

論文題目

Adjusting for Differential Proportions of Second-line Treatment to Test and Estimate the Treatment Arm Effect with Structural Nested Models and Marginal Structural Models in a Clinical Trial of Unresectable Non-small-cell Lung Cancer

切除不能非小細胞肺癌臨床試験における後治療の影響を調整した群間比較
—構造ネストモデルと周辺構造モデルの応用—

氏名 山口 拓洋

1. はじめに

切除不能非小細胞肺癌における標準化学療法を確立するため、前治療のないステージ IIIB/IV 期で PS (Performance Status) が 0-2 の被験者に対して、イリノテカンとシスプラチン併用療法 (CPT-P)、ビンデシンとシスプラチン併用療法 (VDS-P) 及びイリノテカン単独療法を比較する 3 群ランダム化臨床試験が実施された。本試験の主要な目的の 1 つは、CPT-P 群の VDS-P 群に対する延命効果の優越性の検証であったが、仮説と反し、ステージ IIIB においてその優越性は検証できなかった。ステージ III 期の治療方針として、化学療法と放射線療法の併用が American Society of Clinical Oncology (ASCO) の非小細胞肺癌の治療に対するガイドラインにおいて推奨されている。本試験において、後治療として放射線治療が行われた被験者は、CPT-P 群では 49 人中 22 人 (45%)、VDS-P 群では 46 人中 26 人 (61%) であった。担当医はおそらく ASCO のガイドラインを活用し、後治療として放射線治療を行ったと思われ、CPT-P 群の優越性が検証できなかった理由として、放射線治療に効果があり、その実施割合が VDS-P 群で高かったことが原因と考えられた。

通常このような臨床試験において群間比較を行う際には、ITT (Intent To Treat) 解析が標準的な方法であるが、ITT 解析では群間での後治療 (本論文では、胸部への 40Gy 以上の放射線治療と定義) の実施割合の違いを調整していない。この解析方法は、今回の臨床試験で比較されている治療が実臨床の場で用いられた場合の治療効果を推定しており、この結果は臨床的に重要な意味を持つ。一方、我々が比較したいのは、両群の被験者が全員同じ後治療を受けた場合 (あるいは全員後治療を受けなかった場合) の群間での生存曲線の違いである (直接治療効果 direct treatment arm effect)。この疑問に答えるために、治療群を表す変数に加えて、後治療を受けているかどうかを時間依存性共変量とした比例ハザードモデルを用いる方法などが提案されている。しかしながら、後治療実施には被験者の臨床状態や副作用の発現、臨床検査値異常などの要因が関係しており、また、これらの要因は死亡に関するリスク因子でもある。したがって、我々が後治療の効果の推定を行う際には、これらの時間依存性の交絡要因を調整する必要がある。一方、これらの要因はそれ以前の後治療の影響を受けており、いわゆる中間 (結果) 変数となっていることから、これらの要因で調整した解析を行うと、後治療の効果の推定にバイアスが生じる。既存の解析方法では、これらの要因を適切に調整しておらず、結果として CPT-P 群の VDS-P 群に対する直接治療効果の推定にバイアスが生じることになる。中間変数でもある時間依存性の交絡要因が存在する場合の治療効果の適切な推定方法として、構造ネストモデルと周辺構造モデルを用いた方法が提案されているが、今回は治療群の直接効果と後治療の効果の 2 つの治療効果を考慮したモデルが必要である。

本研究の目的は以下の 2 つである。

1. がんの臨床試験において、後治療の影響を調整し治療群の直接効果を検討するのに有用な、構造ネスト加速モデルと周辺構造比例ハザードモデルの2つの因果モデルを提案する。
2. 提案する方法論を切除不能非小細胞肺癌の臨床試験データに適用し、治療群の直接効果と後治療の効果について検討する。

2. データと予備的な解析

本研究では、適格例でステージ IIIB である CPT-P 群 49 人及び VDS-P 群 46 人の被験者を以下の解析対象とした。後治療については、一度放射線治療が行われたらその効果は継続すると仮定した。

