

臨床研究と生物統計

令和6年度 厚生労働省 臨床研究総合促進事業
医師研修
神戸大学医学部附属病院

2024/12/8



臨床研究のデザイン

令和6年度 厚生労働省 臨床研究総合促進事業
医師研修
神戸大学医学部附属病院

2024/12/8



臨床研究の目的

➤研究のゴール

薬事承認 / ガイドライン収載 / 論文発表 / 学位取得 / 学会発表

➤研究の内容における目的

- ・ 因果関係を調べたい（新しい治療方法は有効なのか）
- ・ 予後に関連ある因子を調べたい
- ・ 診断方法を確立したい（感度・特異度など）



臨床研究の目的

➤研究のゴール

薬事承認 / ガイドライン収載 / 論文発表 / 学位取得 / 学会発表

➤研究の内容における目的

- ・ 因果関係を調べたい（新しい治療方法は有効なのか）
- ・ 予後に関連ある因子を調べたい
- ・ 診断方法を確立したい（感度・特異度など）

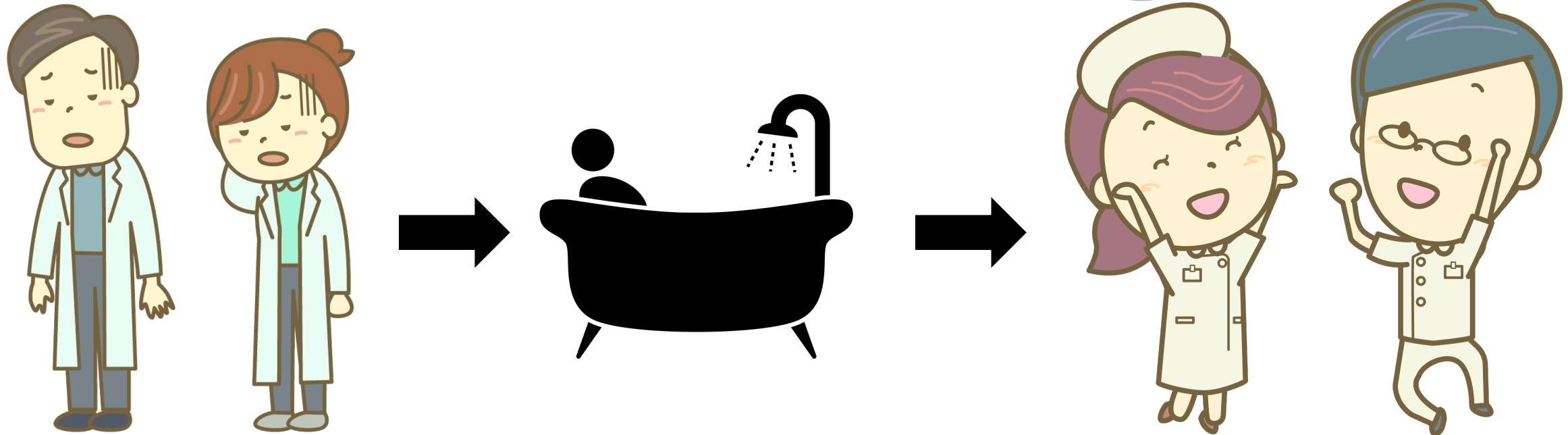
介入研究における
研究目的



治療法が“有効”であるとは

入浴をすると腰痛が改善します。
毎日入浴をして腰痛のない生活を！

入浴が腰痛に対して有効であることを示しているでしょうか？



治療法“有効”であるとは

入浴をすると腰痛が改善します。

毎日入浴をして腰痛のない生活を！・・・

毎日入浴しない人も
腰痛のない生活を
送っている？

