

研究コンセプトの作り方 -統計家の視点から-

国立がん研究センター
山本精一郎

<本講義は、第15回CRCと臨床試験のあり方を考える会議 2015 in KOBEをもとに作成しました> **1**

本講義で使用しますので、スマートフォンやタブレットをお持ちの方はICRwebアプリをダウンロードし立ち上げておいてください。

App Store > ICRwebで検索

アンドロイド版も
できました！



では始めます。

今日はがんの話ではありません。

統計の講義を聞く際に気を付けること

- 例が自分の分野のものでなくても、自分の知っている例に置きなおして考える
- わからない部分があってもめげずに、全体を理解するようにする
 - 細かいことは気にしない
 - 聞いているうちに慣れる

臨床研究における統計家の役割

- 臨床的な命題を、統計的な枠組みに乗せることによって、抽象的な課題を具体化し、検証可能なものとする
- サンプルサイズ計算が仕事ではない
- 本講義では、専門ではない臨床的課題をどのように統計的な枠組みに乗せるかということを紹介することによって、新しい研究を考える際にどのようにコンセプトを練り上げていくかの参考としてほしい

臨床試験のフレームワーク

JCOGのプロトコールコンセプトを
もとに紹介

<http://www.jcog.jp/doctor/tool/concept.html>

臨床試験のフレームワーク

1. クリニカルクエスチョン/臨床的意義

WHY

2. 対象

WHO

3. 標準治療

WHAT

4. 新治療

5. 研究デザイン

HOW

エンドポイント

判断規準

「JCOGコンセプト作成の手引き」より

7

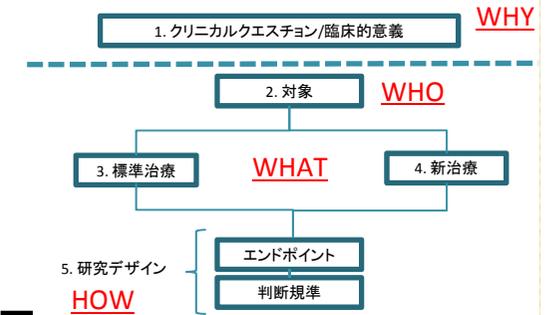
クリニカルクエスチョンと臨床的意義

- **クリニカルクエスチョンの明示ー試験を計画した動機**
 - 対象患者について、現状で何が困っているか？
 - 標準治療の効果がまだ十分でない
 - 標準治療の〇〇毒性が強い、特有の毒性がある
- **結果の意義(Value)の明示ー目指すゴールの具体化**
 - この試験の結果が、**医療現場をどう変えるのか**
 - どんなよいことが見込まれるのか=Value
 - Positive result ならどうか？ Negative result ならどうか？
 - **検証的な結論を得たいのか**
 - =標準治療またはその選択肢を増やしたいのか
 - **探索的な結論を得たいのか**
 - =結果がpositiveであったなら、検証的な研究へ進み、その結果、医療現場がどう変わることを期待しているか

「JCOGコンセプト作成の手引き」より

対象

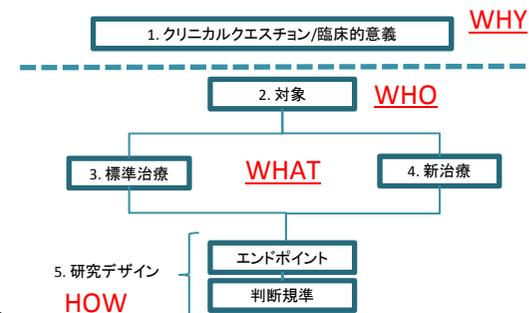
- 試験の対象患者集団の説明
- なぜ、この対象を選んだか？
- なぜ、適格規準をそうしたか？



「JCOGコンセプト作成の手引き」より

標準治療

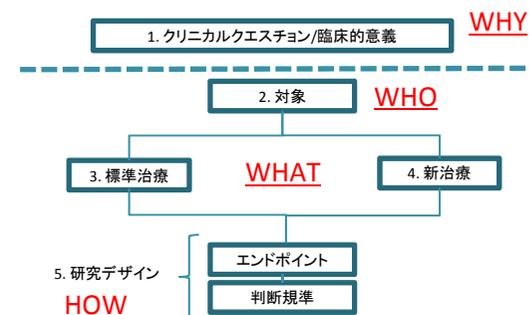
- 対象に対する標準治療の説明
- 標準治療は確定しているか
- 当該疾患の**治療ガイドライン**ではどのような記載がなされているか 推奨レベルや日常診療での普及について
 - 日本での状況
 - 海外のガイドラインとの相違点
 - コミュニティスタンダード



「JCOGコンセプト作成の手引き」より

試験治療

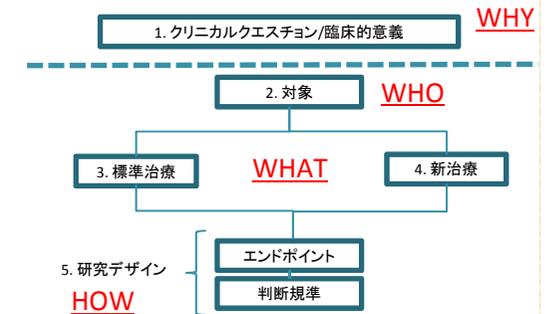
- 試験治療の意義
- Most promisingかどうか、が肝
- 標準治療と試験治療を比較した際の
 1. 「毒性のバランス」または
 2. 「メリット、デメリットのまとめ」を記載（デザインを組み立てる上で重要）



「JCOGコンセプト作成の手引き」より改変

デザイン (≠サンプルサイズ)

- 検証的か探索的か
 - Phase1か2か3か
- エンドポイント
 - 真のエンドポイントか代替エンドポイントか
 - きちんとデータを取れる研究デザイン
- 判断規準
 - どのくらいの値を取れば意義ありとするか
 - 統計的推測に基づくサンプルサイズ設計



「JCOGコンセプト作成の手引き」より改変



いくつかの例をもとに、フレームワークの利用と統計的手法を紹介

実例を基にした、仮想例です。

Case1. 玄米サプリは体にいい？

- 統計家をしていると、よく研究の相談を受けます。私は、予防や疫学の研究もしているので、いろいろな相談を受けます。
- 先日も、当社の玄米サプリが**体にいい**から、それを証明するにはどんな研究をしたらいいのかと聞かれました。

1. クリニカルクエスチョン/臨床的意義
玄米サプリは体にいい/?

WHY

2. 対象
?

WHO

3. 標準治療
?

WHAT

4. 新治療
玄米サプリ

5. 研究デザイン

HOW

エンドポイント
体にいい?

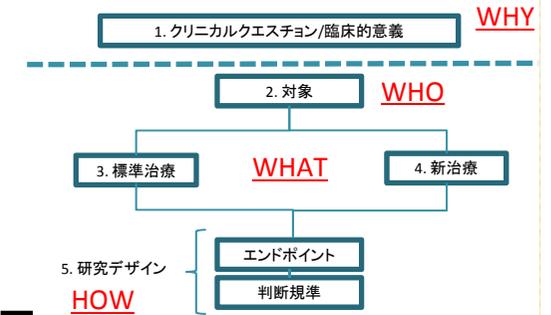
判断規準
?

クリニカルクエスチョンと臨床的意義

- クリニカルクエスチョンの明示ー試験を計画した動機
 - 対象患者について、現状で何が困っているか？
 - ダイエットや運動など、健康行動の変容は難しい。
 - ~~会社としての売り上げにつながる~~
- 結果の意義(Value)の明示ー目指すゴールの具体化
 - この試験の結果が、医療現場をどう変えるのか
 - どんなよいことが見込まれるのか=Value
 - サプリで健康になればお手軽に健康増進に繋がる。
 - 検証的な結論を得たいのか
 - 探索的な結論を得たいのか

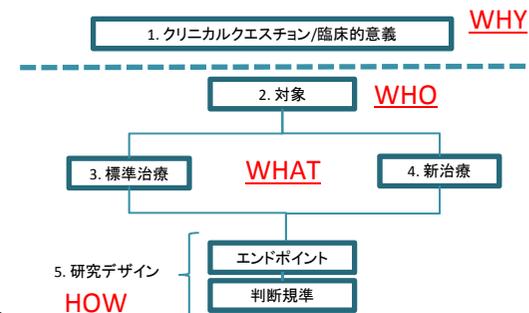
対象

- 試験の対象患者集団の説明
- なぜ、この対象を選んだか？
- なぜ、適格規準をそうしたか？
 - 病気でない人全般？
 - 検査値など異常があったり、境界域の人？
 - 食生活や運動習慣の改善ができない人？



標準治療

- 対象に対する標準治療の説明
- 標準治療は確定しているか
- 当該疾患の**治療ガイドライン**ではどのような記載がなされているか 推奨レベルや日常診療での普及について
 - 食事や運動の指導
 - 異常値や境界域なら投薬



試験治療

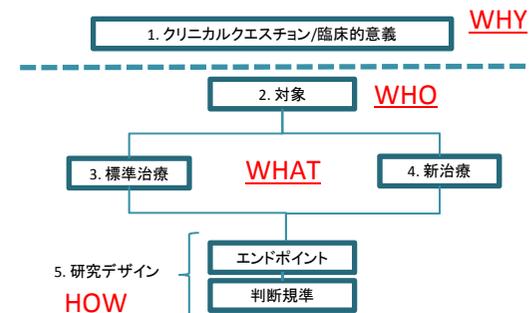
- 試験治療の意義
- Most promisingかどうか、が肝
- 標準治療と試験治療を比較した際の
 1. 「毒性のバランス」または
 2. 「メリット、デメリットのまとめ」を記載（デザインを組み立てる上で重要）

→ 玄米サプリがmost promisingなのか？

メリット：食事や運動より続けやすい

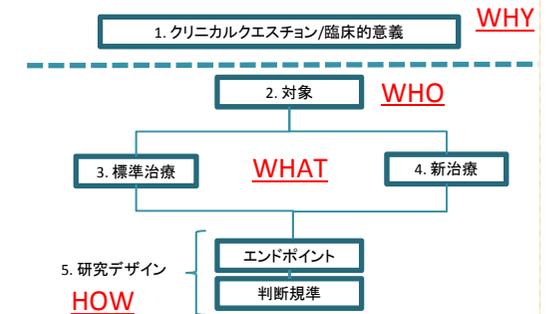
デメリット：お金がかかる

メリットデメリットから考える



デザイン (≠サンプルサイズ)

- 検証的か探索的か
 - Phase1か2か3か→Phase3(RCT, Randomized Controlled Trial)
- エンドポイント→健康って何？
 - 長生き、血圧↓、血糖値↓、BMI↓
 - 真のエンドポイントか代替エンドポイントか
- 判断規準
 - どのくらいの値を取れば意義ありとするか
 - 統計的推測に基づくサンプルサイズ設計



1. クリニカルクエスチョン/臨床的意義
玄米サプリは体にいい/?

WHY

2. 対象
病気の無い人

WHO

3. 標準治療
何もしない

WHAT

4. 新治療
玄米サプリ

5. 研究デザイン

HOW

エンドポイント
寿命

判断規準
?

1. クリニカルクエスチョン/臨床的意義
玄米サプリは体にいい/?

