

因果推論の立場から見たEstimands

- “Per-Protocol” Causal Estimands -

東京大学大学院医学系研究科
公共健康医学専攻 生物統計学
松山 裕

なぜ、補遺を作成？

- ICH E9からかなり時間がたった
 - 当時(1998年)は、まだ議論が十分なされていなかった統計的問題
 - 欠測値、因果推論、感度分析、...
 - ICH E9の一つの大きな目的
 - ランダム化、解析対象集団(FAS)、ITT原則などの概念の導入
- 臨床試験結果と実臨床との乖離
 - 臨床試験から得られる治療効果には実臨床上の意味があるのか？
 - 割付治療の効果
 - レスキュー治療を受けなかったとしたら得られたはずの治療効果など
 - 臨床現場での実際の治療状況(方針)を念頭とした“結果”を、臨床試験(特に、治験)でも求めるべきでは？

2型糖尿病治療薬の例

- 第III相プラセボ対照ランダム化比較試験
 - SGLT-2阻害薬(フォシーガ)
- エンドポイント
 - HbA1cのベースラインからの変化: $\text{HbA1c}_{24\text{week}} - \text{HbA1c}_{0\text{week}}$
 - HbA1cの測定: 0, 4, 8, 12, 16, 24週
- レスキュー治療
 - 試験期間中に血糖値(HbA1c, FPG)がある基準以上を超えた場合には使用可
- 解析対象集団
 - 少なくとも1回は投与を受けて、 $\text{HbA1c}_{0\text{week}}$ とそれ以降の値が少なくとも1回は測定されている全てのランダム化された集団(FAS)
- 統計解析方法
 - LOCF補完した共分散分析
(レスキュー治療後のデータは測定しているが、それらを欠測値とみなす)

FDAのStatistical Reviewerからのコメント

“While FDA has implicitly endorsed LOCF imputation for diabetes trials in the past, there is now more awareness in the statistical community of the limitations of this approach. [...]

My own preferred analysis **simply uses the observed values of patients who were rescued**. This approach may seem counterintuitive if one believes that rescue treatment makes the subsequent outcomes less relevant to evaluation of the test agent. It has the virtue, however, of **respecting the intent-to-treat principle**, in the sense that the analysis is based on the randomized treatment rather than the treatment actually received (i.e., planned treatment plus rescue).”

Holzhauer et al. Pharmaceut. Stat. 2015; 14: 433-47.
Akacha et al. Statist. Med. 2017; 36: 5-19.

LOCF法に対する誤解

- 治療効果に関して保守的な結果が得られる
 - 非保守的な治療効果が得られることもある
 - データのバラツキを無視しているので、 α エラーが担保される保証もない
- レスキュー治療を受けなかった場合の治療効果が得られる
 - 「レスキュー治療後のアウトカムの推移は最後に観察された値のまま変化しない」という非常に強い、非現実的な仮定が必要
 - そのような仮定を保証するだけの生物学的根拠は存在しないのが通常
 - LOCF解析の仮定が合理的と思える状況
 - 症状改善、あるいは治癒したことを理由に脱落した場合

Ware. N Engl J Med 2003; 348: 2136-2137.

Little, Rubin. Statistical Analysis with Missing Data, 2nd ed. John Wiley and Sons, 2001.

感度分析

- スポンサー側
 - Complete-case解析
 - レスキュー治療を受けることなく $\text{HbA1c}_{24\text{week}}$ の測定がなされている対象者に限った解析
 - レスキュー治療以降のデータを欠測値とみなしたMMRM解析
 - MMRM: mixed-effects model for repeated measures

Holzhauser et al. Pharmaceut. Stat. 2015; 14: 433-47.

感度分析

- FDA

- レスキュー治療以降のデータを使用したMMRM解析

- For the purpose of the benefit risk assessment **the smaller treatment difference** between dapagliflozin and placebo resulting from this analysis **should be used**.
 - “The fact that one sensitivity analysis that I will show includes observations made after a rescue may seem counterintuitive. After all, one might reason that the subject’s outcome becomes irrelevant to the evaluation of the original treatment once a rescue treatment is given. However, the widely recognized **intent-to-treat principle says that the statistical analysis should be based on the randomized treatment** rather than the actual, non-randomized treatment that a subject received. [...] Once we attempt to adjust for rescue in any way or exclude the data, **we are endangering the validity of the analysis**”

Holzhauer et al. Pharmaceut. Stat. 2015; 14: 433-47.

ITT信者?

- ITT検定
 - ランダムでない治療法変更があったとしても、sharp causal null に対する**妥当な α レベル検定**
- ITT治療効果
 - 治療方針 (treatment policy) の比較?
 - Treatment policy: “Assign to treatment, then do whatever you want”
 - 正確には、
臨床試験で観察された治療法変更内容・割合、治療法変更に影響を与える要因などが実臨床でも同じだと仮定した場合に観察される効果
 - 仮想的な治療効果
 - **因果治療効果 (causal treatment effect) には一致しない**
 - Biased estimate for causal estimands
 - “Black-box” 治療方針なので、因果効果としては不十分

問題は、ランダム化後の事象 post-randomization events

- ICH E9(R1)では、中間事象 (intercurrent events)
 - 治療開始後に発現し、変数を観測できなくする、又は変数の解釈に影響を与える事象 (R1 用語集より)
 - 例えば、代替治療 (レスキュー薬、治験実施計画書における併用禁止薬、又は後続ラインの治療) の使用、治療の中止、治療の切り替え、状況によっては死亡といった終末事象が含まれる (R1 A.1. 15-18行)
- 本来は、中間事象としてひとまとめにすべきではないと思う
 - プロトコルで規定される治療戦略 (方針・計画) に関わる中間事象
 - プロトコル違反となる中間事象

