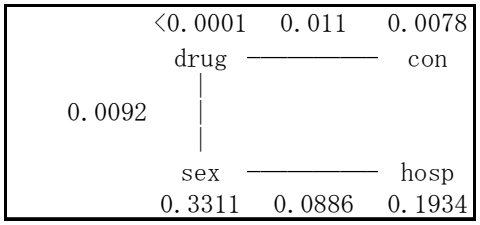
```
proc logistic data=clinical  ;
    class drug age_g hosp sex sev con / param=reference ref=first  ;
    model y123 =  drug | hosp | age_g | sex | sev |con @2
                / expb tech=newton  selection=backward slstay=0.20 ;
    run ;
```

交互作用がある場合の変数選択は，単純でない．追い出し基準が 0.15 の場合に，施設 hosp の P 値は，0.1934 と 0.15 よりも大きいにもかかわらずモデルから追い出されていないのは，性と施設の交互作用の P 値が 0.0086 と 0.15 より小さいので，関連する交互作用の主効果を残してあるのである．このような配慮は，施設 hosp の主効果を除く

と交互作用の正しい P 値が推定できなくなってしまうからである.

表示 4.22　追い出し基準を変えた場合の変数減少法

				変数減少法		
Effect	df	p		p<0.20	p<0.15	p<0.05
drug	1	0.0055	**	0.0003	<0.0001	<0.0001
hosp	3	0.4069		0.1009	0.1934	−
drug*hosp	3	0.3147		0.1764	−	−
age_g	2	0.9619		−	−	−
drug*age_g	2	0.3748		−	−	−
age_g*hosp	6	0.9087		−	−	−
sex	1	0.1358	+	0.1656	0.3311	0.0480
drug*sex	1	0.0013	**	0.0137	0.0092	0.0150
hosp*sex	3	0.0770	+	0.0693	0.0886	−
age_g*sex	2	0.7288		−	−	−
sev	2	0.8547		−	−	−
drug*sev	2	0.4774		−	−	−
hosp*sev	6	0.6868		−	−	−
age_g*sev	4	0.2254		−	−	−
sex*sev	2	0.7010		−	−	−
con	1	0.0141	*	0.0044	0.0078	0.0065
drug*con	1	0.0074	**	0.0057	0.0110	0.0123
hosp*con	3	0.2831		−	−	−
age_g*con	2	0.2701		−	−	−
sex*con	1	0.7671		−	−	−
sev*con	2	0.3911		−	−	−



表示 4.23　追い出し基準が 0.15 の場合の線点図

変数増加法は，selection=foreward で行えるが，結果は減少法の場合と一致する場合も有るが，この例では drug を新たにいれた段階でストップしてしまい，まったく一致しない.

モデルに含める変数の数を k としたときに，1 変数の場合，2 変数の場合，. . . . ，k 変数の中から MODEL ステートメントで指定した変数の中からから総当り的に候補となる変数用いて，あてはまりの良い変数の組み合わせを抽出する機能も LOGISTIC プロシジャには備わっている. 変数選択の方法による差異をなくす方法として優れた方法であ

る．しかしながら，CLASS ステートメントで 3 水準以上の変数をあつかえないこと，交互作用も主効果も区別しないことなどの制約が SAS の Ver9.1 であり，使用範囲は制約される．

4.9. CPMP が提示している留意点

「背景因子の調整」については，積極的に調整を行うべきとの意見と，ランダム化に基づいた基本的な解析を優先すべきとの調整に否定的な意見もある．これらににについて，EMEA（The European Agency for the Evaluation of Medicinal Products）の CPMP（Committee for Proprietary Medicinal Products）が 2003 年に提示した Points to Consider on Adjustment for Baseline Covariates は，この問題に対するヨーロッパの規制当局および製薬団体の共同の意見として尊重すべきである．

Points to consider on adjustment for baseline covariate

V. OVERALL CONCLUSIONS AND RECOMMENDATIONS

Valid analyses of trials that have used simple randomisation generally do not require adjustment for covariates. However, the inclusion of influential covariates may result in more efficient estimates of treatment effects and so their exclusion may result in the failure to detect real treatment effects, if they exist.

Whilst unadjusted analyses are technically simple and easy to interpret, adjusted analyses can be rather complex. Therefore the use of adjusted analyses should be justified and they should be presented in such a manner that the reader or the reviewer can understand them unambiguously. The use of more complex techniques should not be an excuse for the absence of clarity. The following recommendations are made:

・Stratification may be used to ensure balance of treatments across covariates; it may also be used for administrative reasons. The factors that are the basis of stratification should normally be included as covariates in the primary model.