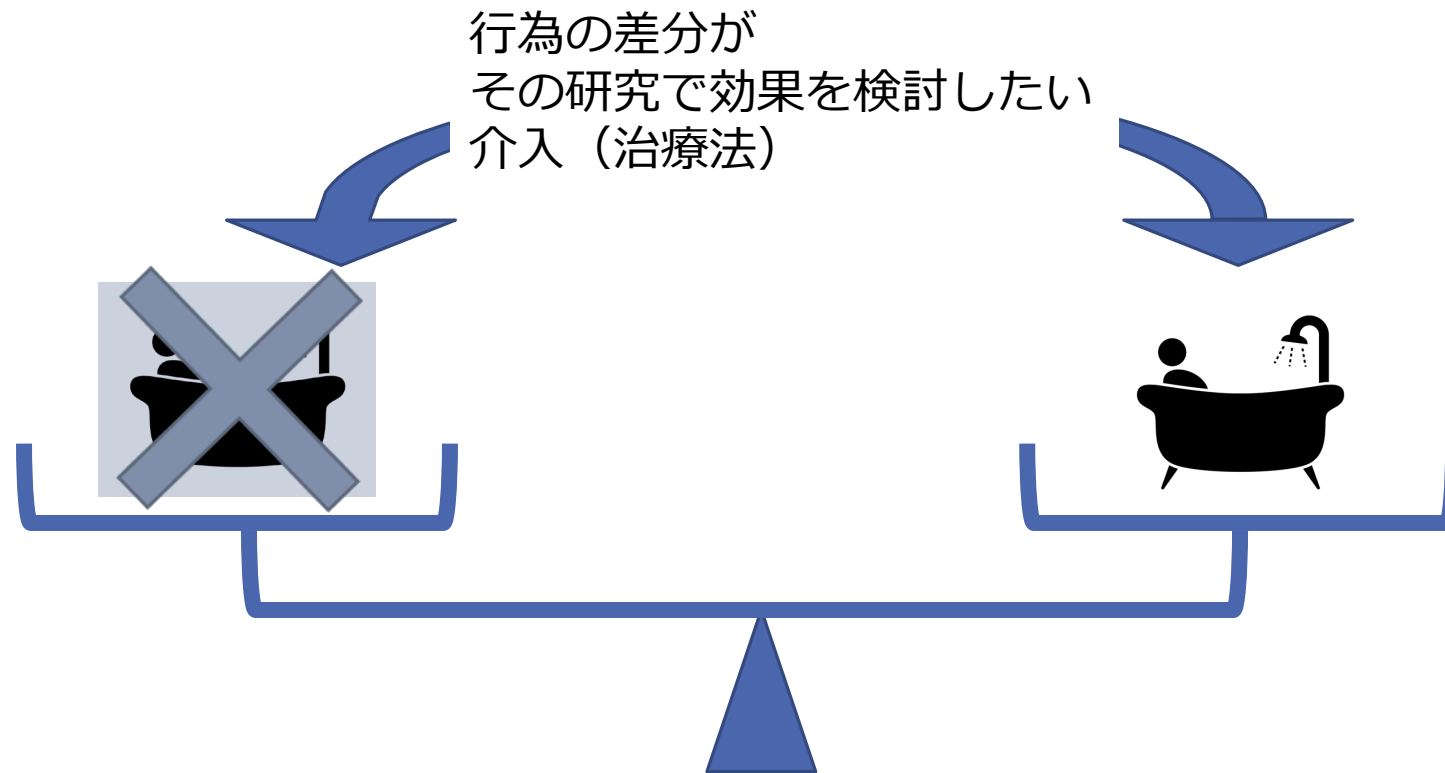


- 入浴をしない場合と比較して、入浴した場合の方がよいことを示す。

※入浴をしない場合：他の方法（清拭、シャワー浴など）
まったく入浴をしない。入浴の条件を変えるなど



比較して検討する



比較対照の種類

臨床研究以前の期間



①臨床研究に参加して
毎日入浴した対象者

臨床研究の期間



比較対照の種類_同じ対象者



①研究に参加する前の
研究対象者における
毎日入浴しない時期

臨床研究以前の期間



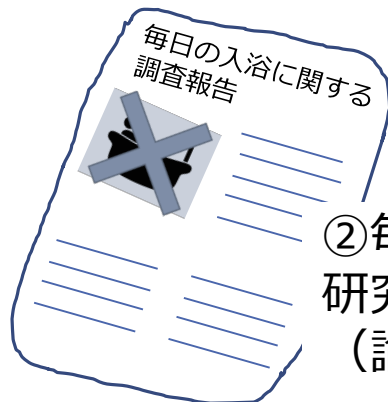
⑩臨床研究に参加して
毎日入浴した対象者

臨床研究の期間



比較対照の種類_既報

臨床研究以前の期間



②毎日入浴しない状態の研究結果
(論文、手元データなど)



①臨床研究に参加して
毎日入浴した対象者

臨床研究の期間



比較対照の種類_研究の中

臨床研究以前の期間



①臨床研究に参加して
毎日入浴した対象者

臨床研究の期間



③臨床研究に参加して
毎日入浴しない対象者



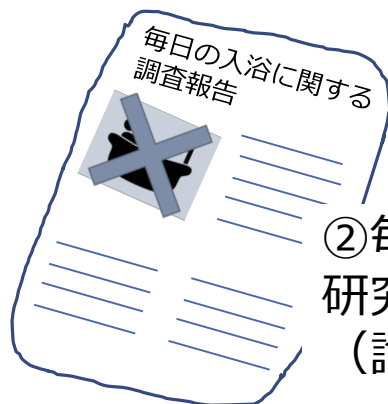
比較対照の種類



単群試験
前後比較試験

①研究に参加する前の
研究対象者における
毎日入浴しない時期

臨床研究以前の期間



単群試験
ヒストリカル
コントロール

②毎日入浴しない状態の
研究結果
(論文、手元データなど)