E9(R1)で提案されているフレームワーク

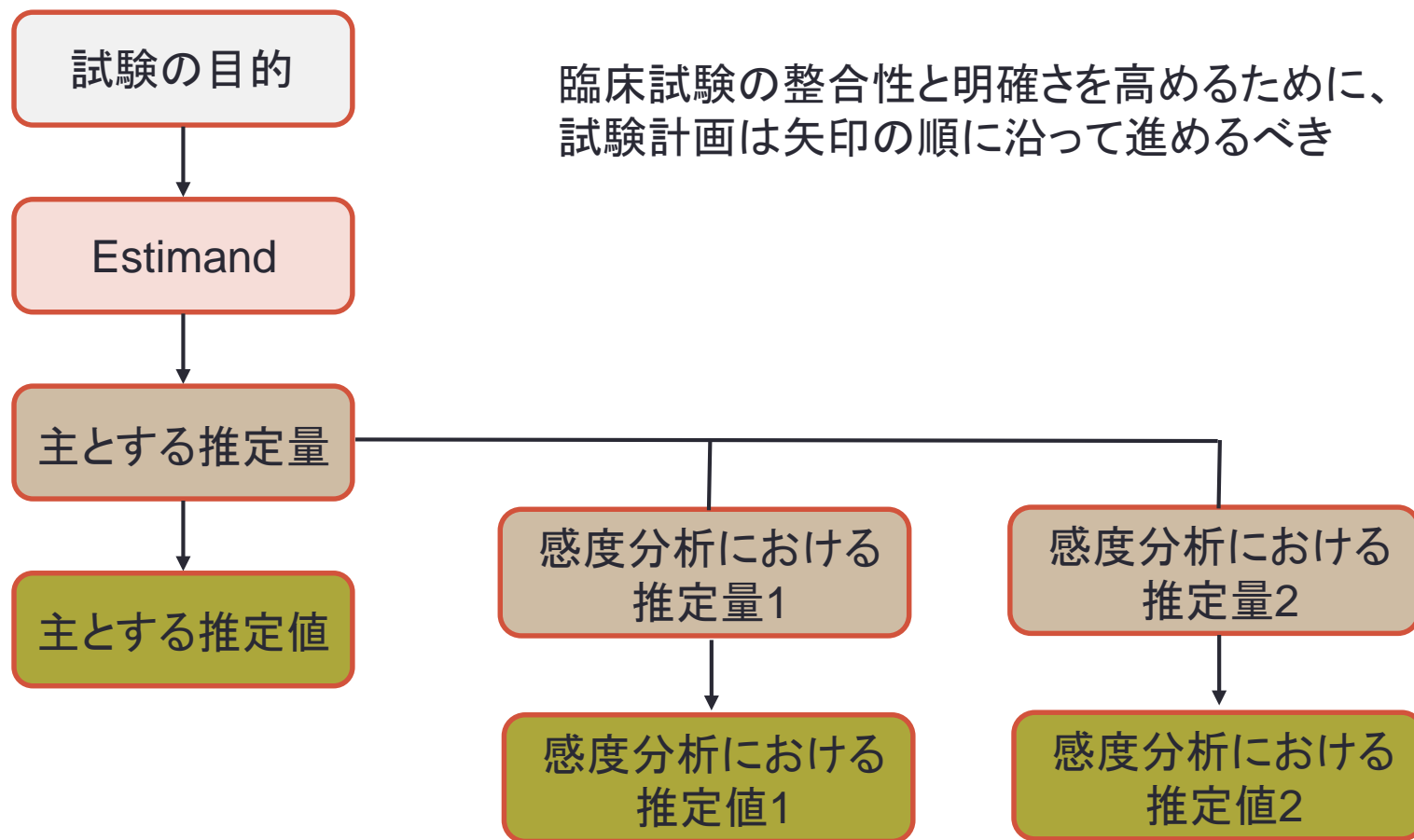


図1. 定められた試験の目的に応じて、推定の対象、推定の方法及び感度分析を整合させる 10

Estimandとは?

- What is to be estimated.
 - ターゲットパラメータ (Target parameter)
- E9(R1)でのestimandの4つの構成要素
 - Population
 - Patients targeted by the scientific question
 - Variable
 - Endpoint to be obtained for each patient, that is required to address the scientific question
 - **Intercurrent events**
 - Specification of how to account for intercurrent events to reflect the scientific question of interest
 - Summary measures
 - Population-level summary for the variable

中間事象に対応するためのストラテジー

ICH E9(R1)では、以下の5つが紹介されている

- 治療方針ストラテジー (Treatment policy)
 - ITT治療効果のこと
 - ただし、treatment policy = ITT では「ない」と思う
- 複合ストラテジー (Composite)
 - エンドポイントの設定の問題ではないか？
- 仮想ストラテジー (Hypothetical)
 - 中間事象がなかった時に得られる治療効果のこと
 - ただし、hypothetical = Causal estimands では「ない」と思う
- 主要層ストラテジー (Principal stratum)
 - 中間事象が起きない仮想的なサブグループでの治療効果
- 治療下ストラテジー (While on treatment)
 - 中間事象の発現前までの治療に対する反応を関心の対象とする
 - AUCやtime-to-failureのようなもの