・Variables known a priori to be strongly, or at least moderately, associated with the primary outcome and/or variables for which there is a strong clinical rationale for such an association should also be considered as covariates in the primary analysis. The variables selected on this

basis should be pre-specified in the protocol or the statistical analysis plan.

- Baseline imbalance observed _post _hoc should not be considered an appropriate reason for including a variable as a covariate in the primary analysis.

- Variables measured after randomisation and so potentially affected by the treatment should not normally be included as covariates in the primary analysis.

- If a baseline value of a continuous outcome measure is available, then this should usually be included as a covariate. This applies whether the primary outcome variable is defined as the `raw outcome' or as the `change from baseline'.

- Only a few covariates should be included in a primary analysis. Although larger data sets may support more covariates than smaller ones, justification for including each of the covariates should be provided.

- In the absence of prior knowledge, a simple functional form (usually either linearity or dichotomising a continuous scale) should be assumed for the relationship between a continuous covariate and the outcome variable.

- The validity of model assumptions must be checked when assessing the results. This is particularly important for generalised linear or non-linear models where mis-specification could lead to incorrect estimates of the treatment effect. Even under ordinary linear models, some attention should be paid to the possible influence of extreme outlying values.

- Whenever adjusted analyses are presented, results of the treatment effect in subgroups formed by the covariates (appropriately categorised, if relevant) should be presented to enable an assessment of the validity of the model assumptions.

- Sensitivity analyses should be pre-planned and presented to investigate the robustness of the primary results. Discrepancies should be discussed and explained. In the presence of important differences that cannot be logically explained - for example, between the results

of adjusted and unadjusted analyses - the interpretation of the trial could be seriously affected.

・The primary model should not include treatment by covariate interactions- If substantial interactions are expected _a _priori, the trial should be designed to allow separate estimates of the treatment effects in specific subgroups.

・Exploratory analyses may be carried out to improve the understanding of covariates not included in the primary analysis, and to help the sponsor with the ongoing development of the drug.

・A primary analysis, unambiguously pre-specified in the protocol or statistical analysis plan, correctly carried out and interpreted, should support the conclusions which are drawn from the trial. Since there may be a number of alternative valid analyses, results based on pre-specified analyses will carry most credibility.

4.10. 共分散分析の無作為化比較臨床試験における留意事項

清見，西田，西島（2005）は，「共分散分析の無作為化比較臨床試験における留意事項」と題して，吉祥寺東急イン第 4 回東京吉祥寺地区月例研究会で，Points to Consider on Adjustment for Baseline Covariates にたいする周辺のコメント・考察、留意事項の講演資料もこの問題の理解に役に立つ.

ノボノルディスクファーマ株式会社，清見文明、西田朋由、西島啓二
2005 年 10 月 31 日（於）吉祥寺東急イン第 4 回東京吉祥寺地区月例研究会

興味をもった発端
検出力が高いとの理由で、2000 年頃（？）から
主要な解析として、共分散分析を使いはじめた。

Points to consider on adjustment for baseline covariate の案が作成された（2001 年）。
（とても強く共分散分析を推奨）
ー本当に検出力が実質的に高いのか

－交互作用があるのでは
－ベースライン共変量をどのように解釈するのか・・・

Points to consider on adjustment for baseline covariate の要約

・層別はバランスを確実にするために利用され, また管理上の理由で用いられる. 通常, 層別因子は主要な解析に共変量として含められるべき.

・主要変数に影響を及ぼすことが既知の変数, and/or 強い臨床的合理性を有する変数は, 主要解析で共変量として考慮されるべき. このことは事前にプロトコル又は SAP に記載すべき.

・事後的に偏りが見られた項目は, それを理由に解析に含めてプライマリーとするべきではない.

・無作為化後に測定される変数は, 潜在的に治療に影響されるので, 主要な解析に共変量として含めるべきではない. 主要変数のベースラインは, 通常共変量に含められるべき（主要変数が raw outcome であろうが変化量であろうが）.

・少数の共変量が主要な解析には含められるべきである. 大容量データが多くの共変量を含めることをサポートするかもしれないが, それら各々の正当性を示さねばならない.

・事前知識がなければ, 共変量と主要変数の間には単純な関数形を仮定すべきである（線形性又は連続変数の 2 値化）.

