

ランダム化と交絡

生物統計部門/JCOGデータセンター

水澤 純基

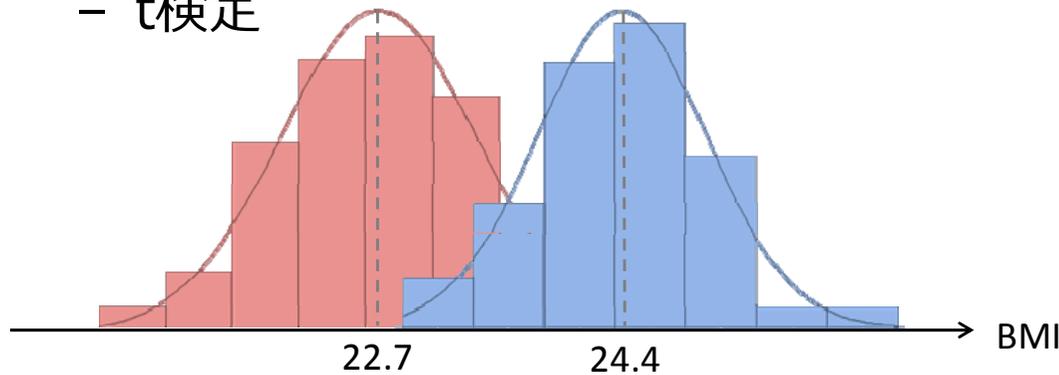
第4回生物統計基礎セミナー

2014. 2.12 (水)

復習：連続量データの解析

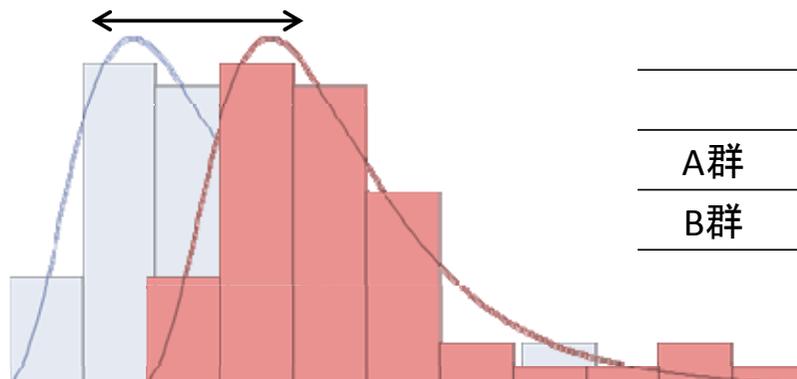
- Step1：ヒストグラムで分布の確認
- Step2：群間比較

- t検定



$$t = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{V\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

- Wilcoxon順位和検定



	BMI (Kg/m ²)				平均値
A群	22.0 [2]	28.3 [8]	19.4 [1]	22.3 [3]	23.0
B群	23.3 [4]	25.1 [7]	24.6 [6]	23.5 [5]	24.1

復習：2値データの解析

- Step1：(2×2) 分割表を作成して頻度集計
- Step2：群間比較

		大腸がん発症		
		あり	なし	合計
アスピリン	あり	129	4932	5061
	なし	87	2440	2527
合計		216	7372	7588

– Fisherの直接確率検定

– 片側P値 = $0.0053 + \dots + 1.2 \times 10^{-106} = 0.0174$

216	4845	5061
0	2527	2527
216	7372	7588

Pr = 2.4×10^{-37}

215	4846	5061
1	2526	2527
216	7372	7588

Pr = 1.3×10^{-35}

214	4847	5061
2	2525	2527
216	7372	7588

Pr = 5.0×10^{-34}

...

129	4932	5061
87	2440	2527
216	7372	7588

Pr = 0.0053

観測データと同じ分割表

...

0	5061	5061
216	2311	2527
216	7372	7588

Pr = 1.2×10^{-106}

復習：2値データの解析

- Step1：(2×2) 分割表を作成して頻度集計
- Step2：群間比較

– カイ2乗検定

		大腸がん発症		合計
		あり	なし	
アスピリン	あり	129	4932	5061
	なし	87	2440	2527
合計		216	7372	7588

観測		期待	
129 (2.5%)	4932 (97.5%)	144 (2.8%)	4917 (97.2%)
87 (3.4%)	2440 (96.6%)	72 (2.8%)	2455 (97.2%)

🐾 観測度数と期待度数の差に基づく検定

- 🥕 観測度数と期待度数の差が大きいほど
帰無仮説の下で稀な結果が観測されたといえる

🐾 統計量

$$\frac{(129 - 144)^2}{144} + \frac{(4932 - 4917)^2}{4917} + \frac{(87 - 72)^2}{72} + \frac{(2440 - 2455)^2}{2455}$$

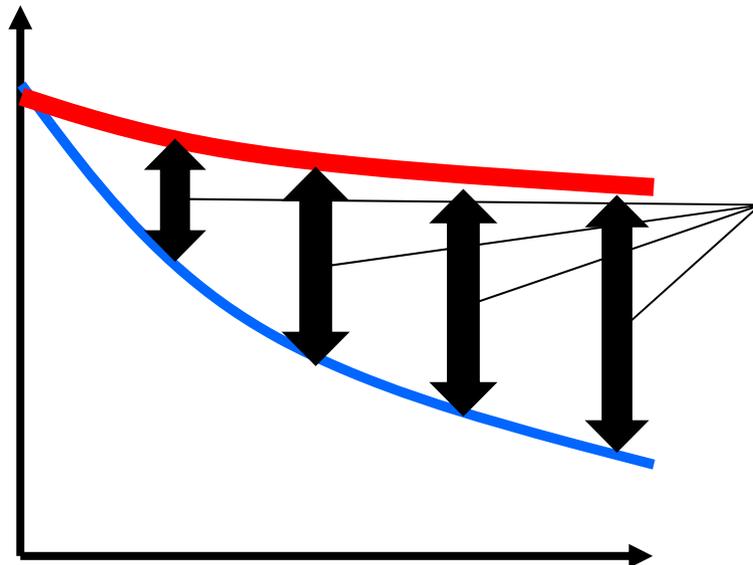
🐾 自由度1のカイ二乗分布に近似的に従う

- 🥕 期待度数が大きいほど近似がよい

🐾 P = 0.0273 (両側5%水準で有意！)

復習：生存時間の解析

- Step1：Kaplan-Meier法により生存曲線の確認
 - 生存期間中央値、年次生存割合の推定
- Step2：群間比較
 - ログランク検定



・ イベント発生時点ごとの生存曲線のズレを足し併せて、曲線全体の差を比較している

・ 単純には、2つの生存曲線が離れているほど、ズレは大きくなる (=p値は小さくなる)

復習 : アウトカムの型と統計手法

	連続量	2値(0 / 1)	生存時間
アウトカムの例	血圧値、 臨床検査値	奏効割合	全生存期間 無増悪生存期間
データの要約	ヒストグラム	分割表	Kaplan-Meier法
群間比較(検定)	t検定、 Wilcoxon検定	カイ2乗検定、 Fisherの 直接確率検定	ログランク検定
モデルのあてはめ	重回帰分析	ロジスティック回帰	Cox回帰

肝転移を有する切除不能の大腸がん患者に手術は有効？

Original article

Long-term survival of patients with unresectable colorectal cancer liver metastases following infusional chemotherapy with 5-fluorouracil, leucovorin, oxaliplatin and surgery

S. Giacchetti,¹ M. Itzhaki,¹ G. Gruia,¹ R. Adam,² R. Zidani,¹ F. Kunstlinger,¹ S. Brienza,¹ E. Alafaci,¹ F. Bertheault-Cvitkovic,¹ C. Jasmin,¹ M. Reynes,³ H. Bismuth,² J. L. Misset¹ & F. Lévi¹

¹Centre de Chronothérapie, Fédération des Maladies Sanguines Immunitaires et Tumorales, Service d'Oncologie Médicale; ²Centre Hépto-biliaire, ³Service d'Anatomie Pathologique, Hôpital Paul Brousse, Villejuif, France

Summary

Context: Long-term survival of patients with metastatic colorectal cancer has been achieved only in patients who underwent complete resection of metastases. Such surgery could be performed in a greater proportion of patients if effective chemotherapy could downstage previously unresectable metastases. This approach has been limited by the low tumor response rate achieved with conventional chemotherapy.

Objective: We studied the outcome of patients with initially unresectable liver metastases from colorectal cancer treated with a three-drug chemotherapy regimen followed by liver metastases surgery whenever possible.

Patients and methods: From March 1988 to June 1994, 151 patients with colorectal liver metastases were considered initially unresectable because of large tumor size (> 5 cm), multinodular (> 4) or ill-located metastases. All patients received fully ambulatory chemotherapy with 5-fluorouracil, leucovorin

and oxaliplatin (chronotherapy in 83% of them). They were periodically reassessed for surgery by a joint medico-surgical team.

Results: In 151 patients, the size of liver metastases decreased by > 50% in 89 patients (59%) and median overall survival was 24 months (95% confidence interval (95% CI): 19–28 months), with 28% surviving at five years (20%–35%). Surgery with curative intent was attempted in 77 patients (51%), complete resection of liver metastases was achieved in 58 patients (38%). The median survival of the 77 operated patients was 48 months (25–71), with a five-year survival rate of 50% (38–61).

Conclusion: This new strategy of combining effective chemotherapy with surgery apparently altered the natural history of unresectable colorectal cancer metastases.

Key words: chronotherapy, colorectal cancer, liver metastases, oxaliplatin, surgery, survival

概要

- 背景

- 転移を伴う切除不能大腸がん患者に対する長期生存は、唯一転移部位の完全切除が施行されることで望むことができる
- ただし切除ができるのは、**化学療法がある程度効いた場合のみ**

- 目的

- 化学療法を施行した肝転移を伴う切除不能大腸がんに対して、肝臓切除を施行した予後を明らかにする

- 対象

- 1988年-1994年の151人の肝転移を伴う切除不能大腸がん患者データを収集
- 83%の患者がFOLFOXを施行

背景因子

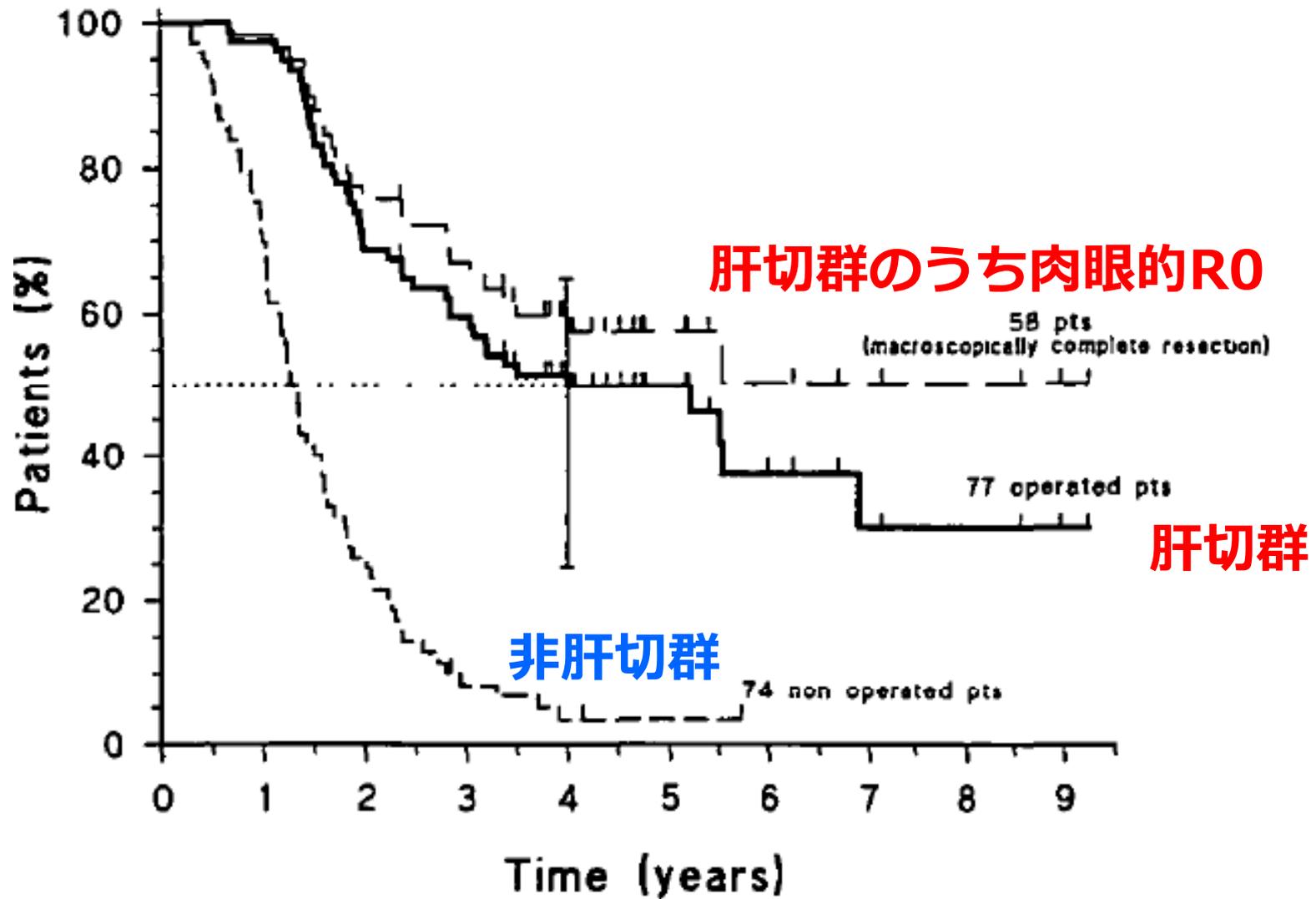
- 肝切群に多いもの
 - 奏効例
 - PR or CR
 - 転移部位の最大腫瘍径が小さい患者
 - 転移臓器個数1個の患者
 - など・・・

最良効果：CR(完全奏効)、PR(部分奏効)、SD(安定)、PD(増悪)、no CX(化学療法なし)

Giacchetti, S., et al. Annals of Oncology 10.6 (1999): 663-669.