中間事象に対応するための戦略

ICH E9(R1)では、以下の5つが紹介されている

- 治療方針戦略 (Treatment policy)
 - ITT治療効果のこと
 - ただし、treatment policy = ITT では「ない」と思う
- 複合戦略 (Composite)
 - エンドポイントの設定の問題ではないか?
- 仮想戦略 (Hypothetical)
 - 中間事象がなかった時に得られる治療効果のこと
 - ただし、hypothetical = Causal estimands では「ない」と思う
- 主要層戦略 (Principal stratum)
 - 中間事象が起きない仮想的なサブグループでの治療効果
- 治療下戦略 (While on treatment)
 - 中間事象の発現前までの治療に対する反応を関心の対象とする
 - AUCやtime-to-failureのようなもの

別々に表記すべきものではない

因果推論の枠組みのもとでは、全てのestimandsは、治療方針の比較であり、仮想的な治療効果

ExplanatoryかPragmaticか

- Explanatoryな(説明的)臨床試験
 - 介入法の作用機序などを解明する目的で、実験条件(施設/対象者選択、データ測定など)をある程度厳しく設定して実施する試験
 - 薬剤の有効性(**efficacy**)を実験室的環境で検証
 - 治療法変更、併用療法などは許さない(**プロトコルで規定**)
- Pragmaticな(実践的)臨床試験
 - 実施条件を緩く設定し、日常診療に近い状況で介入法を評価するために実施する試験
 - 実践医療の中での薬剤の有用性(**effectiveness**)を検証
 - 治療法変更、併用療法などは許される(**プロトコルで規定**)

Piantadosi A (1997), Clinical Trials, Wiley

臨床試験と日常診療の乖離

- 限定された対象者
 - 施設選択
 - 患者選択/除外基準
 - インフォームド・コンセント
 - ...
- 計画されたデータ測定・管理
- プラセボの使用
- ...

環境や状況が大きく異なるのは当然

有効性 (Efficacy) の評価

- 臨床試験 (特に、NDA) においては重要
 - NDA: New Drug Application
- 問題は、
 - プロトコルで規定される許容されない治療法変更・併用薬投与などを遵守 (adhere) できるかどうか
 - 不遵守例 = treatment drop-out (analysis drop-out と同じ扱い)
 - 多くの不遵守例が出る場合には、有効性評価としての臨床試験がそもそも成立していない
 - 得られる結果が実臨床での治療効果になっているか
 - 日常診療では、臨床試験環境下とは異なることが起きうる (治療法の変更などは当然行われる)

2種類のdrop-out

- Analysis drop-out
 - いわゆる "Loss-to-follow-up" のこと
 - 欠測値の問題と捉えられることができ、「対象者の追跡ができた場合」という仮想状況を想定可能
- Treatment drop-out
 - 医学的な理由による治療中止 (withdrawal)
 - 様々な理由が考えられる
 - 状態悪化によるレスキュー/後治療、有害事象発現、併用禁止薬の投与など
 - 中止理由によっては、「そのような中止がなかった場合」という仮想状況を想定することに临床上は疑問が生じることもある

有用性 (effectiveness) の評価

- 日常診療、あるいは医学・倫理的観点を考慮すれば重要
- 問題は、
 - 許容される治療法変更・併用薬投与などの治療戦略(方針・計画)を事前にプロトコルに明確に定義でき、それを遵守できるかどうか
 - Treatment drop-out は基本的には存在しない
 - あるとすれば、併用禁止薬の投与
 - Analysis drop-out は存在する
 - できる場合(疾患領域)もあれば、そうでない場合もある
 - 得られる結果が実臨床での治療効果になっているか
 - 規定された治療戦略が時代とともに古くなることもある

以降では、

- Analysis drop-out については考えない
- Treatment drop-outについて以下の3つを考える
 - レスキュー治療/後治療
 - 開始のための明確な基準設定
 - 実臨床でも実行可能な合意された治療方針
 - 有害事象
 - 軽減のための治療(対処)法が存在するか/しないか
 - 併用禁止薬
 - 有害事象を引き起こす可能性がある薬剤 など
 - ただし、ここでは作用機序の異なる他剤や当該薬剤の効果を増加させるような薬剤は含まない(レスキュー治療に含めて考える)

因果推論の枠組みのもとでは、

- EfficacyとEffectiveness
- どちらも統一的に因果治療効果として表現可能
 - ベースラインデータで識別可能な集団における臨床的に興味のある治療戦略
- 因果治療効果 (Causal estimands/treatment effects)
 - 明確に定義された患者集団において実際に実行可能な治療戦略の対比(比較)
 - Contrasts (comparisons) of treatment strategies that may be realistically implemented in well-defined groups of patients

Hernan and Scharfstein, Ann Intern Med 2018; 168: 515-6.

平均因果効果 (ACE: average causal effects)

- 構成要素のPopulationとVariableが明確に定義されれば、例えばSummary measuresとして、

$$E(Y_{g=a1}) - E(Y_{g=a0})$$

- $E(\cdot)$: 集団全体に対する期待値
- 反事実アウトカム: $Y_{g=a}$
 - ある対象者が治療方針 (general treatment) “ $g=a$ ” に従った場合に観察される潜在アウトカム
- 仮想的な治療効果
 - 識別可能にするためにはデータから検証不能な仮定が必要
 - ただし、対象集団は現実のもの (測定されたベースライン特性で定義可能)

治療戦略？

- プロトコルで事前に明確に規定
 - 規定できれば、いわゆる中間事象の取り扱いが定まる
- これまであまり明確にはなされてこなかった
 - なんとなく暗黙の合意？
 - ITT?
 - 特に、レスキュー治療/後治療を含めた治療計画

治療戦略の例

- 中間事象の内容によらず、ずっと割付治療を投与
 - なんとなく「因果効果」と思われている戦略
 - 全ての中間事象がなかった場合に観察される治療効果
- 有害事象による治療中止がない限り、割付治療を投与
 - 有害事象は実臨床でも避けられないので、「有害事象がなかった場合」を考えることには意味がない
- (プロトコルで規定される) 状態悪化後の治療は任意
 - いわゆるITT効果
- 状態悪化後は治療B (B1, B2) のいずれか
 - 動的治療レジメン (dynamic treatment regime)
- ...

