

多変量解析：重回帰・ロジスティック回帰・比例ハザードモデル

医学統計セミナー2017 第2回目

下川敏雄

医療データサイエンス学
(附属病院臨床研究センター)

11月28日(火)	診断・検査データの解析：ROC曲線を中心に
12月22日(金)	多変量解析：重回帰・ロジスティック回帰・比例ハザードモデル
1月26日(金)	傾向スコアによる解析：マッチングによる解析を中心に
2月23日(金)	メタアナリシス：2値・量的・生存期間データに対するメタ回帰分析
3月23日(金)	臨床試験における必要症例数の計算

多数の共変量に対する影響の評価

名前	応答の形式	例	係数の解釈
重回帰分析	量的	体脂肪率	回帰係数 (標準回帰係数)
ロジスティック回帰分析	2値	治療の成功/失敗	オッズ比
- 名義ロジスティック	名義	疾患の種類	オッズ比
- 比例オッズモデル	順序	疾患の進行程度	オッズ比
Poisson回帰分析	計数	ポリープの検出個数	率比
Cox比例ハザード・モデル	生存時間	がん患者の生存期間	ハザード比

■ 回帰係数は説明変数の尺度に依存するため、すべての変数を標準化したもとで計算する回帰モデルの係数は標準回帰係数(標準化係数)と呼ばれる。標準化係数の絶対値の大きさを利用することで、応答に対する各説明変数の影響を評価できる。

■ ロジスティック回帰、Poisson回帰、Cox比例ハザード・モデルでは、指数関数 $A=\exp(\beta)$ を計算することで、それぞれ、オッズ比、率比、ハザード比を計算できる。それぞれの解釈は下記のとおり：

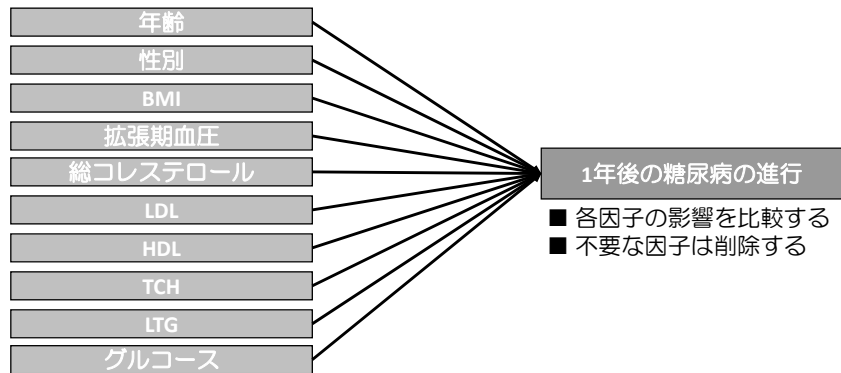
- ・オッズ比：変数 X が1上がると A 倍 $y=1$ になる(例： A 倍治療が成功する)。
- ・率比：変数 X が1上がると A 倍計数が上がる(例： A 倍ポリープが検出される)。
- ・ハザード比：変数 X が1上がると A 倍イベントリスクが高まる(例： A 倍死亡リスクがあがる)。

■ 重回帰分析

Multiple Regression Analysis

重回帰分析

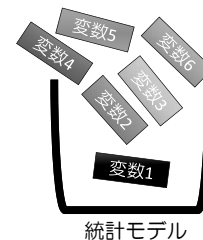
糖尿病患者442名に対して、10変数(ベースライン値)として、
 年齢 性別 BMI 拡張期血圧 総コレステロール
 LDL HDL TCH LTG グルコース
 がとられている。応答Yは、ベースライン時点から1年後の糖尿病の進行程度を数値的に表したものである(Efron et al., 2004)。
 どの因子が、1年後の進行程度に影響を及ぼしていると考えてよいか？



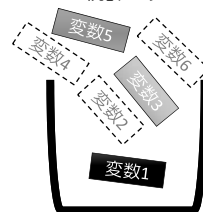
変数選択とは何か？

JMPの手順

- STEP.1 「モデルのあてはめ」を選択し、説明変数、応答変数をそれぞれ「モデル効果の構成」「Y」に追加する。
 STEP.2 「手法」を「ステップワイズ法」に変更(変数選択のため)



多変量解析では、たくさんの説明変数(共変量)を入れるほど情報がたくさんになるので、良い統計モデルになるということ？



必ずしもそうではない。説明変数が多いほど不必要な変数は単なるノイズでしかないわけだし、また、多重共線性の問題などがある。そのため、必要なものだけで統計モデルをつくるのが推奨されることが多い。そのための方法が変数選択である。

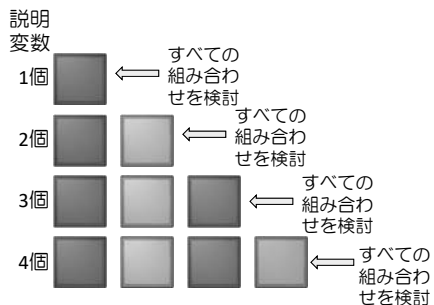


変数選択のアルゴリズムの種類

JMPの手順

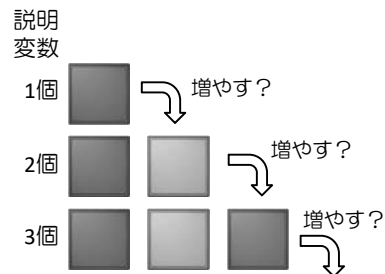
- STEP.1 「停止ルール」で変数選択に用いる測度を設定し、「方向」で変数選択のアルゴリズムを選択する。そして、「実行」ボタンを押す
 注意：変数減少法の場合は、「停止ルール」横の「すべて選択」を押してから実行ボタンを押す。

総当り法



説明変数が1個,2個,...のそれぞれのパターンを計算しその中から最適なものを選び

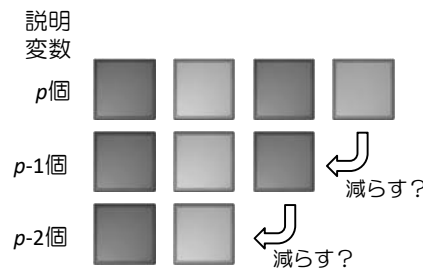
変数増加法 (前進ステップワイズ法)



説明変数が1個の場合からスタートして、変数を追加したほうが良ければ増やし、そうでなければ変数の追加をしない。

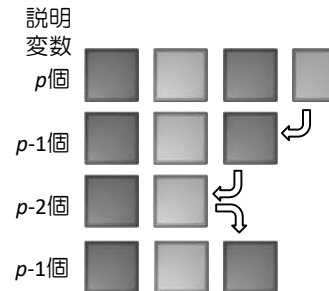
変数選択のアルゴリズムの種類

変数減少法 (後退ステップワイズ法)



全ての説明変数からスタートして、変数を減らしても影響がなければ減らし、そうでなければ変数の削除をしない。

変数増減法 (ステップワイズ法)



変数減少法からスタートするが、変数増減法では変数が削除された場合にその変数を考慮しなかったが、変数増減法では変数の削除と削除した変数の追加の両方を検討しながら各ステップを進める。

