

論文の内容の要旨

論文題目 治療変更を伴う臨床研究データにおける主要層別法に基づく Quality of Life データの解析手法の提案

氏名 川原 拓也

1. 序文

がん臨床試験では、割付治療の開始後、有害事象や腫瘍の増悪に応じて後治療が行われることがある。また、治療の有効性を評価するために、奏効率や生存時間の他に、患者が報告する Quality of Life (QoL) が測定されることが多い。がん臨床試験にて収集された QoL データを解析する際、統計学的方法論上の課題が二点存在する。一点目は、比較を行う治療レジメンを明確に設定することである。割付治療から、患者データに応じた後治療まで明示的に定義された一連の治療レジメンを、動的治療レジメンと呼ぶ。患者全員が仮想的にある動的治療レジメンに従った場合の比較は、intention-to-treat 解析を補足する解析方法として有用である。

二点目は、QoL の欠測である。患者の都合等による欠測と、患者が死亡した後の欠測とは異なる性質をもつ。死亡による欠測が存在する状況で、生存者に限定して解析を行うと、QoL への治療の影響について、誤った推測を導く危険性がある。死亡による欠測へのアプローチのひとつは、主要層別法に基づいた survivor average causal effect (SACE) に着目することである。SACE は、仮想的にいずれの治療を受けた場合にも生存しているであろうサブグループでの治療効果と定義され、死亡による欠測が引き起こすバイアスを回避可能であるという特長がある。SACE 推定に関する方法論の研究の多くは一時点の治療の設定であり、動的治療レジメンの SACE 推定に関する研究はない。

本研究では、ランダム化がん臨床試験を想定し、仮想的にいずれの動的治療レジメンに従って治療を受けても生存しているであろう患者での QoL への治療効果 (SACE) およびその推定方法を提案する。次に、シミュレーション実験により、提案手法の統計学的性能を評価する。最後に、本研究を実施するモチベーションとなった、転移・再発乳がんに対する一次療法を評価したランダム化比較試験、Selection of effective chemotherapy for breast cancer (SELECT BC) 試験に提案手法を適用する。

2. 方法

本研究では、SACE を推定するために、反事実結果変数に対する独立性の仮定をおく。治療が一時点のみのとき、ベースライン共変量を L_0 、治療の影響を受ける共変量を L_1 、QoL を Y 、QoL 測定時点まで生存していることを表す指示変数を S (生存なら 1、死亡なら 0) とする ($S = 0$ のとき Y は定義されない)。ランダム割付される治療 $A = a$ ($a = 0, 1$) に対する反事実生存変数を S^a 、反事実 QoL を Y^a とする ($S^a = 0$ のとき Y^a は定義されない)。 $S^1 = 1$ かつ $S^0 = 1$ である患者での反事実 QoL の治療の平均因果効果と定義される SACE は、 $S \prod S^{1-a} |L_0, L_1, A = a$ 、および L_1-S 間、 $S-Y$ 間、 L_1-Y 間に未測定交絡がない仮定の下で識別可能である。

本研究では、前段落の仮定を経時的に行われる治療の場合に拡張する。時点 $k = 0, \dots, K$ において、追跡打ち切りを表す指示変数 C_k (打ち切りであれば 1、追跡中であれば 0)、生存を表す指示変数 S_k (生存であれば 1、死亡であれば 0)、共変量ベクトル L_k 、治療 A_k が測定され、加えて時点 K では QoL 値 Y が測定されるとする。本研究では、割付治療 $A_0 = a_0$ ($a_0 = 0, 1$) から開始し、共変量および治療履歴に応じて時点 k における治療 A_k を規定する動的治療レジメンのクラスを考える。これは、共変量および治療履歴の関数の集合 $g(a_0)$ として表される。提案する SACE は、「仮想的に $g(a_0)$ (a_0