治療戦略とnon-adherence

治療戦略	関連する中間事象のプロトコルでの規定			non-adherence
	レスキュー/後続治療	有害事象による治療中止	併用禁止薬	
中間事象に関係なく投与	× ¹⁾	△ ²⁾	○	全ての中間事象
有害事象がない限り投与	× ¹⁾	○ ³⁾	○	レスキュー/併用禁止薬使用例
Black-box	△ ⁴⁾	○ ³⁾	○	併用禁止薬使用例
規定された治療	○	○ ³⁾	○	併用禁止薬使用例

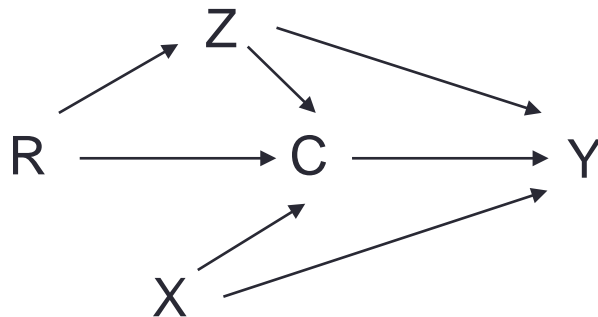
- 1) 併用禁止薬の中で規定される
- 2) 有害事象を軽減するための治療が実臨床では存在する
- 3) 有害事象を軽減するための治療が実臨床では存在しない(プロトコル遵守例なので治療中止後のデータも測定する必要あり)
- 4) 後治療の内容は明確に規定されていない

ACEの推定 中間事象に関係なく投与

- 推定量の意味
 - 全員が予定通りの治療を受け続けた時の治療効果
(レスキュー治療、有害事象、併用禁止薬投与がなかった場合の効果)
- ターゲットとなる対象
 - プロトコル遵守例 (Per-Protocol)
- 推定に際しての問題
 - 観察されたPP集団にはバイアスがある

Post randomization confounding

- ランダム化
 - ベースラインの交絡バイアスを制御可能
 - ランダム化後のイベントによる交絡/選択バイアスは制御不能
- PP解析
 - バイアスがある
 - ランダム化後のイベントによる交絡を調整した解析が必要
 - “PP” causal estimands: XやZによる交絡(選択)バイアス補正済みの効果



R: 割付治療、Y: アウトカム、C: 中間事象、X: ベースラインリスク因子、Z: 時間依存性交絡因子

“PP” Causal Estimands 1

- IPCW 推定量 Inverse Probability of Censoring Weighted
 - Non-adherence時点で打ち切った解析
 - 観察PP集団における重み付き解析
 - 重み: 全対象集団から推定された観察(脱落)確率の逆数
 - Multinomial logistic regression
- 推定のために必要なデータ
 - Non-adherenceに影響する因子(XとZ)
- 仮定
 - 観察確率推定のためのモデルが正しい
 - adherenceに関する未測定の交絡因子が存在しない

“PP” Causal Estimands 2

- g 推定量
 - Randomized g-estimator
 - Observational g-estimator
 - 全対象者に対する因果モデル (SNMs: structural nested models) のあてはめ
- 推定のために必要なデータ
 - Non-adherence後の治療、アウトカムデータ
 - Non-adherenceに影響する因子 (XとZ): Observational g-estimatorの場合
- 仮定
 - SNMsが正しい
 - 未測定の交絡因子が存在しない: Observational g-estimatorの場合

“PP” Causal Estimands 3

- IPTW推定量 Inverse Probability of Treatment Weighted
 - 繰り返し治療の観察研究データだと考える
 - 各治療を受ける確率の逆数で重み付けた解析
 - Multinomial logistic regression
- 推定のために必要なデータ
 - Non-adherence後の治療、アウトカムデータ
 - Non-adherenceに影響する因子(XとZ)
- 仮定
 - 確率推定のためのモデルが正しい
 - 治療法選択に関する未測定の交絡因子が存在しない

Hernan, Lanoy, Costagliola, Robins. Basic Clin Pharmacol Toxicol 2006; 98: 237-42.

ACEの推定 有害事象がない限り投与

- 推定量の意味
 - 有害事象発現がない限り、全員が予定通りの治療を受け続けた時の治療効果(レスキュー治療、併用禁止薬投与がなかった場合の効果)
- ターゲットとなる対象
 - プロトコル遵守例(Per-protocol)
 - 有害事象発現例も含む割付治療を最後まで遵守した症例
 - 有害事象発現例に関してはアウトカムの測定は必要
- 推定に際しての問題
 - 観察されたPP集団にはバイアスがある
- 推定方法
 - IPCW 推定量、g 推定量、IPTW推定量

ACEの推定 Black-box

- 推定量の意味
 - 中間事象以降の治療内容は任意の場合の治療効果（ただし、併用禁止薬投与がなかった場合の効果）
- ターゲットとなる対象
 - プロトコル遵守例 (Per-protocol)
 - レスキュー治療例/有害事象発現例はプロトコル遵守例
- 推定に際しての問題
 - (併用禁止薬使用例を除外した) PP集団にはバイアスがある
- 推定方法
 - IPCW 推定量