変数選択の段階においてモデルの良し悪しをどう評価するか？

情報量規準に基づく方法

モデルの当てはまりの良さを情報量規準(BIC, AIC等)を用いて評価する

検定に基づく方法

変数を追加(or 削除)した場合とそうでない場合のモデルを検定を用いて比較する。

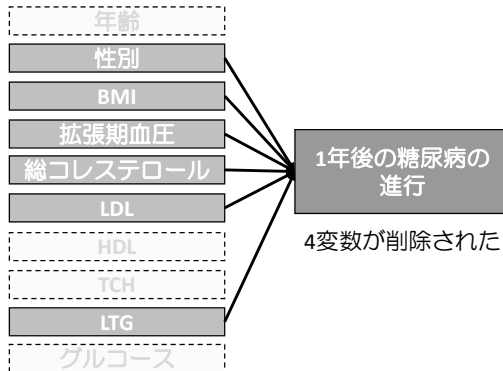
交差検証(確認)法(Cross-validation)に基づく方法

モデルの予測度の高さをコンピュータを使って計算して評価する。

JMPのOutput

事例では(変数減少法, BIC)

現在の推定値			
ロック	追加	パラメータ	推定値
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	切片	-346.15314
<input type="checkbox"/>	<input type="checkbox"/>	年齢	0
<input type="checkbox"/>	<input checked="" type="checkbox"/>	性別(1-2)	10.7955055
<input type="checkbox"/>	<input checked="" type="checkbox"/>	BMI	5.71110674
<input type="checkbox"/>	<input checked="" type="checkbox"/>	血圧	1.12655255
<input type="checkbox"/>	<input checked="" type="checkbox"/>	総コレステロール	-1.0428764
<input type="checkbox"/>	<input checked="" type="checkbox"/>	LDL	0.84327695
<input type="checkbox"/>	<input type="checkbox"/>	HDL	0
<input type="checkbox"/>	<input type="checkbox"/>	TCH	0
<input type="checkbox"/>	<input checked="" type="checkbox"/>	LTG	73.3065264
<input type="checkbox"/>	<input type="checkbox"/>	グルコース	0



推定されたモデルを評価する：寄与率

JMPの手順

STEP.1 「モデルの作成」 → 「OK」 ボタンを押す

重回帰分析では、

$$(\text{応答のバラツキ}) = (\text{予測値のバラツキ}) + (\text{誤差のバラツキ})$$

モデルの当てはまりの良さ モデルの当てはまりの悪さ

に分けることができる。これを**変動分解**という。

寄与率(R2値)とは、

$$(\text{予測値のバラツキ}) / (\text{応答のバラツキ})$$

で、モデルがどれぐらいの

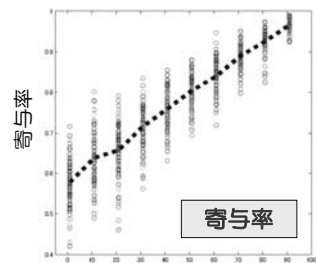
割合で、応答のバラツキを表しているか(説明できているか)を表す統計量である。

JMPのOutput

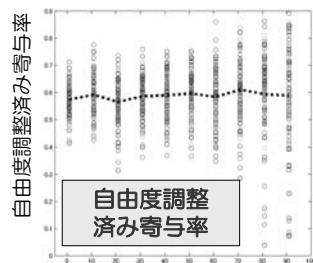
あてはめの要約	
R2乗	0.514884
自由度調整R2乗	0.508193
誤差の標準偏差(RMSE)	54.06454
Yの平均	152.1335
オブザベーション(または重みの合計)	442

JMPでは、R2乗が寄与率を表している。この事例では、全体の0.515の割合(51.5%)の説明能力があることを表している。

寄与率の注意点と自由度調整済み寄与率



寄与率



自由度調整済み寄与率

寄与率では、応答変数に影響しない説明変数であっても、沢山投入すれば、寄与率はそれにつれて上昇する。

JMPのOutput

あてはめの要約	
R2乗	0.514884
自由度調整R2乗	0.508193
誤差の標準偏差(RMSE)	54.06454
Yの平均	152.1335
オブザベーション(または重みの合計)	442

そのため、通常は、自由度調整済み寄与率(自由度調整R2乗)を用いて評価する。その結果、推定された回帰モデルは、全体の0.508(50.8%)を説明する能力があることを示している。

推定されたモデルを評価する：回帰の分散分析

JMPのOutput

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	6	1349515.1	224919	76.9487
誤差	435	1271494.0	2923	p値(Prob>F)
全体(修正済み)	441	2621009.1		<.0001*

重回帰分析では、

$$(\text{応答のバラツキ}) = (\text{予測値のバラツキ}) + (\text{誤差のバラツキ})$$

に分けることができる。

(応答のバラツキ)：全体の平方和、(予測値のバラツキ)：モデルの平方和
(誤差のバラツキ)：誤差の平方和

平均平方：「平方和/自由度」により計算されたもの

F値：(モデルの平均平方)/(誤差の平均平方)により計算されたもの

- F値とは、(自由度で調整されたモデルのバラツキ)/(自由度で調整された誤差のバラツキ)を表している。
- いいかえれば、「当てはまりの良さ/当てはまりの悪さ」を検定したものが、回帰の分散分析であり、回帰モデルに意味があるか否かを検討することができる。

事例では、 $p < 0.001$ なので、回帰モデルに意味があることが示されている。

推定されたモデルを評価する：回帰係数の評価

JMPのOutput
パラメータ推定値

項	推定値	標準誤差	t値	p値(Prob> t)
切片	-346.1531	25.77073	-13.43	<.0001*
性別[1]	10.795506	2.852819	3.78	0.0002*
BMI	5.7111067	0.707262	8.07	<.0001*
血圧	1.1265526	0.215843	5.22	<.0001*
総コレステロール	-1.042876	0.220751	-4.72	<.0001*
LDL	0.843277	0.229754	3.67	0.0003*
LTG	73.306526	7.308256	10.03	<.0001*

帰無仮説 H_0 「回帰係数が0である」に対して、対立仮説 H_1 「回帰係数が0でない」を検定したものが回帰係数の検定である。

モデル選択において、残った変数においても、この検定で有意にならない場合もあるので注意

応答(糖尿病の進行程度)に対する予測式

$$-346.153 + 10.796 \times (\text{性別}) + 5.711 \times (\text{BMI}) + 1.127 \times (\text{血圧}) - 1.043 \times (\text{総コレステロール}) + 0.843 \times (\text{LDL}) + 73.307 \times (\text{LTG})$$



LTGが最も影響が強いといっってよいか?