WHY

2. 対象
病気の無い人

WHO

3. 標準治療
何もしない

WHAT

4. 新治療
玄米サプリ

5. 研究デザイン

HOW

エンドポイント
血圧

判断規準
?

1. クリニカルクエスチョン/臨床的意義
玄米サプリは体にいい/?

WHY

2. 対象
高血圧の人

WHO

3. 標準治療
降圧薬

WHAT

4. 新治療
玄米サプリ

5. 研究デザイン

HOW

エンドポイント
血圧

判断規準
?

1. クリニカルクエスチョン/臨床的意義
玄米サプリは体にいい/?

WHY

2. 対象
高血圧の人

WHO

3. 標準治療
降圧薬

WHAT

4. 新治療
玄米サプリ

5. 研究デザイン

HOW

エンドポイント
BMI

判断規準
?

1. クリニカルクエスチョン/臨床的意義
玄米サプリは体にいい/?

WHY

2. 対象
肥満の人

WHO

3. 標準治療
何もしない

WHAT

4. 新治療
玄米サプリ

5. 研究デザイン

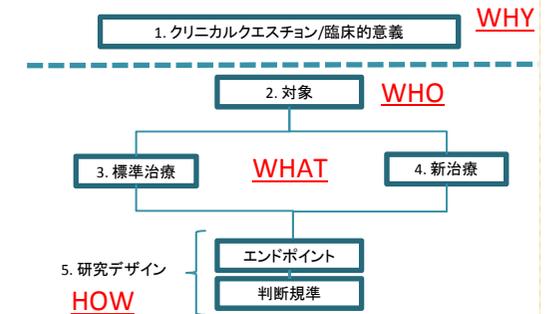
HOW

エンドポイント
BMI

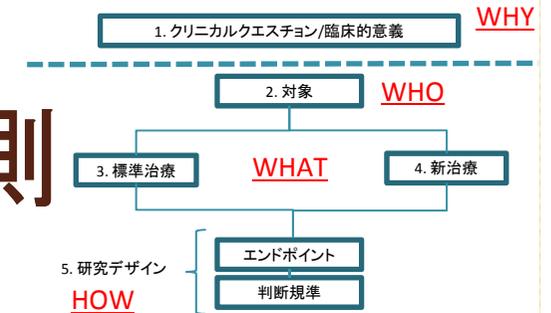
判断規準
?

デザイン (≠サンプルサイズ)

- 検証的か探索的か
 - Phase1か2か3か→Phase3(RCT)
- エンドポイント→健康って何？
 - 長生き、血圧↓、血糖値↓、BMI↓
 - 真のエンドポイントか代替エンドポイントか
- 判断規準
 - どのくらいの値を取れば意義ありとするか
 - 統計的推測に基づくサンプルサイズ設計



判断根拠と統計的推測



- 「玄米サプリをずっと飲むことによって、BMIが下がる」ことをもって、臨床的意義ありとする

- 1?2?3?下がる

BMIが1下がるとは、身長160cmの人で体重が2.5kg下がることに対応

- どのくらい下がりそうか、どのくらい下がると臨床的意味があるか、は異なる

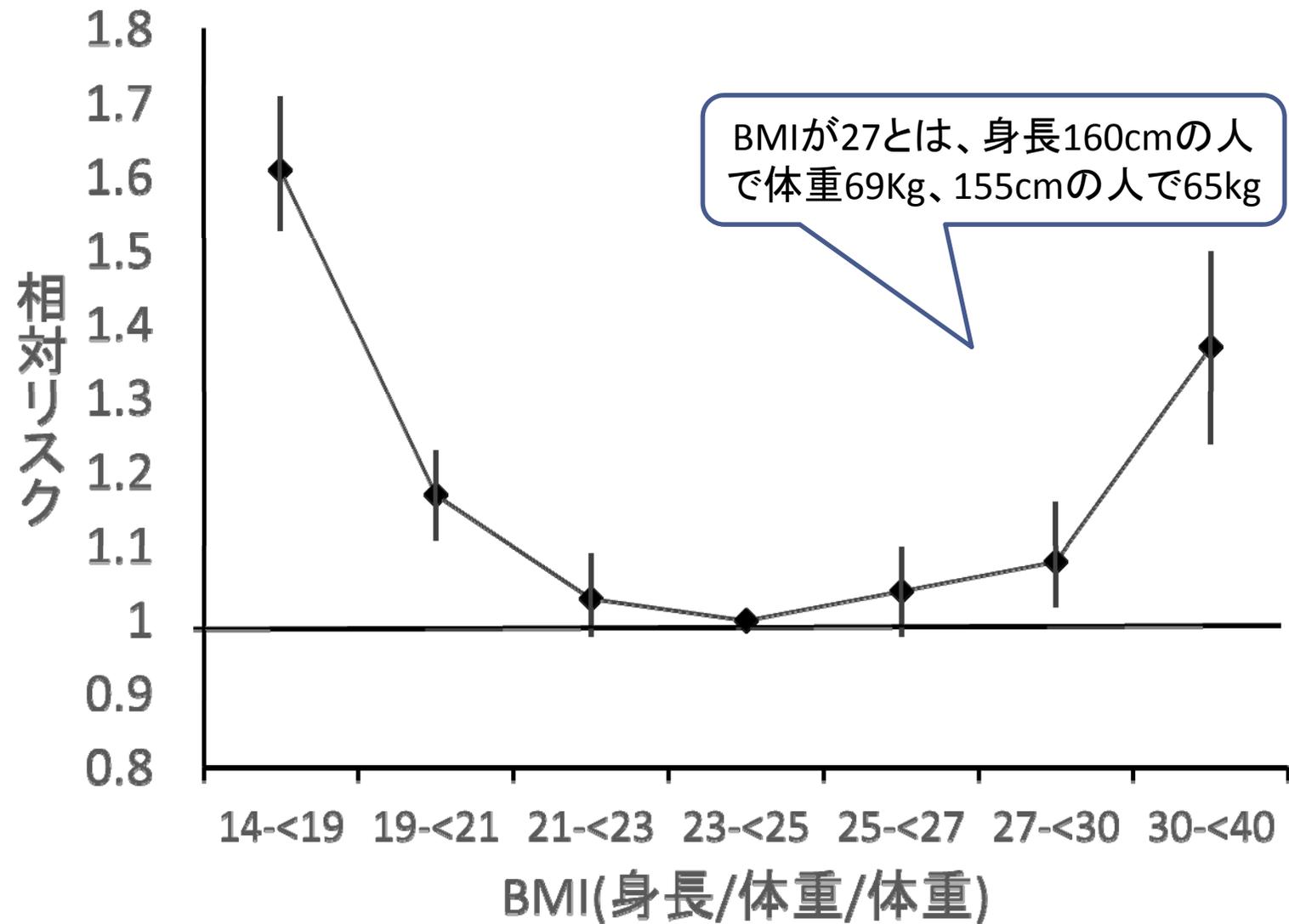
- 臨床的意義のある差を見つけるために試験を組むべき

誰がどのくらい痩せれば臨床的意義があるのか？

臨床的意義とは、デメリット(有害事象など)よりメリット(効果)が上回ること

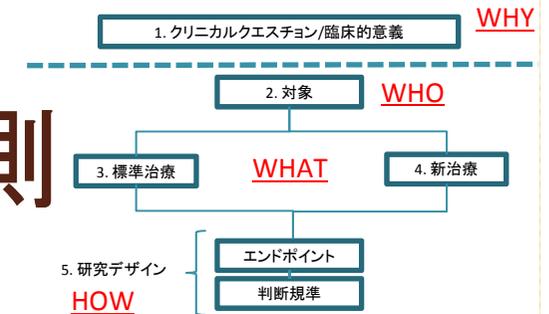
どのくらい痩せればいい？

BMIと全死亡(日本人女性)



Sasazuki S. et al. J Epidemiol 2011;21:417-3028

判断根拠と統計的推測



- 「玄米サプリをずっと飲むことによって、BMIが下がる」ことをもって、臨床的意義ありとする
 - 1下がる、3下がる
 - どのくらい下がりそうか、どのくらい下がると臨床的意味があるか、は異なる
 - 臨床的意義のある差を見つけるために試験を組むべき

検証する仮説:

玄米サプリを1年間飲むことによって、飲まない場合に比べ、BMI>27の人のBMIが2下がる

1. クリニカルクエスチョン/臨床的意義
玄米サプリにより、肥満の人の体重が減少する

WHY

2. 対象
BMI>27の人

WHO

3. 標準治療
何もしない

WHAT

4. 新治療
玄米サプリを
1年飲む

5. 研究デザイン

HOW

エンドポイント
BMI

判断規準
BMIが2下がる

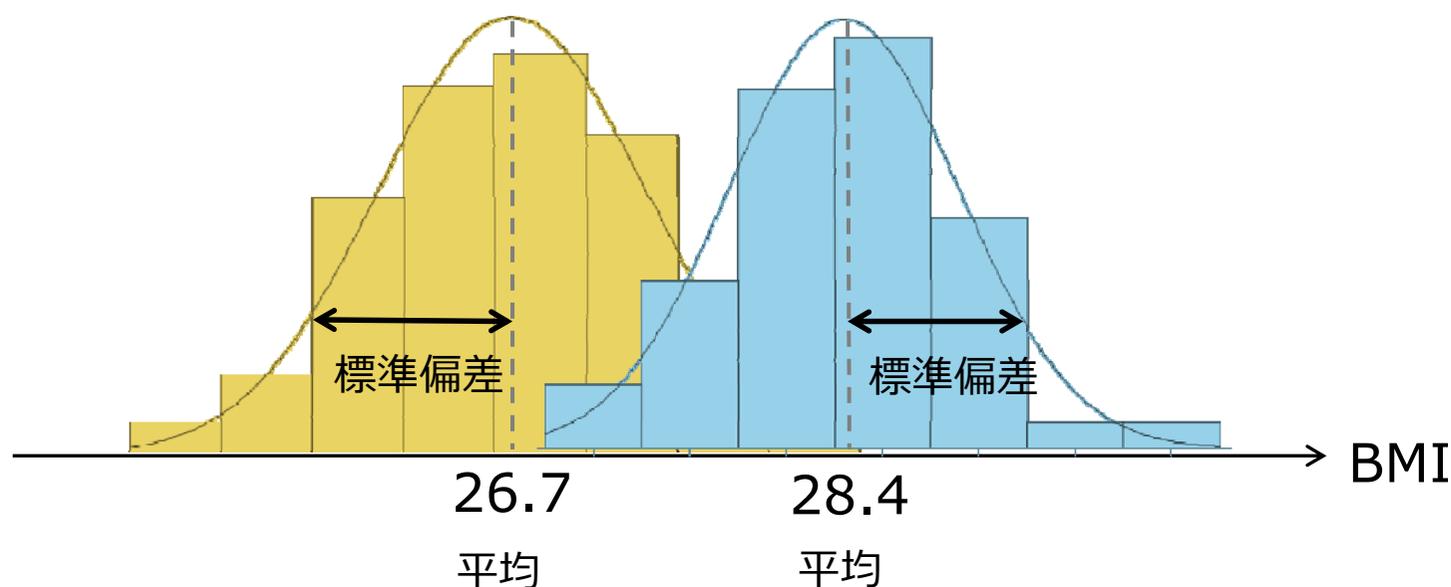
実際に得られるデータ

- 玄米サプリ摂取群と何もしない群で、1年後のBMIが異なるか？
- 帰無仮説：なし群とサプリ群で1年後のBMIが等しい

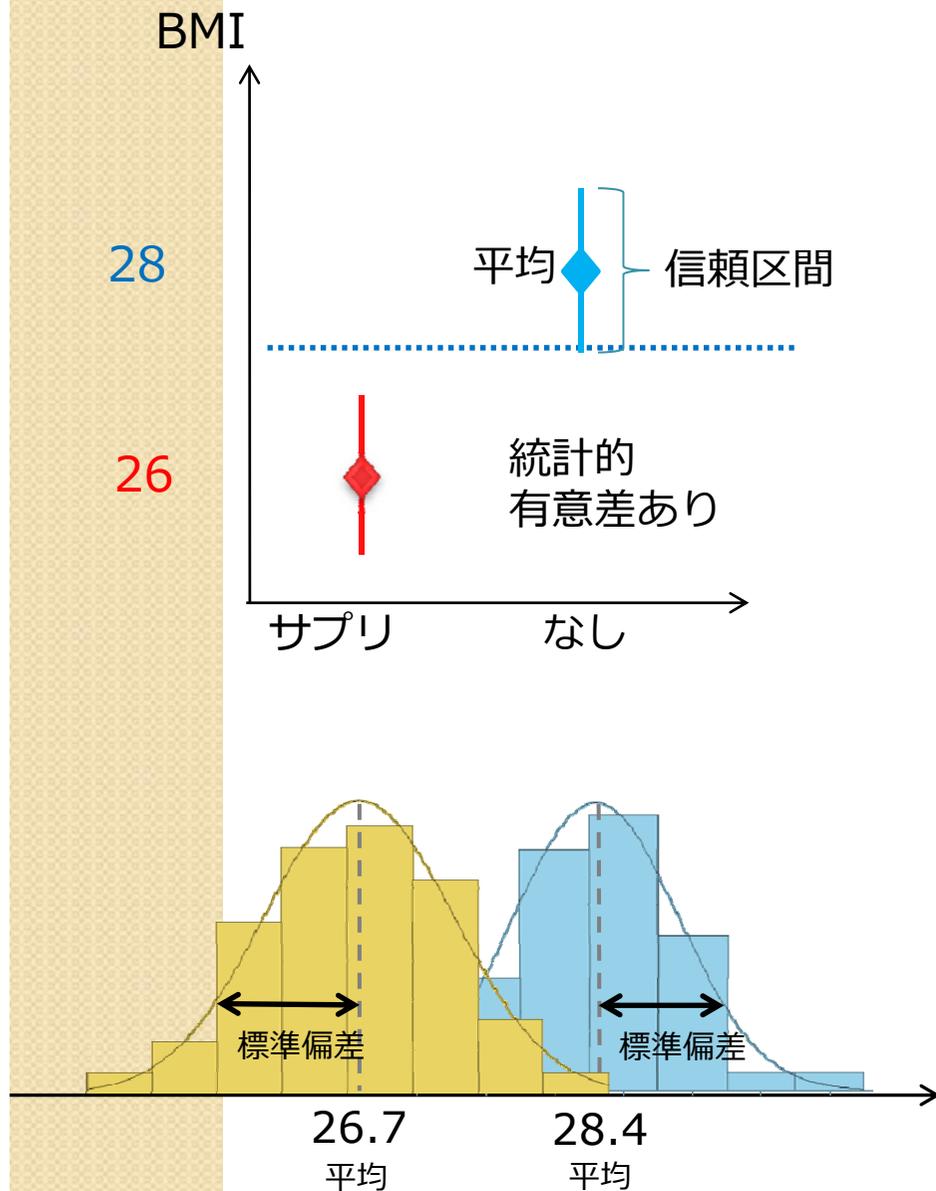
BMI(Kg/m ²)			
なし群 (n=100)	28.3	...	28.5
サプリ群 (n=100)	26.0	...	26.3

まず、ヒストグラムで分布の形を確認

BMI(Kg/m ²)				平均値
なし群 (n=100)	28.3	...	28.5	28.4
サブリ群 (n=100)	26.0	...	26.3	26.7



データの要約

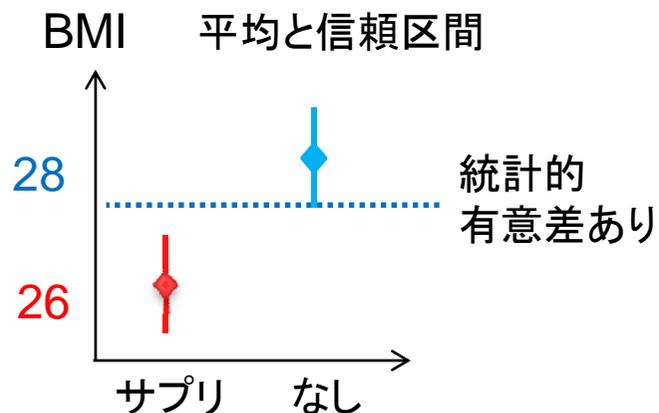


平均：分布の位置を表す代表値

信頼区間：平均値の精度であり、分布の広がり（標準偏差）と測定した人数Nから計算できます。

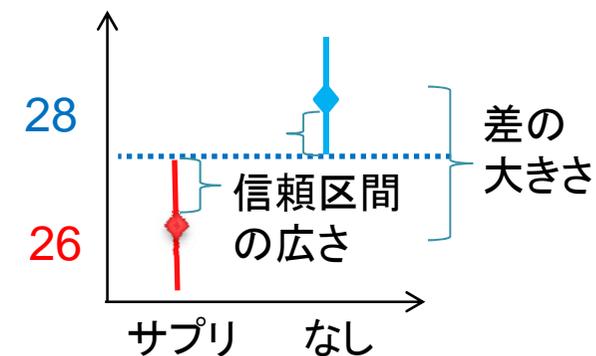
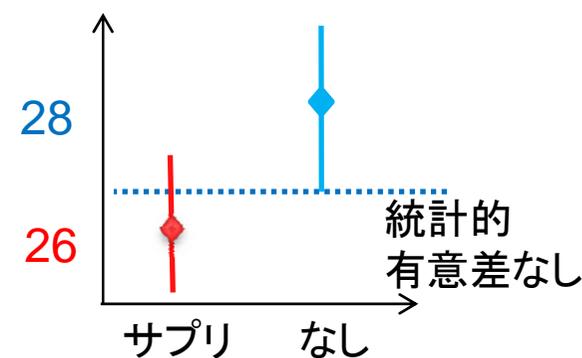
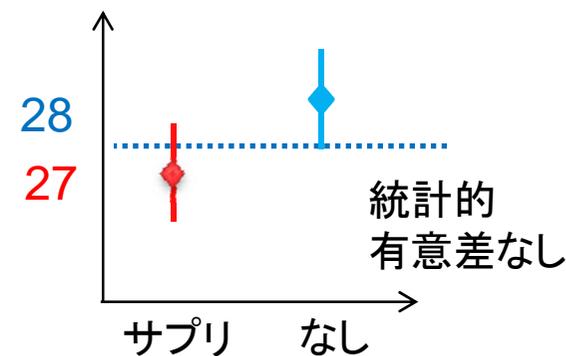
2つの群の平均値の信頼区間が重なっていない時、平均値が異なるといえる（統計的有意）
→t検定

検査値のような連続量の 統計的推測とサンプルサイズ設計



- ・ 差が小さいと有意になりにくい
- ・ Nが小さいと信頼区間が広くなり、有意になりにくい

ちょうど有意になるところは、
差の大きさ＝信頼区間の広さの和、なので
この式を解くと、Nが逆算できる(サンプル
サイズ設計のイメージ)



サンプルサイズ計算を実際にやってみましょう！

- BMIを2下げる
- BMIの標準偏差は女性3.6、男性3.2
- α エラー=0.05、検出力=80%

The screenshots show the following steps:

- Home screen with 'サンプルサイズ計算' selected.
- Calculation parameters: 2 groups, survival period comparison, two-sided alpha (0.05), power (0.8), standard deviation (3.6).
- Calculation progress: '2群の生存期間の比較に基づく'.
- Calculation progress: '2群の生存期間の比較に基づく'.
- Final result table:

	1	2
ABC		
GHI	4	5
JKL		
PQRS	7	8
TUV		
	.	0

連続量のエンドポイントの例

- 糖尿病
 - HbA1c、空腹時血糖値、1,5AG, グルコアルブミン
- 脂質異常症
 - LDL、HDL、総コレステロール、HDL/LDL比
- 高血圧：血圧
- 喘息
 - 発作回数、ピークフロー値
- COPD：1秒率 (FEV1.0%)
- 骨そしょう症
 - 骨量、骨代謝マーカー
- 緑内障：眼圧
- 排尿障害
 - 排尿回数、IPSSスコア、自排尿量、Qmax、Qwave
- クローン病：CDAIスコア
- 腎性貧血
 - Hb値、網赤血球数
- うつ病
 - HAM-Dスコア、MADRSスコア
- 認知症
 - ADAS-J Cog、CIBIC-plus J
- C型慢性肝炎
 - HCV-RNA陰性化率

Case2. パーソナルトレーニングは健康にいい？

- 最近、すごく太ってしまったので、痩せるには運動するのがいいと思って走ってるのですが、全然痩せません。やっぱり基礎代謝を上げるべきかと思い、筋トレしようと思ってネットでジム探しをしていると、あるパーソナルトレーニングを見つけました。
- このパーソナルトレーニングは、ホームページによると、無理なく続けられて、痩せるだけでなく、腰の痛み、高血圧、高コレステロール、高血糖など、様々な体調不良に効くとのこと。

1. クリニカルクエスチョン/臨床的意義
パーソナルトレーニングは健康にいい/?

WHY

2. 対象
?

WHO

3. 標準治療
?

WHAT

4. 新治療
パーソナル
トレーニング

5. 研究デザイン

HOW

エンドポイント
健康?

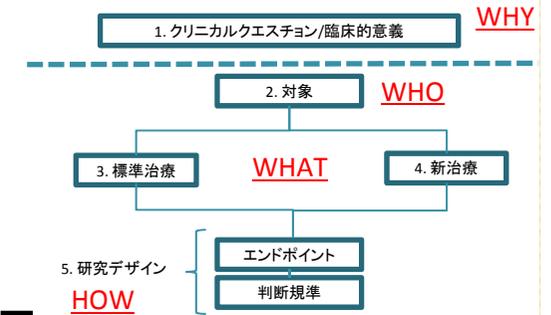
判断規準
?