ACEの推定 Black-box (ITTと呼ばれている)

- 推定量の意味
 - 中間事象以降の治療内容は任意の場合の治療効果
- ターゲットとなる対象
 - FAS (Full analysis set)
 - レスキュー治療例/有害事象発現例/併用禁止薬投与例についてもアウトカム測定
- 推定に際しての問題
 - 特になし (推定値の臨床的解釈には疑問あり)
- 推定方法
 - 標準的手法

ACEの推定 規定された治療計画

- 推定量の意味
 - プロトコルで規定された治療を受けた場合の治療効果（ただし、併用禁止薬投与がなかった場合の効果）
- ターゲットとなる対象
 - プロトコル遵守例 (Per-protocol)
- 推定に際しての問題
 - (併用禁止薬使用例を除外した) PP集団にはバイアスがある
- 推定方法
 - IPCW 推定量

治療計画の明確な規定

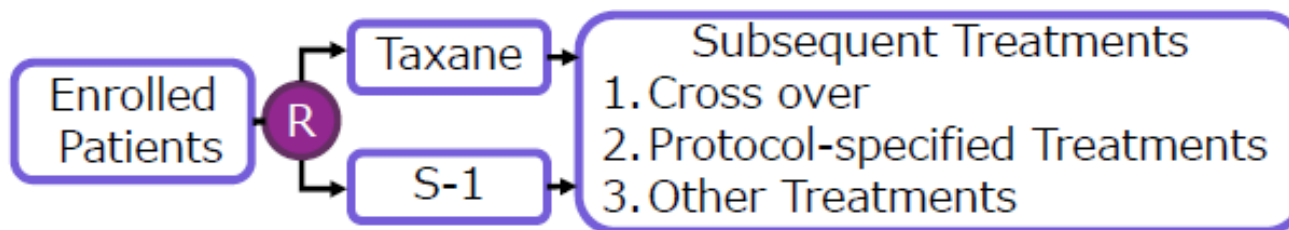
- 動的治療計画 (dynamic/sustained treatment strategies)
 - ベースラインデータで明確に定義可能な患者集団において、実臨床においても実行可能 (治療ガイドラインなどでの規定)
 - 当該の臨床試験においても、大多数の患者が従う計画
- 実臨床を想定した effectiveness estimand
 - Treatment drop-out は基本的には存在しない
 - 併用禁止薬の使用例は多くないのが通常
- Efficacy estimandも検討可能
 - 例えば、プロトコルで規定されるレスキュー治療 (B1, B2) に対して、PP causal estimandとして、
Treat X until progression, then treat B2
vs Treat P until progression, then treat B2

Outcome Trials

- 真のエンドポイントに対する大規模長期イベント試験
 - Pragmatic trials (large simple trial)
 - Treatment drop-out は不可避
- 動的治療計画の明確な規定
 - 可能な場合がある

SELECT BC 試験

- 対象
 - HER2陰性、ホルモン療法抵抗性転移・再発乳がん患者
- 試験デザイン・目的
 - 1次治療としてのタキサン系薬剤に対するTS-1の全生存期間に関する非劣性検証のための第III相ランダム化比較試験
- 疾患増悪後の二次治療に関するプロトコル上の規定
 - 割付治療のクロスオーバー
 - アンスラサイクリン系薬剤など他の抗がん剤



Takashima, et al. Lancet Oncol 2016;17:90-8.

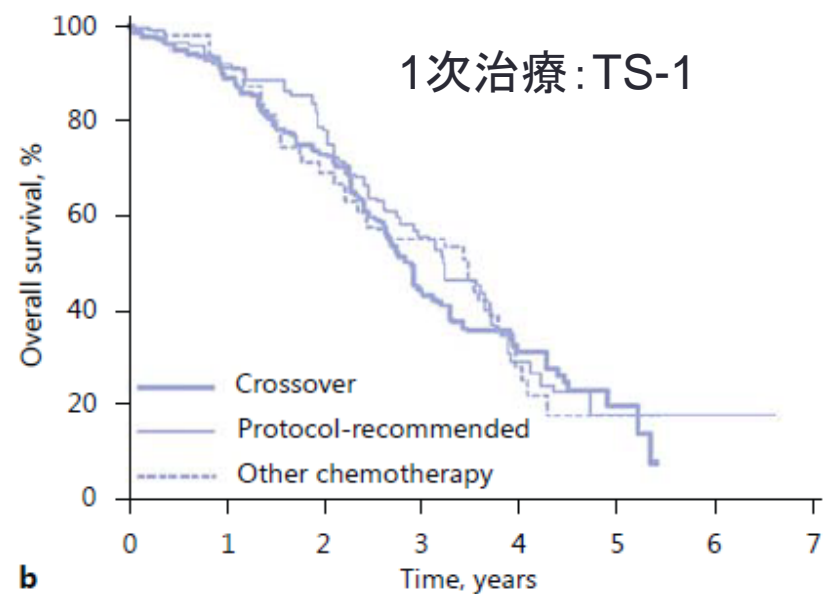
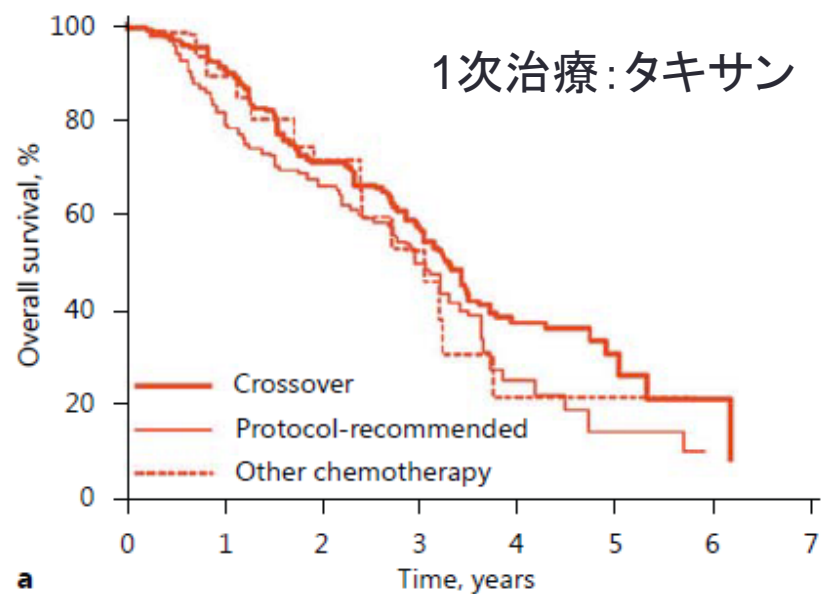
SELECT BC試験での後治療