①臨床研究に参加して
毎日入浴した対象者

臨床研究の期間

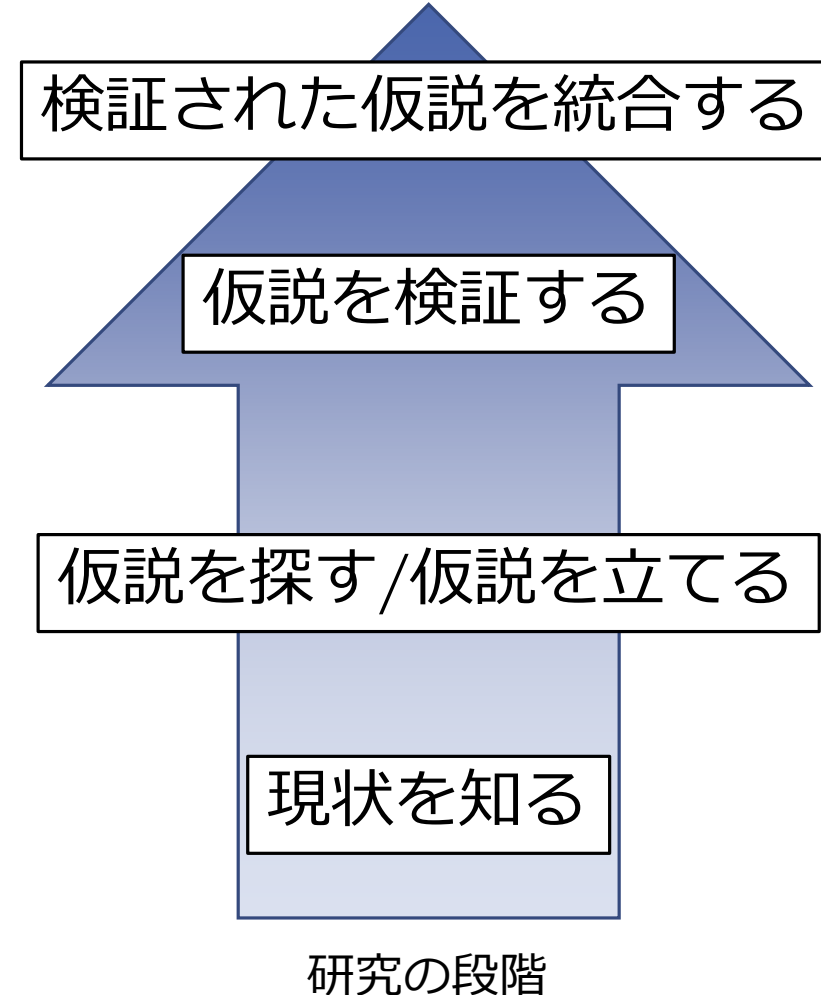
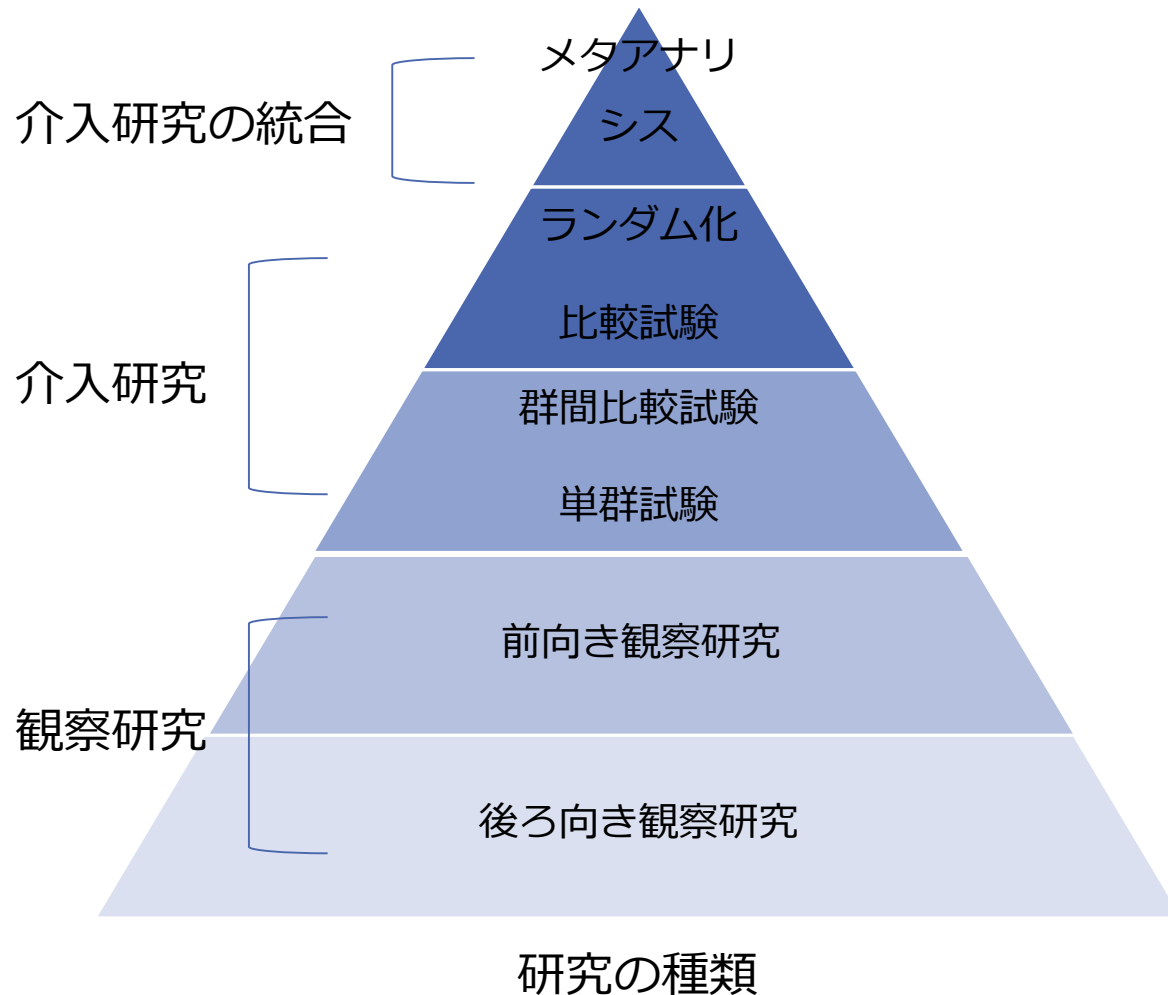


