

博士論文（要約）

構造ネスト制限付き平均損失時間モデルの  $g$  推定法による  
時間依存性治療の効果推定

萩原 康博

## 論文の内容の要旨

論文題目 構造ネスト制限付き平均損失時間モデルの  $g$  推定法による時間依存性治療の  
効果推定

氏名 萩原康博

### 1. 序文

臨床研究を行う主要な目的のひとつは、関心のあるアウトカムに対する治療の効果を推定することである。時間とともに治療が変化する治療は時間依存性治療と呼ばれ、時間依存性治療の効果をバイアスなく推定する方法について研究されてきた。時間依存性治療の効果をバイアスなく推定するためには、治療の影響を受ける時間依存性交絡因子を適切に調整しなければならないが、古典的な層別解析や回帰分析では治療の影響を受ける時間依存性交絡因子を適切に調整することはできないことが知られている。

時間依存性交絡因子を適切に調整し、時間依存性治療の効果をバイアスなく推定可能な手法として、構造ネストモデルの  $g$  推定法が提案されている。この方法には、周辺構造モデルなどのほかの因果解析手法と比較して、いくつか利点がある。生存時間アウトカムに対する構造ネストモデルとして、治療による生存時間の伸縮をモデル化した構造ネスト加速生存時間モデルと、治療によるリスクの増減をモデル化した構造ネスト累積生存時間モデルが提案されている。先に提案された構造ネスト加速生存時間モデルの  $g$  推定法は応用事例も多いが、妥当な  $g$  推定には人口打ち切りと呼ばれる処理が必要であり、この人工打ち切りにより統計的問題が引き起こされることが知られている。一方、構造ネスト累積生存時間モデルの  $g$  推定法は、人工打ち切りが不要という点で改良が加えられたが、モデル化の対象を生存時間ではなくリスクに変更されている。

近年、臨床研究における生存時間アウトカムの要約指標として、制限付き平均生存時間 (restricted mean survival time: RMST) や制限付き平均損失時間 (restricted mean time lost: RMTL) が注目されている。制限時点まで追跡した場合に期待される生存時間である RMST や、制限時点までに期待されるイベント発生により失われる時間 (イベント発生から制限時点までの時間) である RMTL は、(1) 治療効果が生存時間の増減で表され、(2) 比例ハザードモデルなどのモデルに依存せずに推定対象が定義され、(3) 生存確率や生存時間中央値と異なり制限時点までの追跡期間全体の要約になっているなどの利点がある。

以上のような背景を踏まえ、生存時間アウトカムに対する時間依存性治療の効果を推定する手法として、構造ネスト RMTL モデルの  $g$  推定法を提案する。

## 2. 提案手法

時点  $k=0, 1, \dots$  で測定される共変量を  $L_k$ 、治療を  $A_k$  ( $A_k=0$  で対照治療、 $A_k=1$  で試験治療を表す) とし、その履歴を上線で表す (たとえば  $\bar{L}_k=(L_0, \dots, L_k)$ )。生存時間を  $T$ 、制限時点  $m=1, 2, \dots, M$  までの制限付き損失時間を  $m-T(m)=m-\min(T, m)$  とする。ここでは、簡単のために最大の制限時点  $M$  まで生存時間の打ち切りが発生しないとする。静的治療レジメ  $\bar{a}=(a_0, a_1, \dots)$  のもとでの反事実生存時間を  $T^{\bar{a}}$  と定義し、同様の反事実制限付き損失時間を  $m-T^{\bar{a}}(m)=m-\min(T^{\bar{a}}, m)$  と定義する。時点  $k \leq m-1$  に関して、加法および乗法構造ネスト RMTL モデルをそれぞれ

$$\begin{aligned} & E[m-T^{\bar{a}_k, 0}(m) | \bar{L}_k = \bar{l}_k, \bar{A}_k = \bar{a}_k, T > k] - E[m-T^{\bar{a}_{k-1}, 0}(m) | \bar{L}_k = \bar{l}_k, \bar{A}_k = \bar{a}_k, T > k] \\ & = \gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi^*) \end{aligned}$$

および

$$\frac{E[m-T^{\bar{a}_k, 0}(m) | \bar{L}_k = \bar{l}_k, \bar{A}_k = \bar{a}_k, T > k]}{E[m-T^{\bar{a}_{k-1}, 0}(m) | \bar{L}_k = \bar{l}_k, \bar{A}_k = \bar{a}_k, T > k]} = \exp\{\gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi^*)\}$$