	タキサン群 n = 286	TS-1群 n = 306
後治療実施例	241 (84.3%)	266 (86.9%)
割付治療のクロスオーバー	110 (38.5%)	114 (37.3%)
プロトコル指定の後治療	76 (26.6%)	68 (28.1%)
プロトコル指定外の後治療	55 (19.2%)	84 (27.5%)

Mukai et al. Oncology 2017; 3(5): 315-22.

“PP” Causal KM estimates

- 動的治療に対するIPTW推定



Mukai et al. Oncology 2017; 3(5): 315-22.

Table 3. Comparison of different chemotherapies

Second-line treatment	Median overall survival, months	Hazard ratio (95% CI) <i>p</i> value for superiority	
		vs. A3	vs. A2
A Taxane as first-line chemotherapy			
A1: Crossover	39.6	0.71 (0.43–1.18) <i>p</i> = 0.183	0.72 (0.52–0.98) <i>p</i> = 0.037
A2: Protocol-recommended	35.7	0.99 (0.61–1.62) <i>p</i> = 0.971	–
A3: Other chemotherapy	36.9	–	–
Second-line treatment	Median overall survival, months	Hazard ratio (95% CI) <i>p</i> value for superiority	
		vs. B3	vs. B2
B S-1 as first-line chemotherapy			
B1: Crossover	34.3	1.06 (0.74–1.53) <i>p</i> = 0.739	1.16 (0.85–1.60) <i>p</i> = 0.350
B2: Protocol-recommended	38.9	0.91 (0.62–1.34) <i>p</i> = 0.644	–
B3: Other chemotherapy	42.0	–	–
C Either taxane or S-1 as first-line chemotherapy			
C1: Crossover	36.5	0.89 (0.66–1.19) <i>p</i> = 0.423	0.92 (0.74–1.15) <i>p</i> = 0.461
C2: Protocol-recommended	37.8	0.96 (0.72–1.29) <i>p</i> = 0.810	–
C3: Other chemotherapy	38.6	–	–

Symptom-related Trials

- 症状改善を目的とした代替エンドポイントに対する試験
 - 実験条件をある程度厳しく設定した試験
 - Short, small, double-blinded, tightly controlled experiments in highly selected patients
 - プラセボの使用もありえる
 - 薬剤の有効性評価が第一義的な目標
- 治療計画の明確な規定
 - 治療ガイドラインでの規定・合意
 - できる疾患領域もあるかもしれないが、特に治験においては難しい？

2型糖尿病治療薬の例では、

- プラセボを使用しているのだから有効性評価が目的?
 - 有害事象の内容によるが、「有害事象による治療中止がない限り投与」
 - Estimand: Treat X (or P) continuously during follow-up, until severe adverse effects arise (even if progression occurs)
- 中間事象の取り扱い
 - 有害事象発現例以外はtreatment drop-out
- レスキュー治療
 - 観察アウトカム (HbA1c、FPG) に依存して開始
- IPCW解析によるPP causal estimand
 - レスキュー治療を受ける確率: $P_{rescue,t} = f(\text{group}, X, HbA1c_{t-1}, FPG_{t-1})$
 - Complete-caseの重み付き解析

Measure of Intervention Effect

Holzhauser et al. Pharmaceut. Stat. 2015; 14: 433-47.

Treatment difference of the originally assigned treatments at randomization based on

- in case of intake of rescue medication hypothetical values if rescue medication had not been given, because interest is in the treatments assigned at randomization rather than treatment regimens including rescue medication as needed;
- in case of discontinuation of study treatment the actual off-treatment values, because we are interested in the effect of assigning the randomized treatment and patients discontinuing treatment, for example, due to an adverse event would not have continued treatment in real clinical practice, either;
- for losses to follow-up hypothetical values as if patients had continued to take part in the trial including continued intake of randomized treatment.

観察アウトカムのみに基づいてレスキュー

- 欠測メカニズムはMAR (missing at random)
 - Estimand
 - 有害事象がない限り、全員が予定通りの治療を受け続けた時の治療効果
 - 打ち切り時点までの全データを使用した解析は妥当
 - レスキュー治療以降のデータを欠測値とみなしたMMRM解析

Murray and Findlay. *Statist in Med* 1988; 7: 941-6.
Little. *JASA* 1995; 90: 1112-21.

- ただし、HbA1cだけでなくFPGにも依存
 - 2変量MMRMのあてはめ

Matsuyama and Ohashi. *Statist in Med* 1997; 16: 1587-601.
Holzhauer et al. *Pharmaceut. Stat.* 2015; 14: 433-47.