・モデルの妥当性をチェックしなければならない. これは特に線形／非線形モデルに対して重要である. 規格・設定の誤りが（mis-specification）治療効果の不正確な推定を導くことがあるからである. 通常の線形モデルの下でさえ, 外れ値の影響には気を配る必要がある.

・調整解析が提示されるときはいつでも, 共変量のカテゴリ毎（連続変数は意味があればカテゴリ化して）の治療効果も提示すべきである（モデルの仮定の妥当性評価を可能にするために）

・主要な結果の頑健性を調査するために，感度分析を事前に計画し実施すべきである．

・主要な解析には治療と共変量の交互作用を含めるべきではない．実質的な交互作用が事前に予期されるならば，当該治験は部分集団解析を可能にするようにデザインされるべきである．

・共変量を含めない探索的解析が，理解促進のため行われるかもしれない．

・主要な解析とは，プロトコル又は SAP に曖昧さなく事前に特定され，正確に実行して解釈されるものである．この主要な解析が試験から導かれる結論を support すべきである．他の多くの妥当な解析が存在するので，事前に計画された解析に基づく結果が最も信頼されるであろう．

5. SAS コード

```
Title 'clinical_305.sas 2006-1-26 Y.Takahashi ' ;

data clinical ;
    input id hosp drug sex age age3 severity c_drug y123 y122 y112 ;
        if y122=1 then y122r = 'Placebo' ;
        if y122=2 then y122r = 'Active ' ;
        if y112=1 then y112r = 'Placebo' ;
        if y112=2 then y112r = 'Active ' ;

/*症例番号 施設 薬剤 性 年齢 年齢３区分 重症度 併用薬 効果 123 効果 122 効果 112*/
datalines ;
1 1 2 1 63 2 1 1 1 1 1
2 1 2 1 63 2 1 1 3 2 2
3 1 2 1 57 2 1 2 1 1 1
4 1 2 1 22 1 1 2 3 2 2
5 1 2 1 45 2 2 2 2 2 1
6 1 2 1 56 2 3 2 1 1 1
7 1 2 1 46 2 3 2 2 2 1
8 1 2 2 52 2 1 2 1 1 1
9 1 2 2 64 2 2 1 1 1 1
10 1 2 2 69 3 2 1 1 1 1
11 1 2 2 67 3 2 2 3 2 2
12 1 2 2 48 2 3 1 3 2 2
13 1 1 1 56 2 1 2 3 2 2
14 1 1 1 64 2 2 2 2 2 1
15 1 1 1 67 3 2 2 2 2 1
16 1 1 1 67 3 3 2 3 2 2
17 1 1 1 69 3 3 2 3 2 2
18 1 1 1 36 1 3 2 3 2 2
19 1 1 2 76 3 1 1 3 2 2
20 1 1 2 64 2 1 1 3 2 2
21 1 1 2 72 3 1 2 2 2 1
22 1 1 2 45 2 2 1 3 2 2
23 1 1 2 58 2 2 1 3 2 2
24 1 1 2 69 3 3 1 2 2 1
25 1 1 2 47 2 3 1 3 2 2
26 1 1 2 32 1 3 2 1 1 1
27 1 1 2 58 2 3 2 2 2 1
28 1 1 2 33 1 3 2 2 2 1
29 1 1 2 63 2 3 2 3 2 2
30 2 2 1 69 3 1 1 1 1 1
31 2 2 1 61 2 1 1 1 1 1
32 2 2 1 44 1 1 1 1 1 1
33 2 2 1 62 2 1 1 2 2 1
34 2 2 1 56 2 1 1 2 2 1
35 2 2 1 57 2 1 1 2 2 1
36 2 2 1 40 1 1 1 2 2 1
37 2 2 1 61 2 1 1 3 2 2
38 2 2 1 63 2 1 1 3 2 2
39 2 2 1 56 2 1 1 3 2 2
40 2 2 1 42 1 1 2 2 2 1
41 2 2 1 49 2 1 2 2 2 1
42 2 2 1 42 1 1 2 3 2 2
43 2 2 1 58 2 2 1 2 2 1
44 2 2 1 31 1 2 1 2 2 1
45 2 2 1 35 1 2 1 3 2 2
46 2 2 1 48 2 2 2 1 1 1
47 2 2 1 53 2 2 2 2 2 1
48 2 2 1 39 1 2 2 2 2 1
49 2 2 1 38 1 3 1 3 2 2
50 2 2 1 60 2 3 2 1 1 1
51 2 2 1 35 1 3 2 2 2 1
52 2 2 1 61 2 3 2 3 2 2
53 2 2 2 74 3 1 1 1 1 1
54 2 2 2 55 2 1 1 2 2 1
55 2 2 2 41 1 1 1 3 2 2
56 2 2 2 66 3 1 1 3 2 2
57 2 2 2 33 1 1 2 2 2 1
58 2 2 2 55 2 1 2 2 2 1
59 2 2 2 61 2 1 2 3 2 2
60 2 2 2 44 1 1 2 3 2 2
61 2 2 2 51 2 1 2 3 2 2
62 2 2 2 32 1 1 2 3 2 2
63 2 2 2 69 3 2 1 1 1 1
64 2 2 2 60 2 2 1 2 2 1
65 2 2 2 52 2 2 1 2 2 1
66 2 2 2 53 2 2 1 2 2 1
67 2 2 2 61 2 2 1 3 2 2
68 2 2 2 76 3 2 2 2 2 1
69 2 2 2 56 2 2 2 3 2 2
70 2 2 2 55 2 2 2 3 2 2
71 2 2 2 32 1 2 2 3 2 2
72 2 2 2 51 2 3 1 1 1 1
73 2 2 2 40 1 3 1 1 1 1
