

抗がん剤の第III相試験 における生存時間解析に 関する統計的課題

当日のディスカッション結果

Aチーム

全生存期間に対する後治療の影響について

デザインにより後治療を考慮した経験

- PFSをprimary endpoint, OSをkey secondary endpointに設定した
- 倫理面, 実施可能性からクロスオーバーを許容した
- (作用機序などを考慮して) 後治療の種類を限定した
- 1次治療の比較試験において(後治療の調整解析に利用するため) 可能な限り追跡情報を収集した
- 後治療により治療効果が薄まることを見越してeffect sizeを保守的に見積もった
- 後治療の影響について当局相談した

デザインにより後治療を考慮した経験

- 国際共同治験において、各国における後治療として利用可能な薬剤を考慮して、参加国を選定した
- 参加国別に後治療の入りやすさを予測し、どの程度治療効果が薄まる可能性があるか、シミュレーションにより検討した
 - 後治療の入り方も複数あるので、そのパターンの予測や、入ったときの対策を予め検討しておくことは重要。また予測が外れた場合の対策も考えておくことは重要
- OSについて解釈できる結果が得られるように計画立案することは重要

解析により後治療を考慮した経験

- Conventionalな解析
 - 後治療開始時点で打ち切り
 - 後治療の有無別でSPPを評価
 - 後治療（ライン）ごとに記述的に評価
- やや発展的な解析
 - RPSFT法やIPCW法などを適用
 - 調整に利用した共変量を含め，SAPに事前規定した
 - 近年適用事例は増えている

RPSFT法の利用可能性

- 規制上は後治療の影響も含めた治療効果を確認する必要あり (eg, $HR < 1$)。HTA上は後治療の影響を除いた効果も確認する必要あり
- 手法を適用すべきかどうかは状況による。後治療が確立されているか否かでも異なるし、新薬承認と標準治療（治療方針）の確立でも解析の位置付けが異なる
- 仮定の適切性について説明する必要あり（IPCW法における「未測定の変因はない」の仮定やRPSFT法におけるcommon treatment effectの仮定など）
- クロスオーバーを許容していなくても、実際には起こりうることなので、事前に解析を計画しておくべきではないか
- 主たる解析として利用することは難しいと思われる
- どういった解析を実施して、そのためにどういったデータを集める必要があるのか、計画立案時に検討することは重要

競合リスクを考慮した解析

競合リスクを考慮した解析の実施経験

- 胃がん患者を対象とした試験において、PFSのイベントタイプ（新病変あるいは既存病変）について競合リスクを考慮した解析を実施した
 - 新病変の予後が悪く，既存病変の予後が良い状況
- 記述的にイベントタイプ別の評価を実施した
- OSに対する解析の感度解析として実施した。なお競合リスクを考慮した解析を実施する場合，その解析の位置付けも重要
- 疼痛を評価する際に死亡による準競合リスクの問題が生じた
- 適用手法はCIFやGray's test

競合リスクを考慮した解析について

- OSやPFSについて複合エンドポイントとして評価することが殆どで、競合リスクの枠組みでイベントタイプ別の影響についてはあまり評価していない
- 最終的にはOS（全死亡）で評価することが重要。競合リスクを考慮した解析にどこまで重きをおけばよいか
- 造血幹細胞移植における原病悪化とGVHDの競合解析などでは標準的に用いられているものの、臨床試験において一般に実施経験は少ないように思われる。適用経験を増やすことも重要
- 結果の解釈が難しく、また特に統計家以外に如何に分かりやすく説明するかが課題

比例ハザード性が成立しないタイプの生存時間データ

主たる解析の実施経験

- KM曲線が後半で重なり比例ハザード性が不成立。施設間において、患者集団の特性（質）の相違によって治療効果の方向が逆転した。事後的に判明したため、解析手法としてはログランク検定を実施
- 後治療や反応性の異なる集団の影響が示唆された。事後的に判明したこと、バイオマーカーのカットオフ値が決まっていなかったため、解析手法としてはログランク検定を実施
- インフルエンザ症状の治癒を評価項目とした臨床試験。解析手法として一般化Wilcoxon検定を実施
- 比例ハザード性の評価に、二重対数プロットによる図示（プロトコルで事前規定）、時間依存性共変量を含めたCox回帰分析を実施
- 参加者の実施経験はなかったが、Fleming-Harringtonの重み付きlog-rank検定（重みは事前規定）が利用可能

RMSTを指標とした解析の利用可能性

RMSTを指標とした解析について

- 比例ハザード性の成立の如何にかかわらず利用可能。またRMSTを指標とした解析の枠組みで検定と推定を実施したり，あるいは推定のみにも利用することも有用
- 比例ハザード性が成立しない場合でも臨床的に解釈しやすい
 - ただし比例ハザード性が成立しない理由を考えることは重要で，例えば（反応性の異なる集団が混在するなどの理由で）KM曲線が交差する場合は，そもそも全体の臨床的な解釈は困難ではないか
- Supportiveな解析として実施したり，感度解析の観点から境界範囲を変化させることは有用
- 主たる解析として実施する際の主な課題は境界範囲の事前規定と中間解析の実施

RMSTを指標とした解析の利用可能性

RMSTを指標とした解析について（続き）

- 状況に応じたハザード比とRMSTの使い分けを検討すべきではないか
 - KM曲線の形状によっても用いるべき指標は異なる
 - 非劣性試験におけるRMSTの利用可能性（ハザード比の利用には限界あり）
 - 発生イベント数が少ない試験（Adjuvant試験など）や初期で差が付く試験におけるRMSTの利用可能性

境界範囲の設定について

- 治験環境（臨床開発の観点）や薬剤の効果（臨床的意義の観点）を考慮して試験ごとに決定する，がん種ごとに診療ガイドラインなどで規定するなど

RMSTを指標とした解析の利用可能性

境界範囲の設定について（続き）

- 設定した境界範囲の適切性を示す必要あり
- 境界範囲まで追跡すれば（主たる）解析は実施可能。ただし設定した境界範囲の適切性を示すためにはそれ以降の追跡も重要
 - KM曲線が試験後半で重なるような場合には，追加の追跡が必要となる可能性あり
 - 実施上の課題として，ログランク検定の場合，例えば発生イベント数が当初の予想よりも少なかった場合，必要に応じて試験期間を延長しイベント数を増やすことができるが，境界範囲を事前規定した場合，そのような柔軟な対応を取ることができず，検出力の低下を招く可能性あり
- 境界範囲を長くすると長期生存例の影響を受ける