E9(R1)で提案されているフレームワーク

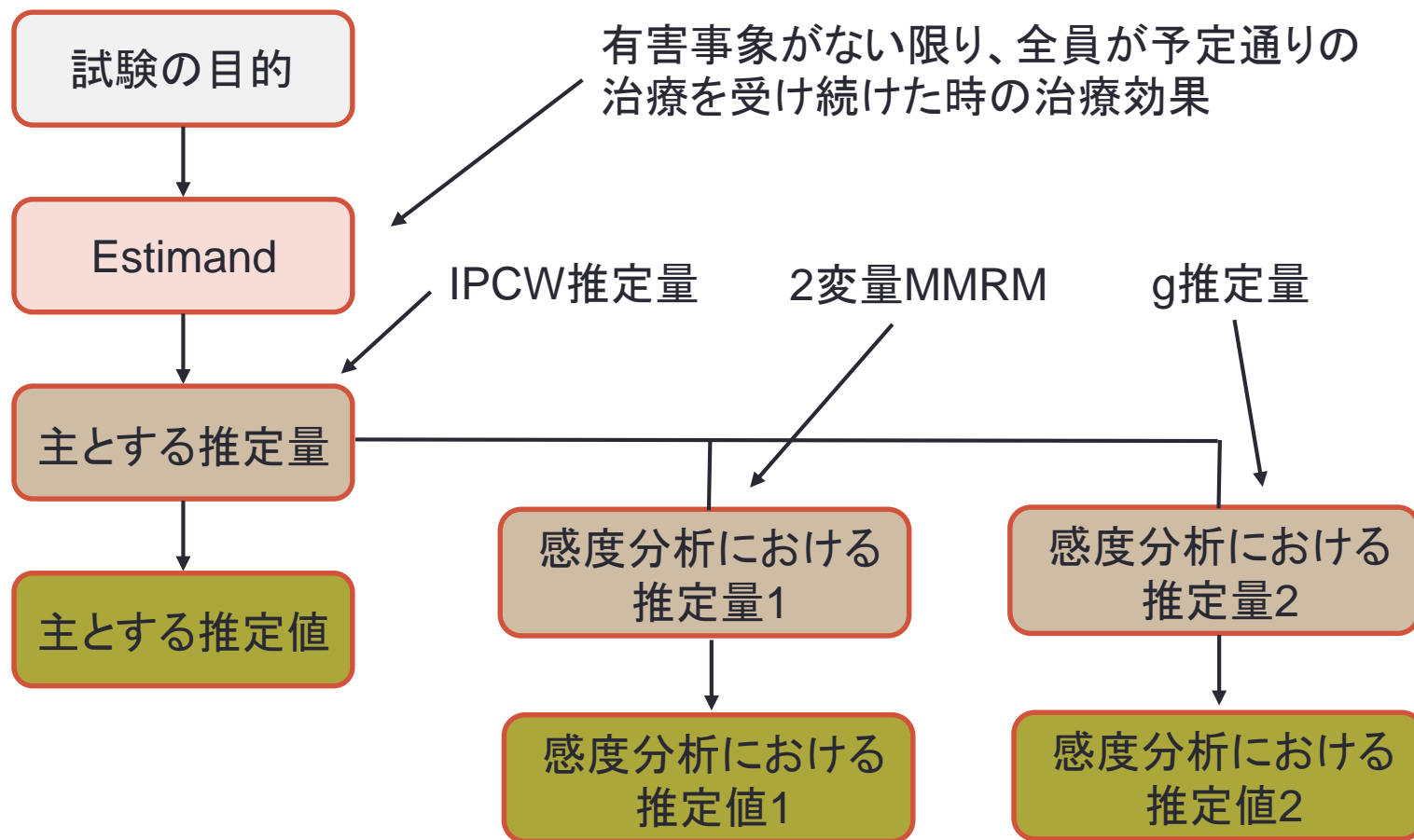


図1. 定められた試験の目的に応じて、推定の対象、推定の方法及び感度分析を整合させる 44

フレームワークで大事なこと

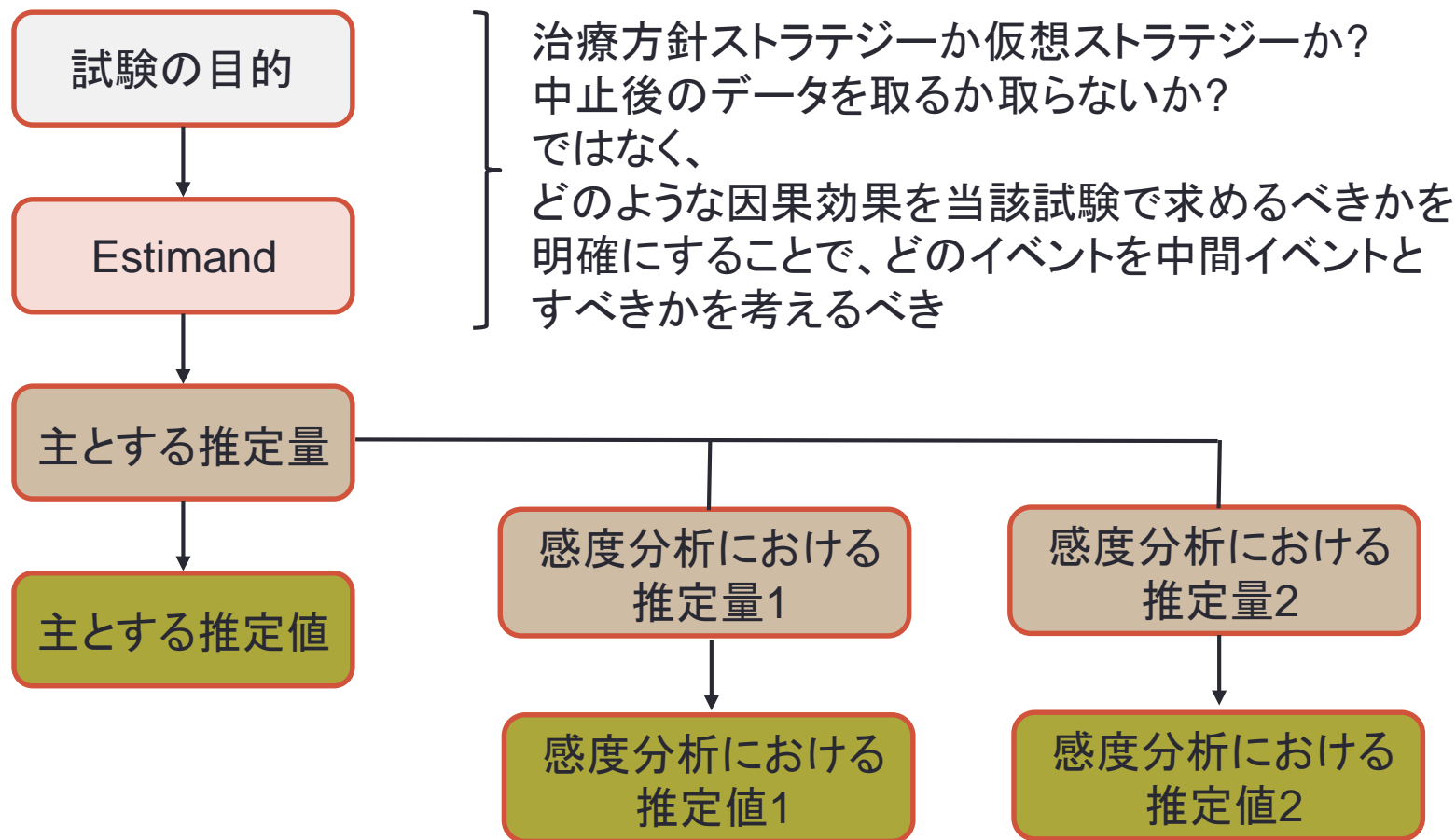


図1. 定められた試験の目的に応じて、推定の対象、推定の方法及び感度分析を整合させる

感度分析について

- 同じEstimandに対する異なる仮定のもとでの統計解析
 - IPCW
 - 観察確率の推定が正しい(adherenceに関する未測定交絡因子が存在せず、観察確率推定のためのモデルが正しい)
 - 二変量MMRM
 - 欠測メカニズムがMAR
 - 二変量MMRMモデルが正しい
- 個々の解析における仮定の頑健性の検討も必要

Estimandの議論

- 臨床試験の計画・実施に関与する臨床家、統計家および他の専門分野の担当者も含む多分野にわたる作業が必要
 - 非常に喜ばしいこと
- 臨床試験の様々な側面からの議論が必要
 - 統計的側面：当然
 - 臨床的側面：試験の性質、疾患の特性、治療体系、中間事象の種類、...
- まずは、共通認識(言語)をもつことが大事
 - 議論が成立しないことがある

最後に、

ICH E9(R1) の final recommendations がどのようなものになるかはわかりませんが、

- ベースラインデータで定義可能なターゲット集団の明確化
- 臨床的に(あるいは、当該試験において)興味ある治療戦略の明確化
 - 中間事象の分類
- その治療戦略に対する質の高い (high adherence, low analysis drop-out, ...) 臨床試験の実施

が大事 !!

参考文献

- Akacha M, Bretz F, Ruberg S. Estimands in clinical trials—broadening the perspective. *Statistics in medicine* 2017; 36: 5-19.
- Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic and Clinical Pharmacology and Toxicology* 2006; 98: 237-42.
- Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials. *New England Journal of Medicine* 2017; 377: 1391-8.
