

[別紙 1]

## 論文の内容の要旨

論文題目 非ランダム化医学研究における治療効果の推定方法に対する検討

指導教官 大橋 靖雄 教授

東京大学大学院医学系研究科

平成 12 年 4 月入学

博士後期課程

健康科学・看護学専攻

氏名 汪 涛

### はじめに

医学研究において治療や曝露などの効果を検証するためにはランダム化研究を行うことがベストであるが、倫理上などの理由により観察研究のような非ランダム化研究に頼らざるを得ない場合もある。非ランダム化研究においては、比較群間(例えば治療群と対照群)で共変量(背景因子、リスク因子)の分布が異なっている可能性があり、治療効果を正確に推定するために共変量の分布に対する調整が必要である。

共変量の分布を調整し治療効果を推定する方法として、共分散分析がよく用いられている。共分散分析は結果(例えば脂質低下の量)－治療割り付け(治療群か対照群か)－共変量という 3 者の関係をモデル化することにより治療効果を推定する。3 者の関係が明らかであり、正確なモデル化が可能な場合、共分散分析による推定はバイアスが生じないあるいはほぼ生じない。共分散分析以外には、Rosenbaum and Rubin により提案されたプロペンシティブスコア(propensity score)を用いる調整法(本研究では PS 調整法と呼ぶ)と Robins and Mark により提案された E 推定法がある。プロペンシティブスコアとは、観察された共変量を与えられた下で治療を受ける条件つき確率である。PS 調整法と E 推定法は共に、1. 治療割り付け－共変量 2 者の関係のみをモデル化することによりプロペンシティブスコアを推定する；2. プロペンシティブスコアを利用して治療効果を推定する、という 2 段階の推定方法である。

PS 調整法は治療効果を推定する第 2 段階において層別あるいはマッチング(本研究では層別のみを検討対象とする)を用いるので、プロペンシティスコアの推定値が連続量であることに起因する残差交絡の問題が存在する。E 推定法は残差交絡の問題がないが、結果変数が 2 値変数である場合に適用不可という欠点がある。

これらの方法を実際の研究でより正確に応用するためには、3 つの推定法がどのような性質(不偏性、精度、モデルの誤特定に対する頑健性)を持つかについての研究が必要であると考えられる。

## 目的

本研究の目指すところは、非ランダム化研究において治療効果の推定を行うために提案されている PS 調整法、E 推定法、共分散分析という 3 つの推定法は、連続結果変数の場合だけでなく、これまで明確されていない 2 値及び計数結果変数の場合にどのような性質を持つかを明らかにすることと、3 つの推定法の比較することである。

具体的な目的は、まずシミュレーションを用い、1. モデルの誤特定がない場合、PS 調整法において残差交絡による治療効果の推定量の不偏性(バイアスの大きさ、方向とそのバイアスに影響を及ぼす要因)と精度について検討することと 3 つの推定方法の不偏性や精度などの性質を比較すること；2. モデルの誤特定がない場合、「free parameters」をモデルに加えることが治療効果の推定量に与える影響について検討すること；3. モデルの誤特定が存在する場合、誤特定により各推定法における治療効果の推定量に与えるバイアスの大きさと方向について検討すること；4. PS 調整法において層別することにより治療効果の推定量に関するバイアスがどの程度減少するかについて検討することである。次に実データを用い、各推定方法による治療効果の推定量の性質に対して検討し、比較することである。

## 方法

シミュレーション研究においては、以上の目的に応じ、3 種類の結果変数の場合それぞれにおいて、治療効果、結果と関連する共変量の効果、治療割り付けと関連する共変量の効果を 2 つずつのレベルを設定し、サンプルサイズ 2500 とサンプルサイズ 100 のシミュレーションを行う。シミュレーションの回数については、サンプルサイズ 2500 のシミュレーションは 50 回、サンプルサイズ 100 のシミュレーションは 1000 回である。推定法の性質に対する評価には、治療効果の推定量の不偏性を評価する PB (percent bias)、精度を評価する STD(standard deviation)及び不偏性と精度を同時に評価する MSE (mean square error) という 3 つの指標を用いる。PS 調整法においては、層別することによるバイアスの減少に対する評価には、PBR (percent bias reduction) を用いる。

実データへの応用研究については、KLIS という高脂血症の治療に関する研究のデータを用い、共変量調整なしの解析と共変量調整ありの解析を行い、その結果を比較する。共変量調整ありの解析においては、3 つの推定法を用い、治療効果を推定し、それらの推定量の不偏性や精度について比較する。

## 結果

サンプルサイズ 2500 のシミュレーション結果：

1. モデルの誤特定がない場合、PS 調整法による推定は少しバイアスがあるのに対し、E 推定法と共分散分析による推定は不偏あるいはほぼ不偏であることが示された。精度については、連続的結果変数の場合には PS 調整法が他の両方法より良くないものの、2 値結果変数と計数結果変数の場合には PS 調整法が共分散分析とほぼ一致することが示された。計数結果変数の場合には E 推定法の精度が他の両方法より良くないことが示された。

2. モデルの誤特定がない場合、「free parameters」をモデルに加えることにより各推定法による推定にバイアスの増加はないことが示された。精度について、「free parameters」を加えることにより、計数結果変数の場合に E 推定法による推定は良くなったことが示さ

れたが、この場合の他の両方法による推定はあるいは他の結果変数の場合の各推定法による推定にはあまり変化がないことが示された。

3. モデルの誤特定が存在する場合、各種類の結果変数の場合において各推定法による推定にバイアスが増加したことが示されたものの、2 値と計数結果変数の場合において PS 調整法による推定には誤特定によるバイアスの増加が他の両方法より小さいことが示された。