= 0, 1) のいずれに従って治療を受けた場合にも、QoL 測定時点まで生存しているであろう患者での反事実 QoL の差の期待値」と定義される。 $g(a_0)$ に従って治療を受けた場合の反事実生存変数、反事実 QoL をそれぞれ $S_K^{g(a_0)}$, $Y^{g(a_0)}$ とする。SACE 推定のために、一般的な因果仮定(一致性、逐次交換可能性、正の治療確率)の他に、反事実結果変数に対する逐次条件付き独立性の仮定、

$$\begin{aligned}
 & S_K^{g(1-a_0)} \prod S_k \left| \bar{L}_{k-1} = \bar{l}_{k-1}, \bar{A}_{k-1} = \bar{a}_{k-1}^{g(a_0)}, S_{k-1} = 1, C_k = 0 \right. \\
 & Y^{g(1-a_0)} \prod S_k \left| \bar{L}_{k-1} = \bar{l}_{k-1}, \bar{A}_{k-1} = \bar{a}_{k-1}^{g(a_0)}, S_{k-1} = 1, C_k = 0, S_K^{g(1-a_0)} = 1 \right. \\
 & S_K^{g(1-a_0)} \prod L_k \left| \bar{L}_{k-1} = \bar{l}_{k-1}, \bar{A}_{k-1} = \bar{a}_{k-1}^{g(a_0)}, S_k = 1, C_k = 0 \right. \\
 & Y^{g(1-a_0)} \prod L_k \left| \bar{L}_{k-1} = \bar{l}_{k-1}, \bar{A}_{k-1} = \bar{a}_{k-1}^{g(a_0)}, S_k = 1, C_k = 0, S_K^{g(1-a_0)} = 1 \right.
 \end{aligned}$$

を提案する。この仮定は、治療が一時点の場合の仮定を拡張したものであり、たとえば図の点線のような時間依存性共変量と QoL の未測定交絡が存在する場合には成立しない。仮定を成立させるためには、時間依存性交絡因子、生存および QoL に影響しうる因子を十分に測定する必要がある。

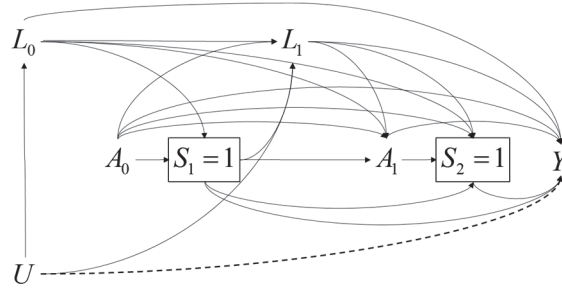


図. $K = 2$ のときの因果ダイアグラム

提案する SACE 推定の手順は、次の通りである。まず、治療確率および打ち切り確率に対して適切なモデルを当てはめ、個人ごとの時間依存性の重みを求める。次に、死亡ハザードに対するモデルの重み付き最尤法により、動的周辺構造モデルパラメータを推定する。最後に、これらのパラメータ推定値を用いて個人ごとの重みを求め、 $g(a_0)$ ($a_0 = 0, 1$) のいずれかに実際に従って治療を受けた患者で観察された QoL 値に対し、重み付き解析を行う。最終解析で用いる重みは、実際の治療歴を受ける確率の予測値の逆数と、もう一方の動的治療レジメンに仮想的に従った場合の生存確率の予測値との積である。信頼区間の構成には、サンドイッチ分散を用いる。

3. シミュレーション実験

3.1 設定

提案手法が妥当であるために必要とする仮定を満たすデータ発生プロセスの下、実臨床試験を想定した小標本 ($n = 500$)、および大標本 ($n = 20\,000$) において、二次治療の傾向および生存確率の変動に応じた提案手法の性能を評価した。仮定の崩れに対する感度を評価するため、図の点線のパスが存在する状況を想定した実験も行った。シミュレーションの繰り返し回数は 2 000 回とした。

3.2 結果

提案手法の点推定値は、一致性を示した。提案手法の信頼区間は名目有意水準を担保していた。興味のある動的治療レジメンに従った対象者の割合が小さくなれば、信頼区間の保守的な傾向が高くなった。被覆確率と生存確率との関連はみられず、推定標準誤差は真の標準誤差を約 1.1 倍から 1.3 倍過大評価した。仮定が崩れた場合のバイアスは、実用上許容される程度の大きさであった。