ITT の原則に基づき群間で生存曲線に違いが見られるかどうかログランク検定を行った結果、有意差はみられなかった ($p=0.31$)。また、CPT-P 群の VDS-P 群に対するハザード (死亡率) 比は、比例ハザードモデルを用いて推定した結果 1.24 (95%信頼区間: 0.81-1.91) であり、CPT-P 群が逆に劣っている傾向であった。一方、後治療開始をイベントとした解析では、VDS-P 群が CPT-P 群に比べてより早く後治療を開始していた (ログランク検定 $p=0.07$)。

後治療実施の有無を時間依存性共変量とした比例ハザードモデルを用いた解析では、CPT-P 群の VDS-P 群に対するハザード比の推定値は 1.24 (95%信頼区間: 0.80-1.91)、後治療実施の有無に関するハザード比は 0.97 (95%信頼区間: 0.62-1.52) であった。

3. 構造ネスト加速モデルと周辺構造比例ハザードモデルを用いた解析及び結果

3.1. 構造ネスト加速モデルを用いた解析及び結果

今回の解析では、治療群及び後治療実施の有無を表す 2 つの治療変数を考慮する必要がある。構造ネストモデルでしばしば用いられる、単一の治療変数のみを考慮したモデルを、2 つの治療変数が扱えるよう拡張し、後治療の影響を調整し治療群の直接効果を検討するのに有用な、構造ネスト加速モデルを提案する。

このモデルでは、被験者 i の実際に観察された生存時間 T_i と、被験者 i が VDS-P 群に割り付けられ、後治療を実施されなかった場合の反事実的な生存時間 T_{i00} (ベースライン生存時間) とを関連づける。CPT-P 療法により生存時間が VDS-P 療法と比較して一定の割合で延びる (あるいは縮まる)、また、後治療実施により生存時間が延びる (縮まる)、というモデルである。

$$T_{i00} = \exp(\beta_r^* R_i) [S_i + \exp(\beta_s^*) (T_i - S_i)]$$

ここで、 R_i は CPT-P 群であれば 1、VDS-P 群であれば 0 をとる変数であり、 S_i は後治療開始の時間を表す (実施しなければ T_i と等しい)。後治療が実施されている時間 $T_i - S_i$ が、後治療により $\exp(\beta_s^*)$ 倍となり、更に、後治療を実施せず治療群が R_i である被験者の生存時間 $T_{i00} = S_i + \exp(\beta_s^*) (T_i - S_i)$ が CPT-P 療法により $\exp(\beta_r^*)$ 倍になるという前提のモデルである。 β_r^* 及び β_s^* はそれぞれ、CPT-P 群の VDS-P 群に対する直接治療効果、後治療の効果を表すパラメータであり、値が負の場合に治療効果があることに相当する。

パラメータ β_r^* 及び β_s^* の推定について、以下の 2 通りの方法を考える。

3.1.1. ランダム化に基づく解析

本試験では被験者に対しランダム割付けが行われているので、この前提のみに基づいて β_r^* 及び β_s^* を同時に推定する。ベースライン生存時間 T_{i00} は、年齢や他のベースラインの共変量と同じく、ランダム化前からもともと決定されており、割り付けられた治療群や後治療の実施状況に依存しない、すなわち、 T_{i00} と R_i は独立であることを利用してパラメータ推定を行う。

結果: CPT-P 群は VDS-P 群と比して 7% 生存を縮める (95% 信頼区間: 74% 生存を縮める ~ 46% 生存を延ばす)、また、後治療実施により生存が 63% 延びるという結果であった。

3.1.2. 観察研究の仮定を組み合わせた解析

後治療の実施の有無と実施時期は担当医の判断に依存しており、被験者の臨床状態（腫瘍の増悪、体重減少など）、副作用や臨床検査値異常の発現の有無、家族の希望などが密接に関係している。よって、見方を変えれば本試験は、後治療に関する観察的研究とも捉えることが可能である。この観点に基づき最初に後治療の効果 β_s^* を推定し、その推定値を用いて CPT-P 群の VDS-P 群に対する直接治療効果 β_d^* の推定を行う。3.1.1.の方法と比べて、より精度が高いパラメータ推定値が得られる。