	肝切あり N=77	肝切なし N=74
年齢(歳)		
中央値(範囲)	59 (32-79)	58 (27-76)
最良効果(化学療法による腫瘍縮小効果)		
PD or no CX	0	17
SD	16	28
PR or CR	61	29
転移部位の最大腫瘍径		
<=5cm	46	33
> 5cm	24	27
不明	7	14
転移臓器個数		
1	14	19
2-4	38	16
>4	20	25
不明	5	14

全生存期間の結果



この結果から言えることは？

1. 肝切除は有効
2. 肝切除は有効とは言えない
3. 肝切除は無効

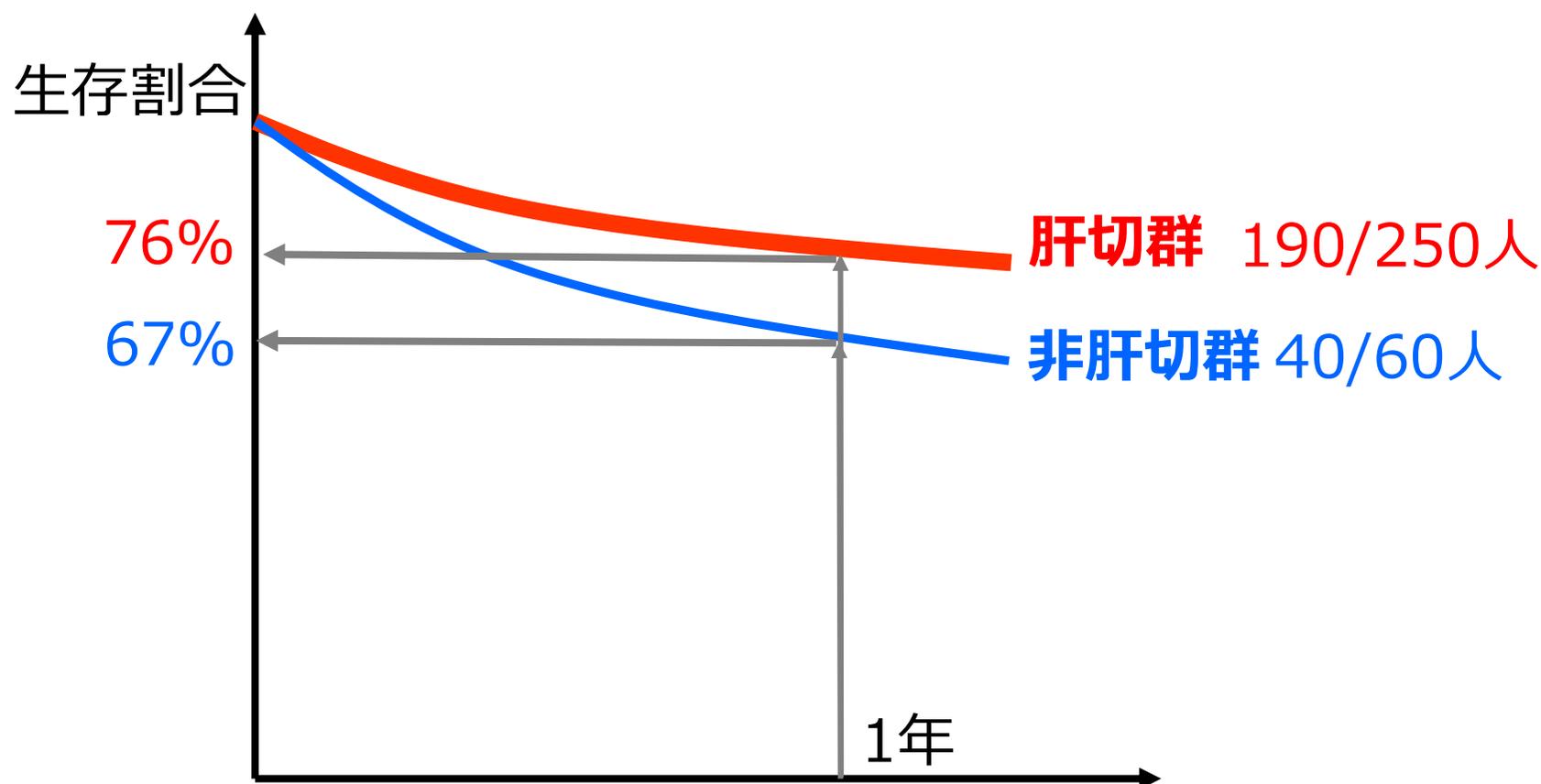
Outline

- 交絡とランダム化
- 交互作用
- サブグループ解析
- 層別解析

交絡とランダム化

国内の学会で見かける発表

- 当院における肝転移を有する切除不能大腸がん患者を**肝切群**(250例)と**非肝切群**(60例)に分けてレトロスペクティブに検討した。
- **肝切群**は**非肝切群**と比較し予後良好であった。

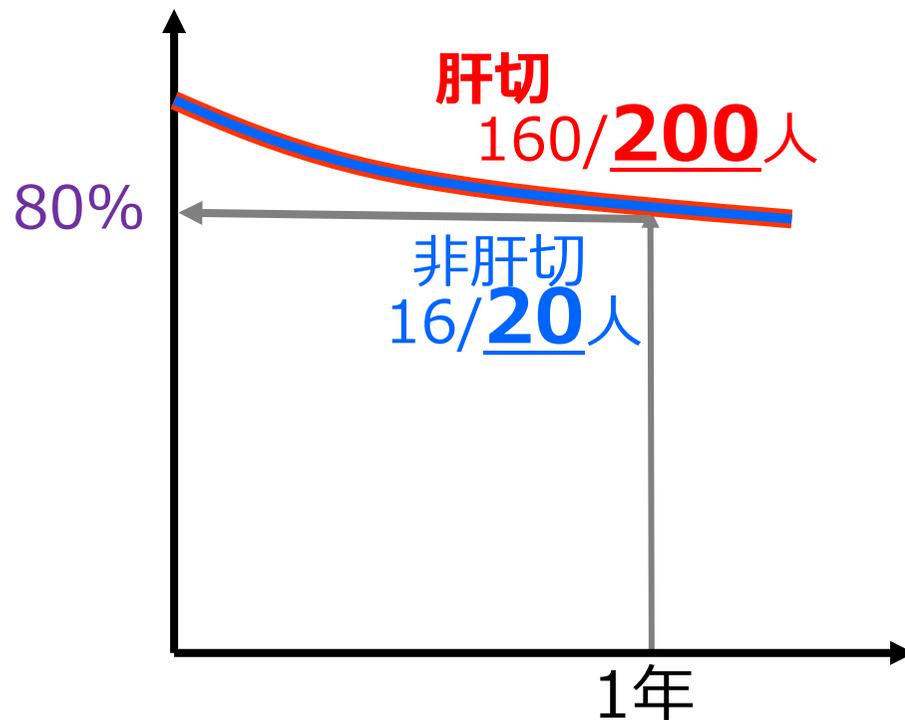


注：仮想例

最良効果で分けた場合の予後

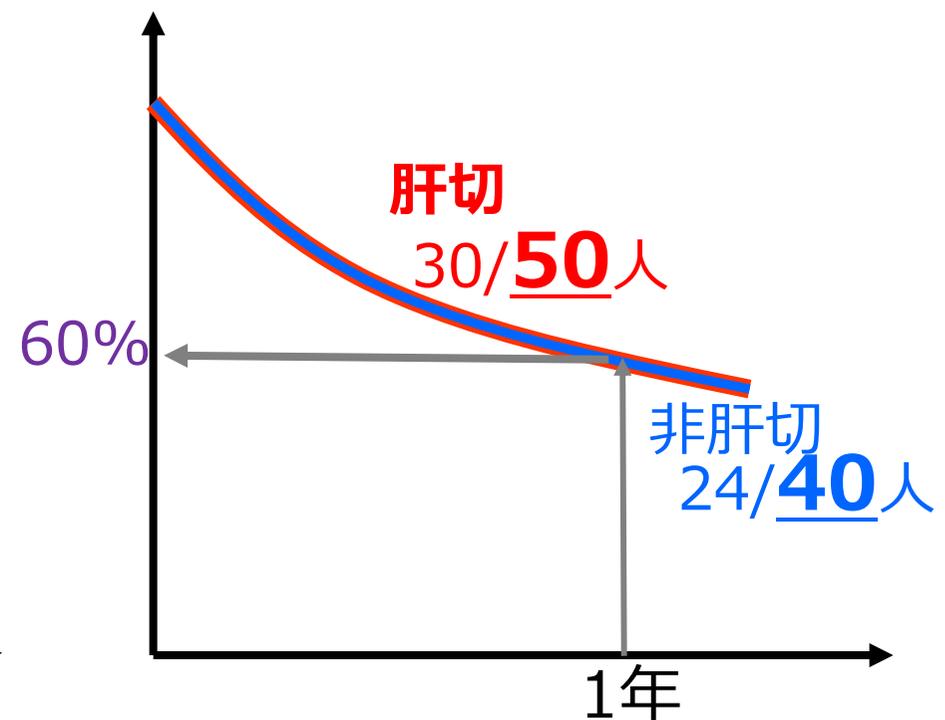
PR or CR
(化学療法に奏効した群)

生存割合



SD or PD
(化学療法に奏効しなかった群)