74 2 2 2 51 2 3 1 2 2 1
75 2 2 2 36 1 3 1 2 2 1
76 2 2 2 36 1 3 1 3 2 2
77 2 2 2 58 2 3 1 3 2 2
78 2 2 2 56 2 3 1 3 2 2
79 2 2 2 59 2 3 2 3 2 2
80 2 1 1 57 2 1 1 2 2 1
81 2 1 1 65 3 1 1 3 2 2
82 2 1 1 44 1 1 2 3 2 2
83 2 1 1 62 2 2 1 2 2 1
84 2 1 1 65 3 2 1 3 2 2
85 2 1 1 41 1 2 1 3 2 2
86 2 1 1 51 2 2 2 1 1 1
87 2 1 1 69 3 2 2 1 1 1
88 2 1 1 53 2 2 2 3 2 2
89 2 1 1 52 2 3 1 3 2 2
90 2 1 1 69 3 3 1 3 2 2
91 2 1 1 41 1 3 2 2 2 1
92 2 1 1 58 2 3 2 3 2 2
93 2 1 1 55 2 3 2 3 2 2
94 2 1 1 48 2 3 2 3 2 2
95 2 1 2 61 2 1 1 1 1 1
96 2 1 2 68 3 1 1 2 2 1
97 2 1 2 46 2 1 1 2 2 1
98 2 1 2 33 1 1 1 3 2 2
99 2 1 2 57 2 1 1 3 2 2
100 2 1 2 42 1 1 2 1 1 1
101 2 1 2 39 1 1 2 2 2 1
102 2 1 2 59 2 1 2 2 2 1
103 2 1 2 69 3 1 2 2 2 1
104 2 1 2 55 2 1 2 3 2 2
105 2 1 2 42 1 2 1 1 1 1
106 2 1 2 51 2 2 1 2 2 1
107 2 1 2 44 1 2 1 2 2 1
108 2 1 2 43 1 2 1 3 2 2
109 2 1 2 66 3 2 1 3 2 2
110 2 1 2 41 1 2 2 1 1 1
111 2 1 2 68 3 2 2 1 1 1
112 2 1 2 63 2 2 2 2 2 1
113 2 1 2 47 2 2 2 2 2 1
114 2 1 2 52 2 2 2 2 2 1
115 2 1 2 43 1 2 2 3 2 2
116 2 1 2 57 2 2 2 3 2 2
117 2 1 2 66 3 2 2 3 2 2
118 2 1 2 58 2 3 1 1 1 1
119 2 1 2 67 3 3 1 1 1 1
120 2 1 2 65 3 3 1 3 2 2
121 2 1 2 53 2 3 1 3 2 2
122 2 1 2 36 1 3 2 1 1 1
123 2 1 2 55 2 3 2 2 2 1
124 2 1 2 53 2 3 2 3 2 2
125 3 2 1 74 3 1 1 1 1 1
126 3 2 1 61 2 1 1 1 1 1
127 3 2 1 70 3 1 1 2 2 1
128 3 2 1 78 3 1 1 3 2 2
129 3 2 1 35 1 1 1 3 2 2
130 3 2 1 71 3 1 2 1 1 1
131 3 2 1 47 2 1 2 1 1 1
132 3 2 1 67 3 1 2 2 2 1
133 3 2 1 59 2 2 1 1 1 1
134 3 2 1 56 2 2 2 2 2 1
135 3 2 1 59 2 2 2 3 2 2
136 3 2 1 54 2 3 1 1 1 1
137 3 2 1 45 2 3 1 1 1 1
138 3 2 1 48 2 3 2 2 2 1
139 3 2 2 67 3 1 1 2 2 1
140 3 2 2 56 2 1 1 2 2 1
141 3 2 2 53 2 1 1 3 2 2
142 3 2 2 55 2 1 1 3 2 2
143 3 2 2 57 2 1 2 2 2 1
144 3 2 2 42 1 1 2 2 2 1
145 3 2 2 60 2 1 2 3 2 2
146 3 2 2 37 1 2 2 3 2 2
147 3 2 2 44 1 2 2 3 2 2
148 3 2 2 56 2 2 2 3 2 2
149 3 2 2 60 2 3 1 2 2 1
150 3 2 2 58 2 3 1 3 2 2
151 3 2 2 40 1 3 1 3 2 2
152 3 1 1 69 3 1 1 3 2 2
153 3 1 1 47 2 1 2 1 1 1
154 3 1 1 62 2 1 2 3 2 2
155 3 1 1 56 2 1 2 3 2 2
156 3 1 1 55 2 2 1 1 1 1
157 3 1 1 40 1 2 1 2 2 1
158 3 1 1 30 1 2 1 3 2 2
159 3 1 1 39 1 2 1 3 2 2
160 3 1 1 53 2 2 2 2 2 1
161 3 1 1 66 3 3 1 3 2 2
162 3 1 1 64 2 3 2 1 1 1
163 3 1 1 51 2 3 2 1 1 1
164 3 1 1 56 2 3 2 2 2 1
165 3 1 1 61 2 3 2 2 2 1
166 3 1 1 53 2 3 2 3 2 2
167 3 1 1 64 2 3 2 3 2 2
168 3 1 2 41 1 1 1 2 2 1
169 3 1 2 66 3 1 1 3 2 2
170 3 1 2 65 3 1 1 3 2 2
171 3 1 2 56 2 1 2 1 1 1
172 3 1 2 43 1 1 2 2 2 1
173 3 1 2 57 2 1 2 3 2 2
174 3 1 2 52 2 2 1 3 2 2
175 3 1 2 20 1 2 2 2 2 1
176 3 1 2 32 1 2 2 3 2 2
177 3 1 2 72 3 3 1 2 2 1
178 3 1 2 65 3 3 1 2 2 1
179 3 1 2 48 2 3 1 3 2 2
180 3 1 2 54 2 3 1 3 2 2