と定義する。ただし、 $\gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi)$  は、 $\psi$  に関して滑らかで、すべての  $\bar{l}_k$ 、 $\bar{a}_{k-1}$ 、 $\psi$  に関して  $\gamma_{k,m}(\bar{l}_k, \bar{a}_{k-1}, \mathbf{0}; \psi) = 0$  を満たす既知関数とし、 $\psi^*$  を  $p$  次元未知パラメータベクトルの真値とする。これらのモデルは、治療  $A_k$  による RMTL の因果的変化を表している。データ解析で有用な 1 パラメータモデルとして、加法モデルでは

$$\gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi) = \{\psi + 2\psi(m-k-1)\}a_k,$$

乗法モデルでは

$$\exp\{\gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi)\} = 1 + \frac{1 - \exp(\psi a_k)}{(m-k)^2} - \frac{2\{1 - \exp(\psi a_k)\}}{m-k}$$

が挙げられる。

構造ネスト RMTL モデルの推定法として、ランダム化にもとづく g 推定法と観察研究の仮定にもとづく g 推定法という、異なる仮定にもとづくふたつの方法を提案する。ランダム化にもとづく g 推定法はランダム化が保証する治療  $A_0$  が予後と独立に決定されているという仮定を利用する一方、観察研究の仮定にもとづく g 推定法は治療  $A_k$  ( $k=0, 1, \dots, M-1$ ) が観察履歴を条件付けたもとで予後と独立に決定されているという仮定を利用する。いずれの手法も、仮定のもとで構造ネスト RMTL モデルに対する漸近正規性を有する一致推定量を与える。ただし、現実のデータ解析で観察研究の仮定にもとづく g 推定法が妥当であるためには、傾向スコアモデルまたはアウトカムモデルを正しく特定する必要がある。

推定された構造ネスト RMTL モデルを用いて、対照治療を継続するときと試験治療を継続するときの周辺 RMTL  $E[m-T^{\bar{0}}(m)]$  と  $E[m-T^{\bar{1}}(m)]$  を推定できる。本研究では、任意の構造ネスト RMTL モデルのもとで周辺 RMTL  $E[m-T^{\bar{0}}(m)]$  の一致推定量を導出した。また、構造ネスト RMTL モデルが  $\gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi) = \gamma_{k,m}(l_0, a_k; \psi)$  のクラスに属するとき、 $E[m-T^{\bar{1}}(m)]$  を  $E[k-T^{\bar{0}}(k)]$ 、 $S^{\bar{1}}(k) = \Pr[T^{\bar{1}} > k]$ 、 $\gamma_{k,m}(l_0, a_k; \psi^*)$  ( $k=0, 1, \dots, m-1$ ) を用いて表現できる。

本研究ではこれを利用して  $E[m - T^{\bar{1}}(m)]$  の一致推定量を導出した。生存確率  $S^{\bar{1}}(k)$  の推定には、 $T^{\bar{1}}$  が区分指数分布に従うという仮定を用いた。加えて、試験治療を基準とした構造ネスト RMTL モデルを再推定することで、任意の  $\gamma_{k,m}(\bar{l}_k, \bar{a}_k; \psi)$  のもとで  $E[m - T^{\bar{1}}(m)]$  を一致推定する推定量も導出した。

最大の制限時点  $M$  までに生存時間の打ち切りが発生するときに用いる IPCW 法を提案した。情報のない打ち切りの仮定などのもとで、上述の提案手法は同様の性質を持つ。

