

多変量解析

<教材提供>
AMED 支援「国際誌プロジェクト」提供

無断転載を禁じます

無断転載禁止

目次

はじめに

多変量解析の意義と無作為化を行わない研究における背景のズレ

多変量解析の種類

多変量解析で検討できる説明変数の数

説明変数の選び方

参考文献

無断転載禁止

はじめに

治療の効果を調べる研究において、治療を受けた患者さんと治療を受けなかった患者さんで死亡などのアウトカムを他の因子の影響を考慮せずに比べてしまうと、一見治療を受けた患者さんの方が、死亡割合が高くなってしまいう結果が出る場合があります。このような現象は治療を行った結果として患者さんの具合が悪くなったのではなく、もともと具合の悪い患者さんが治療を受けていることによって、起こってしまっていることが原因と考えられます。研究対象者を無作為に各群に割りつける「無作為化」を行わない研究では、このように研究対象者の背景が比較群間で不均衡になることで、治療の効果を適切に評価できなくなってしまう「交絡」という現象が起こります。この単元では研究において無作為化を行う意義と、無作為化が実施できない時に、統計的に交絡を除去する**多変量解析**の基本について解説します。

学習目標

本単元を通じてあなたが習得を目指すのは

- 多変量解析の基本と注意点を習得する
- 状況に応じた多変量解析が選べるようになる
- 多変量解析で考慮する変数の選び方の基本を習得する

多変量解析の意義と無作為化を行わない研究における背景のズレ

無作為化を行わず、心臓病ハイリスクの患者さんでアスピリンに延命効果があるかどうかを調べた研究の例を見てみましょう[1]。アスピリンを使用した群(アスピリン使用群)が使用していない群(アスピリン非使用群)と比較して死亡リスクが高くなるという結果が出たとします。研究に参加した患者さんの背景(下の表)をみると、アスピリン使用群では平均年齢がアスピリン非使用群に比べて高く、男性の割合も高く、また糖尿病・高血圧・冠動脈疾患などの既往歴をもつ患者が多いなど、もともと予後の悪そうな患者さんが多いことが分かりました。これではアスピリンに害があったのか、あるいはそもそも予後の悪い人がアスピリン使用群に多かったために間違っアスピリンに害があるかのような結果が出たのか分かりません。このような背景情報のズレによって、治療効果などを適切に評価できなくなる状態を「**交絡(Confounding)**」といいます。

無作為化を行わない研究の患者背景

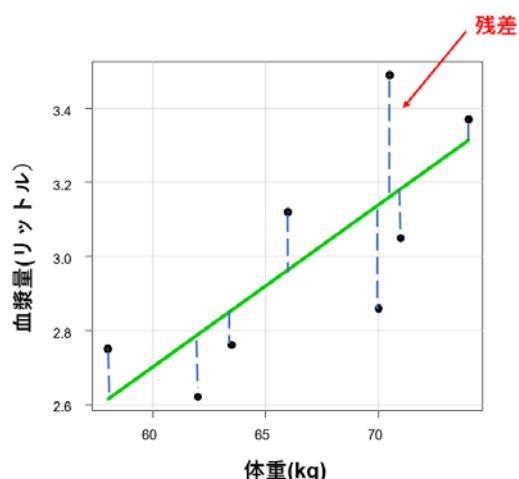
	アスピリン 使用群 (2310 人)	アスピリン 非使用群 (3864 人)
年齢(平均)	62 歳	56 歳
男性	77%	56%
糖尿病歴	17%	11%
高血圧	53%	41%
冠動脈疾患歴	70%	20%

このような状態を防ぐために、治療などを無作為に患者さんへ割り当てる「**無作為化**」が行われることもありま

す。しかし無作為化は治療のような「有益な」暴露因子の効果を見るときは実施可能ですが、「肥満が生活習慣病を誘発する」というような「有害な」可能性のある暴露因子の効果を検証する場合には、害になる事柄に研究対象者を割り付けることは倫理上問題となり実施が不可能になります。無作為化のできない研究で治療効果などを比較する場合、比較群間の背景の違いを解析で考慮し、交絡が起こらないようにすることが求められます。

交絡への対処法としては様々な手法が存在しますが、その代表的なものが**回帰モデル**を用いた多変量解析です。回帰モデルは、日本では中学 1 年生で習う $y = \alpha + \beta x$ のような形式で表現することができます。例えば下の図は、8 人の研究対象者の血漿量と体重との関係をグラフに表したものです。縦軸(y 軸)には血漿量、横軸(x 軸)には体重がプロットされています。緑で描かれている直線を「線形回帰直線」又は「最小二乗(回帰)直線」と呼び、血漿量を y 、体重を x と置くと、回帰モデルをあてはめたモデルは、正確には以下のように表すことができます。

$$y = 0.086 + 0.04x + \text{残差}$$



通常臨床研究では y はアウトカム(被説明変数)として、 x はアウトカムを説明する変数(説明変数)として扱います。この例ではアウトカム(y)である血漿量を、説明変数として選択された体重(x)がどの程度説明できるか、つまり血漿量に体重がどれほど影響を及ぼしているかを表しています。一方で、「残差」は、被説明変数の予測値と観測値の差として計算されるものであり、回帰モデルをあてはめた際に y を x で説明できない量を表しています。上の図では、各データ値から回帰直線までの長さが残差になります。