そうとはいえない。そもそも、単位が異なる。



推定されたモデルを評価する：標準回帰係数

JMPの手順

STEP.1 「▼」のなかの「要因のスクリーニング」→「尺度化した推定値」を選択(標準回帰係数を求めることができる)。

STEP.2 「▼」のなかの「回帰レポート」→「信頼区間をすべて表示」を選択(信頼区間を求めることができる)。

JMPのOutput
尺度化した時のもの

項	R尺度化した推定値	標準誤差	t値	p値(Prob> t)	下側95%	上側95%
切片	151.44961	2.57793	58.75	<.0001*	146.38286	156.51635
性別[1]	10.795506	2.852819	3.78	0.0002*	5.1894828	16.402528
性別[2]	-10.79551	2.852819	-3.78	0.0002*	-16.40253	-5.188483
BMI	69.104392	8.557875	8.07	<.0001*	52.284467	85.924316
血圧	39.992616	7.662436	5.22	<.0001*	24.932615	55.052616
総コレステロール	-106.3734	22.51658	-4.72	<.0001*	-150.6282	-62.11858
LDL	84.665006	23.06726	3.67	0.0003*	39.327872	130.00214
LTG	104.42148	10.41025	10.03	<.0001*	83.960847	124.88212

総コレステロールの影響が最も高く、高くなるほど、糖尿病の進行程度を表す値(応答)が小さくなるのがわかる。一方で、性別の影響は最も低かった。

変数選択の要件 (1/2)

(1) 評価したい要因は変数選択に強制的に加える

ランダム化比較試験の結果を評価する場合、治療群を表す共変量を含まなければ意味をもたない。つまり、このような場合には、背景因子などの他の共変量を調整したうえで治療群(評価変数)を調べることに意義がある。

(2) 変数増加法の落とし穴

標本サイズが小さい場合に、変数増加法を用いて変数選択を行う場合、結果の解釈が困難なモデルを選択することがしばしばある。また、本当は必要な共変量が入り込まれる前に変数選択が終了する場合がある。

(3) 多数の共変量(項目)がある場合の留意点

医学系研究では、多数の調査項目(共変量)を評価に用いることは少なくない。このような場合には、全ての共変量を用いて変数選択を行うのではなく、事前スクリーニングを行うことが推奨される。事前スクリーニングでは、共変量毎に単変量解析(1個の共変量による回帰モデルを推定する)を実施し、その回帰係数に対する検定(回帰係数が0であるか否かを評価する検定)のp値や回帰係数(オッズ比、ハザード比)を用いる。

変数選択の要件 (2/2)

(4) 欠測が多い共変量(項目)には注意が必要である

多変量解析では、共変量のなかで1個でも欠測があれば、その被験者を削除しなければならない。そのため、欠測が多い共変量をモデルに含めると、多くの被験者を削除することになる。また、観測方法が煩雑な場合には、欠測が多くなる傾向にある。そのため、このような共変量は、予め変数選択の候補から覗いておくことが望ましい。

(5) 可能であれば総当たり法を用いる

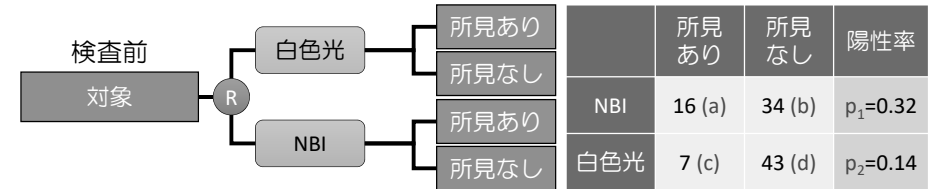
変数増加法や変数減少法が必ずしも最適モデルに到達するとは限らない。最適なモデルを選択できる唯一の方法は、すべての候補モデルを評価する総当たり法のみである。共変量の数が10個の場合、候補となるモデルの数は1,023個である。最近のコンピューターであれば実現可能な数ではない(共変量の数が20個の場合には1,048,575個となり、不可能に近い数値となる)。そのため、臨床的知見あるいは、事前スクリーニングなどを用いて変数選択に用いる共変量を可能な限り少なくし、そのもとの、総当たり法によって変数選択を実施することが考えられる。

■ ロジスティック回帰分析

Logistic Regression Analysis

オッズ比とは何か

いま、ある消化器癌に対する内視鏡検査において、NBIと白色光の比較を行ったときの所見の有無の比較を行う仮想的な試験があったとする。



オッズとは、ある結果が生じる比率とある結果が生じない比率との比である。

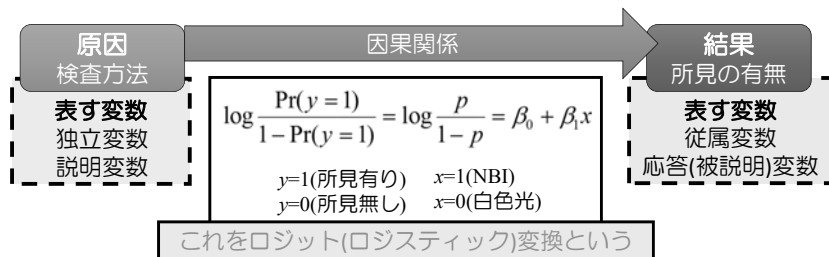
$$\text{NBI} = \frac{0.32}{1-0.32} = 0.471, \quad \text{白色光} = \frac{0.14}{1-0.14} = 0.163$$

オッズ比は、結果(所見の有無)に対してどれぐらい要因(検査法)が寄与しているかを表す。

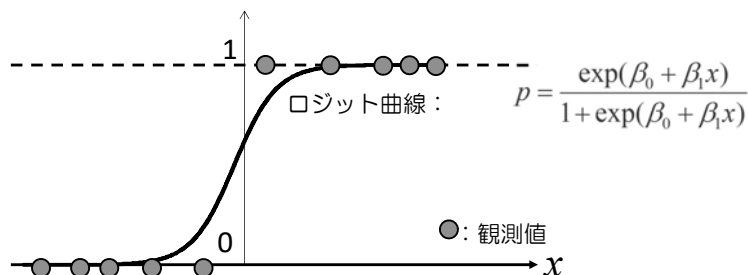
$$\text{オッズ比} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{a \times d}{b \times c} = \frac{16 \times 43}{34 \times 7} = 2.891$$

つまり、NBIは白色光に比べて2.891倍の所見の上昇を与えることができる。

ロジスティック回帰とは



ロジスティック回帰モデルは、対数オッズに対するモデル



ロジスティック回帰分析とオッズ比の関係

$$\log \frac{\Pr(y=1)}{1-\Pr(y=1)} = \log \frac{p}{1-p} = \beta_0 + \beta_1 x \quad \begin{array}{l} y=1(\text{所見有り}) \quad x=1(\text{NBI}) \\ y=0(\text{所見無し}) \quad x=0(\text{白色光}) \end{array}$$

■ 白色光のオッズ $odds_{\text{白}}$ を計算

$$\log \frac{p_{\text{白}}}{1-p_{\text{白}}} = \beta_0 + \beta_1 \times 0 = \beta_0 \Rightarrow odds_{\text{白}} = \frac{p_{\text{白}}}{1-p_{\text{白}}} = \exp(\beta_0)$$

■ NBIのオッズ $odds_{\text{N}}$ を計算

$$\log \frac{p_{\text{N}}}{1-p_{\text{N}}} = \beta_0 + \beta_1 \times 1 = \beta_0 + \beta_1 \Rightarrow odds_{\text{N}} = \frac{p_{\text{N}}}{1-p_{\text{N}}} = \exp(\beta_0 + \beta_1)$$

■ NBI/白色光のオッズ比 OR を計算

$$OR = \frac{odds_{\text{N}}}{odds_{\text{白}}} = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_0 + \beta_1 - \beta_0) = \exp(\beta_1)$$