並行群間比較試験

③臨床研究に参加して
毎日入浴しない対象者

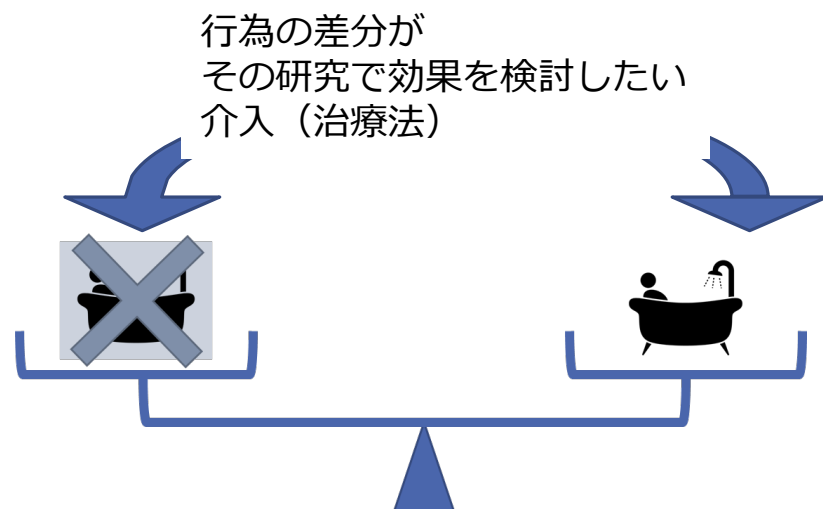


比較の対象とエビデンスレベル

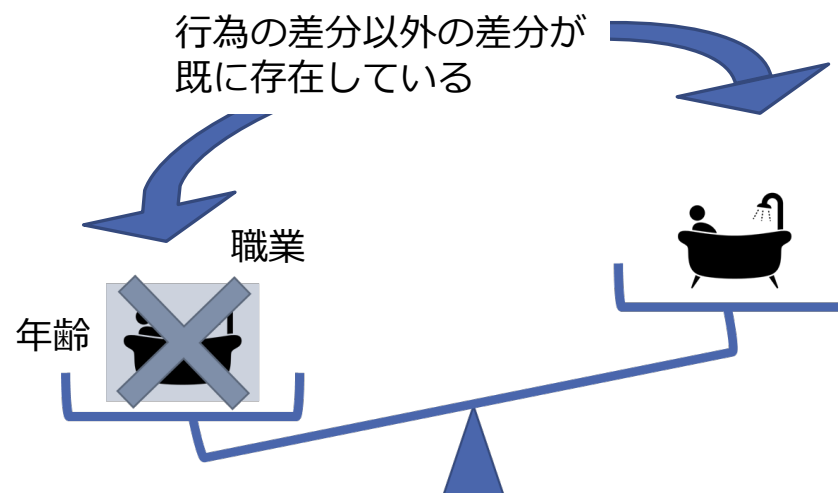


適切に比較をするために、

適切な比較



不適切な比較



比較の背景がそろっている必要がある。

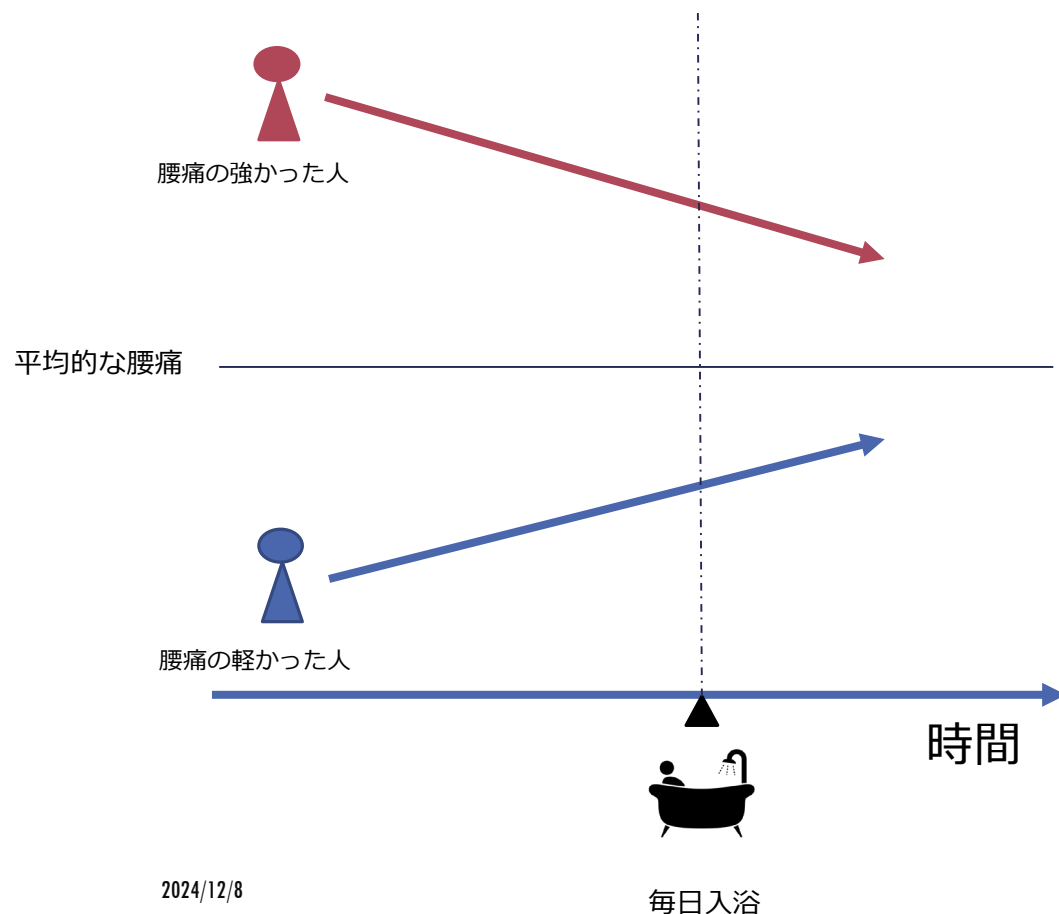


比較の背景をそろえる。

- ・ 個人で揃える ⇒ 同じ人で比較するので背景はおおよそ揃う。
前後比較試験だと、時間的な要素はそろわない。
疾患の状態が同じかどうかは不明。
後の治療効果の評価時に、前の治療効果が残る可能性。
平均への回帰が生じうる（後述）
- ・ 集団で揃える ⇒ いろんな人がいるけれど、毎日入浴ありとなしのグループの平均としてそろえる。
 - ・ 単群試験：比較になるデータが今回の研究と同じ背景の対象者かどうかについては注意が必要。
 - ・ ランダム化比較試験：その試験中での比較の妥当性は担保できる。



