

学位論文内容要旨

北里大学大学院 薬学研究科 臨床医学（臨床統計学）

氏名 大川 慶純

題目

Statistical considerations for design and analysis of comparative clinical trials

（比較臨床試験における試験デザインと解析のための統計的考察）

序論

新薬の有効性と安全性を検証する比較臨床試験において、より良い薬を早期に上市させるためには、臨床試験デザインの工夫が非常に重要である。さらに、臨床試験から得られたデータの統計解析手法は、治療効果を検証する医薬品の特性を考慮して、妥当な解析手法を採用して試験実施計画を立てるべきである。本研究では、比較臨床試験における試験デザインと解析手法に関して2つのテーマについて取り上げ、それぞれのテーマにおいて提案と考察を行った。

研究1. 盲検化モニタリングツール（Blinded Monitoring Tool）の提案

【背景】臨床試験デザインのうち、中間解析を取り入れた群逐次デザイン（group sequential design）は、近年、臨床試験において大変よく用いられている。中間解析は、計画段階に設定した試験完了時期より先行して治療の有効性や安全性について評価を行うための解析であり、試験の早期中止による医薬品開発の効率化や、試験治療への被験者の不要な曝露の予防の観点から、実施するメリットが大きいことが期待されている。しかし、中間解析の計画、実施、中間結果の評価、早期中止などの各種手続きにおいては、しばしばコストとリソースが実施上の負担となる。例えば、中間解析を実施するには、試験の正式な完了前にデータにアクセスする必要がある。特に、試験中、治療割付がマスクされている盲検化試験では、治療割付を開示する操作が伴う。また、もしこの時点での治療効果が小さい場合、試験途中の正式な群間比較は α エラーの無駄な消費に繋がってしまう。そこで、治療割付が盲検化されたままのデータから、治療効果についてある程度妥当な予測を行い、予測された結果から中間解析の実施判断について検討することで、そのような不都合を回避するための手法が期待される。

【目的】中間解析を実施するか否かの意思決定を、治療割付が盲検化されているデータを用いて試験の全体性を損なわずに行うための、新しいモニタリング手法について提案する。

【方法】二値エンドポイント（奏効か否か）について、治療群と対照群の奏効割合を比較する無作為化臨床試験を想定し、1) 優越性、2) 無益性、または 3) その両方による早期中止のための中間解析を試験期間中、1 回計画していると仮定した。盲検化されたデータからは、その時点の両群の症例数 N と両群の奏効数 T の情報を得ることができる。治療群と対照群の奏効数がそれぞれ、 T_1 と T_0 であり、母数 p_1 と p_0 の二項分布に従うと仮定すると、奏効数 T の確率質量関数は2つの二項分布の混合分布として、以下のように書ける。

$$Pr(T = t) = \binom{N}{t} \{qp_1 + (1-q)p_0\}^t \{q(1-p_1) + (1-q)(1-p_0)\}^{N-t}, \quad q \in (0, 1)$$

ここで、 q は治療群への割り付け比率である。対照群の奏効割合 p_0 について十分な確信がある場合、治療群の奏効割合 p_1 は推定可能である。その場合、 \hat{p}_1 と p_0 の差に関して検定統計量を盲検化されたデータから算出することができ、盲検解除後の中間解析の結果が優越性または無効性の中止基準を満たすか否かを予測することができる。この方法により、中止基準を満たすことが予測された場合は中間解析の実施を、そうでない場合は中間解析をスキップすることを試験実施者に推奨することができる。盲検化されたデータを用いて中間解析の意思決定を行う本提案法の有用性について検討するために、様々なシナリオを設定した数値実験を実施した。さらに、**ECOG-ACRIN Cancer Research Group** によって実施された実臨床試験データに適用し、提案法の活用方法について示した。

【結果】数値実験の結果から、提案法を用いることにより、試験全体の症例数や検出力への影響はほとんどないが、治療効果の差が小さいときには中間解析の実施を劇的に減少させることができることが示唆された。このことから、提案法を試験デザインに取り入れることにより、盲検化されたデータの概要についての有用な参考情報を、臨床試験の完全性を損なうことなく、試験期間中に得ることができると考えられた。

【結語】提案法の利用により、不必要な中間解析の実施を回避することができるため、試験にかかるコストとリソースを節約できることが期待される。

研究 2. Restricted Mean Survival Time をエンドポイントとした試験デザインの提案

【背景】臨床試験の成功可能性を向上させるためにデザイン上の鍵となるのは、ターゲット集団に対する治療の有効性を評価するために、適切な臨床エンドポイントを選択し、そのエンドポイントについて妥当な定量的評価を行うことである。例えば、抗がん剤の有効性を検証する臨床試験では、イベント（死亡や増悪）までの時間をエンドポイントに設定することが多く、その定量的な効果指標としてはハザード比が慣例的に用いられている。デザイン段階においては、ハザード比のある一定の効果量（例えば、ハザード比 0.75）を達成するための症例数設計が行われている。ハザード比は慣例的に用いられているものの、その臨床的解釈は容易ではない。特に、前提としている比例ハザードモデルの仮定が成立しない場合は、推定されたハザード比の解釈が非常に難しい。一方

