

氏 名 野中 孝浩

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 2154 号

学位授与の日付 2020 年 3 月 24 日

学位授与の要件 複合科学研究科 統計科学専攻
学位規則第6条第1項該当

学位論文題目 第Ⅲ相臨床試験における治療効果とその予測マーカーを評価するための統計的仮説検定とその基準に関する研究

論文審査委員 主 査 教授 伊藤 陽一

教授 金藤 浩司

教授 二宮 嘉行

准教授 船渡川 伊久子

教授 五所 正彦

筑波大学 医学医療系生物統計学

教授 松井 茂之

名古屋大学 医学系研究科生物統計学分野

(様式3)

博士論文の要旨

氏 名 野中 孝浩

論文題目 第Ⅲ相臨床試験における治療効果とその予測マーカーを評価するための統計的仮説検定とその基準に関する研究

がん患者に対する抗がん剤による薬物治療は、従来、がん細胞のみならず正常細胞を含め、活発に増殖や分裂する細胞内で起こるデオキシリボ核酸 (DNA) やタンパク質の合成などを阻害し、非選択的に細胞毒性を示す殺細胞性薬剤が主流であったが、近年、発がんやがん細胞の増殖などに関連する特定の分子を標的とし、がん細胞の増殖などを選択的に阻害する分子標的薬剤の使用に移行している。分子標的薬剤によって治療効果が得られる患者は、特定の部分集団、すなわち、薬剤の標的となる分子に変異などが認められる患者集団に限定される可能性が高い。したがって、特にがんの分子標的薬剤の開発に際しては、治療の恩恵を受ける患者集団を同定することができる治療効果の予測因子 (predictive marker、以下、当該因子の候補も含め、単にマーカーと呼ぶ) を開発することが重要であると考えられている。

マーカーが存在するものの、マーカーの治療効果予測性能に関する証拠が十分に信頼できる状況にない場合には、検証的試験として位置づけられる第Ⅲ相臨床試験が実施される際に、マーカーで層別するデザインで試験が実施されることが多い。このようなデザインの試験は一般にマーカー層別第Ⅲ相臨床試験などと呼ばれる。第Ⅲ相臨床試験がマーカー層別第Ⅲ相臨床試験として計画される状況下では、一般に、マーカーで定義される部分集団間で治療効果が異なり、不均一性が認められることがより自然な状況であると考えられるため、本研究では、このような自然な状況を仮定した上で、マーカー層別第Ⅲ相臨床試験における統計的仮説検定 (多重検定) の枠組みと方法論を開発する。

本研究では、マーカーが定義する特定の部分集団間での治療効果が異なり、かつ治療効果はマーカー陰性集団と比較してマーカー陽性集団で高いというマーカー仮説のもと、マーカー層別第Ⅲ相臨床試験の計画と解析に関する検討を行い、大きく三つの提案を行った。

一つ目の提案は、マーカーに基づく部分集団間の統計的仮説検定の方法 (多重検定) に対する新しい枠組みの提案であり、第一種の過誤確率の制御に、強い制御ではなく弱い制御を許容する枠組みを提案する。マーカー陽性集団および陰性集団に対する「治療効果なし」の帰無仮説をそれぞれ $H_0^{(+)}$ および $H_0^{(-)}$ で表すとき、マーカー層別第Ⅲ相臨床試験において、「マーカー陽性集団における治療効果の大きさが、マーカー陰性集団における治療効果の大きさと比較して小さくない」という「マーカー仮説」が成立するもとでは、帰無仮説シナリオ 1 (グローバル帰無仮説) ($H_0^{(+)}$ が真、かつ、 $H_0^{(-)}$ が真)、および帰無仮説シナリオ 2 ($H_0^{(+)}$ が偽、かつ、 $H_0^{(-)}$ が真) といった 2 つの帰無仮説シナリオが想定される。この場合、たとえ弱い制御であっても、少なくともマーカー陽性集団に対して治療効果を主張することに対する厳格な第一種の過誤確率の制御は保証されている。その一方で、マーカー陰性集団に対して治療効果があると誤って主張してしまう、すなわち、誤って新規