平均への回帰



「平均への回帰」

⇒「ある時点に特別に良かったり、悪かったりした値は、次に測定するときには平均的な値に近づく」という現象

前後比較試験での問題点：
毎日入浴前後で比較をしているのか、
「平均への回帰」を見ているのか
区別がつかない。

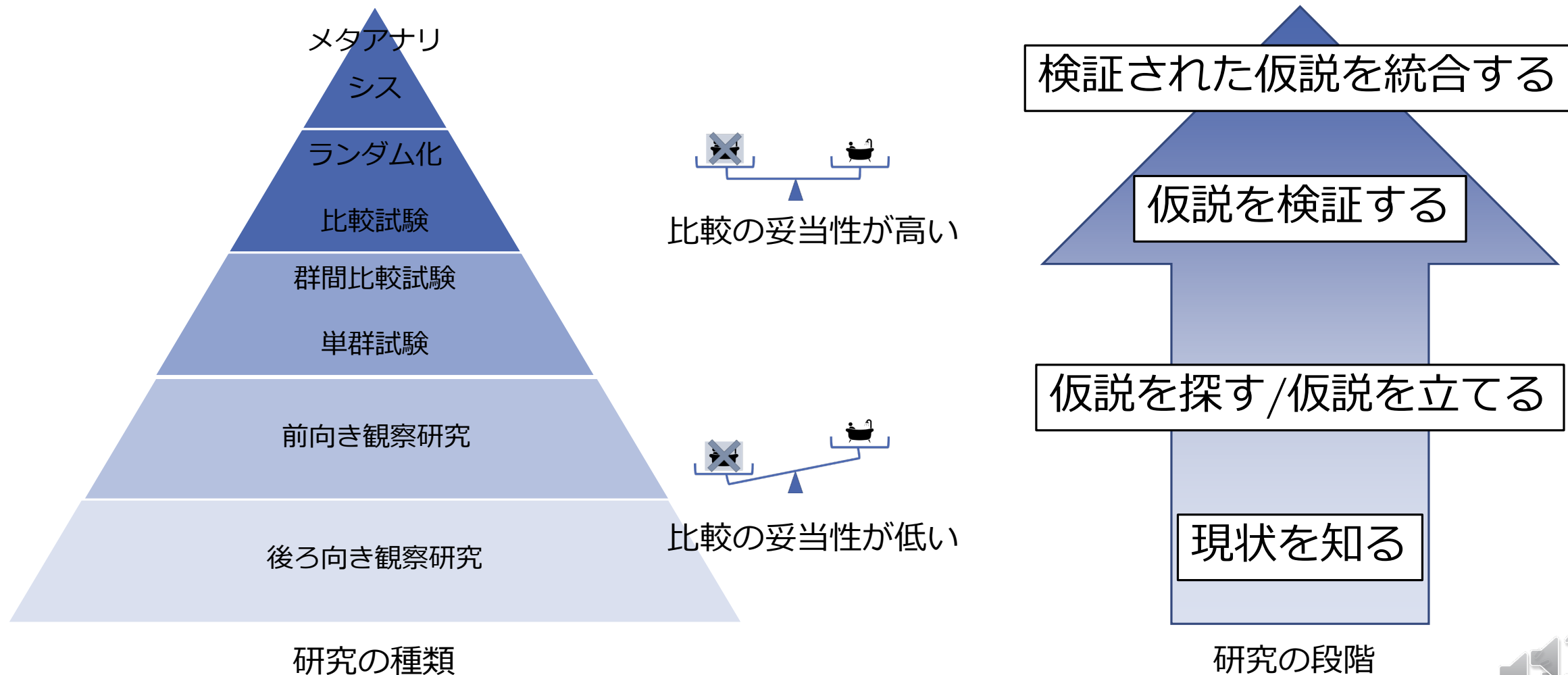


研究デザインと比較の妥当性

	個人毎の背景	平均への回帰	比較の妥当性
単群試験 (前後比較)	○	あり	△
単群試験 (ヒストリカル コントロール)	×	あり	?
並行群間比較試験 (非ランダム化)	×	あるが群間で同じ	△ (確実性に乏しい)
並行群間比較試験 (ランダム化)	×	あるが群間で同じ	○



比較の対象とエビデンスレベル



まとめ

- 治療方法の有効性を検討するには、それ以外の方法の場合と有効性を比較する必要がある。
- 有効性を比較する場合には比較の妥当性を考える必要がある。
(治療法以外の差が生じていないか)
- 比較の妥当性は研究デザインによって異なるためその研究の位置づけから適切な研究デザインを選択する必要がある。



サンプルサイズ

令和 6 年度 厚生労働省 臨床研究総合促進事業
医師研修
神戸大学医学部附属病院

2024/12/8



サンプルサイズに求められること

- 倫理的な観点

対象者の不利益を小さくするサンプルサイズであること

- 科学的な観点

判断可能なサンプルサイズであること

- 実施上の観点

集積可能なサンプルサイズであること



サンプルサイズに求められること

➤倫理的な観点

対象者の不利益を小さくするサンプルサイズであること

➤科学的な観点

判断可能なサンプルサイズであること

➤実施上の観点

集積可能なサンプルサイズであること



対象者の不利益を小さくする

「現在の最も有効な治療法は標準治療」

⇒ 研究に参加することで標準治療を受けられない。

⇒ 研究に参加することで新しい治療に因る未知の有害事象の可能性。

このような不利益を被る患者さんを最小限にしなければならない。

必要最小限の患者さんに研究に参加していただく。



サンプルサイズに求められること

➤倫理的な観点

対象者の不利益を小さくするサンプルサイズであること

➤科学的な観点

判断可能なサンプルサイズであること

➤実施上の観点

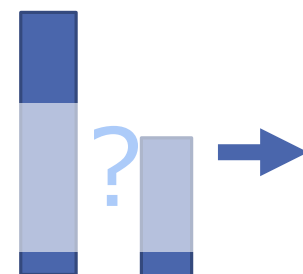
集積可能なサンプルサイズであること



研究における判断

検討したいこと：毎日入浴すると腰痛が改善する。

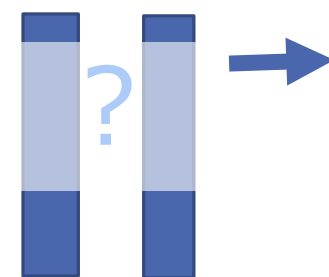
研究実施



研究結果 a)