つまり、ロジスティック回帰では、オッズ比が $\exp(\beta_1)$ で表される。

ロジスティック回帰分析における調整オッズ比

下の表は、ある疾患に対する治療法(既存治療：0, 新規治療：1)と年齢(65歳未満：0, 65歳以上：1)とその成功の有無(成功：1, 失敗：0)に対する2個のクロス集計表である。

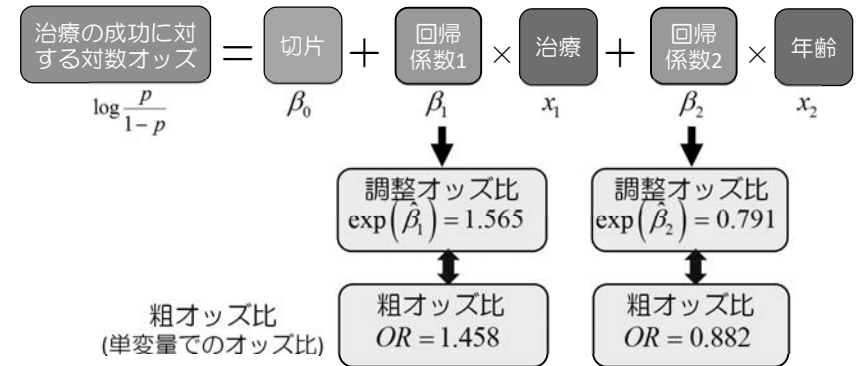
年齢	治療	成功/失敗		
		成功(1)	失敗(0)	合計
65歳以上(1)	新規(1)	444	567	1011
	既存(0)	56	150	206
	合計	500	717	1217
65歳未満(0)	新規(1)	286	324	610
	既存(0)	171	254	425
	合計	457	578	1035

(多重)ロジスティック回帰分析のモデルは

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

既存治療：0
新規治療：1
65歳未満：0
65歳以上：1

調整オッズ比と粗オッズ比



粗オッズ比は、他の要因(共変量)を何も考慮しない。これに対して、調整オッズ比では、

- ・治療のオッズ比では年齢による違いを考慮
- ・年齢のオッズ比では治療による違いを考慮してオッズ比を計算している。

JMPの結果を見ながら解釈する

赤血球沈降測定度(ESR)が20未満(正常), 20以上(異常)がとられた被験者に対して
フィブリノゲン, グロブリン
がとられている (Collett & Jemain, 1985).

ここでの目的は、ESRに対して、二つの血漿蛋白の水準との間に連関が存在する
か否かを検討することにある。

JMPの手順

STEP.1 「モデルのあてはめ」を選択し、説明変数、応答変数をそれぞれ「モデル効果の構成」「Y」に追加する。

注意： 応答が名義尺度の場合には、自動的にロジスティック回帰を選択してくれる。

JMPのOutput

要因	対数値	P値
fibrinogen	1.838	0.01452
globulin	0.766	0.17156

削除 追加 編集 FDR
勾配で収束しました, 5回の反復

相対値とは、 $-\log(p)$ で表すことができ、各共変量の影響の強さを表すことができるが、あまり用いることはない。

ロジスティック回帰モデルの適合度の評価

JMPのOutput

モデル全体の検定				R2乗(U)		
モデル	(-1)*対数尤度	自由度	カイ2乗	p値(Prob>ChiSq)		
差	3.956925	2	7.91385	0.0191*		0.2562
完全	11.485557				AICc	29.8283
縮小	15.442482				BIC	33.3683
					オブザベーション(または重みの合計)	32

■ p値の意味

「完全」のモデル：(切片) + (フィブリノゲン) + (グロブリン)

「縮小」のモデル：(切片)

ここでのp値は、切片のみのモデルと共変量を含めたモデルのあてはまりの違いを比較することで、モデルの適合度(当てはまりの良さ)を評価している。

p値が0.05未満(0.0191)なので、推定されたモデルが適切である。

■ R2値, AICc, BICとは?

ー R2値は、重回帰分析に類似した定義であるが、ロジスティック回帰分析の場合は、かなり小さな値をとるので、あまり参考にならない。

ー AICc, BICは、それぞれ修正赤池の情報量規準, Bayes流情報量規準というもので、複数のモデルを比較するのに用いることができる(詳細は後程)。

ロジスティック回帰モデルの評価

JMPのOutput

あてはまりの悪さ(LOF)				パラメータ推定値			
要因	自由度	(-1)*対数尤度	カイ2乗	項	推定値	標準誤差	カイ2乗 p値(Prob>ChiSq)
あてはまりの悪さ(LOF)	29	11.485557	22.97111	切片	-12.792075	5.7964205	4.87 0.0273*
飽和モデル	31	0.000000	p値(Prob>ChiSq)	fibrinogen	1.91036842	0.9709981	3.87 0.0491*
あてはめたモデル	2	11.485557	0.7779	globulin	0.15577813	0.1195382	1.70 0.1925

推定値は次の対数オッズに対するものです: ESR > 20/ESR < 20

■ LOF(不適合度 Lack of Fit)とは,

飽和モデル: (切片) + (フィブリノゲン) + (グロブリン) + (フィブリノゲン) × (グロブリン)

あてはめたモデル: (切片) + (フィブリノゲン) + (グロブリン)

を比較して、これ以上モデルを複雑にすべきか否かを検討している。

p値が0.05以上(0.7779)なので、推定されたモデルが適切である。

■ パラメータ推定値のp値は、帰無仮説 H_0 「回帰係数が0である」に対して、対立仮説 H_1 「回帰係数が0でない」を検定したものが回帰係数の検定(尤度比検定)である。

グロブリン(globulin)の回帰係数が有意でない

グロブリンを除いた場合との比較

JMPのOutput

グロブリンを含む場合				グロブリンを除いた場合			
モデル	(-1)*対数尤度	自由度	カイ2乗 p値(Prob>ChiSq)	モデル	(-1)*対数尤度	自由度	カイ2乗 p値(Prob>ChiSq)
差	3.956925	2	7.91385 0.0191*	差	3.022304	1	6.044608 0.0139*
完全	11.485557			完全	12.420178		
縮小	15.442482			縮小	15.442482		

グロブリンを含む場合		グロブリンを除いた場合	
指標	値	指標	値
R2乗(U)	0.2562	R2乗(U)	0.1957
AICC	29.8283	AICC	29.2541
BIC	33.3683	BIC	31.7718
オブザベーション(または重みの合計)	32	オブザベーション(または重みの合計)	32

グロブリンを除いた場合のほうが、AICC, BICが小さい

グロブリンを除いたモデルのほうが、AICC, BICともに小さい。情報量規準では、数値が小さいほど当てはまりが良いと解釈されるため、

(切片) + (フィブリノゲン) + (グロブリン)

のようが良いことがわかる。

対応のあるデータでの解析：条件付きロジスティック回帰分析

- 皮膚科等での無作為化ハーフサイド試験
- 1対1マッチングされたケースコントロール研究
- 傾向スコア分析におけるマッチングさせたデータの統計解析

では、共変量が同じで、アウトカムに対応がある。

Case	Control	
●	●	Same Covariate
●	●	Same Covariate
●	●	Same Covariate
⋮	⋮	
●	●	Same Covariate

