

氏名（本籍） まつ うら けんたろう 松 浦 健太郎（茨城県）
学位の種類 博士（工学）
学位記番号 甲第 1153 号
学位授与の日付 2023 年 3 月 19 日
学位授与の要件 学位規則第 4 条第 1 項該当
学位論文題目 **Optimal Adaptive Allocation using Deep Reinforcement Learning in Dose-Finding Studies**
(用量設定試験における被験者の割り付けの深層強化学習による最適化)

論文審査委員 (主査) 教授 寒水 孝司
教授 赤倉 貴子 教授 池口 徹
准教授 入江 豪 教授 橋口 博樹

論文内容の要旨

薬剤の適正な用量は、最終的に臨床試験の結果に基づいて決定される。複数の用量の候補を複数の被験者に投与して、被験者の臨床的な結果（以下、アウトカム）から、適正な用量を探索する試験が用量設定試験である。本論文は、抗がん剤とそれ以外の薬剤（非抗がん剤）の用量設定試験を対象にして、被験者を効率的に割り付ける方法を提案するものである。

臨床試験の第 1 相試験から承認に至る確率は 10～20%という報告がある。その確率の低さの原因として、用量選択が適切でないことが挙げられる。例えば、用量が低すぎると、後期の臨床試験で臨床的に意味のある効果が得られない可能性が大きい。用量が高すぎると、許容できない有害事象が発生する可能性が大きい。用量とアウトカムの関係を精度よく推定して、適切な用量を選択することは薬剤開発の重要な過程の一つである。一般的に用量設定は、非抗がん剤では第 2 相試験、抗がん剤では第 1 相試験で行われる。

非抗がん剤の第 2 相試験の用量設定試験では、候補用量は 3～6 程度であり、被験者数は 50～300 人程度である。アウトカムは有効性の指標である。この相では、後続の第 3 相試験と同様の患者を対象として、用量と有効性の関係（用量反応曲線）を推定し、第 3 相試験で評価する至適用量を決定する。用量反応曲線の推定には、分散分析、multiple comparison procedure – modeling (MCP-Mod) 法、ベイズモデル平均化法などの方法が提案・適用さ

れている。これらの方法では、一般的に、被験者を各用量に均等に割り付けるため、検出力、推定した至適用量の正解率、推定した用量反応曲線の誤差といった評価指標が十分に最適化されない。これに対して、D-optimal法、TD-optimal法、aMCP-Mod法などの、被験者を適応的に割り付ける方法が提案されている。これらの方法の性能は漸近的に保証されるため、実際の臨床試験の参加者数のもとでは、評価指標が最適化されるとは限らない。

抗がん剤の第1相試験の用量設定試験では、候補用量は3~8程度であり、被験者数は20~40人程度である。アウトカムは毒性の指標である。この相では、標準治療で効果がなかった患者を対象として、用量と毒性の関係を推定し、忍容性のある最大用量であるMaximum Tolerated Dose (MTD)を推定する。MTDは、後続の第2相試験以降で評価する用量の基準になる。第1相試験では、薬剤をはじめてヒトに投与するため、安全性の観点から、最初に被験者を最も低い用量に割り付けて、徐々に用量を増やしていく。用量の漸増方法とMTDの推定には、3+3デザイン、Continual Reassessment Method (CRM)法、Bayesian Optimal Interval (BOIN)法などが使われている。しかしながら、いずれの方法のもとでも、MTDを正しく選択する確率 (proportion of correct selection: PCS)は十分に高くない。

非抗がん剤と抗がん剤の用量設定試験に共通する課題として、既存の方法のもとでは、評価指標は十分に最適化されない。この課題を解決するため、本研究では、それぞれの用量設定試験のもとで、深層強化学習を用いた被験者の割り付け方法を提案する。

(1) 非抗がん剤の用量設定試験

前述の課題を解決するため、深層強化学習を用いて評価指標を直接最適化する適応的な割り付け法を構築した (Matsuuraら, 2022)。具体的には、適切に定義された状態、報酬、環境をもつ深層強化学習を用いて、適応的な割り付けルール(状態 s において用量 k に割り付ける離散的な確率分布)を数値的に導出した。状態 s は、各用量に割り付けられた被験者の応答変数の平均と標準偏差、被験者数の割合から成る多次元変数とした。報酬は、最適化する評価指標を0から1の範囲に変換した値とした。環境は、用量反応曲線の推定方法における仮定に従ってデータを生成するものとした。

Bornkampら(2007)とDragalinら(2010)と同様の条件のもとで、数値実験によって提案法と既存法の性能を比較した。提案法では、最適化する評価指標の性能が向上すること、特に、平均絶対誤差を最適化する評価指標にすると、他の指標の性能も向上しやすいこと、学習に用いた環境から外れたシナリオでも性能を維持しやすいことが明らかになった。