研究における判断

a) 毎日入浴すると、毎日入浴しない場合に比べて腰痛が改善する。



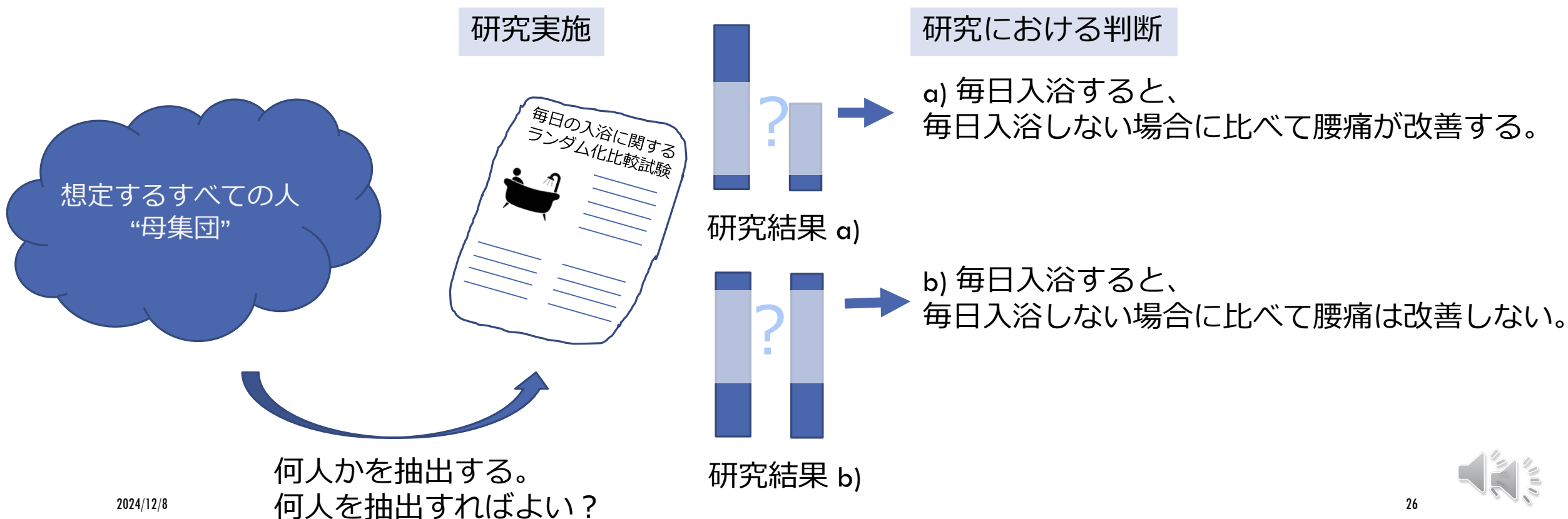
研究結果 b)