181 3 1 2 58 2 3 2 1 1 1
182 3 1 2 56 2 3 2 2 2 1
183 4 2 1 76 3 1 1 1 1 1
184 4 2 1 42 1 1 1 1 1 1
185 4 2 1 62 2 1 1 1 1 1
186 4 2 1 37 1 1 1 2 2 1
187 4 2 1 44 1 1 1 2 2 1
188 4 2 1 60 2 1 1 2 2 1
189 4 2 1 41 1 1 1 2 2 1
190 4 2 1 52 2 1 1 3 2 2
191 4 2 1 51 2 1 1 3 2 2
192 4 2 1 54 2 1 1 3 2 2
193 4 2 1 53 2 1 1 3 2 2
194 4 2 1 86 3 1 2 3 2 2
195 4 2 1 69 3 2 1 1 1 1
196 4 2 1 69 3 2 1 1 1 1
197 4 2 1 41 1 2 1 2 2 1
198 4 2 1 41 1 2 1 2 2 1
199 4 2 1 55 2 2 1 3 2 2
200 4 2 1 53 2 2 1 3 2 2
201 4 2 1 67 3 2 2 1 1 1
202 4 2 1 48 2 2 2 1 1 1
203 4 2 1 40 1 2 2 2 2 1
204 4 2 1 60 2 2 2 2 2 1
205 4 2 1 45 2 2 2 3 2 2
206 4 2 1 54 2 2 2 3 2 2
207 4 2 1 45 2 2 2 3 2 2
208 4 2 1 65 3 2 2 3 2 2
209 4 2 1 53 2 2 2 3 2 2
210 4 2 1 81 3 3 1 2 2 1
211 4 2 1 44 1 3 1 2 2 1
212 4 2 1 53 2 3 1 2 2 1
213 4 2 1 75 3 3 1 3 2 2
214 4 2 1 57 2 3 2 1 1 1
215 4 2 1 61 2 3 2 2 2 1
```

```
216  4  2  1  50  2  3  2  2  2  1
217  4  2  1  61  2  3  2  3  2  2
218  4  2  1  59  2  3  2  3  2  2
219  4  2  2  61  2  1  1  1  1  1
220  4  2  2  46  2  1  1  1  1  1
221  4  2  2  84  3  1  1  1  1  1
222  4  2  2  42  1  1  1  2  2  1
223  4  2  2  39  1  1  1  2  2  1
224  4  2  2  52  2  1  1  3  2  2
225  4  2  2  65  3  1  1  3  2  2
226  4  2  2  59  2  1  1  3  2  2
227  4  2  2  58  2  1  2  1  1  1
228  4  2  2  46  2  1  2  1  1  1
229  4  2  2  56  2  1  2  1  1  1
230  4  2  2  55  2  1  2  1  1  1
231  4  2  2  42  1  1  2  2  2  1
232  4  2  2  57  2  1  2  2  2  1
233  4  2  2  63  2  1  2  2  2  1
234  4  2  2  60  2  1  2  3  2  2
235  4  2  2  53  2  1  2  3  2  2
236  4  2  2  47  2  2  1  1  1  1
237  4  2  2  79  3  2  1  2  2  1
238  4  2  2  77  3  2  1  3  2  2
239  4  2  2  58  2  2  2  1  1  1
240  4  2  2  64  2  3  1  1  1  1
241  4  2  2  47  2  3  1  3  2  2
242  4  2  2  65  3  3  1  3  2  2
243  4  2  2  59  2  3  1  3  2  2
244  4  2  2  53  2  3  2  1  1  1
245  4  2  2  55  2  3  2  2  2  1
246  4  2  2  73  3  3  2  3  2  2
247  4  2  2  44  1  3  2  3  2  2
248  4  1  1  35  1  1  1  1  1  1
249  4  1  1  54  2  1  1  3  2  2
250  4  1  1  64  2  1  2  1  1  1
251  4  1  1  36  1  1  2  1  1  1
252  4  1  1  43  1  1  2  2  2  1
253  4  1  1  54  2  1  2  2  2  1
254  4  1  1  39  1  1  2  2  2  1
255  4  1  1  64  2  1  2  3  2  2
```

```
256  4  1  1  62  2  1  2  3  2  2
257  4  1  1  66  3  2  1  3  2  2
258  4  1  1  51  2  2  1  3  2  2
259  4  1  1  46  2  2  1  3  2  2
260  4  1  1  68  3  2  2  3  2  2
261  4  1  1  69  3  2  2  3  2  2
262  4  1  1  36  1  2  2  3  2  2
263  4  1  1  58  2  3  1  2  2  1
264  4  1  1  36  1  3  1  3  2  2
265  4  1  1  22  1  3  1  3  2  2