このような場合に利用できるロジスティック回帰モデルが条件付きロジスティック回帰分析である。

■ 条件付ロジスティック回帰分析

Conditional Logistic Regression Analysis

条件付きロジスティック回帰モデルの図示

条件なしロジスティック回帰(通常のロジスティック回帰分析)

$$\text{logit}(x) = \text{切片(共通)} + \text{回帰係数} \times \text{共変量}$$

条件付きロジスティック回帰(通常のロジスティック回帰分析)

$$\begin{aligned} \text{logit}(x_1) &= \text{1番目の切片} + \text{回帰係数} \times \text{共変量} \\ \text{logit}(x_2) &= \text{2番目の切片} + \text{回帰係数} \times \text{共変量} \\ \text{logit}(x_3) &= \text{3番目の切片} + \text{回帰係数} \times \text{共変量} \\ &\vdots \\ \text{logit}(x_n) &= \text{n番目の切片} + \text{回帰係数} \times \text{共変量} \end{aligned}$$

条件付きロジスティック回帰分析では、共通の共変量をもつペア毎に切片が異なるが、共変量の影響は共通である。

JMPの結果を見ながら解釈する

子宮がん患者および健常者126名に対して、2変数として、

胆嚢疾患の有無、高血圧

がとられている。応答Yは、子宮がんの有無である(Breslow & Day, 1980)。

ここでの目的は、高血圧の影響を加味しながら、胆嚢疾患が子宮がんに対して、どの程度のリスクを与えるかを評価することにある。

JMPのData set

	組を表す変数		アウトカムを表す変数		共変量	
	ペア	アウトカム	胆嚢疾患	高血圧		
1	1	1	0	0	0	0
2	1	0	0	0	0	0
3	2	1	0	0	0	0
4	2	0	0	0	0	0
5	3	1	0	1	0	1
6	3	0	0	1	0	1
7	4	1	0	0	0	0
8	4	0	1	0	0	0
9	5	1	1	1	0	0
10	5	0	0	1	0	1

JMPの手順

- STEP.1 「消費者調査」→「選択モデル」を選択する。
- STEP.2 「プロファイルID」にアウトカム、「グループ」に組を表す変数を選択する。
- STEP.3 共変量を「モデル効果の構成」に入力する。
- STEP.4 「▼」のなかの「信頼区間」を選択する。

JMPの結果を見ながら解釈する

JMPのOutput (胆嚢疾患の有無, 高血圧)

項	推定値	標準誤差	下側95%	上側95%
胆嚢疾患	0.9704079430	0.5307308967	-0.012753	2.1161848
高血圧	0.3480656863	0.3769634561	-0.383275	1.1115359

AICc: 86.987843
BIC: 91.074112
(-2)*対数尤度: 82.787843

JMPのOutput (胆嚢疾患の有無)

項	推定値	標準誤差	下側95%	上側95%
胆嚢疾患	0.9555112499	0.5262347888	-0.019594	2.0922555

AICc: 85.719141
BIC: 87.796702
(-2)*対数尤度: 83.653567

高血圧を除いた場合のほうが、AICc, BICが小さい

JMPでは、オッズ比の出力ができないが、回帰係数を用いれば、EXCELで簡単に計算できる。例えば、胆嚢疾患の有無でのオッズ比は、

$$\text{オッズ比} : \exp(0.9555) = 2.59997$$

$$95\% \text{信頼区間} : [\exp(-0.020), \exp(2.092)] = [0.980, 8.101]$$

である。すなわち、胆嚢疾患が有る被験者は、無い被験者に比べて、子宮がんに対するリスクが、2.5997倍であることがわかる。

Poisson回帰分析

Poisson Regression Analysis

カウントデータに対する回帰モデル：Poisson回帰分析

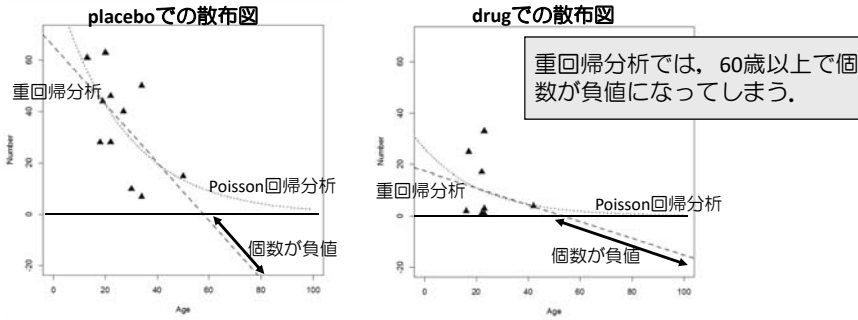
- ポリープの個数などを調べた研究
- 疫学調査において一定期間内に特定の疾患を発症した症例数

カウントデータを取り扱うPoisson回帰モデルでは、次式

$$\log \mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}$$

で定義される。ここで μ_i は、共変量 x_i をもつ個体において、平均的に得られる個数(例えば、ポリープの個数)を表している。

家族性大腸線腫症治療における非ステロイド性抗炎症剤のプラセボ対照試験の結果である。アウトカムは、大腸ポリープの個数である(Piantadosi, 1997)。ここでの共変量は、薬剤(placebo, drug)及び年齢(age)となっている。



JMPの結果を見ながら解釈する

JMPの手順

- STEP.1 「モデルのあてはめ」を選択し、モデルを指定する。
 STEP.2 「手法」を「一般化線形モデル」、「分布」を「Poisson」、「リンク関数」を「対数」にする。

JMPのOutput

▲モデル全体の検定				
モデル	(-1)*対数尤度	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
差分	99.5594678	199.1189	2	<.0001*
完全	133.938413			
縮小	233.49788			
適合度統計量				
	カイ2乗	自由度	p値(Prob>ChiSq)	
Pearson	182.3731	17	<.0001*	
デビアンس	179.5408	17	<.0001*	
AICc				
	275.3768			

注意：Poisson回帰分析のもととなっているPoisson分布は、平均と分散が同じであることが仮定されている。一方で、Poisson回帰分析では、平均<分散となる場合がある。これを、過分散(超過変動)と呼ぶ。

JMPの結果を見ながら解釈する

JMPの手順(過分散を考慮)

- STEP.1 「モデルのあてはめ」を選択し、モデルを指定する。
 STEP.2 「手法」を「一般化線形モデル」、「分布」を「Poisson」、「リンク関数」を「対数」にする。
 STEP.3 「過分散に基づく検定と信頼区間」にチェックする。

JMPのOutput

適合度統計量	カイ2乗	自由度	p値(Prob>ChiSq)
Pearson	182.3731	17	<.0001*
デビアンス	179.5408	17	<.0001*

過分散 10.7278
 1を大きく超えた場合には過分散とされ、この事例では、10.7278ということで過分散になる。ここでは、分散を調整した場合で標準誤差を計算する。

共変量の影響の強さ

JMPのOutput (過分散を考慮した場合)

▲パラメータ推定値						
項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	下側信頼限界	上側信頼限界
切片	3.8494827	0.4799589	62.32016	<.0001*	2.9190042	4.8067085
treat[drug]	-0.679542	0.1926633	15.782373	<.0001*	-1.092302	-0.326924
age	-0.03883	0.0195056	4.5692482	0.0326*	-0.079695	-0.003022