生存割合



最良効果によらず**肝切**と**非肝切**の予後は変わらない

比較したいのは**肝切**と**非肝切**の違いだから

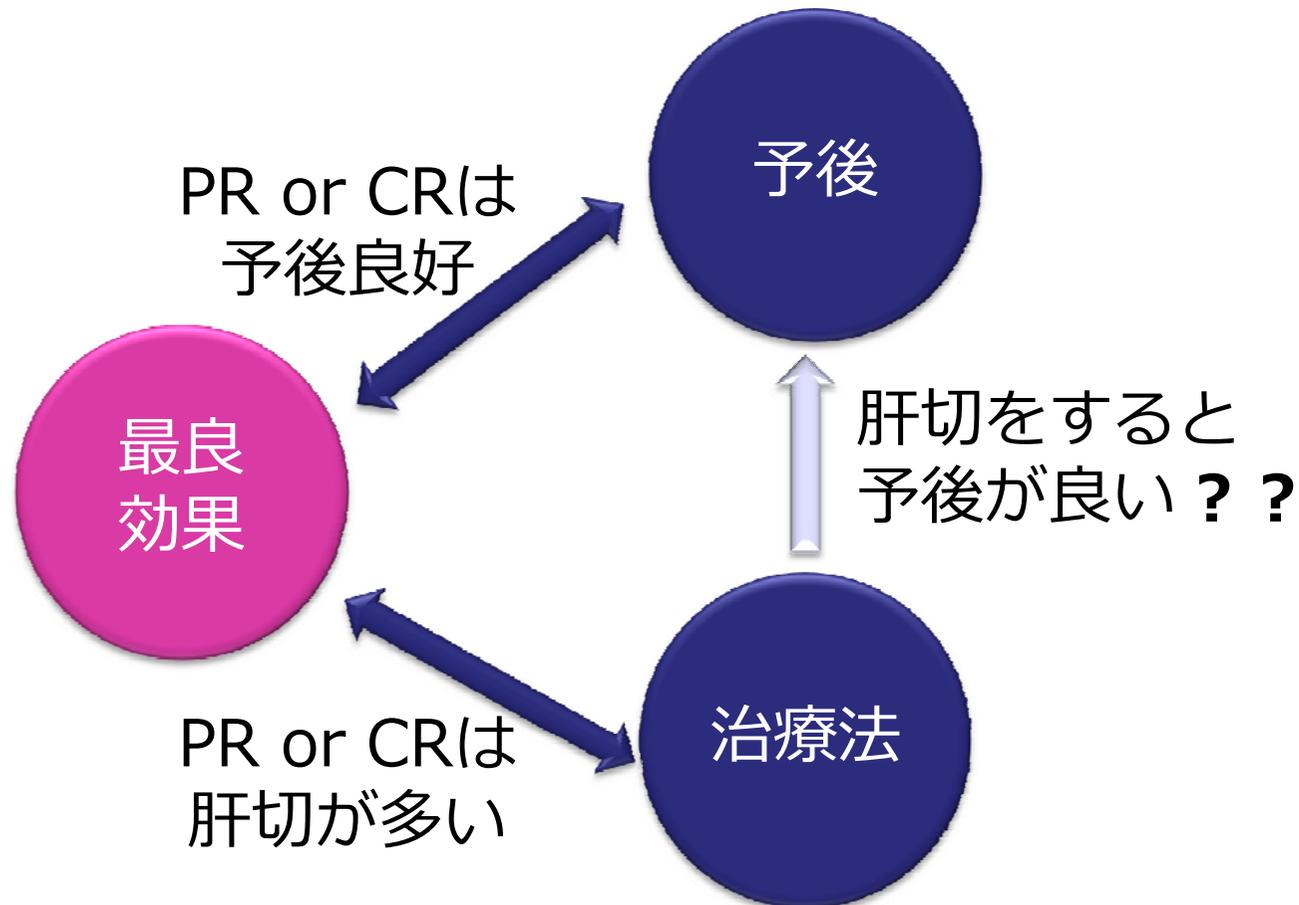
- 治療法以外の予後に影響する要因の条件が同じでなければ、“比較”にならない！！

治療法	PR or CR		SD or PD	合計
肝切	200人 (80%)	>>	50人	250人
非肝切	20人 (33.3%)	<<	40人	60人

- 肝切**は**非肝切**と比べ「**PR or CR**」の割合が高い
- 最良効果によって予後が異なる(PRorCRは予後良)

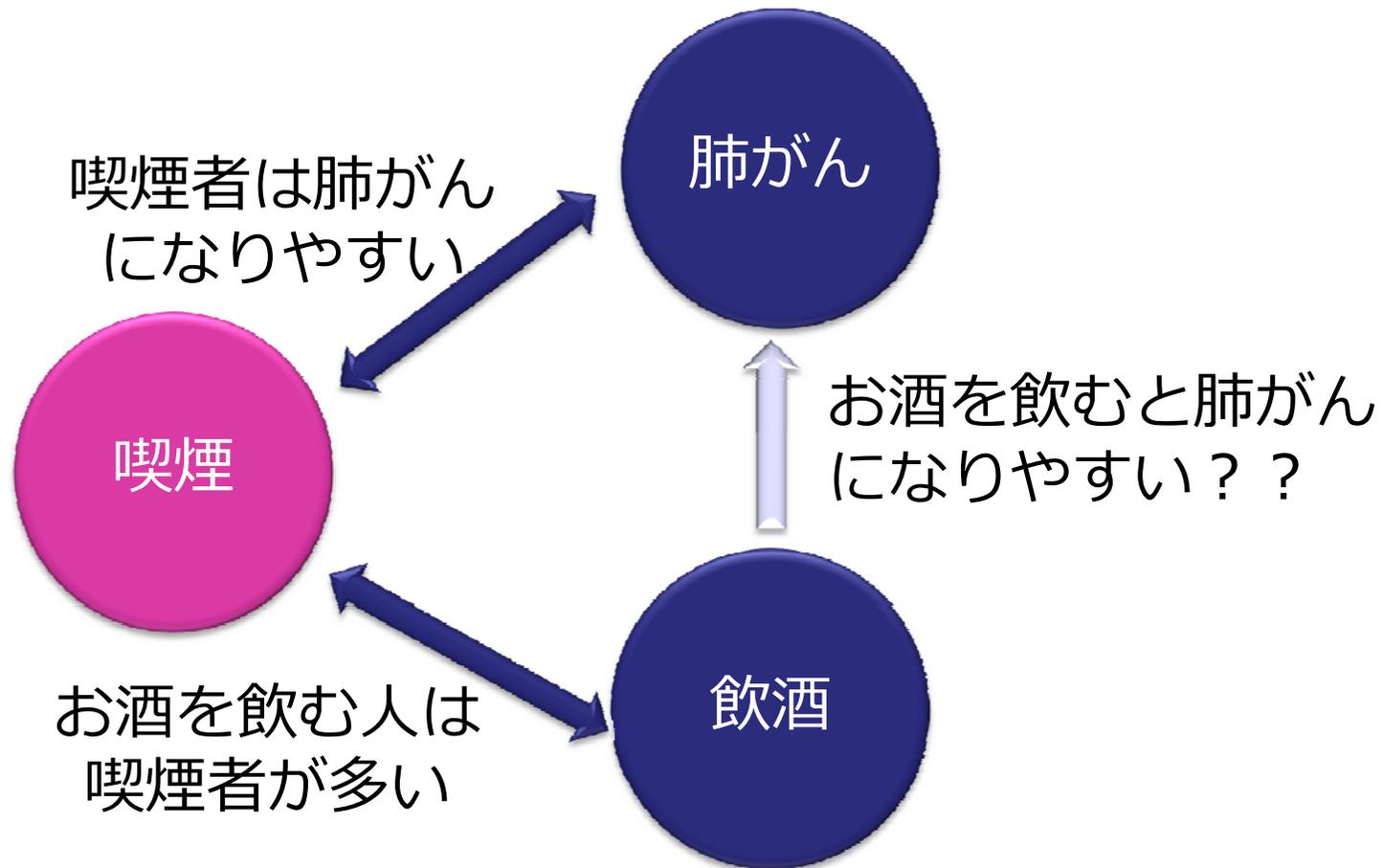
交絡についてのまとめ

- 治療法と予後に関連する第3の因子（最良効果）によって見かけ上の関連が生じてしまう現象のこと
 - 交絡を引き起こす因子(=最良効果)のことを交絡因子という



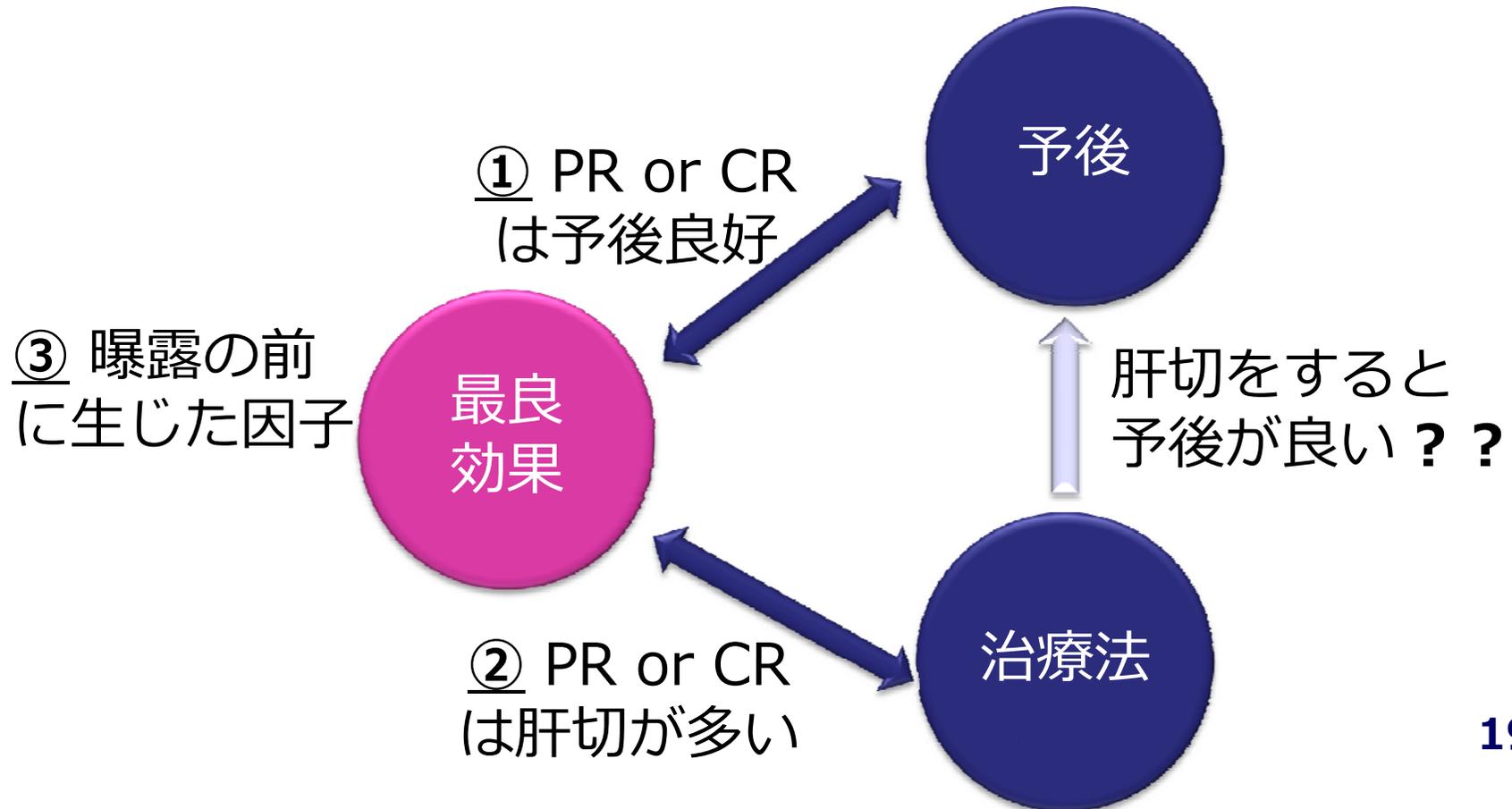
その他の交絡の例

- 飲酒が肺がんを起こすわけではないのに、そう見えてしまう
 - この場合、喫煙が交絡因子



交絡の必要条件

- ① アウトカム(予後)と関係がある
- ② 曝露(治療法)と関係がある
- ③ 中間変数ではない (治療法の前に判明している事象)



交絡がないことを保証するには

- 治療群間で予後に関係する背景因子を揃える
 - 年齢
 - Stage
 - PS
 - その他（未知の因子も含めて）

因子がたくさんある・未知の因子があるために
全てを考慮できない



ランダムに決める

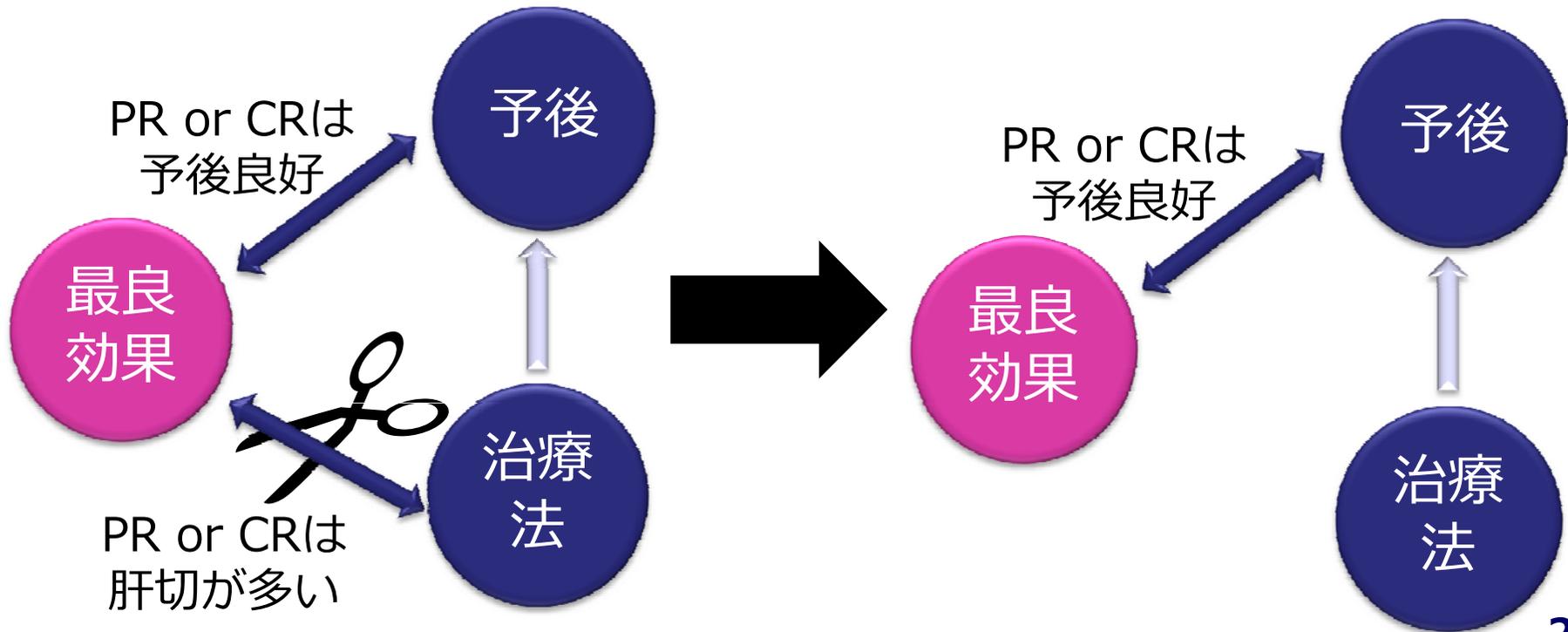
ランダム化 randomization

- 医師あるいは患者の意思によらず、確率に基づいて各治療群に患者を割り付ける
- 予見による患者選択の偏りの防止
 - 状態の良い患者は新薬に割り付けられやすくなる、などを防ぐ
- 比較可能性（内的妥当性）が担保される
 - 治療法以外は等しい集団 → 効果に差があれば治療法の違い



