

交互作用を含む観察データにおける傾向スコア解析手法による結果の差異

加葉田 大志朗^{1,2)}, 山本 紘司¹⁾, 新谷 歩²⁾

1) 大阪大学 大学院医学系研究科 臨床統計疫学寄附講座

2) 大阪市立大学 大学院医学研究科 医療統計学講座

1. はじめに

観察研究により得られたデータの解析では、交絡等のバイアスへの対処が必須となる。特に近年医学分野においては、交絡の除去を目的として、傾向スコアを利用した解析手法が普及している。傾向スコアは治療に割り当てられる確率を背景情報などから計算したものである。現在提案されている主流な傾向スコア解析としては、類似した傾向スコアで対照群を選ぶマッチング法、傾向スコアの逆数を用いて重み付けするIPTW(Inverse Probability of Treatment Weighted)法、IPTW法と多変量回帰モデルを組み合わせたDR(Doubly Robustness)法がある。特にマッチング法は解釈がしやすく、臨床で利用されることも多い。しかしデータによっては各手法も使い分けが必要となるが、その詳細は知られていない。

2. ケーススタディ

本研究では、舌癌患者における予防的リンパ郭清の効果を検証する目的で得られた実際の観察研究データの解析より着想を得た。この解析では、傾向スコアによるマッチング法と、3因子交互作用項を含む多変量回帰モデルを用いて解析を行った。その結果傾向スコアでは治療が有効であることが示されたが、多変量回帰モデルにおいては治療効果と年齢・腫瘍サイズの間交互作用が確認された。つまり治療効果は術前の年齢と腫瘍サイズにより変化し、中年層(40~60歳)で腫瘍サイズが5mm以上の症例において治療は有効であることが示唆された。一方、高齢あるいは若年層で5mm以上の症例においては、予後を悪化させる可能性があることが示唆された。(図1)

この解析では感度解析として、IPTW法とDR法を用いた解析も行ったが、どちらもマッチング法のように治療効果があるという結果は得られなかった。

3. 検証内容

上記の観察研究データ解析におけるマッチング法は、傾向スコアが最も近い症例同士をマッチングするSimple nearest neighbor matchingを用いた。当該手法による解析では基本的にATT(Average Treatment effect for Treated)、主に治療が行われた症例における治療効果を推定することが知られている¹⁾。今回のマッチング法を用いた解析では、中間年齢層で腫瘍サイズが大きい、つまり臨床的に最も治療をうける確率が

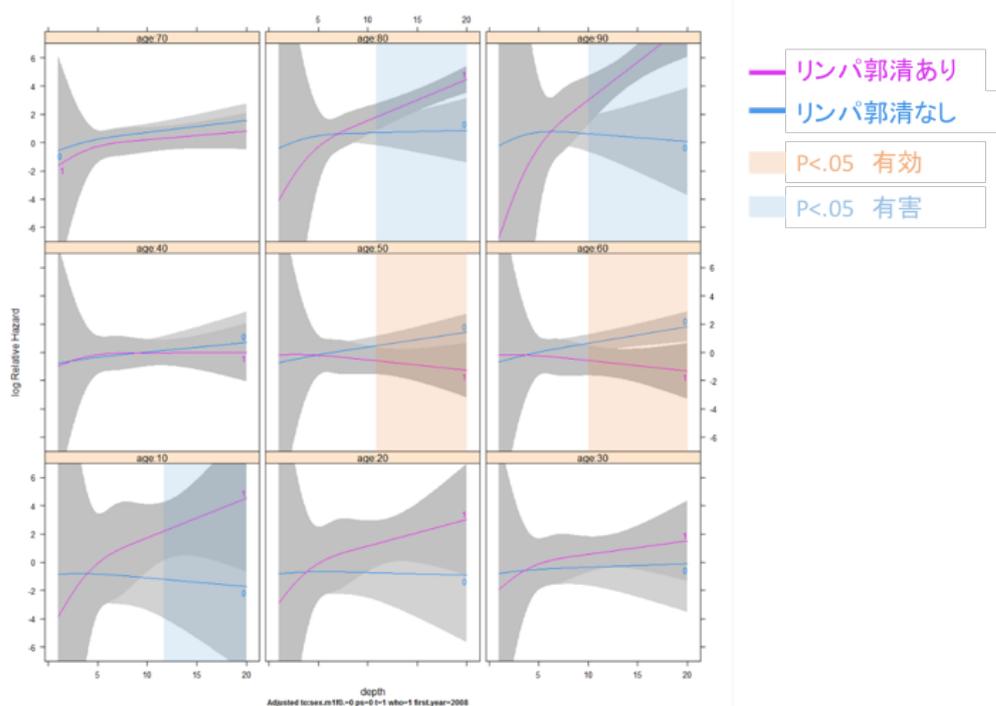
高い症例における平均的な治療効果が反映されたと考えることができる。

IPTW法とDR法はマッチング法とは異なり解析に用いる症例数が減少しないため、最近では臨床論文でも用いられることが増えてきた。しかし基本的にIPTW法とDR法はATE(Average Treatment Effect)すなわち、集団全体の平均的な効果を検証している。そのため当事例のように治療効果が対象症例の背景によって変化する状況においては、集団全体の平均的な治療効果を推定すると、治療を受けた症例における治療効果が十分に検出されないことが考えられる。またIPTW法とDR法においても重みを調整することで、ATTと同様の解析結果を得られることが提唱されているが²⁾、実際にそれらを用いた臨床論文は目にすることが少ない。

傾向スコアを用いた手法比較論文においてはATEに着目した報告が行われることが多く、当事例のように、治療効果が交互作用を持つデータにおける解析手法については言及されることが少ない。しかし医学領域において治療効果が患者背景により変化することは多くのケースで考えられ、その影響を考慮して解析手法を選択することは重要であると考えられる。

本研究では当事例のように治療効果との交互作用を含むデータにおいて、各種傾向スコア解析手法を適用した際の結果の違いを検証し、目的に応じた適切な解析手法を提案することを目的とする。本報告では、当事例に類似した擬似データを発生させ、モンテカルロシミュレーションにより各種傾向スコア解析手法を適用した結果の比較を行い報告する。

(図1) 3way-interactionを含む多変量回帰モデルによる解析結果



- 1) Stuart, E. A., Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science : A Review Journal of the Institute of Mathematical Statistics* 25(1): 1–21, 2010.
- 2) Hirano, K., and Imbens, G. W., Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services and Outcomes Research Methodology* 2(3–4): 259–278, 2001.