クリニカルクエスチョンと臨床的意義

- クリニカルクエスチョンの明示ー試験を計画した動機
 - 対象患者について、現状で何が困っているか？
 - 病気ではないので、薬などを飲んだりしないで、健康を維持したい
 - ダイエットや運動など、健康行動の変容は難しい。
- 結果の意義(Value)の明示ー目指すゴールの具体化
 - この試験の結果が、医療現場をどう変えるのか
 - どんなよいことが見込まれるのか=Value
 - パーソナルとレーニングにより、様々な健康状態が改善するか、疾患予防となるか
 - 運動すれば様々な疾患予防になることはわかっているのだから、運動習慣がつくかどうかでいいのでは
 - 検証的な結論を得たいのか
 - 探索的な結論を得たいのか

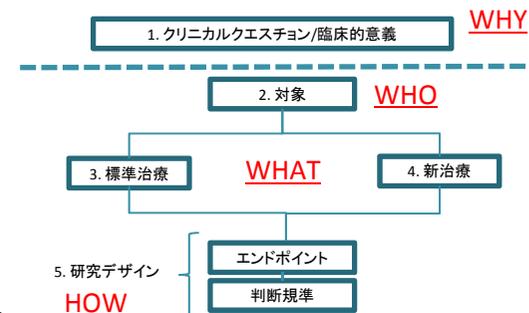
対象

- 試験の対象患者集団の説明
- なぜ、この対象を選んだか？
- なぜ、適格規準をそうしたか？
 - 病気でない人全般
 - 運動したいと思っているが、運動がなかなか続けられない人
 - 運動なんか絶対嫌だと思っている人



標準治療

- 対象に対する標準治療の説明
- 標準治療は確定しているか
- 当該疾患の**治療ガイドライン**ではどのような記載がなされているか 推奨レベルや日常診療での普及について
 - 運動の指導
 - 他の一般的なジム
 - 何もしない



試験治療

- 試験治療の意義
- Most promisingかどうか、が肝
- 標準治療と試験治療を比較した際の
 1. 「毒性のバランス」または
 2. 「メリット、デメリットのまとめ」

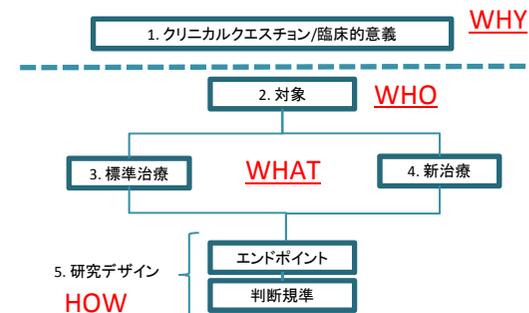
を記載（デザインを組み立てる上で重要）

→運動のためにパーソナルトレーニングがmost promisingか？

メリット：自分でやるより分かりやすい

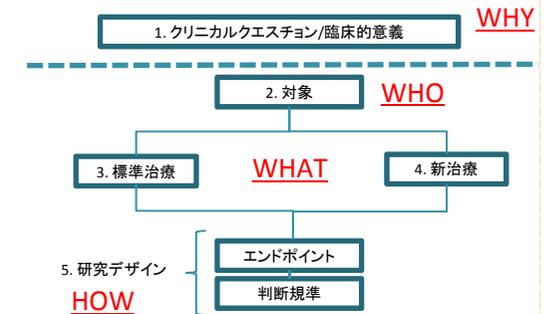
デメリット：お金がかかる

メリットデメリットから考える



デザイン (≠サンプルサイズ)

- 検証的か探索的か
 - Phase1か2か3か→P3(RCT)
- エンドポイント→健康って何？
 - 血圧↓、血糖値↓、BMI↓、骨折↓
 - 運動習慣↑
 - 真のエンドポイントか代替エンドポイントか
- 判断規準
 - どのくらいの値を取れば意義ありとするか
 - 統計的推測に基づくサンプルサイズ設計



エンドポイント(ものさし)

無効

有効



症状緩和度



5年生存率



QOLスコア



奏効率

主なものさし : プライマリ・エンドポイント(主要評価項目)

その他のものさし : セカンダリ・エンドポイント(副次的評価項目)

エンドポイントのタイプ

- 真(True)と代替(Surrogate)
 - 真のエンドポイント(True endpoint)
 - 患者さんの benefit を直接反映しているエンドポイント
 - 代替エンドポイント(Surrogate endpoint)
 - 真のエンドポイントより簡便、早くわかる、少ない患者でわかる・等の利点あり
 - 真のエンドポイントと評価の結果が一致
- ハードとソフト
 - ハードなエンドポイント
 - 誰が見ても評価が一致
 - 何回見ても一致
 - ソフトなエンドポイント
 - 見る者によって評価が異なる・時々によって違う
 - 他の影響を受けやすい

連続量のエンドポイントの例

- 代替エンドポイント
- 真のエンドポイント

- 糖尿病
 - HbA1c、空腹時血糖値、1,5AG, グルコアルブミン
- 脂質異常症
 - LDL、HDL、総コレステロール、HDL/LDL比
- 高血圧: 血圧
- 喘息
 - 発作回数、ピークフロー値
- COPD: 1秒率 (FEV1.0%)
- 骨そしょう症
 - 骨量、骨代謝マーカー
- 緑内障: 眼圧
- 排尿障害
 - 排尿回数、IPSSスコア、自排尿量、Qmax、Qwave
- クローン病: CDAIスコア
- 腎性貧血
 - Hb値、網赤血球数
- うつ病
 - HAM-Dスコア、MADRSスコア
- 認知症
 - ADAS-J Cog、CIBIC-plus J
- C型慢性肝炎
 - HCV-RNA陰性化率

デザイン (≠サンプルサイズ)

- 検証的か探索的か

- Phase1か2か3か→P3(RCT)

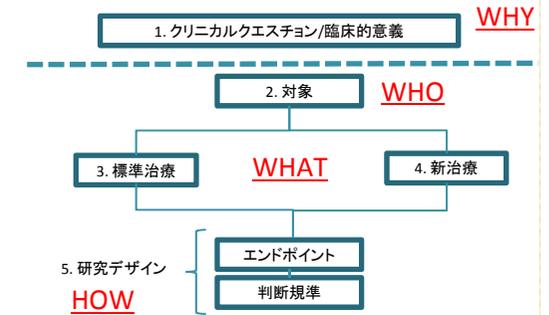
- エンドポイント→健康って何？

- 血圧↓、血糖値↓、BMI↓、骨折↓
- 身体活動量↑

- 真のエンドポイントか代替エンドポイントか

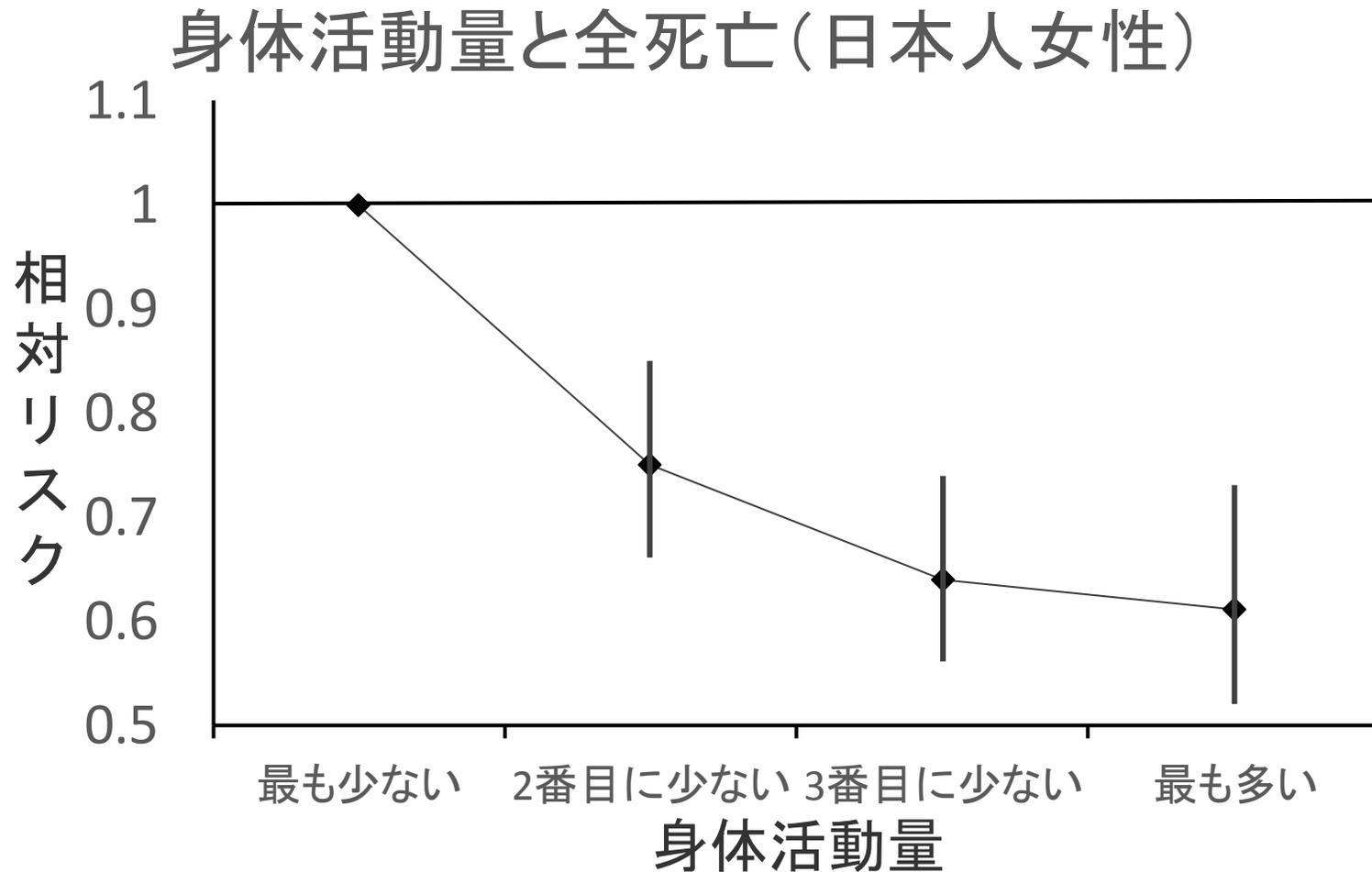
- 判断規準

- どのくらいの値を取れば意義ありとするか
- 統計的推測に基づくサンプルサイズ設計



- 代替エンドポイント
- 真のエンドポイント

身体活動量は健康の代替エンドポイントか？



Inoue M. et al. Ann Epidemiol 2008;18:522-30

1. クリニカルクエスチョン/臨床的意義
パーソナルトレーニングは健康にいい/?

WHY

2. 対象
?

WHO

3. 標準治療
?

WHAT

4. 新治療
パーソナル
トレーニング

5. 研究デザイン

HOW

エンドポイント
健康?

判断規準
?

1. クリニカルクエスチョン/臨床的意義
パーソナルトレーニングは健康にいい/?

WHY

2. 対象
すべての人

WHO

3. 標準治療
中央区のジム

WHAT

4. 新治療
パーソナル
トレーニング

5. 研究デザイン

HOW

エンドポイント
1年後身体活動量

判断規準
?

1. クリニカルクエスチョン/臨床的意義
パーソナルトレーニングは健康にいい/?

WHY

2. 対象
高齢者

WHO

3. 標準治療
何もしない

WHAT

4. 新治療
パーソナルトレ
ーニング5年間

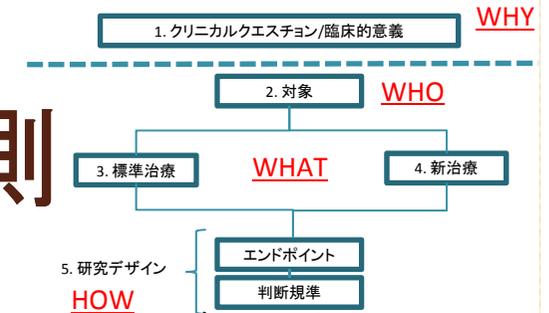
5. 研究デザイン

HOW

エンドポイント
骨折

判断規準
?