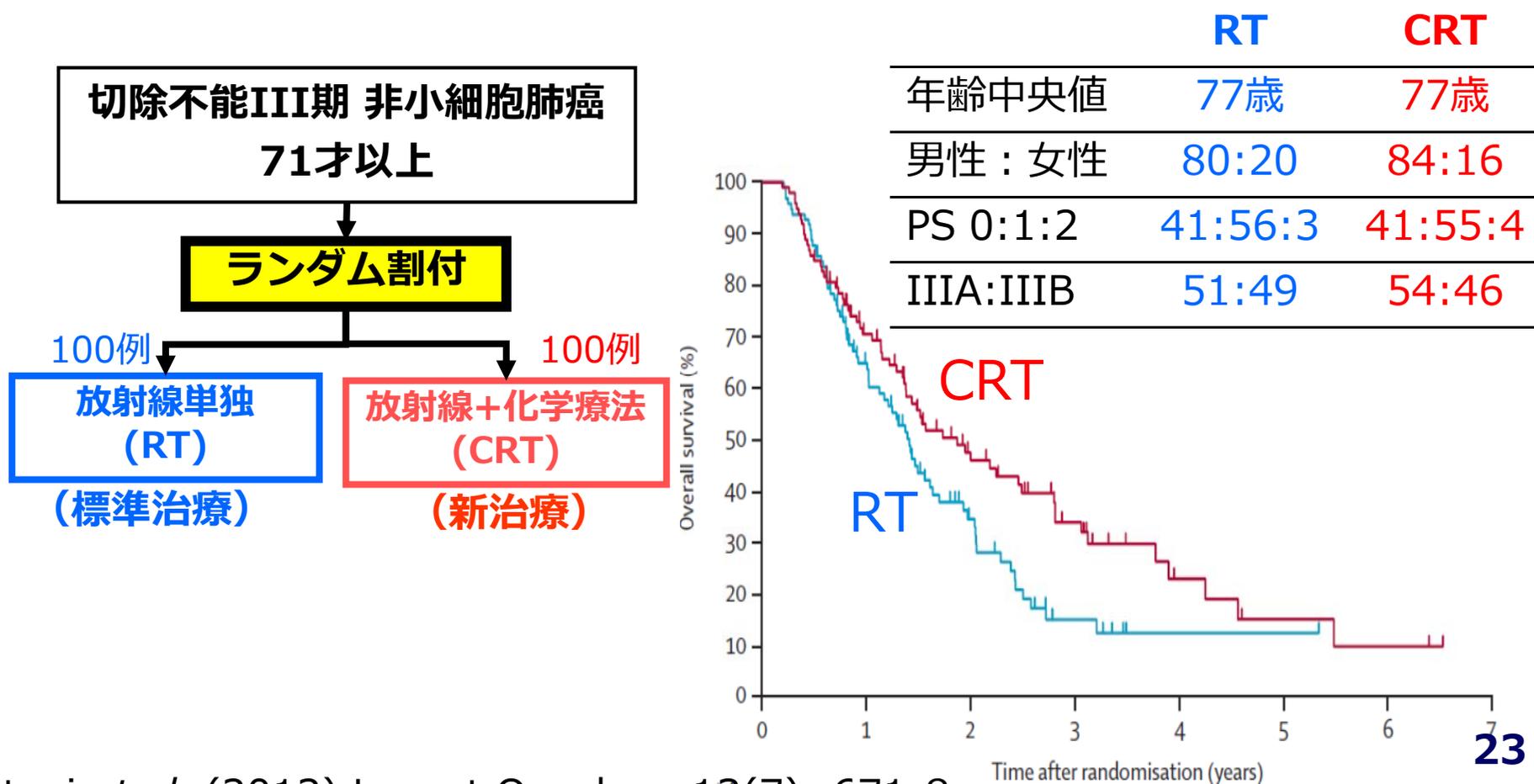
ランダム化の意義

- 治療法と最良効果の関連を除去できる
 - 最良効果による交絡は除去されるので、治療法と予後の関係性を評価できる
 - 注：予後と最良効果の関係性は残ったまま



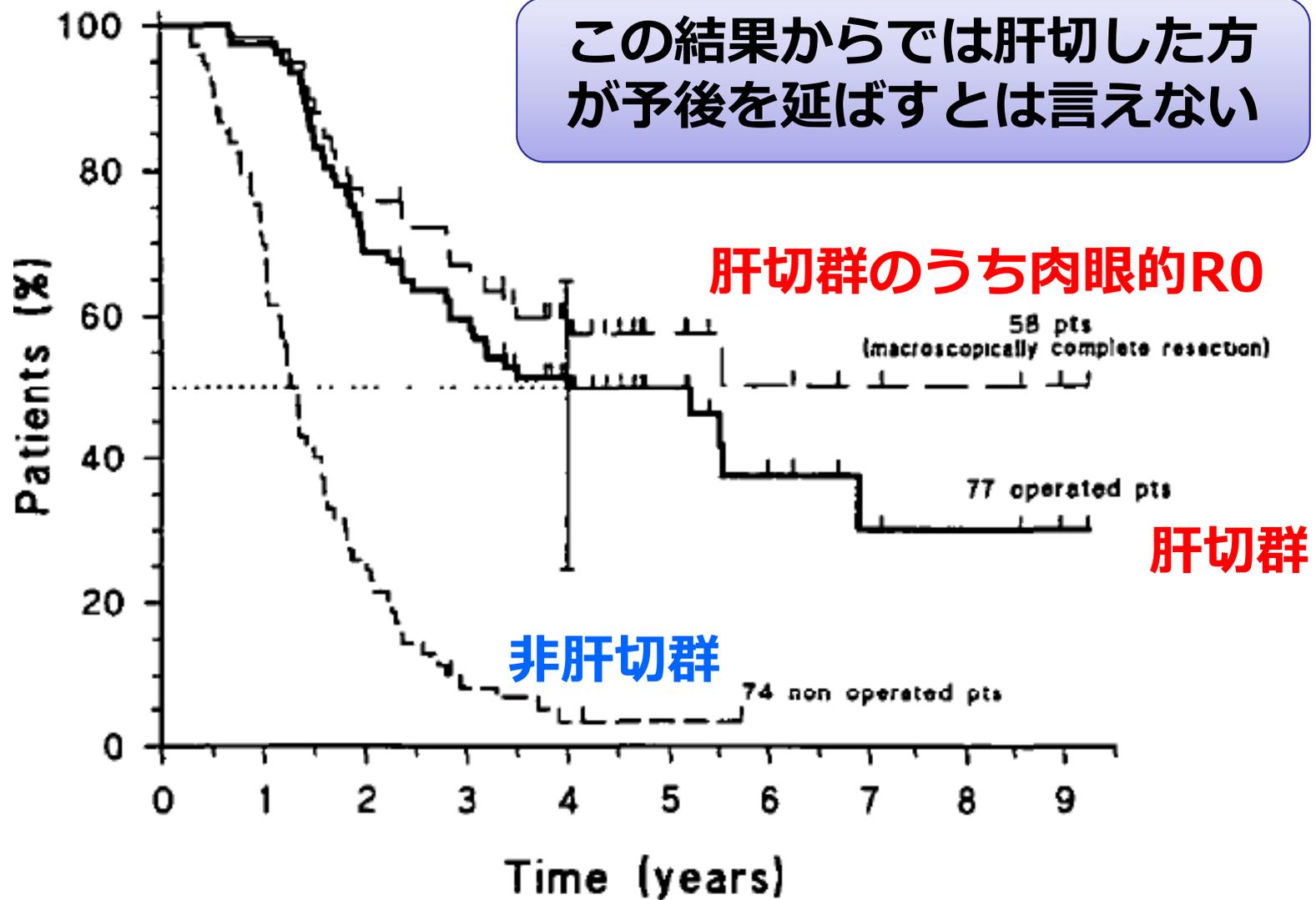
JCOG0301の場合

- **RT**と**CRT**を比較するために**ランダムに割り付けた**
 - 治療法以外の背景因子は平均的に治療群間で同じ
 - 生存曲線の違いは治療法による違いであると期待できる



Atagi et al. (2012) Lancet Oncology 13(7): 671-8.

再掲：肝切有無に関する全生存期間の結果



交絡と誤解しやすい統計的概念

～交互作用～

交互作用とは？

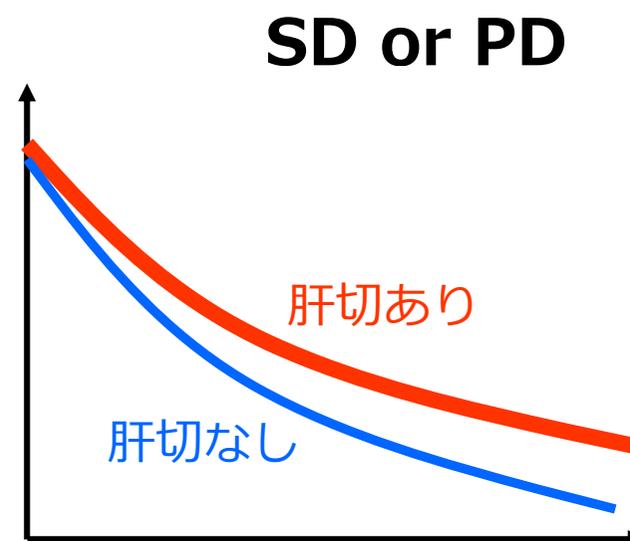
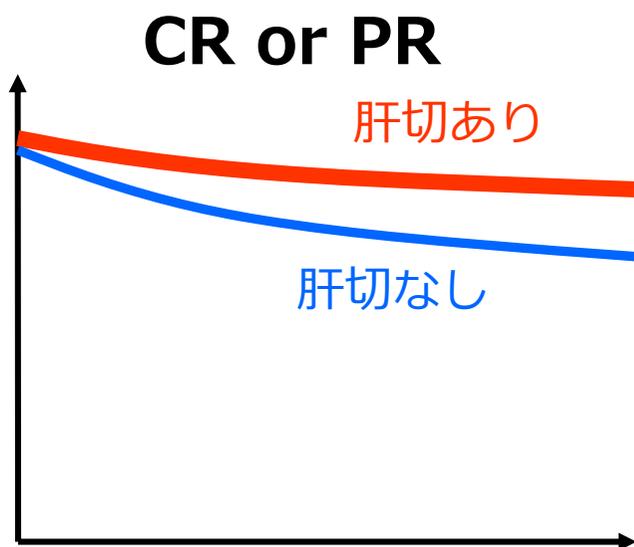
治療効果の大きさが、
サブグループ毎に異なるとき
“交互作用がある”という