266  4  1  1  66  3  3  2  1  1  1
267  4  1  1  71  3  3  2  2  2  1
268  4  1  1  67  3  3  2  3  2  2
269  4  1  1  62  2  3  2  3  2  2
270  4  1  1  60  2  3  2  3  2  2
271  4  1  1  57  2  3  2  3  2  2
272  4  1  2  47  2  1  1  1  1  1
273  4  1  2  39  1  1  1  3  2  2
274  4  1  2  42  1  1  1  3  2  2
275  4  1  2  60  2  1  1  3  2  2
276  4  1  2  71  3  1  1  3  2  2
277  4  1  2  31  1  1  1  3  2  2
278  4  1  2  42  1  1  2  1  1  1
279  4  1  2  51  2  1  2  1  1  1
280  4  1  2  52  2  1  2  1  1  1
281  4  1  2  66  3  1  2  2  2  1
282  4  1  2  68  3  1  2  2  2  1
283  4  1  2  57  2  1  2  2  2  1
284  4  1  2  69  3  1  2  2  2  1
285  4  1  2  80  3  1  2  3  2  2
286  4  1  2  64  2  1  2  3  2  2
287  4  1  2  65  3  1  2  3  2  2
288  4  1  2  90  3  1  2  3  2  2
289  4  1  2  73  3  2  1  1  1  1
290  4  1  2  65  3  2  1  2  2  1
291  4  1  2  54  2  2  1  3  2  2
292  4  1  2  51  2  2  2  1  1  1
293  4  1  2  50  2  2  2  1  1  1
294  4  1  2  64  2  2  2  2  2  1
295  4  1  2  73  3  2  2  2  2  1
```

```
296  4  1  2  37  1  2  2  3  2  2
297  4  1  2  78  3  3  1  1  1  1
298  4  1  2  65  3  3  1  2  2  1
299  4  1  2  41  1  3  1  3  2  2
300  4  1  2  63  2  3  1  3  2  2
301  4  1  2  32  1  3  1  3  2  2
302  4  1  2  53  2  3  1  3  2  2
303  4  1  2  40  1  3  2  2  2  1
304  4  1  2  45  2  3  2  3  2  2
305  4  1  2  47  2  3  2  3  2  2

;
proc print data=clinical ;
run ;

title2 '<<< 共変量の同定 >>>' ;
proc tabulate data=clinical noseps
format=10. ;
     class drug hosp sex age3 severity
c_drug  y123 ;
          table all hosp sex age3
severity c_drug, y123  all  / rts=15 ;
          run ;

title2 '<<< 累積ロジットによる共変量の
同定 >>>' ;
proc logistic data=clinical  ;
     class age3 / param=reference
ref=first ;
          model y123 = age3 severity
/ expb tech=newton  ;
          run ;

title2 '<<< 治療法のみの場合 >>>' ;
proc logistic data=clinical  ;
     class drug / param=reference
ref=first ;
          model y123 = drug / expb
tech=newton  ;
          run ;
```

```
Title 'clinical_305b.sas  2006-1-26  Y.Takahashi ' ;

/*data clinical ;
     input id hosp drug sex age age3 severity c_drug y123
y122 y112 ;
          if y122=1 then y122r = 'Placebo' ;
          if y122=2 then y122r = 'Active ' ;
          if y112=1 then y112r = 'Placebo' ;
          if y112=2 then y112r = 'Active ' ;

症例番号 施設 薬剤 性 年齢 年齢３区分 重症度 併用
薬 効果 123 効果 122 効果 112
datalines ;
1  1  2  1  63  2  1  1  1  1  1
2  1  2  1  63  2  1  1  3  2  2
3  1  2  1  57  2  1  2  1  1  1

304  4  1  2  45  2  3  2  3  2  2
305  4  1  2  47  2  3  2  3  2  2 */

;
proc print data=clinical ;
run ;

title2 '<<< 共変量の同定 >>>' ;