### 3. シミュレーション実験

ふたつの実験を行った。実験 1 では、時点数の少ないデータで構造ネスト RMTL モデルのパラメータ推定に関して観察研究の仮定にもとづく  $g$  推定法の統計学的性能を評価した。実験 2 では、時間依存性交絡を伴う治療変更が発生するランダム化比較試験を想定したデータで、1 パラメータ構造ネスト RMTL モデルの  $g$  推定法により、ITT (intention-to-treat) 解析や PP (per-protocol) 解析によるバイアスをどの程度低減させることができ、またそのときの提案手法の推定精度はどの程度かを評価した。実験 1 では、(1) 回帰分析にはバイアスが認められ、(2) 傾向スコアモデルまたはアウトカムモデルが正しく特定されているとき、提案手法によるパラメータ推定値にバイアスは認められず、(3) 正しく特定されているかを問わずアウトカムモデルを用いることで推定精度が向上し、(4) アウトカムモデルを用いないと、パラメトリックモデルの推定を考慮しない標準誤差の推定量は標準誤差を過大評価し、(5) アウトカムモデルを正しく特定すると、特に治療モデルを誤特定したときに、パラメトリックモデルの推定を考慮する標準誤差の推定量は、標準誤差を過小評価する傾向にあった。実験 2 では、周辺 RMTL の差に関して、提案手法は ITT 解析や PP 解析のバイアスを低減可能であることが示された。帰無の治療効果以外のシナリオでは多くの場合、提案手法の平均二乗誤差 (バイアスと推定精度の両方を考慮した指標) は ITT 解析と PP 解析より小さかった。ただし、乗法モデルを用いた  $E[m - T^{\bar{1}}(m)]$  の推定では、治療効果が有害な場合、極端な推定値が得られることがあった。

### 4. 提案手法の MEGA study への適用

冠動脈疾患に対するプラバスタチンの 1 次予防効果を検証したランダム化比較試験である MEGA study データに提案手法を適用した。1 パラメータ加法モデルと乗法モデルに対してランダム化にもとづく  $g$  推定法と観察研究の仮定にもとづく  $g$  推定法を適用し、プラバスタチンをまったく服用しなかった場合と継続して服用した場合の周辺 RMTL を推定した (表 1)。観察研究の仮定にもとづく  $g$  推定法よりランダム化にもとづく  $g$  推定法で大きな治療効果を表す推定値が得られたが、ともに ITT 解析による治療効果の推定値と大きくは違わなかった。

## 5. 考察

本研究では、生存時間アウトカムに対する新しい構造ネストモデルとして構造ネスト RMTL モデルを提案し、このモデルの推定法として、ベースライン以後の治療メカニズムに仮定を置かないランダム化にもとづく g 推定法と、ベースライン以後の治療メカニズムに仮定を置くことで推定効率の向上と識別可能なモデルの拡大が可能な観察研究の仮定にもとづく g 推定法を提案した。提案した g 推定法は、既存の生存時間アウトカムに対する構造ネストモデルの g 推定法と同様の仮定のもとで、漸近正規性を持つ一致推定量を与えることを示した。加えて、推定した構造ネスト RMTL モデルを用いた周辺 RMTL の推定手法をふたつ提案した。シミュレーション実験では有限標本での提案手法の統計的性能を評価し、提案手法は仮定のもとで構造ネスト RMTL モデルのパラメータが表す治療効果をバイアスなく推定し、治療変更を伴うランダム化比較試験では ITT 解析や PP 解析のバイアスを低減できることが示された。提案手法を MEGA study データに適用した結果、ITT 解析と提案手法による結果に大きな差異は見られなかったが、提案手法による実データ解析方法を示し、実際の臨床研究データにも提案手法を適用可能であることを示した。以上の成果により、提案手法は生存時間アウトカムに対する時間依存性治療の効果を推定する手法として有用と考えられる。今後の研究課題には、(1) 構造ネスト RMTL モデルに対する包括的なモデル選択手法の開発、(2) より推定効率の高い augmented-IPCW 法の応用、(3) 上下界のない構造ネスト RMTL モデルの g 推定法の開発が含まれる。

表 1 1 パラメータモデルによる 60 か月までの周辺 RMTL の推定結果

方法	食事療法単独	プラバスタチン併用	差	比
ITT	0.729 (0.556, 0.902)	0.500 (0.354, 0.646)	-0.229 (-0.455, -0.003)	0.69 (0.47, 1.00)
加法モデル				
ランダム化	0.745 (0.546, 0.925)	0.477 (0.300, 0.667)	-0.268 (-0.564, 0.035)	0.64 (0.36, 1.06)
観察研究の 仮定	0.709 (0.529, 0.879)	0.516 (0.363, 0.684)	-0.193 (-0.438, 0.050)	0.73 (0.47, 1.09)
乗法モデル				
ランダム化	0.750 (0.547, 0.970)	0.487 (0.337, 0.672)	-0.262 (-0.573, 0.034)	0.65 (0.39, 1.06)
観察研究の 仮定	0.713 (0.532, 0.893)	0.513 (0.368, 0.689)	-0.200 (-0.450, 0.055)	0.72 (0.46, 1.10)

データは推定値 (95%信頼区間)。RMTL の単位は月。