Poisson回帰分析の場合も、ロジスティック回帰分析と同様に回帰係数の推定値の指数をとったもの(ロジスティック回帰ではオッズ比)を用いて解釈できる。これを、率比(rate ratio)という。

JMPでは、率比の出力ができないが、回帰係数を用いれば、EXCELで簡単に計算できる。例えば、薬剤(drug, placebo)での率比は、

$$\text{オッズ比} : \exp(-0.6795) = 0.5069$$

$$95\% \text{信頼区間} : [\exp(-1.0923), \exp(-0.0030)] = [0.3354, 0.9970]$$

である。すなわち、薬剤を投与することで、プラセボに比べてポリープの出現を半分以下に抑えることができる。

▲パラメータ推定値 標準誤差が異なるため、p値及び信頼区間に影響する。						
項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	下側信頼限界	上側信頼限界
切片	3.8494827	0.1465373	668.56005	<.0001*	3.5634664	4.138063
treat[drug]	-0.679542	0.0588225	169.3106	<.0001*	-0.797745	-0.566877
age	-0.03883	0.0059553	49.018115	<.0001*	-0.050743	-0.027393

▲パラメータ推定値						
項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	下側信頼限界	上側信頼限界
切片	3.8494827	0.4799589	62.32016	<.0001*	2.9190042	4.8067085
treat[drug]	-0.679542	0.1926633	15.782373	<.0001*	-1.092302	-0.326924
age	-0.03883	0.0195056	4.5692482	0.0326*	-0.079695	-0.003022

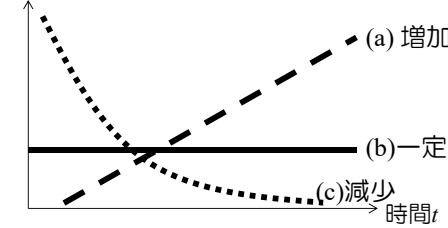
■ 比例ハザード・モデル

Proportional Hazard Model

ハザードの意味

被験者が時点 t まで生存しているという仮定のもとで微小区間 $t \leq T < t + \Delta t$ で死亡する確率

ハザード関数 $h(t)$

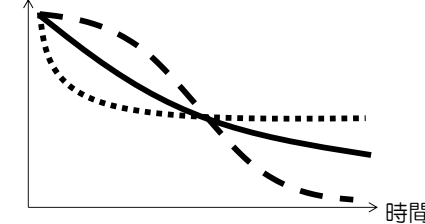


(a) の状況
登録時点では死亡(あるいはイベント)のリスクは低いものの時間経過とともに増大することを意味する。

(b) の状況
全時間上でハザード(すなわちリスク)が一定であることを意味する。

(c) の状況
登録時点では死亡(あるいはイベント)のリスクは高いものの時間経過とともに減少することを意味する。一般に、進行がんに対して見られる。

生存確率 $S(t)$



比例ハザード・モデル

■ モデルの記述：
$$h(t, \mathbf{x}) = h_0(t) \exp\{\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p\}$$
$$= h_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}\}$$

記法 — $h_0(t)$: 基線ハザード(Baseline hazard)
共変量に依存しない潜在的なハザード
 \mathbf{x} : $p \times 1$ 共変量ベクトル
 $\boldsymbol{\beta}$: $p \times 1$ パラメータ・ベクトル

上式を模式化すると下記のようになる。因みに、ハザードがわかれば、生存曲線を推定することは可能

ハザード = 基線ハザード $\times \exp$ $\left\{ \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \right\}$

時間 t のみに依存 (係数 $1 \times$ 共変量 $1 +$ 係数 $2 \times$ 共変量 $2 + \dots$ 共変量 \mathbf{x} のみに依存)

- 比例ハザード・モデルとは、時間のみに依存する基線ハザード(ハザードの形状)と共変量のみ依存する線形項で構成される。
- ハザードの形状は、共変量の状況に依存しないので「比例ハザード」である。

比例ハザード・モデルの利用

■ ハザード比

いま、 $x_1=1$ のとき新規治療、 $x_1=0$ のとき既存治療とするときに $\exp(\beta)$ がハザード比 HR を表している。

■ 交絡因子の(事後的な)調整

いま、 $x_1=1$ のとき新規治療、 $x_1=0$ のとき既存治療とする。また、対象の予後が腫瘍径に依存することがわかっているときに、 x_2 が腫瘍径を表しているとする。このとき、モデル

$$h(t, \mathbf{x}) = h_0(t) \exp\{\beta_1 x_1 + \beta_2 x_2\}$$

であらわされるときに $\exp(\beta_2)$ は、腫瘍径の影響を調整したもとの治療効果によるハザード比 HR を表している。

■ 交互作用の検討(詳しくは後述)

上の設定のもとで

$$h(t, \mathbf{x}) = h_0(t) \exp\{\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2\}$$

であらわされるときに $\exp(\beta_3)$ は、腫瘍径と治療効果の交互作用 HR を表している。

JMPで比例ハザード・モデルを実行する

ここでは、サンプルデータVA Lung Cancer.jmpを用いる。このデータは、肺がんに対する無作為化比較試験の予後因子の評価を実施するデータである。

セル：名義尺度(Adeno, Large, Small, Squamous),
前治療：名義尺度(Yes, No)

処置：名義尺度(Standard, Test)
年齢・月数・一般全身状態(KPS)：連続尺度

JMPの手順 (比例ハザード・モデルの実行)

- STEP.1** 「分析」→「信頼性/生存時間分析」→「比例ハザードのあてはめ」を選択する。
- STEP.2** 「生存日数」を選択して「イベントまでの時間」を押す。
- STEP.3** 「セル」、「処置」、「前治療」、「年齢」、「月数」、「KPS」を選択して「モデル効果の構成」のなかの「追加」を押す。
- STEP.4** 「打ち切りの有無」を選択して「打ち切り」を押す(打ち切りの値がCensoredになっていることを確認する)。

調整ハザード比の計算

- STEP.1** ▼をクリックする。
- STEP.2** 「リスク比」を選択する。

JMPのアウトプット

モデル全体				
	AICC	BIC		
イベントの数	128	967.484	989.719	
打ち切りの数	9			
合計数	137			
モデル	(-1)*対数尺度	カイ2乗	自由度	p値(Prob>ChiSq)
差分	30.7046	61.4091	8	<.0001*
完全	475.1794			
縮小	505.8840			

共変量がないモデル(null)モデルに対する比例ハザード・モデルの当てはまりの良さを検定で評価した結果(<0.001なのでモデルに意味がある)。

推定された回帰係数と95%信頼区間

パラメータ推定値				
項	推定値	標準誤差	下側95%	上側95%
セル[Adeno]	0.57719588	0.1849861	0.2146297	0.939762
セル[Large]	-0.2114757	0.1741197	-0.552744	0.1297927
セル[Small]	0.24538322	0.1592518	-0.066745	0.557511
処置[Standard]	-0.1449679	0.1036051	-0.34803	0.0580943
前治療[No]	-0.0361633	0.1160663	-0.263649	0.1913224
年齢	-0.0085494	0.0093042	-0.026785	0.0096864
月数	-0.000092	0.0091251	-0.017977	0.0177929
KPS	-0.0326217	0.0055052	-0.043412	-0.021832