で、 Kaplan-Meier 曲線の特定期間の曲線下面積は、その期間中の平均的な生存時間と同値であり、生存曲線に関する妥当で定量的な要約指標である。この指標は、制限付き平均生存時間 (restricted mean survival time; RMST) と呼ばれる。RMST は、直感的で臨床的な解釈が可能であり、特別な仮定をせずに RMST の群間差を推定できる。生存時間解析においては、各治療群の効果プロファイルを全体的に捉えることのできる単一の指標は存在しない。しかし、臨床試験のデザインと解析においては群間比較のために、妥当で、モデルの仮定に依存せず、臨床的解釈が可能な 1 つの主要な要約指標が必要である。

【目的】近年実施された、肺がん免疫治療を評価した臨床試験を例に、ハザード比の代替指標として RMST を用いた臨床試験デザインの組み方とその解析手法について提案する。

【方法】従来用いられている典型的な臨床試験デザインについて説明するために、CheckMate057 試験 (治療歴を有する進行期非小細胞肺癌患者を対象にドセタキセルに対するニボルマブの優越性を評価した第 III 相無作為化臨床試験) を例として用いた。CheckMate057 試験では、全生存期間 (overall survival; OS) を主要評価項目とし、計画段階で期待されていた 3.1 ヶ月の OS 中央値の差 (ニボルマブ群 11.1 ヶ月、ドセタキセル群 8.0 ヶ月) について、ハザード比 $8/11.1=0.72$ に換算し、このハザード比を検出するための症例数設計が行われていた。本研究では、このような従来どおりの試験デザインの組み方について問題点を提起し、また、試験の結果から得られたハザード比の推定値の妥当性について検討した。さらに、ハザード比の推定に代替となる統計手法について整理したうえで、試験計画時に期待されている臨床的効果について、ハザード比に換算せずに、群間差についての直接的な推論をするためのアプローチとして、RMST の差を効果量の指標とした試験デザインの代替案について考案した。

【結果】CheckMate057 試験から得られた OS のハザード比の推定値は 0.73 (95%信頼区間: 0.62-0.88) であり、ニボルマブがドセタキセルよりも統計的に有意に生存時間を延長することを示唆した。しかし、Kaplan-Meier 曲線から、比例ハザード性の仮定が成立していないことが確認され、この推定値について臨床的に解釈することは困難であった。一方、2 群の 24 ヶ月 RMST の差を用いた場合は、1.7 (95%信頼区間: 0.4-0.31) ヶ月の差と推定され、この指標においても統計的に有意に生存時間を延長することを示した。モデルの仮定に依存しない RMST の差の推定値は、臨床的に解釈可能で、本解析においてはより妥当な推定値であると考えられた。そこで、RMST に基づいた試験デザインの詳細について、CheckMate057 試験の状況に倣って、新たに提示した。

【結語】RMST に基づいた試験デザインと統計解析手法は、比較臨床試験においてハザード比を用いた臨床試験デザインの代替案として有用であることが示唆された。また、RMST による群間比較は統計的な仮定なしに臨床的解釈を可能にするため、今後、実施される臨床試験での生存時間解析において、主要に用いることが期待される。

考察

臨床試験デザインにおいては、その試験の対象となる集団や治療の臨床的ベネフィットに応じて、統計解析手法を詳細に事前規定する必要がある。本研究における2つの提案はどちらも今後の臨床試験デザインを組む上で有用であると考えられる。臨床的に有効な治療薬の開発をより早く、効率的に実施するためには、有効性が認められない治療や安全性の問題が示唆される治療の開発を早期に中止し、コストとリソースを有効な治療に集約することが重要である。さらに、臨床試験において用いられる統計手法やそこから得られる結果は、臨床家や患者にとって有用で解釈可能であるものであることは、臨床試験の社会的意義の観点からも重要である。

総括

本研究では、臨床試験の中間解析の実施におけるモニタリングの工夫と、統計的・臨床的に妥当なエンドポイントである **RMST** を用いた試験デザインについて提案した。これら2つのデザインの提案により、コストとリソースの問題、治療効果の指標の解釈の問題という、既存の臨床試験における課題を、それぞれ統計学的観点から打開できる試験実施計画を提案できることが期待される。