交互作用

- ・ 質的交互作用
- ・ 量的交互作用

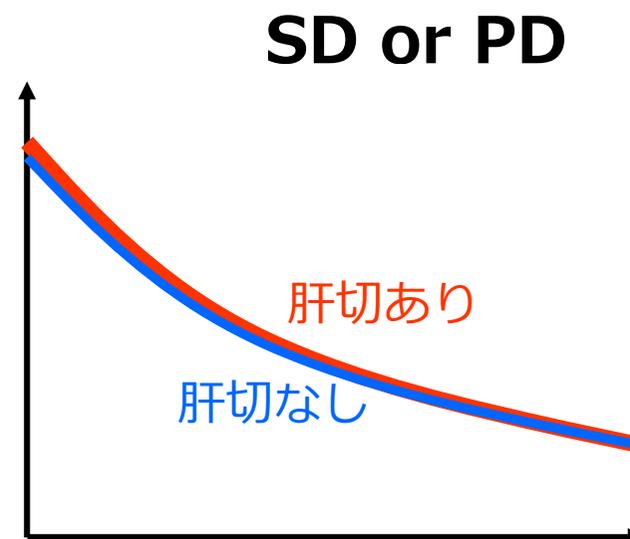
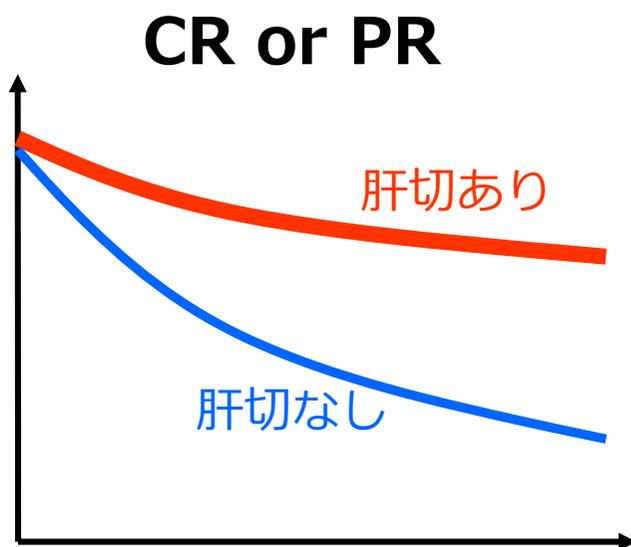
交互作用なし

- 最良効果CR or PR とSD or PDは予後が異なる
- CR or PRとSD or PDのどちらも**肝切あり**は**肝切なし**に対して同程度の治療効果の上乗せがある
 - サブグループで治療効果の違いはない
 - 最良効果は**予後因子**(prognostic factor)



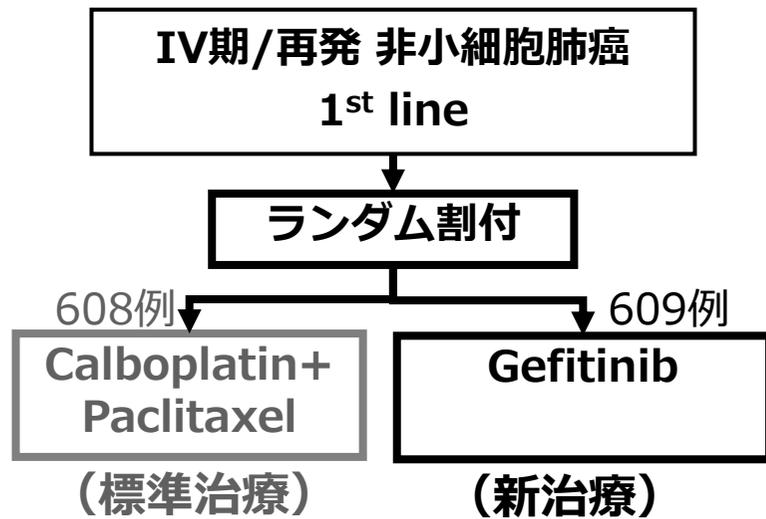
交互作用あり

- CR or PRとSD or PDの**肝切なし**群の予後は同じ
- CR or PRとSD or PDで、治療効果が異なる
 - CR or PRのみ**肝切**の効果がある
 - サブグループで治療効果の違いがある
 - 最良効果は**肝切**の予測因子(predictive factor)

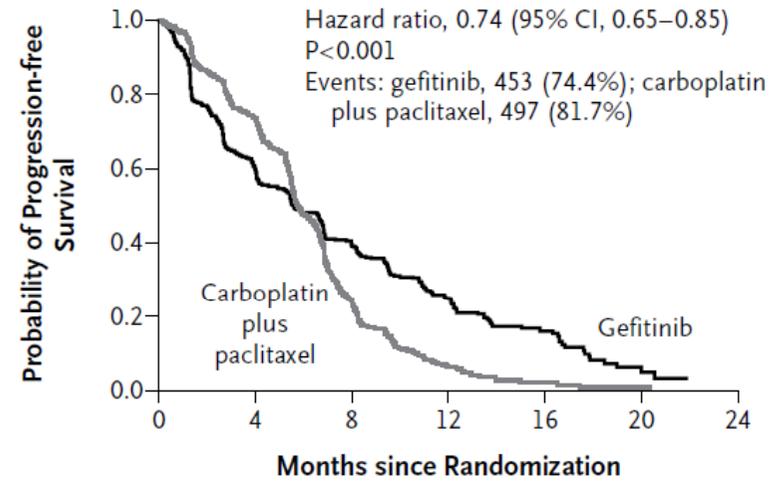


質的交互作用=治療選択すべき予測因子

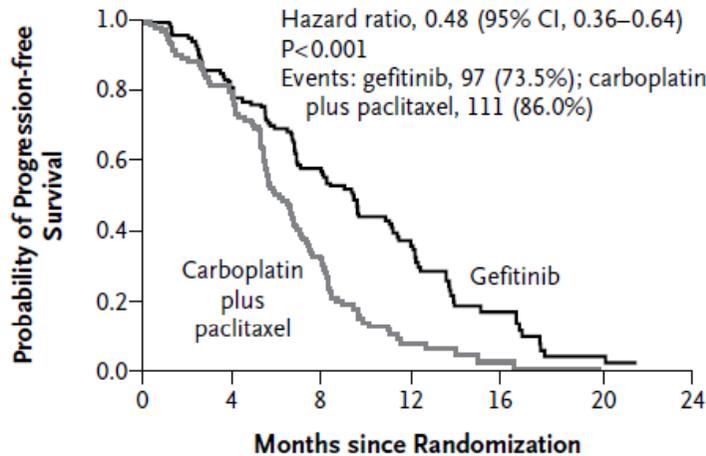
Mok, Tony S., et al. *NEJM* 361.10 (2009): 947-957.



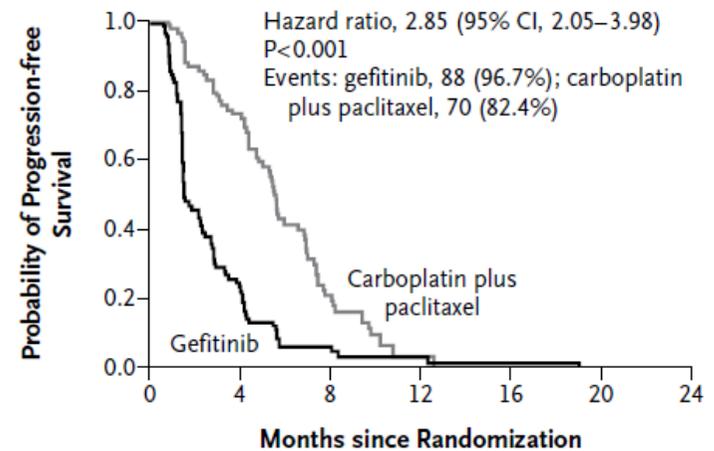
A Overall



B EGFR-Mutation-Positive

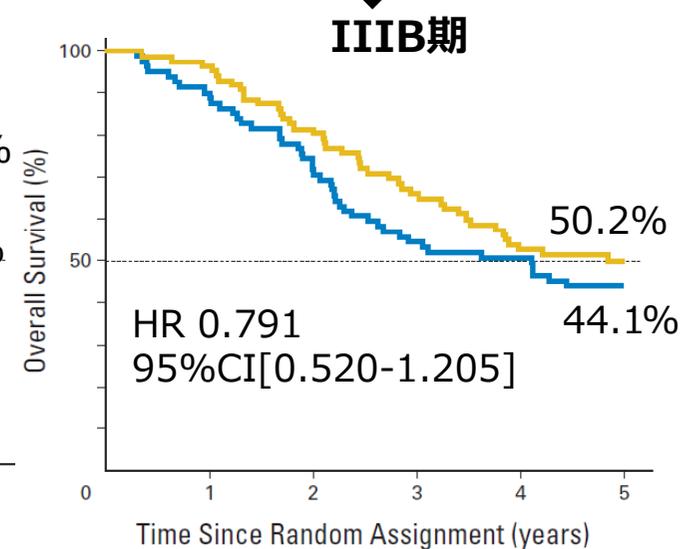
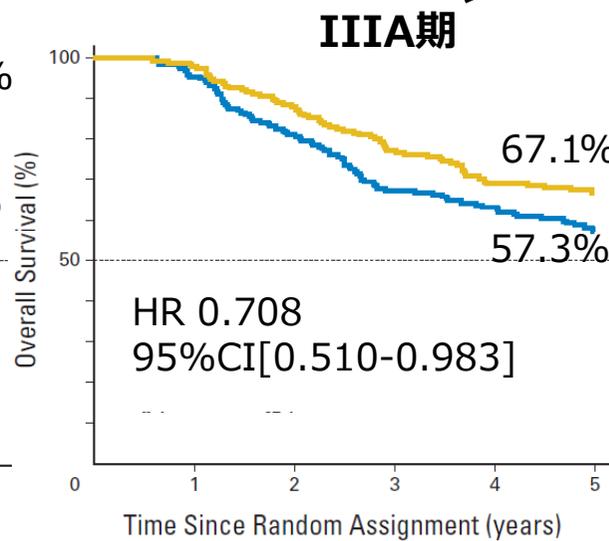
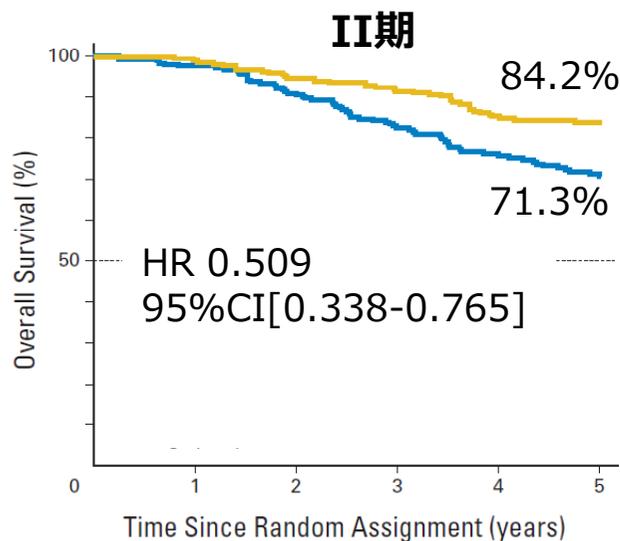
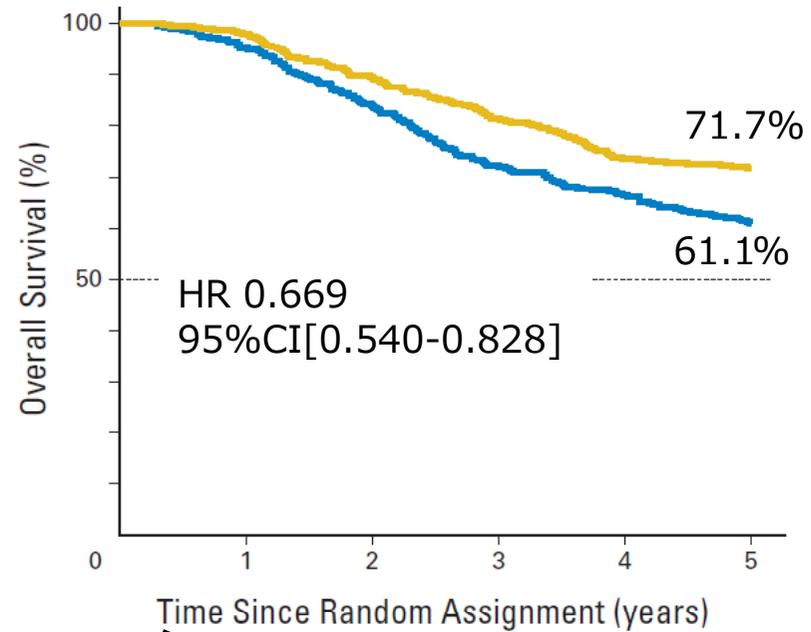
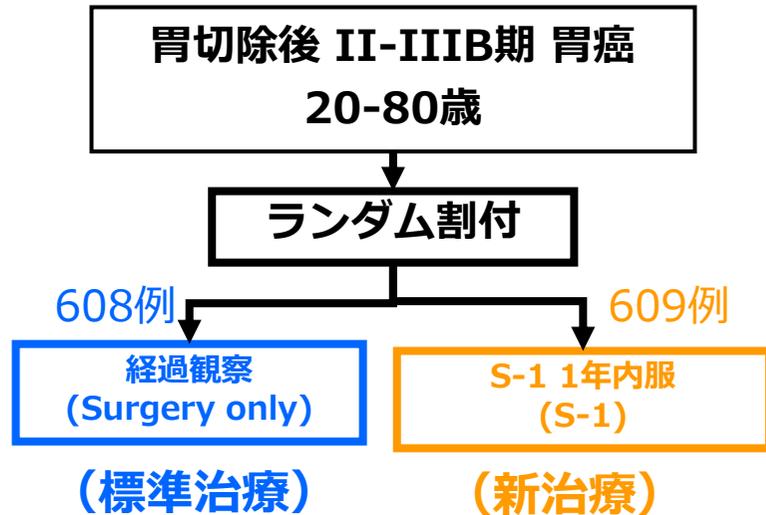