proc tabulate data=clinical noseps format=10. ;
     class  drug hosp sex age3 severity c_drug  y123 ;
          table  all hosp sex age3 severity c_drug, y123
all  / rts=15 ;
          run ;

title2 '<<< 累積ロジットによる共変量の同定 >>>' ;
proc logistic data=clinical  ;
     class  age3  / param=reference ref=first ;
          model  y123 = age3 severity / expb
tech=newton  ;
          run ;

title2 '<<< 治療法のみの場合 >>>' ;
proc logistic data=clinical  ;
     class  drug  / param=reference ref=first ;
          model  y123 = drug / expb tech=newton  ;
          run ;


title2 '<<< 治療群と各共変量との累積反応率についての
解析 >>>' ;
proc transpose data=clinical  out=clinical_T ;
     var  hosp sex age3 severity c_drug ;
     by  id  drug y123 y122r y112r ;
          run ;

data clinical_T ;
     set clinical_T ;
          var_name = _NAME_ ;
     value  =  COL1 ;
          if var_name = 'hosp'  then var_name =
'1:hosp  ' ;
          if var_name = 'sex'  then var_name =
'2:sex  ' ;
          if var_name = 'age3'  then var_name =
'3:age3  ' ;
          if var_name = 'severity'  then var_name =
'4:severity' ;
          if var_name = 'c_drug'  then var_name =
'5:c_drug  ' ;
          drop _NAME_ COL1 ;
     run ;

proc sort data=clinical_T ;
     by var_name id ;
          run ;

proc print data=clinical_T ;
run ;

proc logistic data=clinical_T  ;
     class  drug value / param=reference ref=first ;
```

```
        model   y123 = drug value drug*value/ expb
tech=newton     ;
        by   var_name ;
        run ;

proc logistic data=clinical_T   ;
    class   drug value / param=reference ref=first ;
        model       y123    =   drug    value   /   expb
tech=newton     ;
        by   var_name ;
        run ;


title2 '<<< 治療群別の++反応率とオッズ比 >>>' ;
proc tabulate data=clinical noseps format=10. ;
    class   drug hosp sex age3 severity c_drug   y123 y122
y112 ;
        table    all hosp sex age3 severity c_drug,
drug*(y122 all)   / rts=15 ;
        run ;

proc freq data=clinical_T ;
    table   var_name*value*drug*y122r / nopercent nocol
norow relrisk ;
        run ;

title2 '<<< 治療群別の+以上反応率とオッズ比 >>>' ;
proc tabulate data=clinical noseps format=10. ;
    class   drug hosp sex age3 severity c_drug   y123 y122
y112 ;
        table    all hosp sex age3 severity c_drug,
drug*(y112 all)   / rts=15 ;
        run ;
proc freq data=clinical_T ;
    table   var_name*value*drug*y112r / nopercent nocol
norow relrisk ;
        run ;

title2 '<<< 治療法と施設の主効果モデルとした解析
>>>' ;
proc logistic data=clinical   ;
    class   drug hosp / param=reference ref=first ;
        model       y123    =   drug    hosp    /   expb
tech=newton     ;
        run ;

title2 '<<< 施設と治療法の交互作用 >>>' ;