b) 毎日入浴すると、毎日入浴しない場合に比べて腰痛は改善しない。



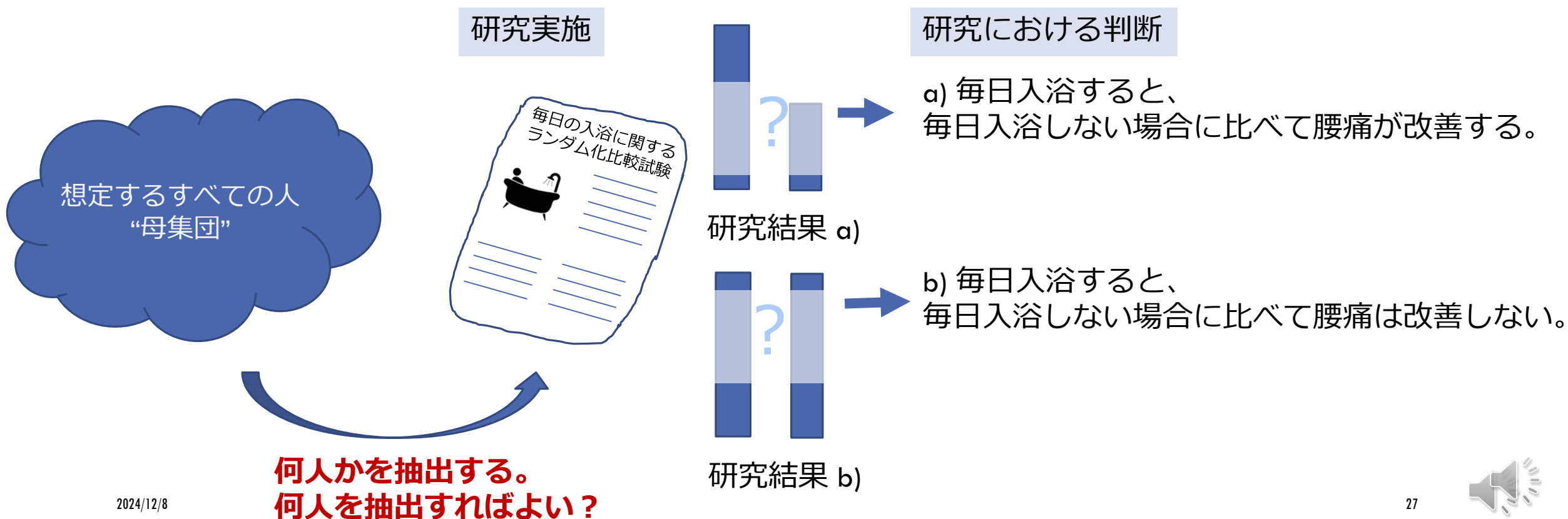
研究における判断

検討したいこと：毎日入浴すると腰痛が改善する。



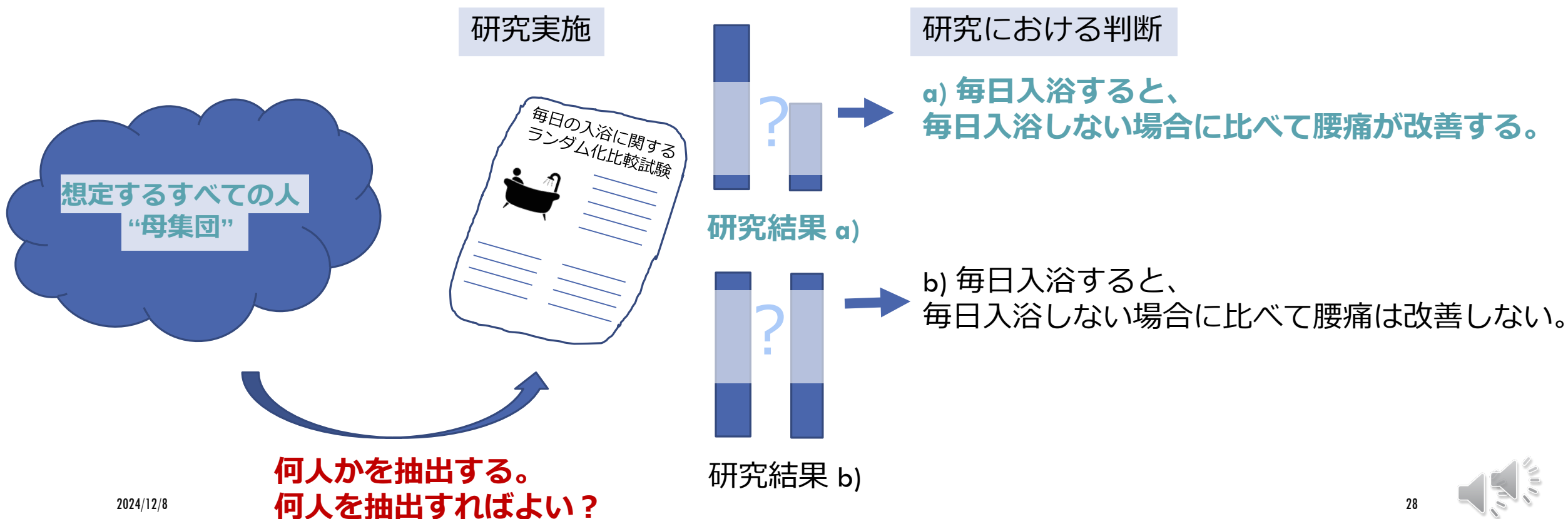
研究における判断

検討したいこと：毎日入浴すると腰痛が改善する。



研究における判断

検討したいこと：毎日入浴すると腰痛が改善する。



サンプルサイズを検討する意味

- サンプルサイズを検討する



- 想定した母集団の下で、期待する研究結果を得、それに基づき判断が可能なサンプルサイズを検討する



サンプルサイズの計算な因子

- 母集団：期待する効果とデータのばらつき
- 研究結果： β エラー（第2種の過誤）
（ $1 - \beta$ エラー=検出力）
- 判断： α エラー（第1種の過誤）



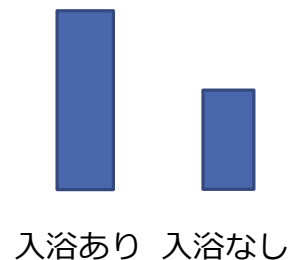
母集団と研究結果

想定するすべての人
“母集団”

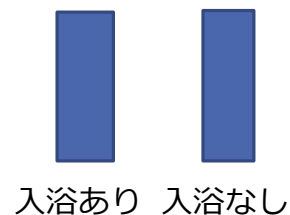
期待する効果とデータのばらつき

- **連続値の時**
平均値の群間差
測定値の標準偏差
- **カテゴリー値の時**
各群の割合
- **生存時間の時**
生存時間中央値または
時点生存割合

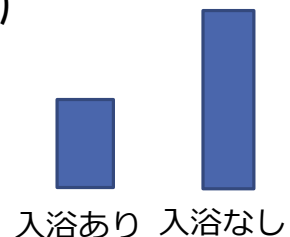
a)



b)



c)



研究結果

同じ母集団から同じ人数で
研究を行っても、
研究結果が毎回母集団の平均と
同じになるとは限らない。

どれくらいの確率で
期待する結果に**ならない**ように
するかを決める。
($\Rightarrow \beta$ エラー)



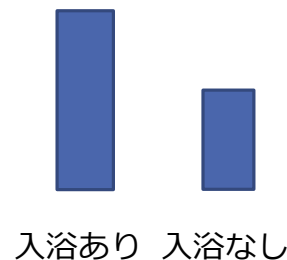
母集団と研究結果

想定するすべての人
“母集団”

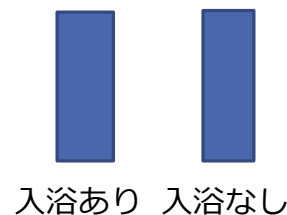
期待する効果とデータのばらつき

- **連続値の時**
平均値の群間差
測定値の標準偏差
- **カテゴリー値の時**
各群の割合
- **生存時間の時**
生存時間中央値または
時点生存割合

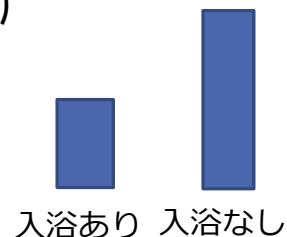
a)



b)



c)