治療法の適応を全集団に広げてしまう第一種の過誤確率については厳格に制御すること（強い制御）は求めない。

二つの目の提案は、マーカー層別第Ⅲ相臨床試験では一般に、マーカーで定義される特定の部分集団で治療効果が示されることが期待され、実際に特定の部分集団のみで新薬などの新規治療法の治療効果が検証された場合には、当該新規治療法の適応を特定の部分集団に制限するためにマーカーが用いられることになる。したがって、マーカーの治療効果予測能、すなわち、マーカーの臨床的妥当性（clinical validity）の評価もマーカー層別第Ⅲ相臨床試験の中で併せて行われることが期待される。そこで、本研究では、マーカーの臨床的妥当性評価を、明確な基準に基づいて行う統計的仮説検定の方法として、マーカー陰性集団に対する無効基準に基づく方法を提案する（提案法）。これより、マーカー陽性集団で治療効果が検証となった際に、マーカーに基づいて集団を限定することの臨床的妥当性についても明示的な根拠が得られていることになる。

三つ目の提案は、全集団、マーカー陽性集団それぞれ、あるいは、いずれかに対して治療効果を主張する確率として三種類の確率、すなわち、全集団に対して治療効果を主張する確率 P_{overall} 、マーカー陽性集団のみに対して治療効果を主張する確率 P_{subgroup} 、および、全集団、マーカー陽性集団のいずれかに対して治療効果を主張する確率 P_{success} を導入し、これらに基づいて、第Ⅲ相臨床試験が満たすべき統計的仮説検定の作用特性の基準も提案する。その上で、提案する作用特性の基準に基づいて提案法（すなわち、マーカー陰性集団に対する無効基準に基づく方法）と既存の様々な統計的仮説検定の方法（固定順序法、フォールバック法、ハイブリッド法（緩やかな有意水準を用いた MaST 法）、交互作用検定に基づく方法）との比較を数値評価で行った。全集団における治療効果の検定にはマーカー層別ログランク検定を、マーカー陽性集団における治療効果の検定には通常の（非層別）ログランク検定を想定した。交互作用に基づく方法での交互作用検定は、マーカーに基づく部分集団内での対数ハザード比（治療効果）の差異に基づいて行った。フォールバック法、ハイブリッド法、交互作用検定に基づく方法、および、提案法では、検定統計量間での相関関係を考慮して有意水準を設定した。なお、参照のため、伝統的な第Ⅲ相臨床試験での治療効果の検定（有意水準 α を用いた全集団での治療効果の検定を一回だけ行う）も併せて実施した（この場合は、 $P_{\text{success}} = P_{\text{overall}}$ ）。その結果、第Ⅲ相臨床試験をデザインする段階において、マーカーの治療効果予測能の信頼性が比較的高い場合には固定順序法およびハイブリッド法が推奨された一方、低い場合には保守的な無効基準を用いた提案法を選択することが妥当と考えられた。

今後の研究の発展として、マーカー陰性集団に対して無効基準を使用することは有効であることから、これを中間解析のモニタリングで使用することが考えられる。この方向性では、アダプティブな適応患者の選択など、いくつかの提案があり、さらに、最終解析で、アダプティブな適応患者の選択と様々な統計的仮説検定（本研究で提案した仮説検定の方法を含む）の組み合わせを検討することも可能である。このような複合的な仮説検定の方法の評価については今後の課題にしたい。

Results of the doctoral thesis screening

博士論文審査結果

Name in Full
氏名 野中 孝浩

Title
論文題目

第Ⅲ相臨床試験における治療効果とその予測マーカーを評価するための統計的仮説検定とその基準に関する研究

(論文審査結果) [2020年1月20日実施]