判断根拠と統計的推測



- 「パーソナルトレーニングのほうが、中央区のジムよりも身体活動量が増える」ことをもって、臨床的意義ありとする

- たくさん増えた人と増えなかった人を混ぜて平均をとって比べても意味ないのでは？
- 人ごとに効果を見たい
 - 増えた人の割合とか、運動習慣がついた人の割合
- どのくらい増えそうか、どのくらい増えれば臨床的意味があるか、は異なる
 - 臨床的意義のある差を見つけるために試験を組むべき

この場合に、メリットがデメリットを上回るとは？

1. クリニカルクエスチョン/臨床的意義
パーソナルトレーニングは健康にいい/?

WHY

2. 対象
運動していない人

WHO

3. 標準治療
中央区のジム

WHAT

4. 新治療
パーソナル
トレーニング

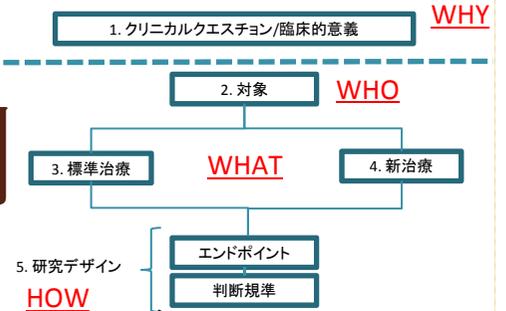
5. 研究デザイン

HOW

エンドポイント
1年後の運動継続

判断規準
?

判断根拠と統計的推測

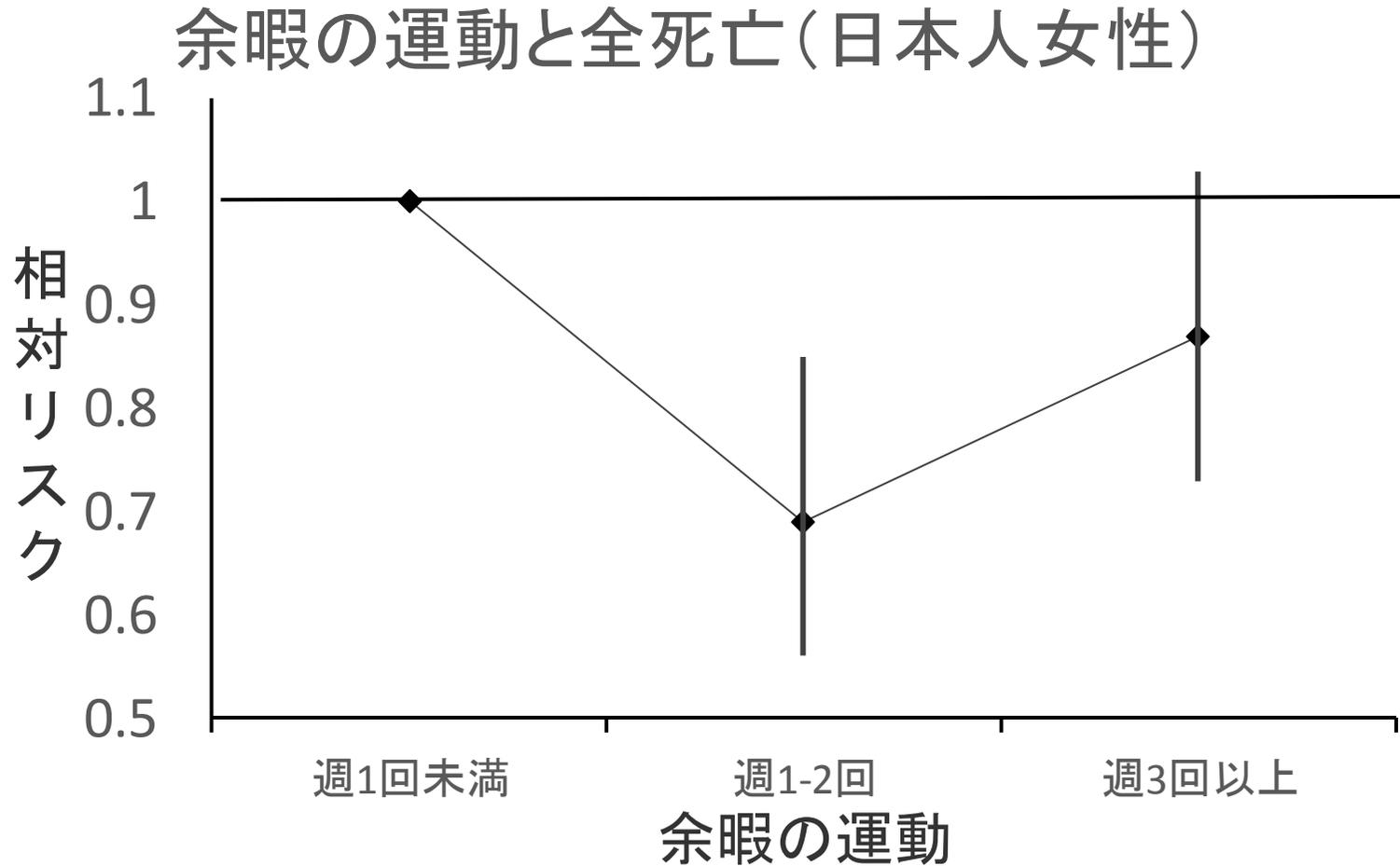


- 「パーソナルトレーニングのほうが、中央区のジムよりも身体活動量が増える」ことをもって、臨床的意義ありとする
 - 運動習慣者の増加へ変更
 - 運動習慣の定義は？
 - メリットがデメリットを上回るとは？

運動習慣の定義は？

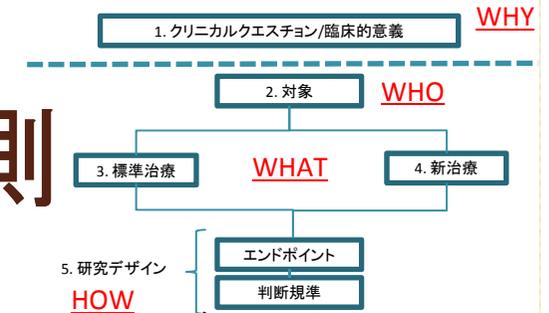
臨床的に意義のある差は？

メリット？
デメリット？



Inoue M. et al. Ann Epidemiol 2008;18:522-30

判断根拠と統計的推測



- 「パーソナルトレーニングのほうが、中央区のジムよりも身体活動量が増える」ことをもって、臨床的意義ありとする
 - 運動習慣者の増加へ変更
 - 運動習慣の定義は？
 - メリットがデメリットを上回るとは？

検証する仮説:

パーソナルトレーニングのほうが、中央区のジムよりも、1年後、週1回以上運動している人の割合が20%高い

1. クリニカルクエスチョン/臨床的意義
パーソナルトレーニングをやると
運動継続が期待できる

WHY

2. 対象
運動していない人

WHO

3. 標準治療
中央区のジム

WHAT

4. 新治療
パーソナル
トレーニング

5. 研究デザイン

HOW

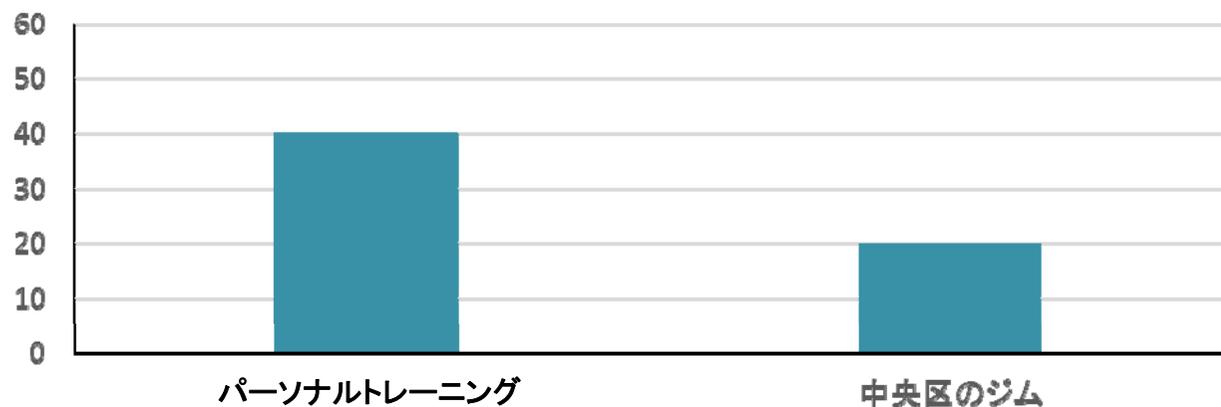
エンドポイント
1年後の運動継続

判断規準
20%高い

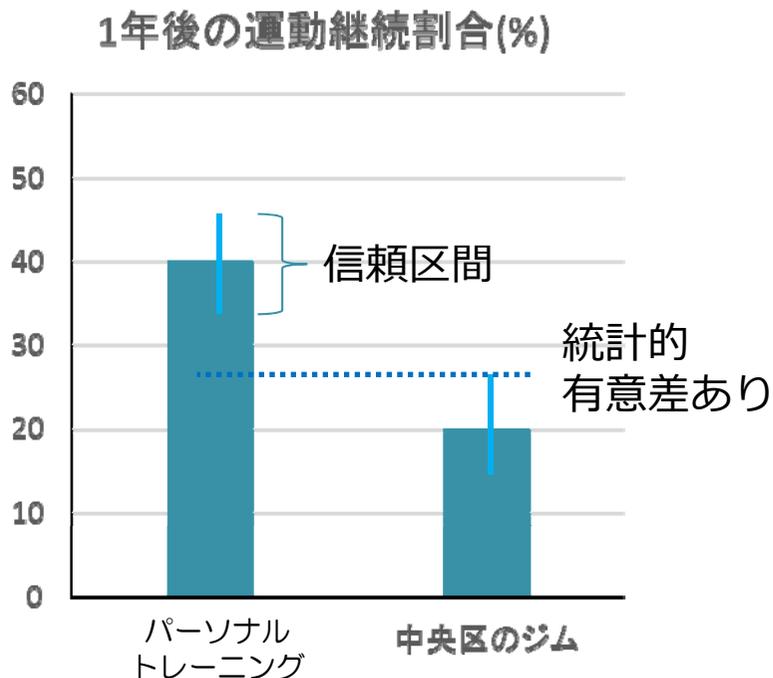
得られるデータ

1年後の運動継続			
	あり	なし	合計
パーソナルトレーニング	40人	60人	100人
中央区のジム	20人	80人	100人
合計	60人	140人	200人

1年後の運動継続割合(%)