C EGFR-Mutation-Negative

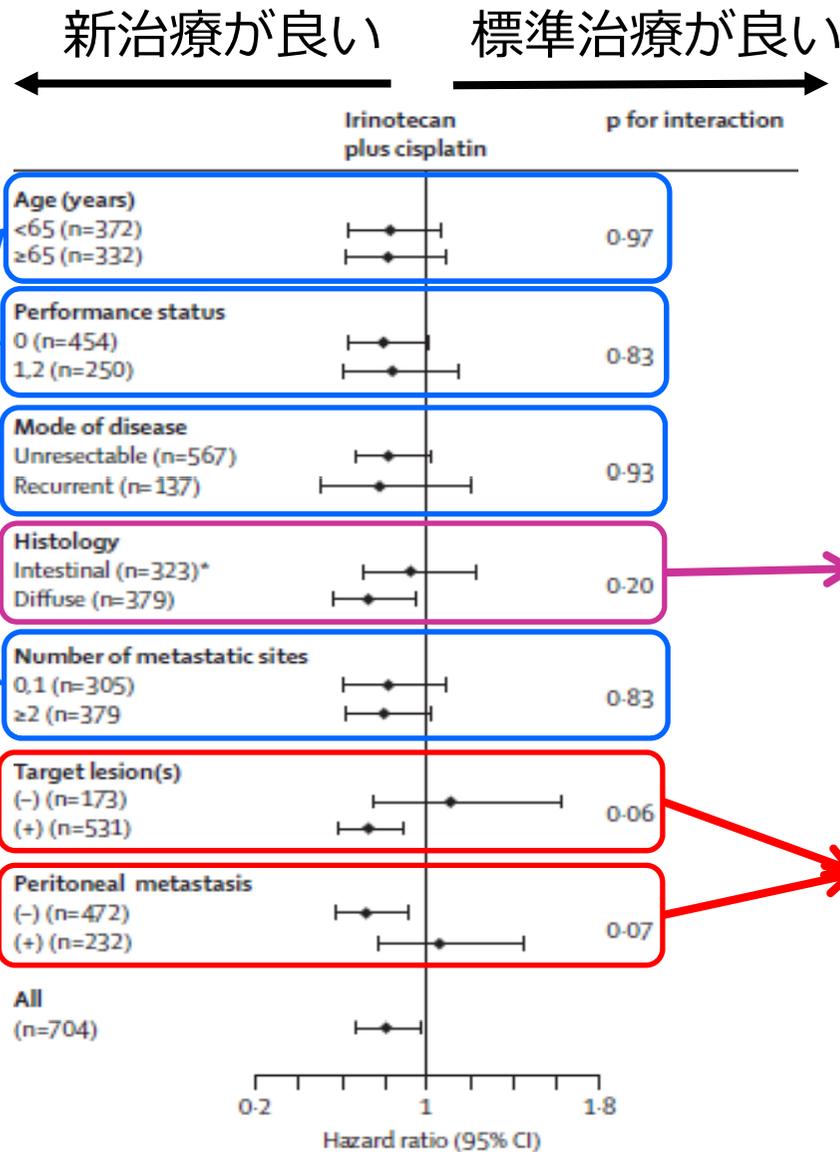


量的交互作用=治療選択を考慮すべき

予測因子



フォレストプロットで交互作用の確認



バイアスと違い、
 ・ 除去できるもの
 ・ 除去すべきもの
 ではないので、
 結果を報告する

交互作用
 はなさそう

量的交互作用が
 ありそう

質的交互作用が
 ありそう

遺伝子-環境相互作用

- UGT1A6遺伝子型に変異を持つ人はアスピリンによる大腸腺腫発症の予防効果がある

Table 1. Risk for colorectal adenoma according to regular aspirin use stratified by UGT1A6 genotype*

Genotype	Nonregular users	Regular users
<u>All UGT1A6 genotypes</u>		
No. of case participants/no. of control participants	373/349	157/183
Age-adjusted OR (95% CI)	1.0 (referent)	0.78 (0.60 to 1.02)
Multivariable OR (95% CI)†	1.0 (referent)	<u>0.76 (0.58 to 0.99)</u>
<u>Wild-type UGT1A6 genotypes</u>		
No. of case participants/no. of control participants	149/154	68/75
Age-adjusted OR (95% CI)	1.0 (referent)	0.90 (0.60 to 1.34)
Multivariable OR (95% CI)†	1.0 (referent)	<u>0.93 (0.60 to 1.44)</u>
<u>Variant UGT1A6 genotypes</u>		
No. of case participants/no. of control participants	224/195	89/108
Age-adjusted OR (95% CI)	1.0 (referent)	0.71 (0.50 to 0.99)
Multivariable OR (95% CI)†	1.0 (referent)	<u>0.66 (0.45 to 0.95)</u>

遺伝子-環境相互作用

- アスピリンによる用量反応関係（量的相互作用）

Table 2. Risk for colorectal adenoma according to dose of aspirin intake, stratified by UGT1A6 genotype*

Genotype	Tablets of aspirin per wk			<i>P</i> _{trend}
	0	0.5–7	>7	
<u>All UGT1A6 genotypes</u>				
No. of case participants/no. of control participants	264/216	207/235	59/81	
Age-adjusted OR (95% CI)	1.0 (referent)	0.72 (0.55 to 0.93)	0.58 (0.40 to 0.86)	.01
Multivariate OR (95% CI)†	1.0 (referent)	0.74 (0.56 to 0.97)	0.51 (0.34 to 0.77)	.004
<u>Wild-type UGT1A6 genotypes</u>				
No. of case participants/no. of control participants	99/96	92/102	26/31	
Age-adjusted OR (95% CI)	1.0 (referent)	0.83 (0.55 to 1.24)	0.76 (0.42 to 1.38)	.44
Multivariate OR (95% CI)†	1.0 (referent)	0.88 (0.57 to 1.37)	0.74 (0.39 to 1.41)	.50
<u>Variant UGT1A6 genotypes</u>				
No. of case participants/no. of control participants	165/120	115/133	33/50	
Age-adjusted OR (95% CI)	1.0 (referent)	0.64 (0.46 to 0.92)	0.48 (0.29 to 0.80)	.003
Multivariate OR (95% CI)†	1.0 (referent)	0.63 (0.43 to 0.91)	0.41 (0.24 to 0.71)	.001

交絡と交互作用のまとめ【臨床の場合】

• 交絡（Confounding）

- 予後に関係する背景因子が治療群間で偏ることで、**見かけ上の関連が生じてしまう or 見かけ上関連が消えてしまう現象のこと**
- バイアスの一種なので、可能な限り排除することが必要
 - ランダム化によって排除できる。不可能な場合は多変量解析などを実施

• 交互作用（Interaction）

- **治療効果の大きさがサブグループ毎に違い**があること
 - EGFR変異有無によりゲフィチニブの効果異なる
- バイアスではないので、存在するかどうかを確認する
 - ランダム化によって排除することはできない
- 結果を適用できる対象者（範囲）の限定
 - 適格規準の検討や生物学的考察などに使う

交絡と交互作用のまとめ【疫学の場合】

• 交絡（Confounding）

- アウトカムに関係する背景因子が曝露群間で偏ることで、**見かけ上の関連が生じてしまう or 見かけ上関連が消えてしまう現象のこと**
- バイアスの一種なので、可能な限り排除することが必要
 - ランダム化によって排除できる。不可能な場合は多変量解析などを実施

• 交互作用（Interaction）

- **曝露の効果の大きさがサブグループ毎に違い**があること
 - UGT1A6遺伝子型によりアスピリンの効果が異なる
- バイアスではないので、存在するかどうかを確認する
 - ランダム化によって排除することはできない
- 結果を適用できる対象者（範囲）の限定
 - 予防介入の対象者の検討や生物学的考察などに使う

ランダム化なしで交絡を 可能な限り除去するには？

～比較可能性を高める方法～

【再掲】 背景因子

- 肝切群に多いもの
 - 奏効例
 - 転移部位の最大腫瘍径が小さい患者
 - 転移臓器個数1個の患者
 - など・・・

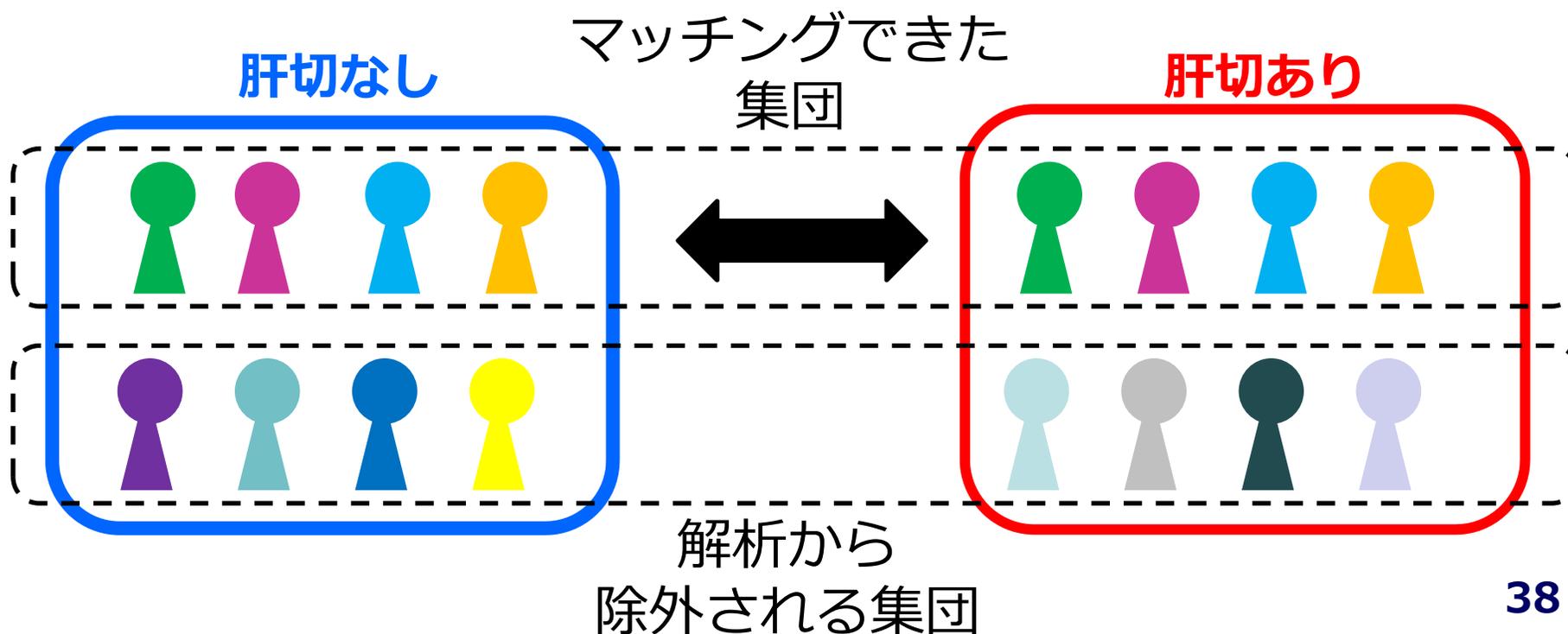
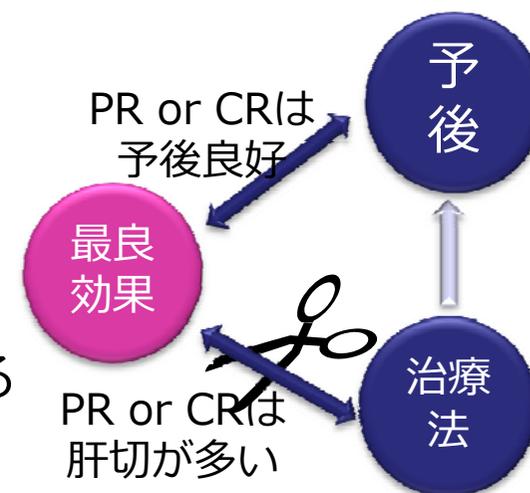
	肝切あり N=77	肝切なし N=74
年齢(歳)		
中央値(範囲)	59 (32-79)	58 (27-76)
最良効果		
PD or no CX	0	17
SD	16	28
PR or CR	61	29
転移部位の 最大腫瘍径		
<=5cm	46	33
> 5cm	24	27
不明	7	14
転移臓器個数		
1	14	19
2-4	38	16
>4	20	25
不明	5	14

Giacchetti, S., et al. Annals of Oncology 10.6 (1999): 663-669.