4. PS 調整法において層別することによる治療効果推定量に関するバイアスの減少については、連続と計数結果変数の場合にはすべての場合に減少したことが示された。その減少の程度(PBR)が連続の結果変数の場合に 88.22%~89.43%、計数結果変数の場合に 70.34%~93.11%であることが示されたが、2 値結果変数の場合には減少不能の場合もあることが示された(PBR が 12.77%~75.94%であった)。

サンプルサイズ 100 のシミュレーション結果：

モデルの誤特定がない場合には、PS 調整法による推定が他の両方法による推定よりバイアスが少し大きいものの、各推定法による推定はいずれもバイアスを持つことが示された。これ以外の場合の結果については、3 つの評価指標がサンプルサイズ 2500 のシミュレーションの場合より大きくなったことを除けば、同様のことが示された。

PS 調整法において層別することによる治療効果推定量に関するバイアスの減少については、2 値結果変数の場合に逆にバイアスが増加してしまっただことが示された(4 層別する時の PBR が約 -8.51%~14.02%、5 層別する時が -9.09%~14.44% であった)。

実データを用いた解析結果：

調整する前に比較群間でリスク因子の分布の違いに対する検定は有意であることが示された。調整なしの解析で対照群に対する pravastatin 介入群の 6 ヶ月時点の脂質低下効果は TC(総血清コレステロール)：11.383mg/dl、LDL-C(低比重コレステロール)：13.539mg/dl であり、P 値が共に<0.0001 であった。

プロペンシティスコアで 5 層別した後、各層において比較群間でほとんどのリスク因子の分布の違いに対する検定は有意ではないこととなった。層別解析により、対照群に対する pravastatin 介入群の 6 ヶ月時点の脂質低下効果は TC：18.075mg/dl、LDL-C：19.296mg/dl であり、P 値が共に<0.0001 であった。

また、共分散分析を用いた解析により、脂質低下効果は TC：18.378mg/dl、LDL-C：19.365mg/dl であり、P 値が共に<0.0001 であった。E 推定法を用いた解析により、脂質低下効果は TC：14.385mg/dl、LDL-C：14.485mg/dl であり、P 値が共に<0.0001 であった。

推定量の標準誤差については、TC：共分散分析 0.968 < PS 調整法 1.048 < E 推定法 3.152、LDL-C：共分散分析 0.975 < PS 調整法 1.112 < E 推定法 2.345 であった。

## 考察

シミュレーションの結果によって、正確なモデル化が可能な場合、共分散分析あるいは E 推定法による推定が不偏あるいはほぼ不偏であるのに対し、PS 調整法による推定は残差交絡のため、少しバイアスを持つものの、実データを用いた解析結果によって、それほど大きな問題はないことが示された。PS 調整法において、治療効果の推定量に関するバイアスの大きさを推定するため、図や t 検定あるいはカイ二乗検定を用いて層別後各層において重要な交絡要因の分布の均一性に対する検証を行うことが役に立つと考えられる。この方法はモデルの誤特定が存在するかどうかに対する検証にも役に立つ。非ランダム化研究において、特に疫学調査のような観察研究の場合において、共変量が多くかつ変数間の関係が複雑であることは常であり、多くの変数間の関係を明らかにすることは容易ではないので、モデルを誤特定する可能性が高い。この意味で、PS 調整法の実研究における応用性は高いのではないかと考えられる。精度については、PS 調整法は他の両方法より少し低いことが示された。これについて 2 つの原因が考えられる。1 つは比較群間のプロペンシティスコア

の分布がどの程度重なっているかということが考えられる。もう 1 つは治療効果の推定量の計算方式が考えられる。

モデルの誤特定が存在する場合には、特に計数結果変数の場合、PS 調整法は他の両方法よりモデルの誤特定に対する頑健性がよいことが示された。3 つの推定法のうち、E 推定法はモデルの誤特定に対する頑健性が最も悪いことが示され、治療割り付け—共変量間の明らかな関係に対する要求が高いと考えられる。

変数選択問題については、PS 調整法も共分散分析も、変数選択により、精度がよくなったことが示された。しかしながら、モデルの誤特定を防止する意味で、PS 調整法における変数選択を慎重に行うことが望ましい。

本研究では、3 つの推定法、特に PS 調整法の性質と実用性について、いくつかの有用な結果を得たが、不十分あるいは明確にしていない点もあり、これらの点について、今後の課題として以下のような研究が必要と考えられる。

PS 調整法において、プロペンシティブスコアで層別することにより治療効果を推定する場合、層の数や各層の境、重みによる影響に関する研究である。

2 値結果変数の場合においては、シミュレーションの結果により、PS 調整法による推定は共分散分析による推定とほぼ一致することが示されたが、層別法を用いてプロペンシティブスコアに関するバイアスを減少させることにより治療効果の推定量に関するバイアスが必ずしも減少することはなかった。この点について、もっと明らかにするよう、解析的な研究である。

## 結論

モデルの誤特定がない場合、PS 調整法による推定には残差交絡のため少しバイアスが伴うものの、サンプルサイズが充分である場合、実研究における応用にはそれほど大きな問題はない。PS 調整法は、特に計数結果変数の場合にはモデルの誤特定に対する頑健性が他の両方法よりよい。一方、E 推定法はモデルの誤特定に対する頑健性が 3 つの推定法の中で最も悪い。また、PS 調整法において、簡単な図や t 検定あるいはカイ二乗検定を用いて比較群間で共変量の分布の均一性に対する検証が可能なので、実研究における応用性が高いと考えられる。