データの要約



割合：代表値

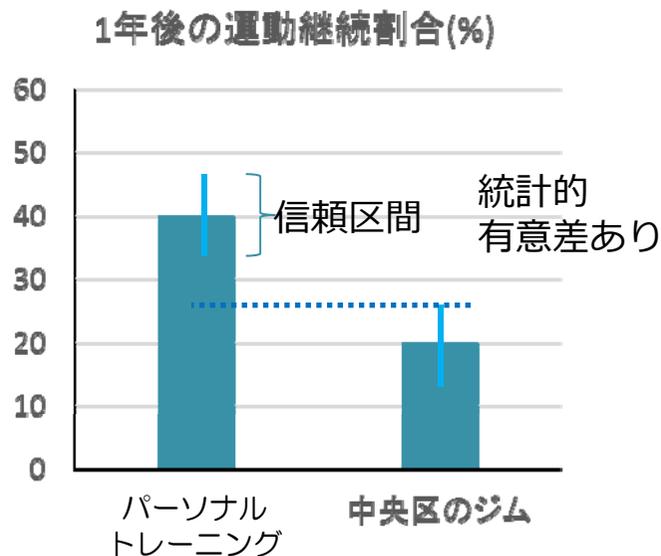
信頼区間：割合の精度であり、割合の値と測定した人数Nから計算できます。

2つの群の割合の信頼区間が重なっていない時、割合が異なるといえる（統計的有意）
→ χ^2 検定(カイ二乗検定)

割合の差 $40-20=20\%$ 継続しやすい
割合の比 $40/20=2$ 倍継続しやすい

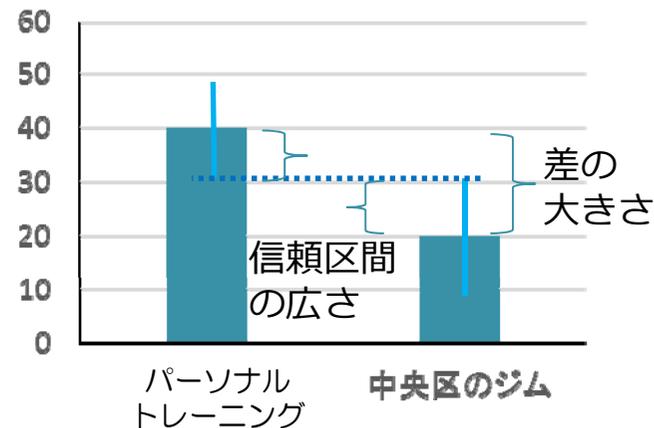
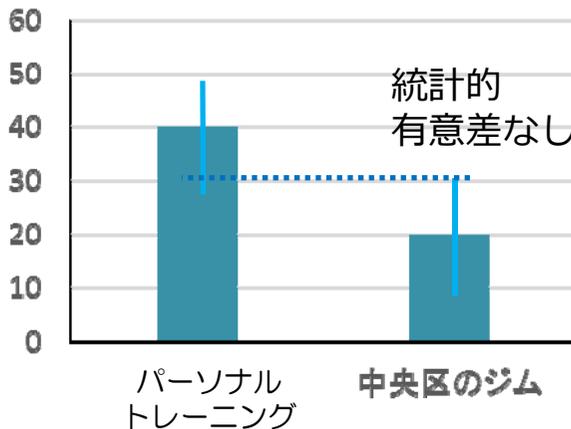
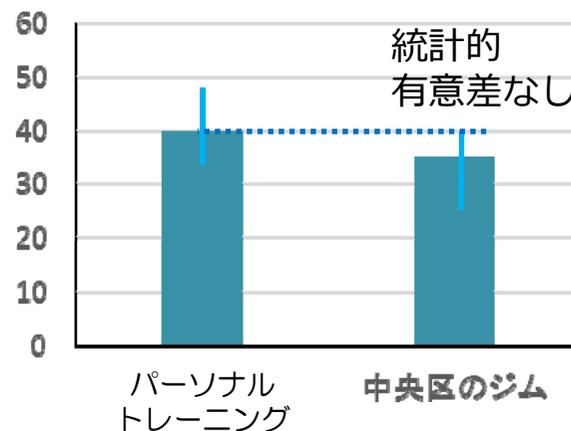
割合の信頼区間が重なるかどうかは、差の信頼区間が0、比の信頼区間が1を含むかどうかに対応します。
したがって、 χ^2 検定は、割合の差が0、比が1かどうかを検定するのと同じ、と言えます

割合の統計的推測とサンプルサイズ設計



- ・ 差が小さいと有意になりにくい
- ・ Nが小さいと信頼区間が広くなり、有意になりにくい

ちょうど有意になるところは、
 差の大きさ = 信頼区間の広さの和、なので
 この式を解くと、Nが逆算できる(サンプル
 サイズ設計のイメージ)



サンプルサイズ計算を実際にやってみましょう！

- 1年後の継続割合が20%高い
- 40-59歳の運動継続者20% (30-39歳は13%)
- α エラー=0.05、検出力=80%

ICR臨床研究入門

研究倫理指針

サンプルサイズ計算

p値計算

検定・推定

講座一覧

ICRからのお知らせ

履修状況をPCで確認

PC版サイト

HOME 講座一覧 DL一覧 ツール

ツール

サンプルサイズ計算

2群の生存期間の比較 (優越性に基づく)

2群の生存期間の比較 (非劣性に基づく)

2群の生存期間の比較 (優越性に基づく)

2群の生存期間の比較 (非劣性に基づく)

2群の生存期間の比較 (優越性に基づく)

2群の生存期間の比較 (非劣性に基づく)

割合の閾値・期待値

2群の平均値の差の比較 (優越性を仮定したt検定に基づく)

2群の割合の差の比較 (優越性検定に基づく)

サンプルサイズ 検出力

検出力

結果表示

2群を合わせたサンプルサイズ

1	2	
4	5	
7	8	
.	0	

2群の割合の差の比較 (優越性、カイ2乗検定に基づく)

検出力

結果表示

2群を合わせたサンプルサイズ

オッズ比

リスク比

割合のエンドポイントの例

- 骨そしょう症

- 骨折割合

真のエンドポイント、かなりハード

- がん

- 腫瘍縮小割合（奏効率）

代替エンドポイント、かなりソフト

Case3. そんなにいいなら

- 玄米サプリとパーソナルトレーニングがそんなにいいなら、もっと積極的に病気の人に使ってはどうか。
- たとえば、糖尿病の人の心血管イベントの予防や、乳がん患者の再発予防など、真のエンドポイントを防げないでしょうか。

1. クリニカルクエスチョン/臨床的意義
玄米サプリとパントレで糖尿病患者の
心血管イベント発症リスクを下げる

WHY

2. 対象
糖尿病患者

WHO

3. 標準治療
なし(通常治療
のみ)

WHAT

4. 新治療
(通常治療に加えて)
玄米サプリとパントレ

5. 研究デザイン

HOW

エンドポイント
心疾患イベント

判断規準
10%減

1. クリニカルクエスチョン/臨床的意義
玄米サプリとパントレで乳がん患者の
再発リスク

WHY

2. 対象
乳がん患者

WHO

3. 標準治療
なし(通常治療
のみ)

WHAT

4. 新治療
(通常治療に加えて)
玄米サプリとパントレ

5. 研究デザイン

HOW

エンドポイント
再発

判断規準
10%減

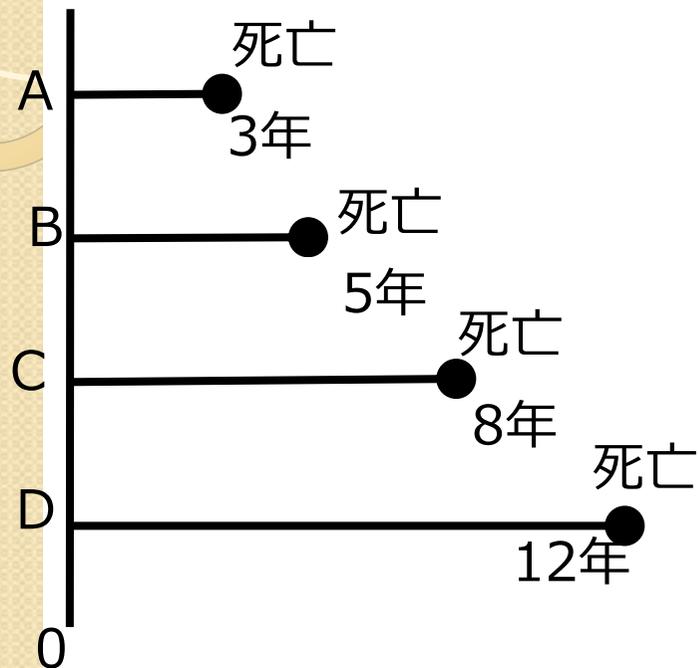
得られるデータ

ID	登録日	治療群	最終確認日	イベント
1	2015/9/10	玄米サプリ・パントレ	2016/5/3	あり
2	2015/9/15	介入なし	2016/10/1	あり
3	2015/10/1	玄米サプリ・パントレ	2017/5/1	なし
4	2015/10/10	玄米サプリ・パントレ	2018/6/1	あり
5	2015/10/25	介入なし	2020/8/31	なし
6	2015/11/7	玄米サプリ・パントレ	2020/8/31	なし
7	2015/11/10	介入なし	2020/8/31	なし
・	・
・
・

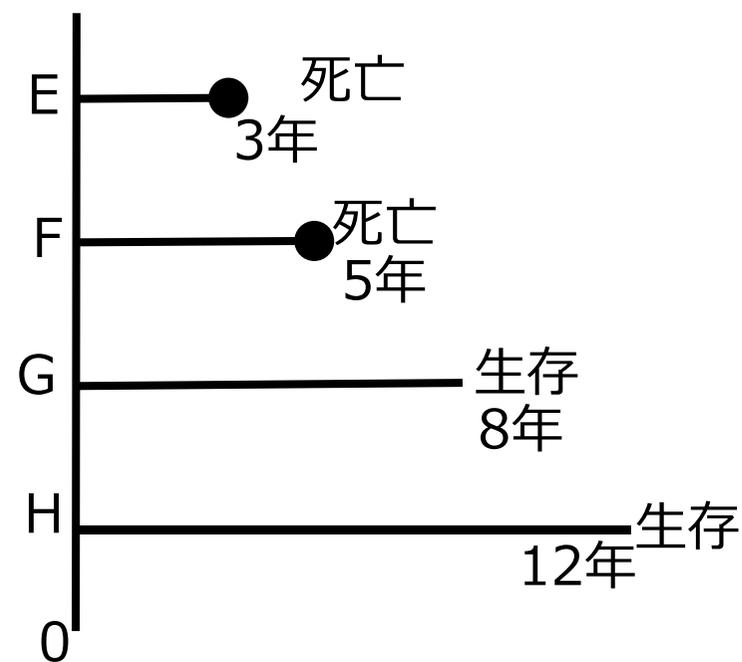
研究終了日 = 最終追跡日にイベントなし

問題1: 平均値 (t検定) で良いでしょうか？

なし群



玄米サプリ・パソトレ群

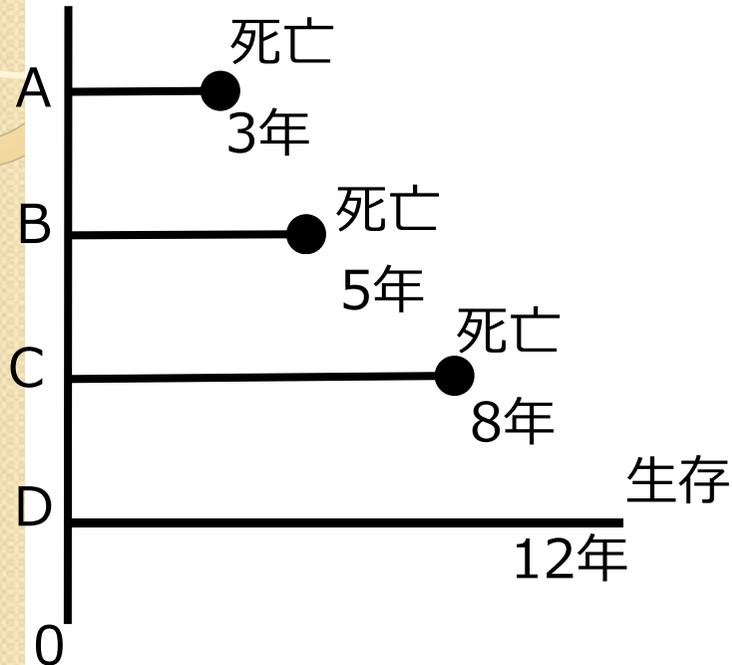


- なし群の平均値: $(3+5+8+12)/4 = 7$ 年
- 玄パ群の平均値: $(3+5+8+12)/4 = 7$ 年 ??