研究結果

同じ母集団から同じ人数で
研究を行っても、
研究結果が毎回母集団の平均と
同じになるとは限らない。

どれくらいの確率で
期待する結果になるように
するかを決める。
($\Rightarrow 1 - \beta$ エラー = 検出力)



期待する効果とサンプルサイズ の関係

- ▶ 連続値の場合（例：血圧、検査値など）
 - 群間の差⇒差が小さいほど必要サンプルサイズは多い
 - 標準偏差⇒標準偏差（ばらつき）が大きいほど必要サンプルサイズは多い
- ▶ カテゴリー値の場合（例：有効/無効など）
 - 割合の差⇒差が小さいほど必要サンプルサイズは多い
 - 各群の割合⇒50%に近くなると必要サンプルサイズは多い
- ▶ 連続値をある基準でカテゴリー値に変換するのであれば、カテゴリー値の方がサンプルサイズは多い
- ▶ 生存時間の場合
 - 観察期間に発生するイベント数でサンプルサイズを検討する
 - ⇒フォロー期間を延ばすことでサンプルサイズは少なくなる



期待する効果とサンプルサイズ の関係

主要評価項目が決まらなければ
効果の大きさも決まらない。

- ▶ 連続値の場合（例：血圧、検査値など）
 - 群間の差⇒差が小さいほど必要サンプルサイズは多い
 - 標準偏差⇒標準偏差（ばらつき）が大きいほど必要サンプルサイズは多い
- ▶ カテゴリー値の場合（例：有効/無効など）
 - 割合の差⇒差が小さいほど必要サンプルサイズは多い
 - 各群の割合⇒50%に近くなると必要サンプルサイズは多い
- ▶ 連続値をある基準でカテゴリー値に変換するのであれば、カテゴリー値の方がサンプルサイズは多い
- ▶ 生存時間の場合
 - 観察期間に発生するイベント数でサンプルサイズを検討する
 - ⇒フォロー期間を延ばすことでサンプルサイズは少なくなる



サンプルサイズの計算な因子

- 母集団：期待する効果とデータのばらつき
- 研究結果： β エラー（第2種の過誤）
（ $1 - \beta$ エラー=検出力）
- 判断： α エラー（第1種の過誤）



判断

研究の判断を「検定」の p 値に基づく判断を行う場合を考えます。
検定における判断の手順。

「否定したい状況（帰無仮説）の下で
研究結果が起こる確率 p 値を計算する。
 p 値がある値（有意水準）以下であれば、
母集団は否定したい状況ではない、と判断する。」

否定したい状況（帰無仮説）：毎日入浴の腰痛に対する効果は〇〇と等しい。
期待する結果（対立仮説）：毎日入浴の腰痛に対する効果は〇〇と比べて異なる。



判断

研究の判断を「検定」の p 値に基づく判断を行う場合を考えます。

検定における判断の手順。

「否定したい状況（帰無仮説）の下で
研究結果が起こる確率 p 値を計算する。」

p 値がある値（有意水準）以下であれば、
母集団は否定したい状況ではない、と判断する。」

サンプルサイズを検討するときには、
 α エラー（第1種の過誤）



臨床研究で起こり得る間違い

		母集団	
研究の判断		効果に差がある	効果に差がない
	p値 ≤ 有意水準 効果に差があると判断する	期待する結果であり正しい判定	α エラー(第1種の過誤)
	p値 > 有意水準 効果に差がないと判断する	β エラー(第2種の過誤)	期待する結果ではないが 正しい判定



臨床研究で起こり得る間違い

研究の判断	母集団	
	効果に差がある	効果に差がない
	期待する結果であり正しい判定	α エラー(第1種の過誤)
p値 $\leq$ 有意水準 効果に差があると判断する	期待する結果であり正しい判定	α エラー(第1種の過誤)
p値 $>$ 有意水準 効果に差がないと判断する	β エラー(第2種の過誤)	期待する結果ではないが正しい判定



臨床研究で起こり得る間違い

研究の判断	母集団	
	効果に差がある	効果に差がない
p値 $\leq$ 有意水準 効果に差があると判断する	期待する結果であり正しい判定	α エラー(第1種の過誤)
p値 $>$ 有意水準 効果に差がないと判断する	β エラー(第2種の過誤)	期待する結果ではないが 正しい判定



臨床研究で起こり得る間違い

		母集団	
研究の判断		効果に差がある	効果に差がない
	p値 ≤ 有意水準 効果に差があると判断する	期待する結果であり正しい判定	α エラー(第1種の過誤)
	p値 > 有意水準 効果に差がないと判断する	β エラー(第2種の過誤)	期待する結果ではないが 正しい判定

これらのエラーを含んだうえで
研究結果に基づき何らかの判断を下すために、
サンプルサイズを検討しています。

α エラー：一般に両側5%
 β エラー：一般に20%
検出力として80%



サンプルサイズに求められること

➤倫理的な観点

対象者の不利益を小さくするサンプルサイズであること

➤科学的な観点

判断可能なサンプルサイズであること

➤実施上の観点

集積可能なサンプルサイズであること



集積可能であること

研究開始後に集積できなければ、

研究体制の変更

少ない症例数で中止

延々と期間延長

研究開始前にできること

適格基準にあう対象者数の推定

実施体制の構築

主要評価項目の再検討

研究の位置づけの再検討

研究デザインの再検討



まとめ

- ▶ 対象者の不利益を小さくするサンプルサイズである必要がある。
- ▶ 一定の間違いを含んだうえで研究結果から判断をくだせるようにサンプルサイズの検討を行う。
- ▶ 研究開始前に集積可能かどうかを十分に検討し、
困難が予想される場合には研究計画を再検討する。