β_s^* の推定は「測定されていない交絡要因はない」という仮定に基づき行う。この仮定は、後治療実施に関係する全ての要因（後治療と生存との因果関係を歪める時間依存性の交絡要因）、例えば上述したような腫瘍の増悪や副作用の発現の有無など、が測定されている下では、ある時点において「後治療を受けた被験者がもし後治療を受けなかった場合の生存時間」と「後治療を受けなかった被験者（が後治療を受けなかった場合）の生存時間」とが統計的に独立である（生存時間が短い被験者ばかりが後治療を受けることはなく、比較可能性が成り立つ）、ということの意味している。後治療実施に関係すると考えられる多くの要因を考慮し近似的にこの仮定が成り立つ条件の下で、後治療を受けるかどうかと生存時間 T_{ir0} が独立となることを利用して β_s^* の推定を行う（G推定 G-estimation）。次に、 β_s^* の推定値を用いて 3.1.1.と同様にランダム化が行われているという前提に基づき β_d^* の推定を行う。

G 推定では、ある時点における後治療実施の有無と生存時間 T_{ir0} との関連について調べるのに、その直前までに得られているベースライン変数や時間依存性の交絡要因、後治療歴については調整を行うが、それらの影響を受けた可能性のあるその時点以降に測定される変数については調整を行わないことから、「測定されていない交絡要因はない」の仮定の下でバイアスのない推論を行うことと同等となる。

結果：後治療実施により生存が 50%延び（95%信頼区間：2~294%生存を延ばす）、また、CPT-P 群は VDS-P 群と比して 9%生存を縮める（95%信頼区間：39%生存を縮める~42%生存を延ばす）結果であった。

3.2. 周辺構造比例ハザードモデルを用いた解析及び結果

周辺構造モデルでは、実際に観察された生存時間 T でなく、ある後治療歴 \bar{a} を有する場合の反事実的な生存時間 $T_{\bar{a}}$ に対してモデル化が行われる。生存時間データを区間データとして取り扱った解析はしばしば行われているが、本論文ではそのような取り扱いをせず、 $T_{\bar{a}}$ に対して比例ハザードモデルをあてはめた周辺構造比例ハザードモデルを提案する。

$$\lambda_{T_{\bar{a}}}(t|R, \bar{a}(t)) = \lambda_0(t) \exp[\beta_1 R + \beta_2 a(t)]$$

左辺は治療群が R で後治療歴 $\bar{a}(t)$ を有する集団での時点 t における周辺ハザード関数を表し、 $a(t)$ は時点 t における後治療実施の有無を表す。パラメータ β_1 は全ての被験者が同じ後治療を受けた場合の CPT-P 群の VDS-P 群に対する因果(対数)ハザード比を表し、関心のある直接治療効果である。パラメータ β_2 は全ての被験者が同じ群に属した場合の、後治療実施の有無に関する因果(対数)ハザード比を表す。通常の時間依存性の比例ハザードモデルとは異なり、この方法では、「測定されていない交絡要因はない」という条件の下で各被験者が受けた後治療を受ける確率を、後治療実施に関係する要因を用いて推定し、その逆数で被験者自身を重み付けした集団に対してモデルをあてはめパラメータ推定を行う。例えば、ある被験者はランダム化後 100 日目で後治療を受け 200 日目で死亡したとする。このような後治療歴（100 日目まで後治療を受けずそれ以降後治療を実施）を有する確率が 0.25 と推定された場合、対応する被験者の寄与は 4 (=1/0.25) 人分となる。このように各被験者を重み付けした擬似集団においては、後治療実施とそれに関係する要因との関連がなくなり、交絡がなくなることから、「測定されていない交絡要因はない」の前提の下でバイアスのない推論を行うことと同等となる。