- Hernán MA, Scharfstein D. Cautions as regulators move to end exclusive reliance on intention to treat. *Annals of internal medicine* 2018; 168: 515-6.
- Holzhauer B, Akacha M, Bermann G. Choice of estimand and analysis methods in diabetes trials with rescue medication. *Pharmaceutical statistics* 2015; 14: 433-47.
- Little RJA. Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association* 1995; 90: 1112-21.
- Little RJA, Rubin DB. (2001). *Statistical Analysis with Missing Data*, 2nd ed. John Wiley and Sons.
- Matsuyama Y, Ohashi Y. Mixed models for bivariate response repeated measures data using Gibbs sampling. *Statistics in Medicine* 1997; 16: 1587-1601.
- Mukai H, Hagiwara Y, Imi K, Isaka H, Watanabe K, Matsuyama Y for the SELECT BC Study Group. The impact of treatment preferences in second-line chemotherapy on the prognosis of HER2-negative metastatic breast cancer. *Oncology* 2017; 93: 315-22.
- Takashima T, Mukai H, Hara F, Matsubara N, Saito T, Takano T, et al. Taxanes versus S-1 as the first-line chemotherapy breast cancer (SELECT BC): an open-label, non-inferiority, randomized phase 3 trial. *Lancet Oncology* 2016; 17: 90-8.
- Ware JH. Interpreting incomplete data in studies of diet and weight loss. *New England Journal of Medicine* 348, 2003; 2136-7.