デザイン段階で対処する方法

- マッチング

- アウトカムに強い影響を与えると考えられる背景因子が一致する患者どうしを合わせる



解析段階で比較可能性を高めるための方法

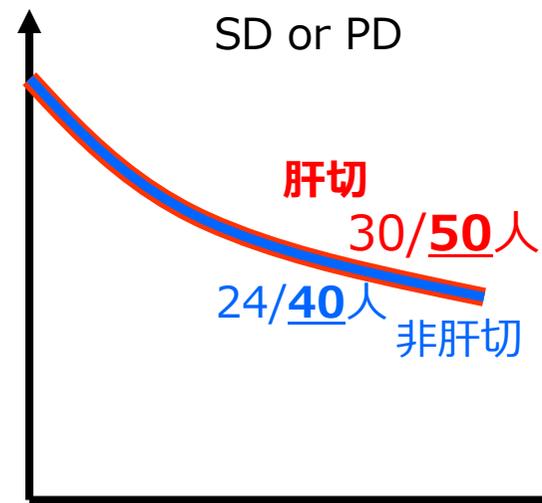
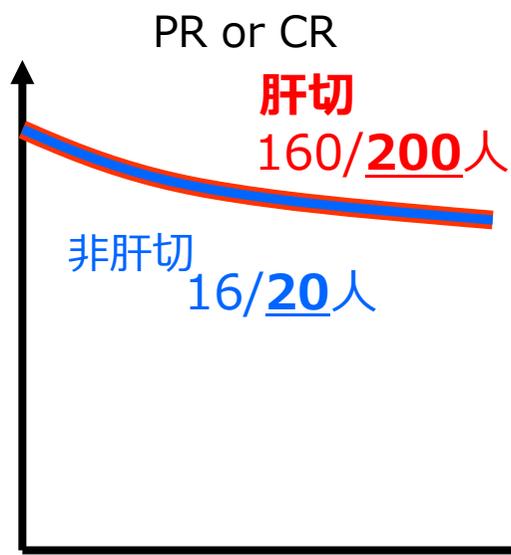
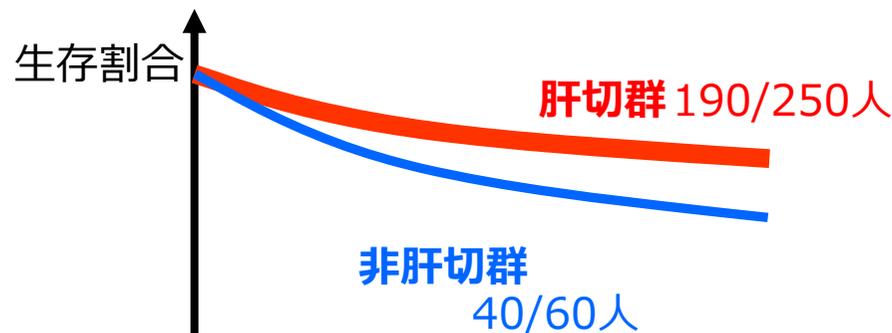
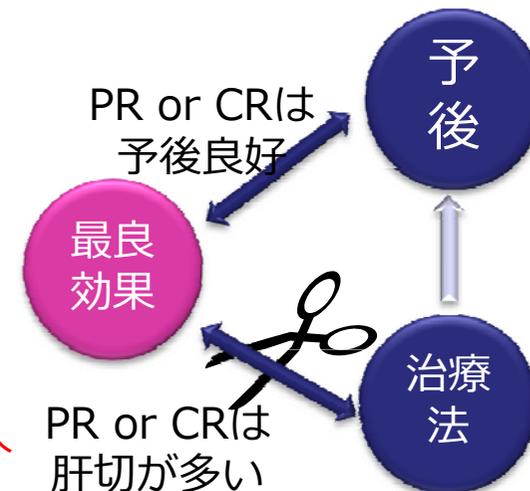
今回のtopic !

- **サブグループ解析(subgroup analysis)**
 - サブセット解析、部分集団解析
 - サブグループ毎に治療効果を見る
- **層別解析(stratified analysis)**
 - サブグループ毎の結果を統合(重み付き平均)して、1つのp値、1つの治療効果を求める
- **モデルを用いた多変量解析(multivariate analysis)**
 - Cox回帰やロジスティック回帰を行う
- **傾向スコアを用いた解析(propensity score analysis)**
 - ある患者がある治療に割り付けられる確率を求め調整する

注：仮想例

サブグループ解析

最良効果別に治療と予後の関係を検討する

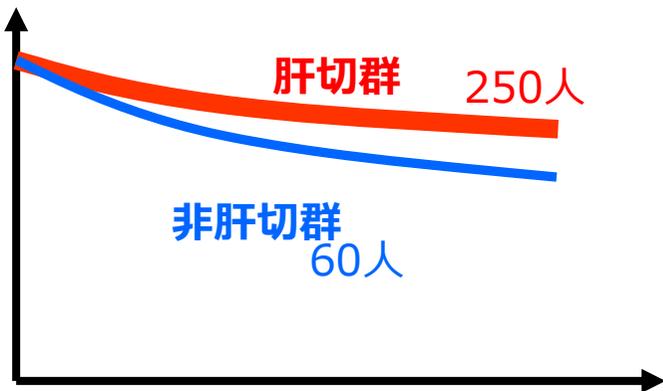


40

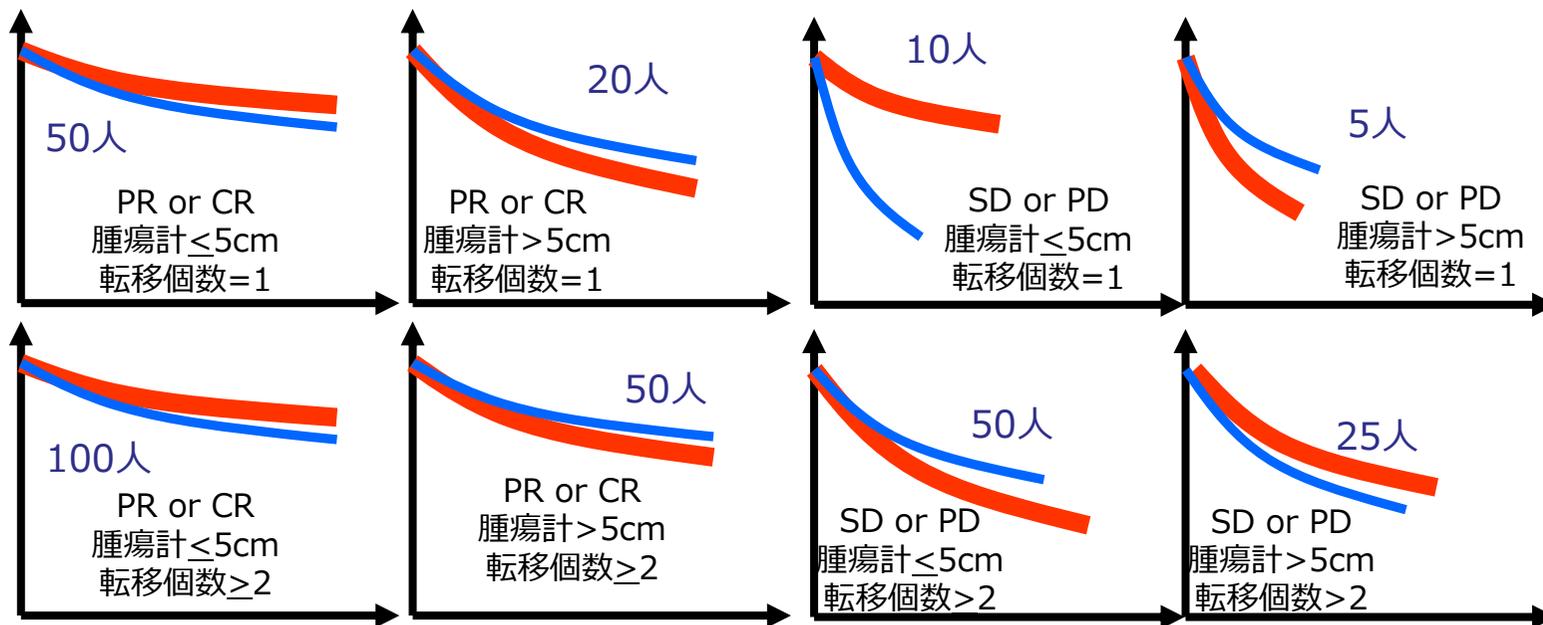
注：仮想例

サブグループ解析の欠点

最良効果、腫瘍径、転移臓器個数別に
治療と予後の関係を検討する



交絡因子が複数の時、各サブグループの
サンプルサイズが小さくなり過ぎる！



サブグループ解析の利点・欠点

- 利点

- 簡単でわかりやすい
 - 各サブグループで治療効果の検討をすればよいだけ
 - 統計的仮定が少ない

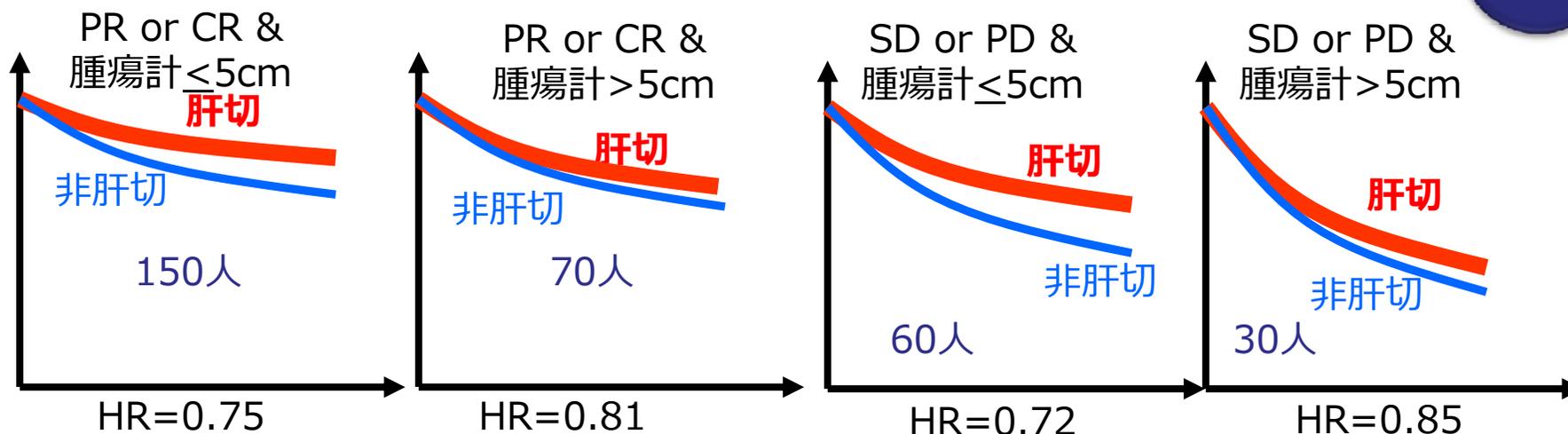
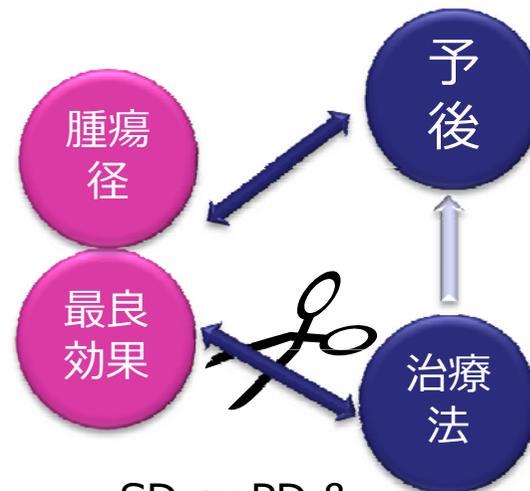
- 欠点