(2) 抗がん剤の用量設定試験

前述の問題を解決するため、深層強化学習を用いてPCSを直接最適化する適応的な割り付け法を構築した (Matsuuraら, 投稿中)。具体的には、適切に定義された状態、報酬、環境をもつ深層強化学習を用いて、適応的な割り付けルール(状態 s において行動 a を選択する離散的な確率分布)を数値的に導出した。状態 s は、直前の被験者に用いた用量、各用量に割り付けられた累積被験者数と毒性が生じた累積被験者数から成る多次元変数とした。行動 a

の選択枝は、「試験継続して用量を減量する/維持する/増量する」または「試験を中止して用量のいずれかを MTD として選択する」とした。報酬は、選択した MTD が正解の場合に 1、不正解の場合に 0 とした。環境は、MTD が明確な仮想的なシナリオに従ってデータを生成するものとした。

Zhou ら (2018a) と同様の条件のもとで、数値実験によって、提案法と既存法の性能を比較した。提案法では、毒性が多く発生するものの、既存法よりも PCS が高くなることが明らかになった。特に、実際の臨床試験で想定されるような、低用量では MTD の出現確率が低く、高用量に MTD が位置付けられるシナリオのもとで、提案法の性能が大きく向上した。

これらの研究成果は、用量設定試験における評価指標の向上を通じて、薬剤の承認に至る確率を向上させることに寄与するものと考えられる。

論文審査の結果の要旨

本論文では、学長からの審査付託を受けて、標記 5 名の審査委員で構成する審査委員会を組織し、提出された学位論文について審査を行った。

審査委員会では、学位申請者に対して、学位論文の内容や前回審査における指摘事項の対応結果について説明を求めた。その上で、質疑応答の内容を踏まえて、博士論文として満たすべき条件や必要な修正点を確認した。

第 1 回審査では、学位申請者から、学位論文の概要について説明があった。

本論文は「Optimal Adaptive Allocation using Deep Reinforcement Learning in Dose-Finding Studies (用量設定試験における被験者の割り付けの深層強化学習による最適化)」と題し、全 5 章より構成されている。

薬剤の投与量 (用量) は、最終的に臨床試験の結果に基づいて決定される。複数の用量の候補を複数の被験者に投与して、被験者の臨床的な結果 (以下、アウトカム) から、適正な用量を探索する試験が用量設定試験である。本論文は、抗がん剤とそれ以外の薬剤 (非抗がん剤) の用量設定試験を対象にして、深層強化学習を用いて被験者を効率的に評価すべき用量に割り付ける方法を提案し、それらが実務上有用であることを示している。

第 1 章では、抗がん剤と非抗がん剤の用量設定試験の概要が述べられ、既存法では評価指標が十分に最適化されないという統計的課題を説明している。

第 2 章では、非抗がん剤の用量設定試験を対象にして、深層強化学習を用いて評価指標を最適化する適応的な割り付け法を提案している。数値実験によって提案法と既存法の性能を比較し、提案法では、最適化する評価指標の性能が向上することを示している。

特に、平均絶対誤差を最適化する評価指標にすると、他の指標の性能も向上しやすいこと、学習に用いた環境から外れたシナリオでも性能を維持しやすいことを述べている。

第3章では、抗がん剤の用量設定試験を対象にして、深層強化学習を用いて至適用量の正解率を最大化する用量漸増法を提案している。数値実験によって提案法と既存法の性能を比較し、提案法では、毒性が多く発生するものの、既存法よりも至適用量の正解率が高くなることを示している。特に、実際の臨床試験で想定されるような、高用量に至適用量が位置付けられるシナリオのもとで、提案法の性能が良いことを述べている。

第4章では、提案法を実際の臨床試験に適用する上での課題を整理している。

第5章では、研究成果を簡潔にまとめた上で、第2章と第3章で提案した方法の結論とその意義を述べている。

第2回審査では、第1回審査における指摘事項の対応について説明があった。特に、第4章に、バンディット理論を含めた既存法の性能が低く、提案法の性能が高いことの原因について考察が述べられた。

第3回審査では、公聴会を兼ねて審査を実施し、公聴会の参加者からの質問に対して適切に回答していることを確認した。

本論文では、用量設定試験で評価指標を改善する新しい方法を提案し、それらが実務的に有用であることを示している。本論文で得られた研究成果は、用量設定試験における評価指標の向上を通じて、薬剤の承認に至る可能性を向上させることに寄与するものである。

以上により、本論文は、博士（工学）の学位論文として十分に価値あるものと認められる。