4. データ解析

4.1 方法

SELECT BC 試験は、転移・再発乳がん患者を対象とし、一次治療として、S-1 のタキサン系薬剤に対する全生存期間での非劣性を検証した臨床試験である。本研究では、Quality of Life Questionnaire Core version 3 により測定された全般的健康度をアウトカムに設定した。

治療開始前に QoL 値が測定された 386 人(タキサン群 176 人、S-1 群 210 人、この集団における 1 年以内の死亡は約 14%)を解析対象とした。ベースライン共変量として、年齢、ホルモン受容体有無、肝転移有無、手術から再発までの時間、リンパ節転移個数、補助化学療法歴としてタキサン系薬剤、5-fluorouracil 系薬剤、内分泌療法を考慮した。時間依存性共変量として、治療開始後の最新の Performance Status、治療開始後の増悪、および QoL 値の変動を考慮した。治療確率、死亡ハザードに対して、それぞれ多項ロジスティックモデルとプールド・ロジスティックモデルを当てはめた。提案手法を適用し、「二次治療が必要となれば割付治療間のクロスオーバーを行う」という動的治療レジメンの、治療開始後 1 年後の全般的健康度への影響を検討した。

4.2 結果

二次治療の選択に影響する因子として、増悪や全身状態だけでなく、QoL 値の変動が抽出された。いずれの動的治療レジメンに従った場合でも 1 年間生存する対象者の割合は、86.3%と推定された。全生存者を用いた割付群間の差 (intention-to-treat 解析) よりも、提案手法の方が大きい群間差が得られた(表)。

5. 考察

本研究では、割付治療($a_0 = 0, 1$) から治療を開始するクラスの動的治療レジメンに限定した。これは、SELECT BC 試験データの解析では自然であると考えられる。なお、より広い動的治療レジメンのクラスを含むように提案手法を拡張できると考えられるが、解析上データの複製が必要となる場合があることに注意する必要がある。

群間差が統計学的に有意であるか否かは、症例数の多寡に依存する。したがって、がん臨床試験の QoL データの解析では、得られた群間差の点推定値が臨床的に有用な差であるか否かを検討する必要がある。過去に提案されたガイドラインを参照すると、全生存者を用いた解析結果はささいな差に分類されたが、提案手法による解析結果は臨床的に有用な小さな差に分類された。このことから、SACE が対象とする集団においては、臨床的に有用な群間差が存在する可能性が示唆された。

表. 通常解析と提案手法による解析間の比較

推定目的	期待値	群間差	95%CI
全生存者を用いた解析			
S-1 群 ($n = 131$)	61.9	2.7	(-3.4, 8.8)
タキサン群 ($n = 105$)	59.2		
提案手法による解析結果(条件付きランダムな欠測の仮定)			
S-1 から開始するレジメン $g(1)$	61.1	6.6*	(-4.5, 17.7)
タキサンから開始するレジメン $g(0)$	54.5		

*Survivor average causal effect の推定値. CI, confidence interval.

シミュレーション実験により、提案手法の点推定値には漸近的にバイアスがないことと、サンドイッチ分散による信頼区間は名目の有意水準を保つものの、保守的であることが確認された。ブートストラップ分散は、サンドイッチ分散よりもより良い性能を示す可能性がある。

本研究の限界のひとつは、シミュレーション実験の範囲である。本研究では時間依存性交絡因子と QoL との間の未測定交絡の存在に対する感度を評価したが、他の仮定の崩れも含めて、より広い範囲でのシミュレーション実験が必要であろう。また、提案手法の限界として、二つの動的治療レジメンの比較に限定されることが挙げられる。三群比較の状況において SACE を推定する試みがあり、本研究にも応用可能か検討が必要である。

6. 結論

治療変更を伴うがん臨床試験において測定される QoL データに対して、いずれの動的治療レジメンに従った場合にも生存する主要層での動的治療レジメンの治療効果、およびその推定方法を提案した。シミュレーション実験により、提案手法の統計学的性能(点推定値の一致性・信頼区間の妥当性)を確認した。SELECT BC 試験データを用いて提案手法を実データへ適用する詳細な手順を述べ、QoL データに関して通常の解析よりも大きな、臨床的に有用な群間差が存在する可能性が示唆された。提案手法は、がん臨床試験における QoL データの解析方法として有用である。