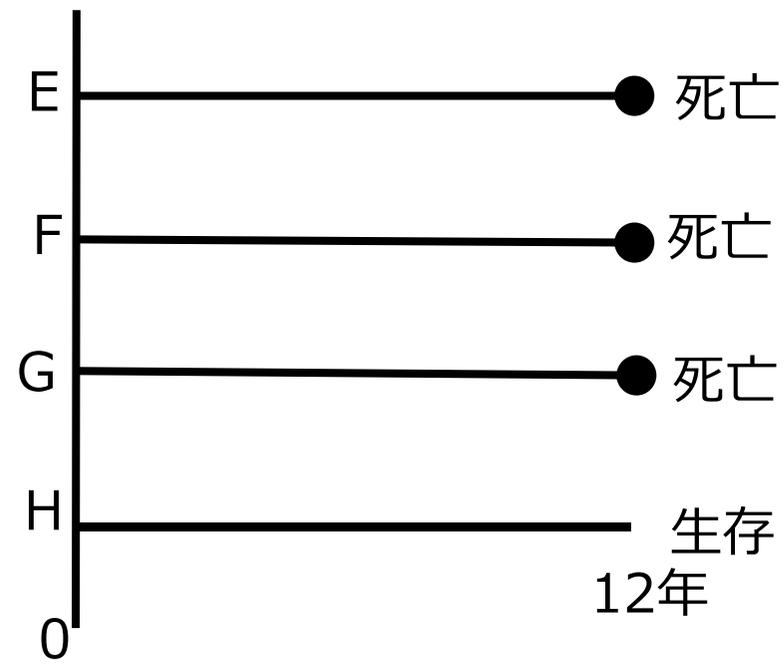
玄パ群の方が明らかに生存している時間は長いはず

問題2:12か月時点の生存割合の比較(カイ2乗検定)で良いでしょうか？

なし群



玄パ群



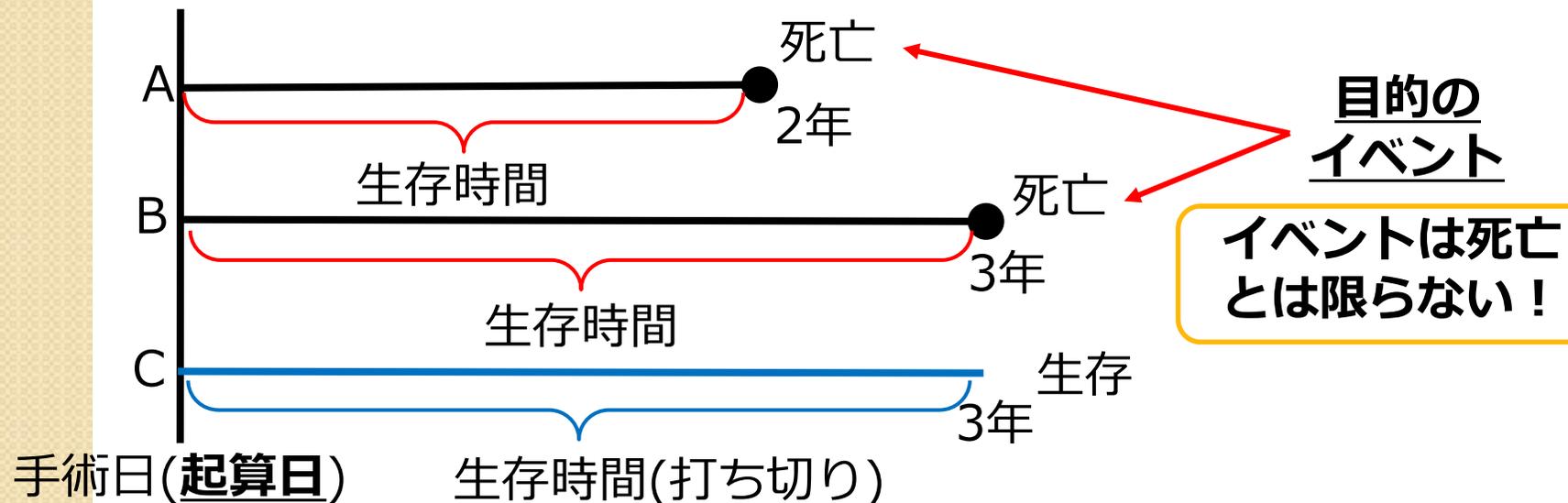
- **なし群**の死亡割合: $3/4 = 75\%$
- **玄パ群**の死亡割合: $3/4 = 75\%$??

玄パ群の方が明らかに生存している時間は長いはず

生存時間解析(survival analysis)とは？

- 生存時間とは

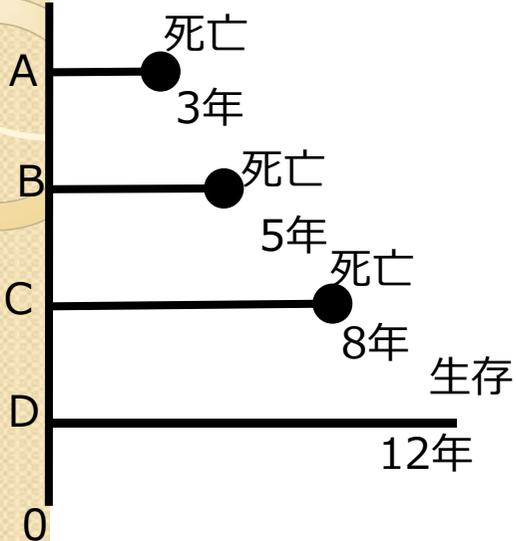
- ある特定の時点(起算日)から、目的のイベントが起きるまでの時間のこと
 - 臨床試験に登録されてから死亡日まで
 - 手術日から再発日まで



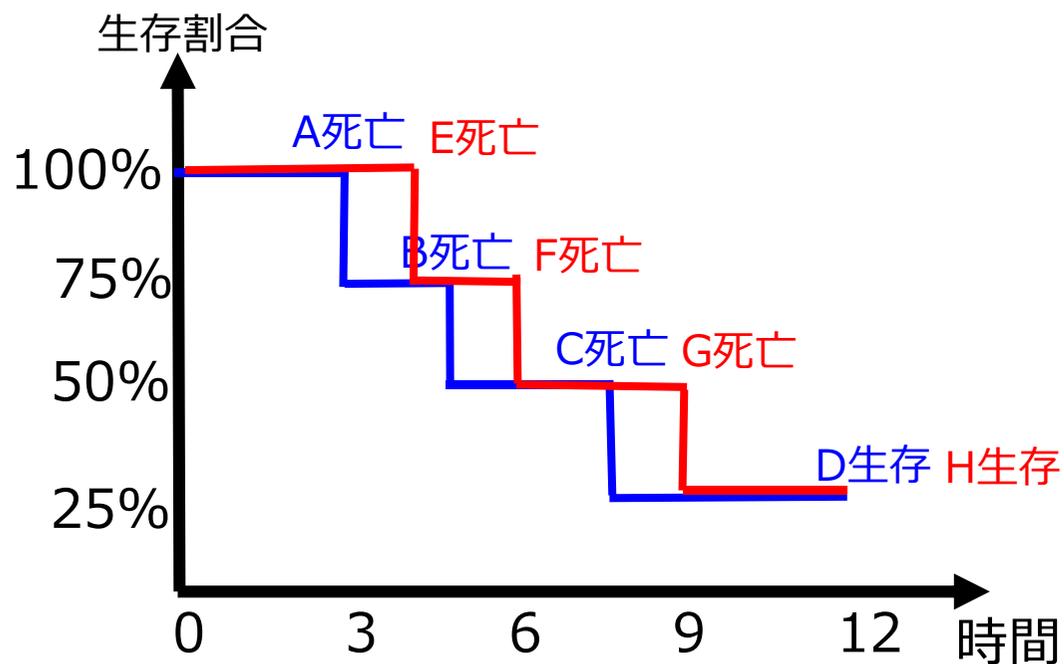
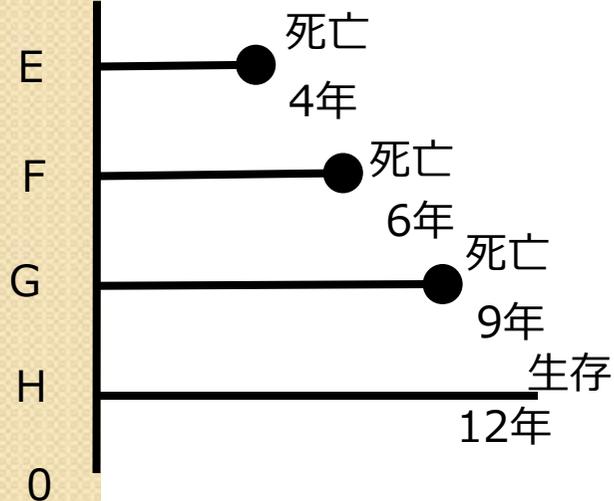
生存時間に興味のある場合に用いる解析手法のことを生存時間解析という