proc logistic data=clinical   ;
    class   drug hosp / param=reference ref=first ;
        model    y122 = drug hosp drug*hosp/ expb
tech=newton     ;
        run ;

proc logistic data=clinical   ;
    class   drug hosp / param=reference ref=first ;
        model    y112 = drug hosp drug*hosp/ expb
tech=newton     ;
        run ;

proc logistic data=clinical   ;
    class   drug hosp / param=reference ref=first ;
        model    y123 = drug hosp drug*hosp/ expb
tech=newton     ;
        run ;


title2 '<<< 薬剤別の患者の背景因子の分布 >>>' ;
proc tabulate data=clinical noseps format=10. ;
    class   drug hosp sex age3 severity c_drug   y123 y122
y112 ;
        table   all hosp sex age3 severity c_drug, drug
all / rts=15 ;
        run ;

proc freq data=clinical ;
    table   hosp*drug / nopercent nocol norow chisq ;
    table   sex *drug / nopercent nocol norow chisq ;
    table   age3*drug / nopercent nocol norow chisq ;
    table   severity*drug / nopercent nocol norow chisq ;
    table   c_drug   *drug / nopercent nocol norow chisq ;
        run ;


title2 '<<< 重症度別・治療法別の反応 >>>' ;
proc tabulate data=clinical noseps format=6. ;
    class   drug hosp sex age3 severity c_drug   y123 y122
y112 ;
        table   severity, drug*(y123 all) / rts=15 ;
        run ;

proc freq data=clinical ;
    table severity*drug*y123 / nopercent nocol norow cmh
score=table ;
    table severity*drug*y122 / nopercent nocol norow cmh
score=table ;
    table severity*drug*y112 / nopercent nocol norow cmh
score=table ;
        run ;


title2 '<<< 追い出し基準を変えた場合の変数減少法
>>>' ;
proc logistic data=clinical   ;
    class   drug   hosp   sex   age3   severity   c_drug   /
param=reference ref=first   ;
        model y123 =   drug | hosp | age3 | sex | severity
|c_drug @2
                / expb tech=newton     ;
        run ;

proc logistic data=clinical   ;
    class   drug   hosp   sex   age3   severity   c_drug   /
param=reference ref=first   ;
        model y123 =   drug | hosp | age3 | sex | severity
|c_drug @2
                / expb tech=newton   selection=backward
slstay=0.20 ;
        run ;

proc logistic data=clinical   ;
    class   drug   hosp   sex   age3   severity   c_drug   /
param=reference ref=first   ;
        model y123 =   drug | hosp | age3 | sex | severity
|c_drug @2
                / expb tech=newton   selection=backward
slstay=0.15 ;
        run ;

proc logistic data=clinical   ;
    class   drug   hosp   sex   age3   severity   c_drug   /
param=reference ref=first   ;
        model y123 =   drug | hosp | age3 | sex | severity
|c_drug @2
                / expb tech=newton   selection=backward
slstay=0.05 ;
        run ;
```