野中孝浩氏の博士論文審査委員会は、2020年1月20日15時00分から約2時間にわたり開催された。申請者による1時間にわたる公開發表による概要説明と質疑応答、さらに約1時間の審査員のみによる審査と口述による学力の確認を踏まえた上での審議の結果、以下に述べる理由により審査委員会は全員一致で提出された博士申請論文は学位授与の水準に達していると判断した。

【論文の概要】

申請論文は全8章68頁からなり日本語で書かれている。第1章は、序論である。医薬品開発においては近年、バイオマーカー、特に治療の恩恵を受ける患者集団を同定することができる治療効果の予測因子（以下、予測マーカー）を開発することが注目されている。とりわけ、がん治療に用いる抗がん剤の開発の主流は、発がんおよびがん細胞の増殖に関連する特定の標的分子を阻害する分子標的薬の使用に移行しており、治療薬の開発に併せて予測マーカーを開発することが期待されている。本論文では、マーカー陰性集団における治療効果を評価するための新たな仮説検定法やマーカーに基づいて部分集団間で治療効果を評価する際の作用特性の基準が提案されている。第2章では、予測マーカーで層別する第Ⅲ相試験における主要な統計的解析（プライマリー解析）のための枠組みが述べられている。その中で、帰無仮説の設定および第一種の過誤確率の制御に関する詳細な理論的根拠が示されている。第3章では、既存の統計的仮説検定の方法に関する概略が述べられている。第4章では、予測マーカーの臨床的妥当性の評価の基準を含む統計的仮説検定の方法が述べられ、マーカー陰性集団に対する無効中止基準に基づく方法が提案されている。第5章では、予測マーカーに基づく部分集団別に治療効果を主張する確率が導入されたもとで仮説検定法の作用特性の基準が提案されている。第6章では、サンプルサイズ設計に提案する基準を適用し、統計的仮説検定の方法を数値実験によって比較を行っている。最後に、第7章は考察および結論であり、第8章は、謝辞である。

【論文の評価】

本論文の貢献としては、予測マーカーが規定する部分集団間での治療効果の不均一性に関するマーカー仮説のもと、予測マーカーで層別する検証的試験（第Ⅲ相試験）の計画と解析に関する検討を行い、大きく三つの提案を行った点が挙げられる。一つ目は、マーカー部分集団間の多重検定に対する新しい枠組みの提案である。二つ目は、マーカーの治療効果予測能（臨床的妥当性）の評価を組み込んだ仮説検定法の提案である。これより、マーカー陽性集団で治療効果が有意となった際、マーカーを使って集団を限定することの臨床的妥当性についても明示的な根拠を与えることになる。三つ目の提案は、全集団、マーカー陽性集団それぞれ、あるいは、いずれかに対して治療効果を主張する確率を導入し、それに基づいて、臨床的に意味のある治療効果の観点から、試験の作用特性を規定することである。以上のように、本研究は、予測マーカーを用いた治療効果の検証的解析をその枠組みから再構築した上で、有効かつ実用的な統計解析の方法・基準の開発に成功しており、個別化医療の臨床試験に役立つ医療統計学の方法論研究として十分価値のあるものと考えられる。なお、申請論文の第 4 章、第 5 章、第 6 章の内容が査読付き国際雑誌「Pharmaceutical Statistics」に掲載されている。

（試験結果）[2020 年 1 月 20 日実施]

総合研究大学院大学における論文博士の学位授与に係る論文審査等の手続き等に関する規程第 7 条に定める学力の確認を行うため、総合研究大学院大学学位規則第 11 条に基づき、口述による学力の確認を実施した。

口述による学力の確認を実施した結果、申請者は、その博士論文を中心としてそれに関連がある専門分野、その基礎となる分野及び英語の能力について博士（統計科学）の学位を授与するに十分な学識を有するものと判断し、合格と判定した。