生存曲線を描く！

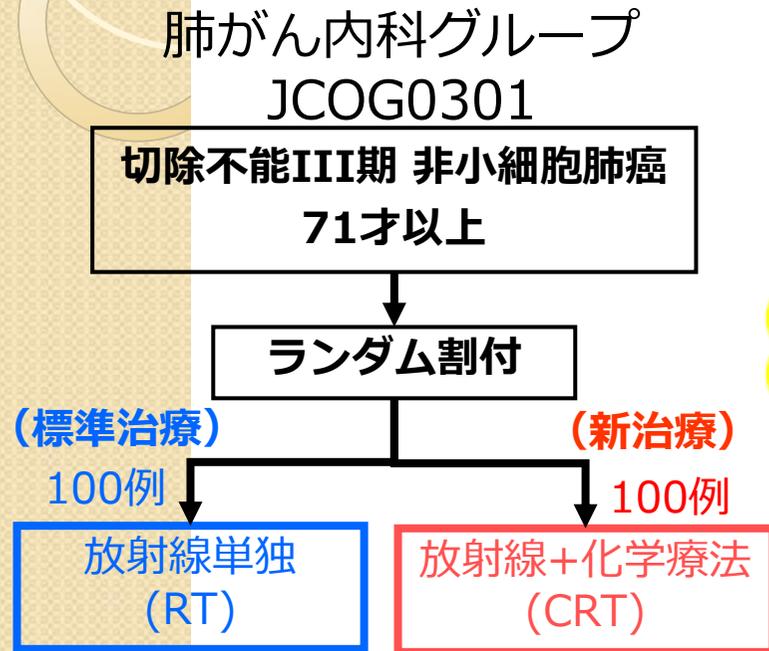
なし群



玄パ群

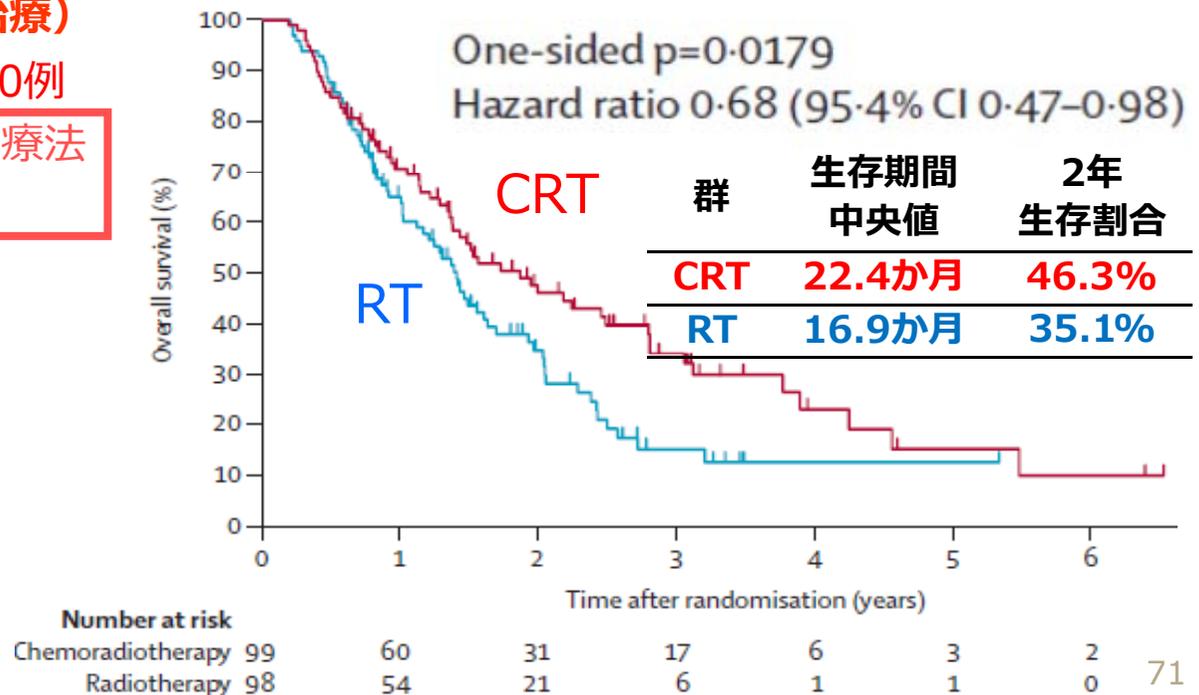


臨床研究で用いられる生存時間解析(がん)



Statistical analysis

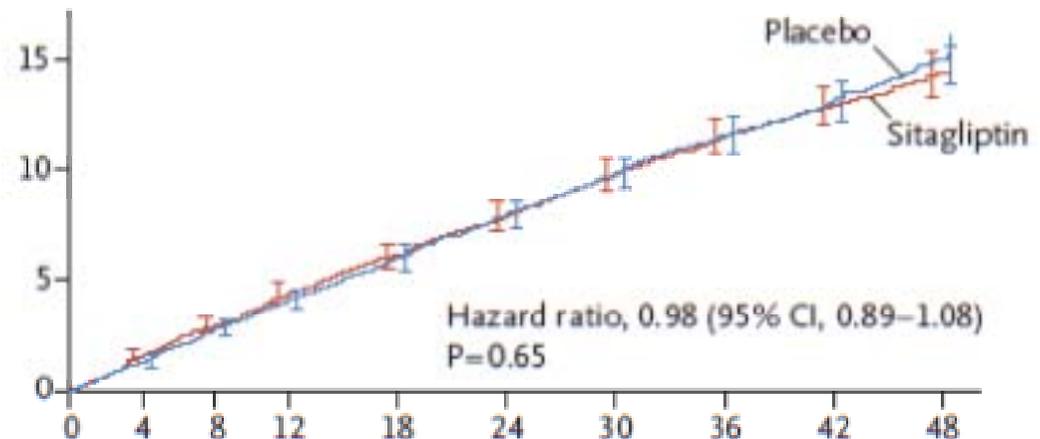
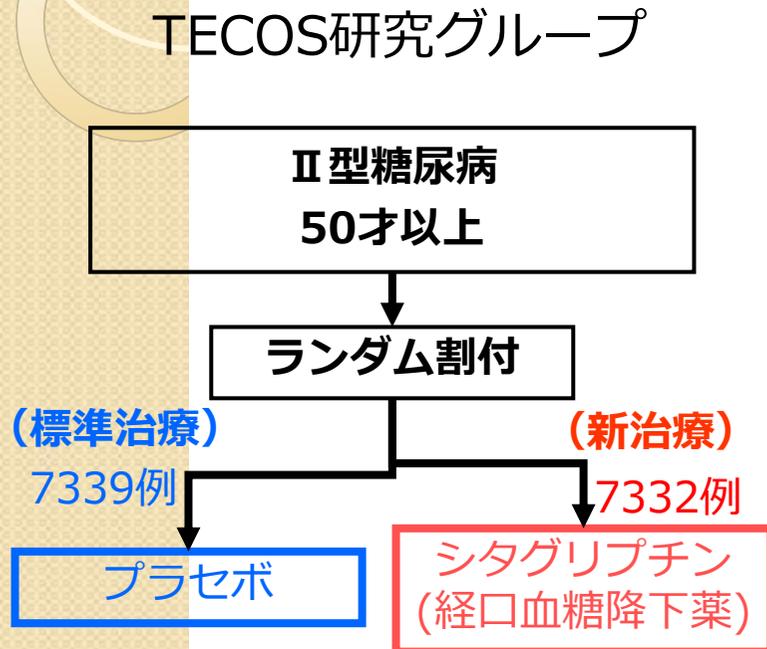
The primary endpoint of overall survival was analysed by the stratified **log-rank test** for eligible patients with performance status (0 vs 1-2 [because of the small number of patients with performance status of 2]) and stage (IIIA vs IIIB) as stratification factors. Overall survival and PFS were estimated using the **Kaplan-Meier method**; hazard ratios (HRs) were calculated by **Cox regression**, and response rates were compared by the



臨床研究で用いられる生存時間解析(糖尿病)

Statistical analysis

We assessed the primary noninferiority hypothesis by determining whether the upper boundary of the two-sided 95% confidence interval of the hazard ratio for the risk of the primary composite cardiovascular outcome did not exceed 1.30 in the sitagliptin group, as compared with the placebo group, in the per-protocol population, with a key supporting analysis in the intention-to-treat population. The statistical analysis plan



サンプルサイズ計算！

- 10年後の再発を10%下げる
- 何もしない時の10年無再発割合を80%
- 5年登録、10年追跡
- α エラー=0.05、検出力=80%

The image displays five sequential screenshots of a mobile application interface for sample size calculation. The app is titled 'ICR臨床研究入門' (ICR Clinical Research Introduction).

- Screenshot 1 (13:15):** Home screen with a navigation menu. The 'サンプルサイズ計算' (Sample Size Calculation) option is highlighted.
- Screenshot 2 (13:12):** The 'ツール' (Tools) menu with 'サンプルサイズ計算' selected.
- Screenshot 3 (13:12):** The 'サンプルサイズ計算' screen showing a list of comparison methods. The first option, '2群の生存期間の比較 (優越性に基づく)' (Comparison of survival periods between two groups based on superiority), is selected.
- Screenshot 4 (2:21):** The input screen for the selected comparison method. Fields include: 'サンプルサイズ' (Sample size), '検出力' (Power), '登録期間(年)' (Registration period in years), 'フォローアップ期間(年)' (Follow-up period in years), '考慮する生存割合の年数(何年時点での生存割合か?)' (Number of years to consider survival ratio), '対照群の生存割合' (Control group survival ratio), and '試験群の生存割合' (Treatment group survival ratio). A '結果表示' (Show results) button is at the bottom.
- Screenshot 5 (2:21):** The output screen showing the calculated results for the input parameters: '検出力' (Power) = 0.05, '登録期間(年)' (Registration period) = 5, 'フォローアップ期間(年)' (Follow-up period) = 10, '考慮する生存割合の年数' (Number of years) = 10, '対照群の生存割合' (Control group survival ratio) = 0.8, '試験群の生存割合' (Treatment group survival ratio) = 0.9, 'ハザード比' (Hazard ratio) = 0.4722, '1群あたりのサンプルサイズ' (Sample size per group) = 171, '2群を合わせたサンプルサイズ' (Total sample size) = 342, and '2群を合わせた期待イベント数' (Total expected events) = 62.

生存時間タイプのエンドポイントの例

- 糖尿病、脂質異常症、高血圧領域の大規模臨床試験
 - 心血管発症リスク(致死性・非致死性心筋梗塞、冠動脈疾患、致死性・非致死性脳卒中、全死亡)
- がん
 - 全死亡、無再発生存時間、無増悪生存時間

まとめ

- 臨床的課題を解決するためには、
 - 本講義のフレームワークに落とし込み、
 - 臨床的仮説を統計的仮説に変換し、
 - 研究のデザインを作る
- その際、重要なのは
 - クリニカルクエスチョンの重要性の吟味
 - 新治療、標準治療
 - エンドポイント(とそのタイプ)
 - どのくらいの差があれば臨床的に意味があるか

臨床試験のフレームワーク

1. クリニカルクエスチョン/臨床的意義

WHY

2. 対象

WHO

3. 標準治療

WHAT

4. 新治療

5. 研究デザイン

HOW

エンドポイント

判断規準

自分のプロトコールやクリニカル
クエスチョンをフレームワークに
落とし込み、サンプルサイズを
計算してみましょう。

サンプルサイズはICRwebのスマホア
プリで計算できます(iOS, Androidとも)。



おさらい

統計の講義を聞く際に気を付けること

- 例が自分の分野のものでなくても、自分の知っている例に置きなおして考える
- わからない部分があってもめげずに、全体を理解するようにする
 - 細かいことは気にしない
 - 聞いているうちに慣れる

今日の統計の話をもう少しきちんと勉強したい人はICRWEBの生物統計基礎セミナーをご覧ください。

謝辞

本日の講義作成にあたり、慶応大学の松嶋由紀子先生にがん以外の事例を教えてくださいました。また、本日の講義のいくつかはJCOG臨床試験セミナーや国立がん研究センター生物統計部主催のセミナーで講師の先生方が作成されたものです。謹んでお礼を申し上げます。



おしまい