Squamousを基準としている

回帰係数に対する検定

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)
セル	3	3	18.6003364	0.0003*
処置	1	1	1.96251687	0.1612
前治療	1	1	0.09639285	0.7562
年齢	1	1	0.82813759	0.3628
月数	1	1	0.00010177	0.9920
KPS	1	1	34.616163	<.0001*

尤度比検定：推定された回帰係数での(部分)尤度と帰無仮説($\beta=0$)の(部分)尤度の差が評価される

Wald検定：推定された回帰係数と帰無仮説($\beta=0$)での回帰係数の違いが評価される。

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)
セル	3	3	17.9163867	0.0005*
処置	1	1	1.95786132	0.1617
前治療	1	1	0.09707858	0.7554
年齢	1	1	0.84434423	0.3582
月数	1	1	0.00010165	0.9920
KPS	1	1	35.1124491	<.0001*

結果の相違があった場合には、そのうちの1つの検定を選ぶ必要がある。 Hosmer et al.(2013)は、尤度比検定の利用を推奨している。

Thank you for your kind attention



■ 補遺

Appendix

- 交互作用(interaction)について
- 共変量が順序カテゴリカル尺度の取扱い
- 共変量が量的尺度の場合の2値化の留意点

交絡と交互作用

交絡	交互作用
<ul style="list-style-type: none"> 重要な予後因子が群間で不均衡 試験計画そのものが問題である(割付因子として組み込む必要性) 交絡がないということは試験計画の妥当性がいえる。 	<ul style="list-style-type: none"> 薬剤の効果が、サブグループ間で異なる。 薬剤の適応患者像に関する制限につながる。 交互作用がないということは治療の一般可能性のあることがいえる。

「浜田知久馬(1999). 学会・論文発表のための統計学, 新興交医書出版部」より引用・改稿

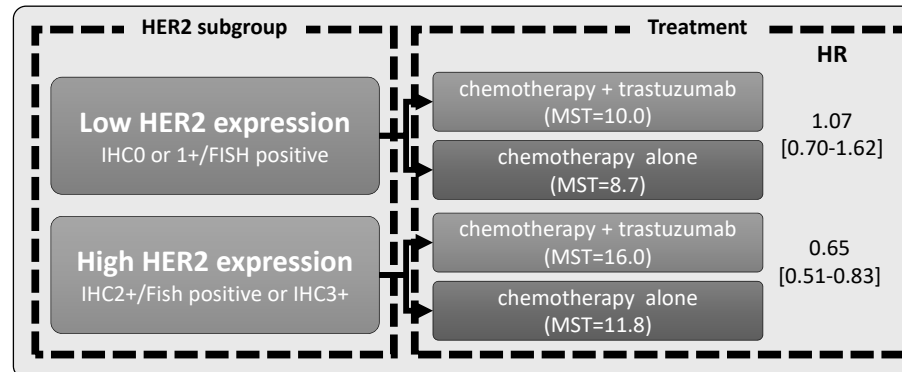
- ICH E5ガイドライン「臨床試験の統計的原則」では、多施設共同試験では、薬剤と施設の交互作用を検討することが要求される。これは、薬剤の一般可能性を証明するためである。
- 交互作用は、回帰モデル(生存時間解析では、比例ハザードモデル)に治療(薬剤)のダミー変数と予後因子の積による項を追加し、その回帰係数の有意性により検出する。一方で、交互作用の構成については、Cox(1984)によって詳細に吟味されている。

Cox, D.R.(1984). Interaction, *International Statistical Review*, 52(1), 1-31.

交互作用とは(1/2) : ToGA study

ToGA study(Bang et al.: *The Lancet*, 376, 687-697, 2010)の一文

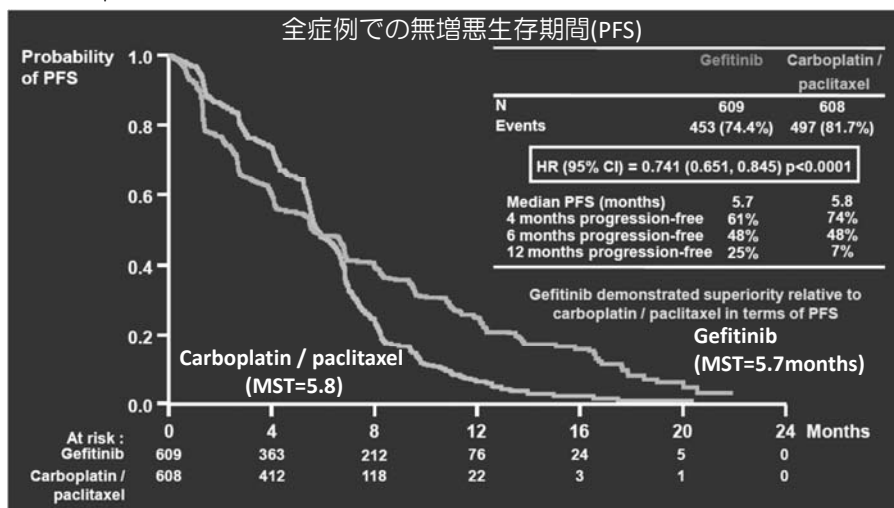
There was evidence of a **significant interaction test** ($p=0.036$) between treatment and the two HER2 subgroups (high HER2 expression vs low HER2 expression).
 切除不能な局所進行性、再発 and/or 転移性のHER2陽性の患者に対する化学療法単独群 (5-FU+ CDDP (FP療法)あるいはCapecitabine +CDDP(XP療法)) と化学療法+Trastuzumab併用群を比較する国際共同無作為化非盲検試験



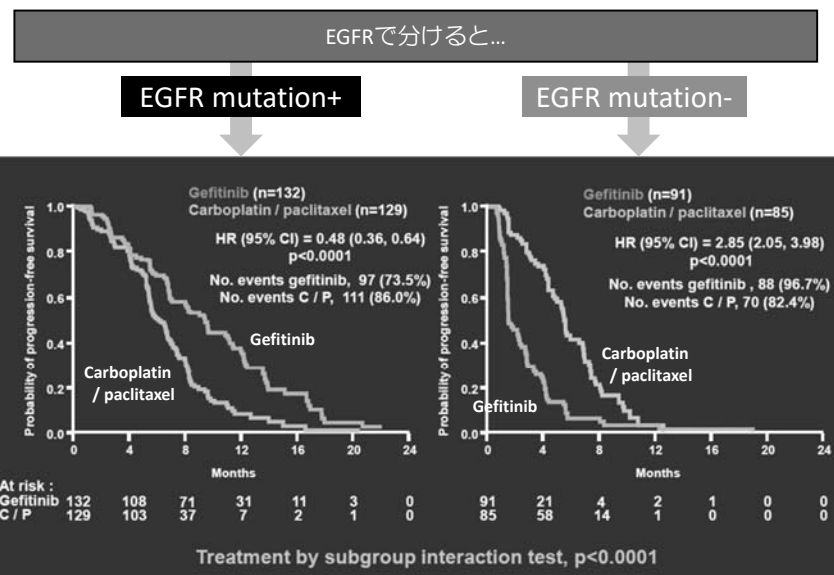
Low HER2ではハザード比(HR)が有意でないが、High HER2では有意に chemotherapy + trastuzumabのハザード比が低い。