結果：CPT-P 群の VDS-P 群に対するハザード比の推定値は 1.20 (95%信頼区間: 0.64-2.28) であり、また、後治療実施の有無に関するハザード比の推定値は 0.85 (95%信頼区間: 0.44-1.62) であった。

4. 構造ネストモデルと周辺構造モデルを用いた解析の比較

提案した周辺構造モデルは時間依存性共変量の解析で標準的に用いられる比例ハザードモデルと類似していることから、直感的にわかりやすい解析方法と考えられる。推定される効果の指標としては、構造ネストモデルでは生存時間が何倍になるか、あるいは、生存が何%延びるかという公衆衛生学的にしばしば関心のある指標であり、一方、周辺構造モデルでは臨床研究でよく用いられるハザード比が得られる。ただし、ベースライン生存時間がワイブル分布に従っていれば、構造ネストモデルを用いてもハザード比を得ることが可能である（今回のデータでは、CPT-P 群の VDS-P 群に対するハザード比の推定値は 1.10、後治療の効果は 0.50 となる）。最も重要な点は、観察研究の仮定を組み合わせた構造ネストモデルや周辺構造モデルを用いた解析の前提である「測定されていない交絡要因はない」という仮定に関する問題である。この前提を近似的にでも保証するためには、後治療実施に関係するより多くの要因を考慮する必要があるが、一方で、非常に多くの共変量を考慮するとパラメータ推定が不安定になるなどの問題があり、また、周辺構造モデルにおいてはより安定した重みを推定する必要があることなどから、解析に用いる要因の選択や後治療実施の有無のモデル化は重要である。これらの方法と比較すると、精度は多少落ちるが、構造ネストモデルを用いたランダム化に基づく解析は、ランダム化しているという前提のみで推論が行えるという利点がある。

5. 考察

時間依存性の比例ハザードモデルを用いた解析では、後治療実施の有無に関するハザード比の推定値は 0.97 であった。この解析結果は、後治療実施に関係する被験者の臨床状態や副作用の発現、臨床検査値異常などの時間依存性交絡要因を適切に調整していないことから、バイアスが生じている可能性がある。今回 2 つの因果モデルを用いた結果、後治療の効果のハザード比は約 0.5-0.9 であり、後治療として実施された放射線療法に効果がある可能性が示唆された。一方、一番の関心である CPT-P 群の VDS-P 群に対する直接治療効果は、今回の解析結果ではハザード比で 1.1-1.2 程度であり、ITT 解析の結果と比較して CPT-P 療法の効果がある方向にほんの少し値が変化しているものの、結論は変わらず CPT-P 療法の延命効果に関する優越性は認められなかった。後治療の効果がおそらく小さいためこのような結果が得られたと推察されるが、本研究のサンプルサイズは大きくないことから更なる検討が必要である。

後治療を実施する理由は様々であり、最初の治療の効果が無いという理由だけで後治療を実施するとは限らない。よって、今回試みたような、被験者全員が同じ後治療を受けた場合の群間比較には意味があると思われる。もちろん、ITT 解析には重要な意味があり主要な解析であることは異論がないが、本試験のように後治療に効果があると予想され、かつその実施割合が群間で異なっているような状況では、ITT 解析に加えて本研究で提案した方法を感度解析として適用すべきであり、結果として、後治療の効果が大きく最初の治療に関する群間比較に影響があると考えられた場合には、今後行われる類似の臨床試験のデザインにこの情報を活用できる（例えば、後治療の影響を受けにくい評価項目を考える）など、臨床試験方法論に貢献できると考える。

6. 結論

1. がんの臨床試験において、後治療の影響を調整し治療群の直接効果を検討するのに有用な、構造ネスト加速モデルと周辺構造比例ハザードモデルを提案した。
2. 提案した解析方法を切除不能非小細胞肺癌の臨床試験データに適用した結果、後治療の影響を調整しても群間差は認められないと推察され、また、後治療に効果がある可能性が示唆された。