- 集団全体での治療効果がわからない
- サブグループが多すぎると、各サブグループのサンプルサイズが小さくなりすぎる
 - 交絡因子が5個あれば、少なくとも $2^5=32$ 個のサブグループ
 - 交絡因子が連続量の場合は、カテゴリー化してからでなければサブグループ解析できない
- 交絡因子自体(最良効果がSDorPDに対するPRorCR)の効果の大きさがわからない

注：仮想例

層別解析(stratified analysis)

最良効果、腫瘍径のサブグループ別に
求めた治療効果を統合



個々のサブグループの治療効果(HR)を
重み付き平均して1つに統合

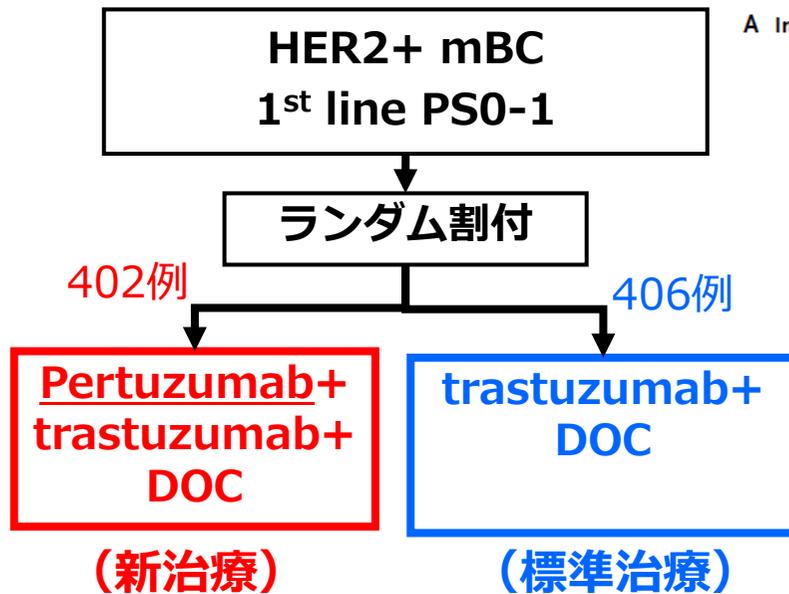
$$\text{全体の効果(HR)} = 0.75 \times \frac{150}{310} + 0.81 \times \frac{70}{310} + 0.72 \times \frac{60}{310} + 0.85 \times \frac{30}{310} = 0.77$$

重みには、サンプルサイズや推定値のバラツキなどを用いる

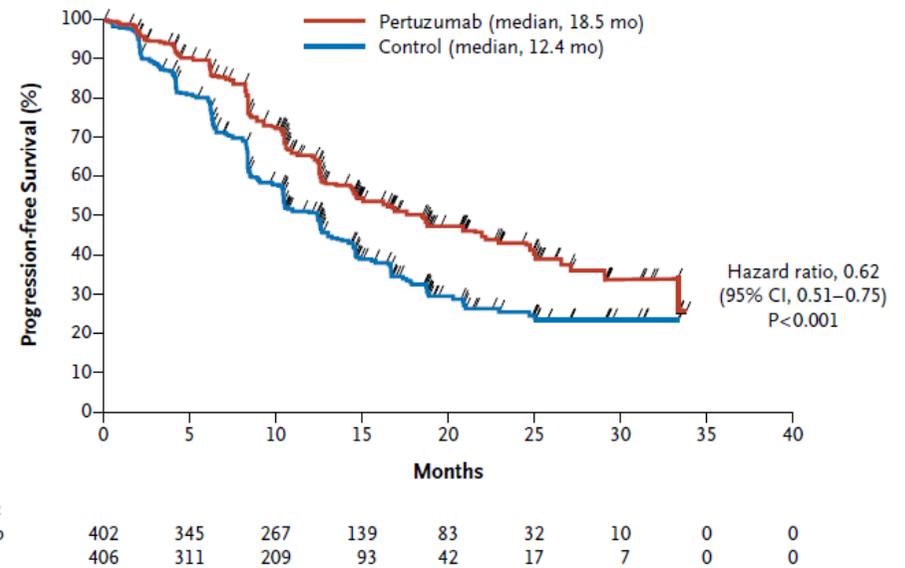
層別解析の論文での使用例

ORIGINAL ARTICLE

Pertuzumab plus Trastuzumab plus Docetaxel for Metastatic Breast Cancer



A Independently Assessed Progression-free Survival



層別解析に関する統計記載

tention-to-treat population (all patients who underwent randomization). The log-rank test, with stratification according to prior treatment status and region, was used to compare independently assessed progression-free survival between the two groups. The Kaplan–Meier approach was used to estimate the median independently assessed progression-free survival in each group. A Cox proportional-hazards model, with stratification according to prior treatment status and region, was used to estimate the hazard ratio and 95% confidence intervals.

free survival. The objective response rate was analyzed on data from patients who had independently assessed measurable disease at baseline and was compared between the groups with the use of the Mantel–Haenszel test, with stratification according to prior treatment status and region. Adverse events were evaluated descriptively

前治療歴(無/有)、
地域(北米/欧/南/亜)



8つのサブグループ
解析の結果を重み付き
平均して1つに統合した
層別解析

PFSについて

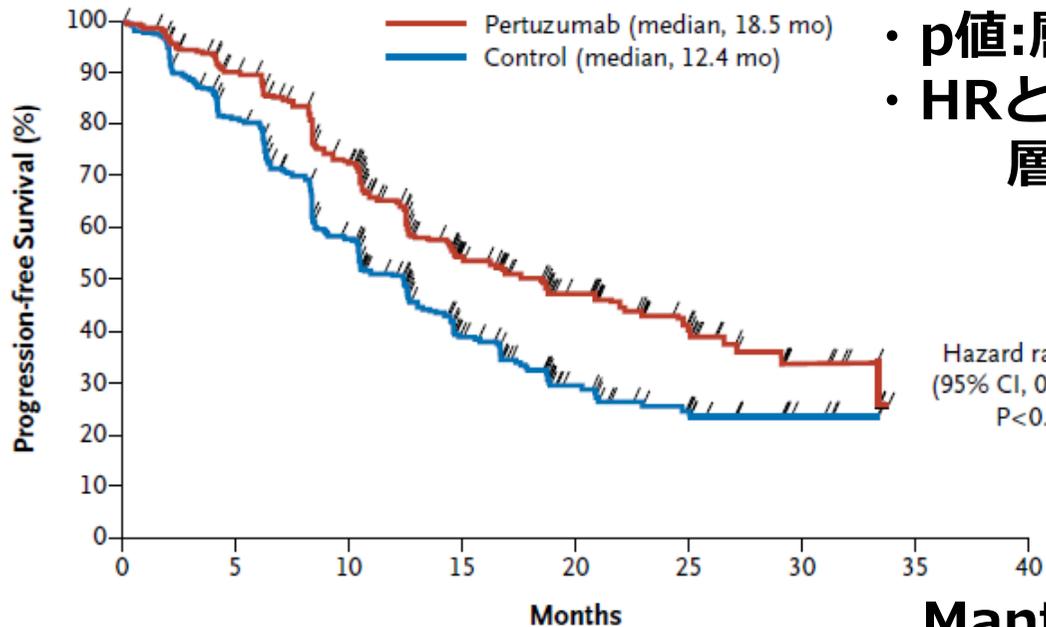
- ・ p値:層別log-rank検定
- ・ HRと95%CI:
層別Cox回帰

奏効割合について

- ・ Mantel-Haenszel検定

層別解析の結果確認

A Independently Assessed Progression-free Survival



- p値:層別log-rank検定
- HRと95%CI:
層別Cox回帰

Hazard ratio, 0.62
(95% CI, 0.51–0.75)
P<0.001

Mantel-Haenszel検定

The difference in response rates was 10.8 percentage points (95% CI, 4.2 To 17.5; **P = 0.001**)

46

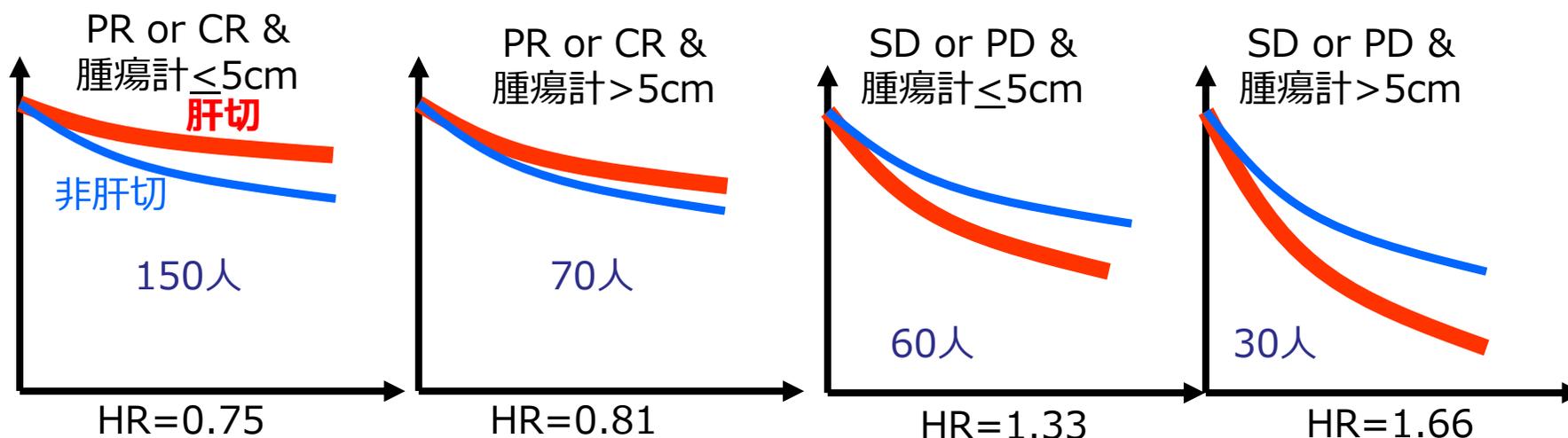
Table 2. Overall Response, as Assessed at an Independent Review Facility.*

Response	Placebo plus Trastuzumab plus Docetaxel (N=336)	Pertuzumab plus Trastuzumab plus Docetaxel (N=343)
	number (percent)	
Objective response	233 (69.3)	275 (80.2)
Complete response	14 (4.2)	19 (5.5)
Partial response	219 (65.2)	256 (74.6)
Stable disease	70 (20.8)	50 (14.6)
Progressive disease	28 (8.3)	13 (3.8)
Not assessable	2 (0.6)	2 (0.6)
No assessment performed	3 (0.9)	3 (0.9)

注：仮想例

層別解析の前提

各サブグループで治療効果が同程度だとみなせること
(大きな交互作用がない)



$$\text{全体の効果(HR)} = 0.75 \times \frac{150}{310} + 0.81 \times \frac{70}{310} + 1.33 \times \frac{60}{310} + 1.66 \times \frac{30}{310} = 0.96(??)$$

層別解析の利点と欠点

- 利点
 - 集団全体の治療効果を求めることができる
 - (モデルを用いた解析と比較して) 仮定が少ない
- 欠点
 - サブグループが多すぎると、各サブグループのサンプルサイズが小さくなりすぎる
 - 交絡因子が5個あれば、少なくとも $2^5=32$ 個のサブグループ
 - 交絡因子が連続量の場合は、カテゴリー化してからでなければサブグループ解析できない
 - 交絡因子自体(最良効果がSDorPDに対するPRorCR)の効果の大きさがわからない

交絡を除去する解析法のまとめ

今回のtopic !

- **サブグループ解析(subgroup analysis)**

- サブセット解析、部分集団解析
- サブグループ毎に治療効果を見る

簡単だが、同時に扱える因子の数の制限がある。連続量の因子が扱えない

- **層別解析(stratified analysis)**

- サブグループ毎の結果を統合(重み付き平均)して、1つのp値、1つの治療効果を求める

- **モデルを用いた多変量解析(multivariate analysis)**

- Cox回帰やロジスティック回帰を行う

- **傾向スコアを用いた解析(propensity score analysis)**

- ある患者がある治療に割り付けられる確率を求め調整する

まとめ

- 予後に関係する背景因子が治療群間で偏ることで、見かけ上の関連が生じてしまう or 見かけ上関連が消えてしまう現象のことを交絡という
 - 交絡はランダム化によって排除できる
- サブグループ間で治療効果が異なることを交互作用があるという
 - 交互作用はランダム化によって排除できない
- ランダム化なしで比較可能性を高める方法(今日紹介したもの)
 - マッチング
 - サブグループ解析
 - 層別解析