交互作用とは (2/2) : IPASS Study

IRESSA Pan-Asia Study(IPASS)試験は、アジア地域における非盲検ランダム化並行群間比較試験である。ここでは、進行非小細胞肺癌患者1,217例を対象に1次治療としてのGefitinib(イレッサ)の有効性・安全性および忍容性をCarboplatin+Paclitaxel併用化学療法と比較した試験である。

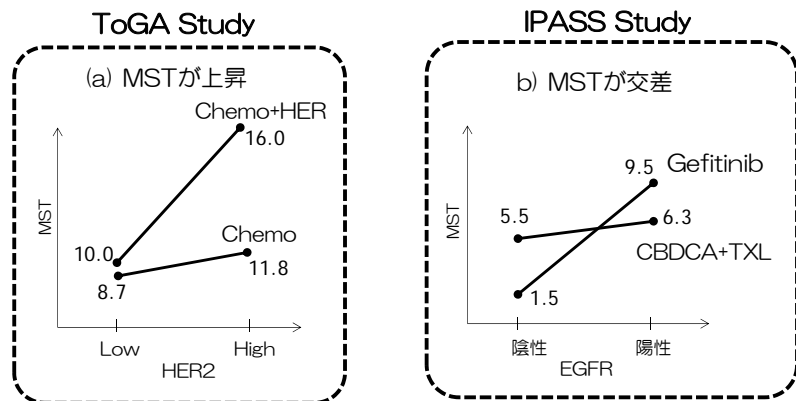


Mok, T.S. et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma, *N.Engl.J Med.*, 361(10), 947-57, 2009.



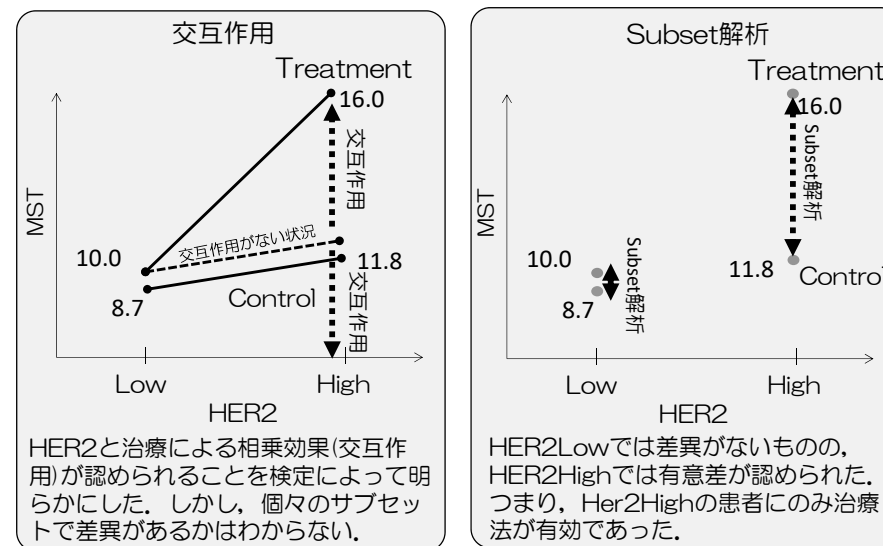
Mok, T.S. et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma, *N.Engl.J Med.*, 361(10), 947-57, 2009.

交互作用とは？



何らかの要因によって、抗がん剤の効果が逆転したり、相乗効果によって、より効果が上昇するような現象を交互作用(interaction)という。交互作用は、治療方法の一般可能性に繋がる。

交互作用の評価とサブセット解析の違い



共変量が順序カテゴリカル変数の場合

ECOG Performance Status(PS)による全生存期間への影響を評価する場合(被験者の適格基準がPSが0~2であるとする)

$$\lambda = \lambda_0(t) \cdot \exp\{\beta_1 \times (PS1) + \beta_2 \times (PS2)\}$$

$\exp\{\beta_1\}$: PS1/PS0のHR, $\exp\{\beta_2\}$: PS2/PS0のHR



ハザード比は、基準となるカテゴリあるいはペアワイズに計算しなければならない。一方で、PSが全生存期間に影響するかどうかは、回帰パラメータ・ベクトル β が有意であるか否か、すなわち

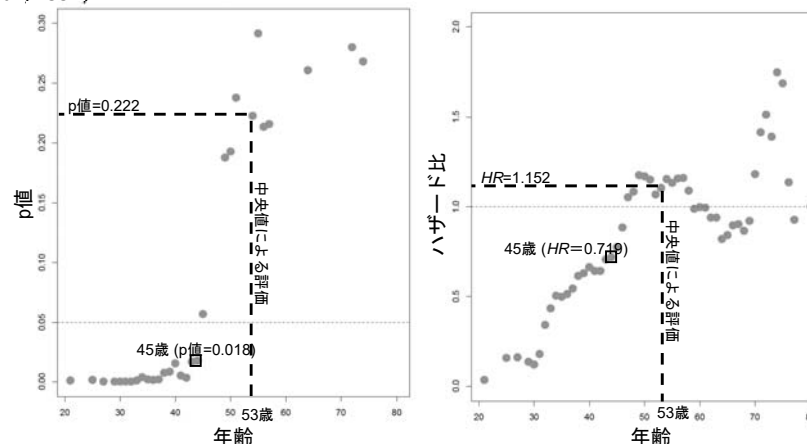
$$H_0: (\beta_1, \beta_2)^T = (0, 0)^T, H_1: H_0 \text{ではない}$$

を検定すればよい(JMPでは、計算できる)。

この検定では、共変量全体では有意でないにも関わらず、個々のカテゴリで見ると有意になる場合がある(解釈が困難になる)。

共変量が量的尺度の場合の2値化の留意点

ドイツ乳がん研究グループによって実施されたランダム化比較試験のデータ(Schumacher et al., 1994)



変量の2値化において用いられることが多い中央値(53歳以上, 53歳未満)をカットオフ値にした場合、p値は0.222である(有意でなかった)。一方で、45歳(45歳以上, 45歳未満)をカットオフ値にした場合、p値は0.018である(有意だった)。

Schumacher, M., et al. : Randomized 2x2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. Journal of Clinical Oncology, 12, 2086-2093, 1994.

共変量が量的尺度の場合の2値化の留意点

45歳(<45/≥45)の場合

高年齢群(45歳以上)のほうが
低年齢群(45歳未満)よりも死
亡リスクが低い (HR=0.719)

解釈が逆



53歳(<53/≥53)の場合

高年齢群(53歳以上)のほうが
低年齢群(53歳未満)よりも死
亡リスクが低い (HR=1.152)

カットオフ値の選定に関するRule of Thumb

- (1) 中央値を用いる
- (2) 臨床的にリーズナブルなカットオフ値を選定する

臨床的にリーズナブルとは、今回の事例の場合には、例えば疫学的調査などから得られた乳癌患者の平均年齢を用いたり、あるいは、閉経年齢を用いることを意味する。いずれにしても、カットオフ値選定におけるゴールド・スタンダードは存在しないため、試行的に決定する必要がある。