

主論文の要約

**Correcting Estimation Bias in Randomized
Clinical Trials With a Test of Treatment-by-
Biomarker Interaction**

〔 予測マーカーを用いたランダム化臨床試験における
治療効果のバイアス補正 〕

名古屋大学大学院医学系研究科 総合医学専攻
臨床医薬学講座 生物統計学分野

(指導：松井 茂之 教授)

豊泉 樹一郎

【目的】

医薬品を含めた新しい治療法の臨床開発では、治療効果を期待できる患者集団の同定が極めて重要であり、治療効果の予測マーカー（predictive marker）の開発も同時に試みられる。マーカーの効果予測能の評価は、検証的試験である第三相試験の前に完了していることが理想的であるが、実際にはそうでない場合も多い。このときの第三相試験は、マーカー値によらず全患者集団を組み込んだものとなり、マーカー予測能の評価として、マーカー陽性集団がマーカー陰性集団より大きな治療効果を有するかの統計的仮説検定が実施される。この検定では、たとえマーカーによる治療効果の差が本当は存在しないときであっても、データに含まれる偶然的変動によって、差が有意となることが一定の確率で起こる（第一種の過誤）。これより、マーカーによる治療効果の差が統計的に有意な場合に限ってマーカー陽性集団での治療効果を推定すると、見かけ上、治療効果が大きく推定されるという、推定バイアスの問題が生じる。

本研究では、このような推定バイアスを補正するための統計的方法を新たに開発する。

【方法】

マーカーによる治療効果の差とマーカー陽性集団における治療効果の推定値の同時分布から、マーカーによる治療効果の差と推定バイアスの関係式（バイアス関数）を理論的に導出した。バイアス関数に基づく最も単純なバイアス補正法（単純補正法）は、マーカーによる治療効果の差の観測値を真の値とみなしてバイアス関数に代入し、補正量を求める方法である。しかし、マーカーによる治療効果の差の推定値自体がバイアスを有するため、補正法として有効でないと考えられる。そこで、バイアス関数をマーカーによる治療効果の差の推定値の一次関数もしくは二次関数を用いて近似する補正法（一次近似法と二次近似法）を新たに開発した。

バイアス補正法の性能評価はモンテカルロ・シミュレーションにより行った。症例数が 250 もしくは 350 例の臨床試験を想定し、全患者集団におけるマーカー陽性集団の割合は 30%または 60%とした。マーカー陽性集団および陰性集団での治療効果（対照治療に対する新治療のハザード比）の組み合わせとして、(1.0, 1.0)、(0.7, 0.7)、(0.5, 1.0)、(0.6, 0.8) の 4 つのシナリオを想定した。マーカーによる治療効果の差に対して有意水準 10%の片側検定を実施し、この検定が有意な場合に限ってマーカー陽性集団の治療効果を推定した。

最後に、転移性結腸直腸がんの患者を対象にパニツムマブの FOLFOX4 への上乗せ効果を調べた臨床試験（Douillard et al. (2010)）と、非小細胞肺癌の患者を対象に化学療法とエルロチニブの比較をした臨床試験（Gregorc et al. (2014)）にバイアス補正法を適用し、解析結果の違いを考察した。

【結果】

シミュレーションの結果を表 1、2 に示す。なお、平均二乗誤差は、マーカー陽性集

団における治療効果の真の値と推定値の差の二乗の平均値であり、治療効果の推定値がどれだけ真値に近いかを表す（小さな値の方が望ましい）。補正を行わなかった場合（従来法）は、推定値の分散は小さいものの、推定バイアスと平均二乗誤差が大きくなる傾向が認められた。推定バイアスの補正においては、二次近似法、一次近似法、単純補正法の順に優れていた。しかし、推定値の分散においては、その逆の順番となった。一方、平均二乗誤差においては、一次近似法が概ね優れるという傾向があった。この傾向は、全患者集団を通して治療効果が一定の場合に顕著であった。

実際の臨床試験への適用結果を表 3 に示す。マーカーによる治療効果の差が十分に大きかった Douillard et al. (2010) では、いずれのバイアス補正法を用いても解析結果はほぼ同じであった。一方、マーカーによる治療効果の差が比較的小さかった Gregorc et al. (2014) では、従来法による推定ではハザード比が 0.58 であったのに対し、一次近似法および二次近似法では、それぞれ、0.75、0.76 と治療効果が小さく推定された。

【考察】

一次近似法の平均二乗誤差が小さくなった理由として、推定バイアスと分散の両方が過度に大きくならなかったことが考えられる。このことは、一次近似法はバイアス関数の形状を比較的柔軟に捉えることができるという性質をもつ一方で、二次近似法よりも係数の数が少ないため推定誤差の影響を受けにくいことによると考えられる。一次近似法は、すべての治療効果のシナリオにおいて比較的良好な推定値が得られる方法と考えられる。

【結論】

第三相ランダム化臨床試験の結果は、治療法の臨床的な位置付けを検討する上で最も重要なエビデンスとなる。特に予測マーカーがよく使われるがん領域では、薬価が高額であり患者、医療財政への負担も大きいことから、適切な患者選択及び治療効果に関する正確なエビデンスに基づいて処方が決定されるべきである。今回開発したバイアス補正法（特に、一次近似法）は、治療効果のより正確な評価を可能とすることから、治療法に関する検証と意思決定において有効な統計的方法と考えられる。