ここでの回帰モデルは、線形回帰モデルと呼ばれたり、説明変数が 1 個なので単回帰モデル、線形単回帰モデルと呼ばれたりもします。また、説明変数が 1 個の場合の回帰モデルを用いた解析は、単純に単変量解析、単回帰解析とも呼ばれます。一方、血漿量は体重だけでなく年齢にも依存するような場合は、説明変数に体重と年齢を入れて、血漿量を説明することになります。説明変数が複数個ある場合の回帰モデルは、単純に線形回帰モデルと呼ばれることもあれば、説明変数が複数個であることを明示すべく重回帰モデル、多変量回帰モデルとも呼ばれます。英語の表記としては、単変量(univariate)、多変量(multivariate)という表記が慣例的には多く用いられてきましたが、単変数(univariable)、多変数(multivariable)という言い方が正しいと多くの海外誌で主張されています。

話をアスピリンの例に戻します。背景を無視して、アスピリンありの群とアスピリンなしの群で死亡イベントの発生割合を(ログランク検定と呼ばれる検定方法を用いて)他の因子を考慮せずに比較すると、アスピリンありの群

の方で死亡リスクが高いかのように見える交絡が起こります。一方、死亡と関連のある因子である年齢をアスピリンと同時に説明変数として回帰モデル(例えば、比例ハザードモデルと呼ばれるモデル)をあてはめると、死亡イベントのデータのうち、どのくらいが年齢で説明できるかを数学的に計算し、年齢で説明できる部分を差し引いたうえで、アスピリンありとなしの群で死亡割合を比較することが可能になります。年齢で説明できる部分を差し引くということは、年齢による背景のズレをなくした(調整した)うえでアスピリンの効果を調べることが可能となることを意味します。

多変量解析の種類

無作為化の行われていない研究では、多変量解析などを用いて比較群間の背景を揃えることが大変重要となりますが、多変量解析で用いる回帰モデルは、①アウトカムの型、②1人の研究対象者からデータが繰り返し採取されているか、等によって異なります。例えば、上の血漿量と体重の関連を調べた回帰モデルは、アウトカムである血漿量は連続変数であり、研究対象者から1時点のみで採取されたデータを用いるので、線形回帰モデルが使用できます。アウトカムが死亡・生存のような2値のデータで、1時点のみで収集されている場合は、ロジスティック回帰モデルを使用するのが一般的です。また死亡・生存のような2値のアウトカムでも、死亡した時間が分かっている場合は、比例ハザード回帰モデルなどを用いることができます。アウトカムが血漿量のように連続変数で表されている場合も、異なる数時点で繰り返し測定されている場合には、線形回帰モデルを用いることはできず、その場合は線形混合効果モデルなど、測定の繰り返しを考慮できるモデルを用いる必要があります。アウトカムに繰り返し測定された2値変数が利用される場合は、一般化推定方程式を用いた回帰モデルなどを利用する必要があります。

このように、回帰モデルは、以下の二つの質問に回答することである程度まで選定できます。

質問1. アウトカムの型(変数の種類)は何か？

質問2. アウトカムの測定は1人あたり1回行われているか、又はデータがくり返し採取されているか？

以下に対応表をまとめてみます。

アウトカムの型	アウトカムの測定は 1人あたり1回行われている	アウトカムの測定が 1人あたり繰り返し行われている
連続変数	線形回帰モデル	線形混合効果モデル 一般化推定方程式を用いた回帰モデル (GEE)
2値変数(打ち切りなし)*	ロジスティック回帰モデル	一般化推定方程式を用いた回帰モデル (GEE)
2値変数(打ち切りあり)	比例ハザード回帰モデル	クラスターを考慮した比例ハザード回帰 モデル

*打ち切りとは、死亡などのアウトカムが観察される前に、追跡終了してしまうことを指します。例えば、入院から30日以内に死亡したかどうかをアウトカムとすると、30日より前に退院するなどして、30日後に死亡しているかどうかを確認できない場合を指しています。詳しくは生存時間解析の単元にて説明します。

多変量解析で検討できる説明変数の数

無作為化の行われていない研究では背景のズレをいかに調整できているかが、解析結果の適切性を担保するうえで重要となります。それ故、回帰モデルを用いて全ての背景因子を説明変数に入れて検討したいところです。しかし、回帰モデルで検討できる説明変数の数は、どのような回帰モデルを用いるか、標本サイズ(症例数)又はイベント数などによっても変わります。症例数又はイベント数が多いほど回帰モデルで検討できる説明変数の数は多くなります。用いる回帰モデル、症例数又はイベント数、モデルに入れることのできる説明変数の数の関係について、おおよその目安を以下に示します。

回帰モデルで検討できる説明変数の数

回帰モデル	検討できる説明変数の数
線形回帰モデル	全症例数を 15 で割った数
ロジスティック回帰モデル	イベントありとなしの人数のうち小さい方を 10 で割った数
比例ハザード回帰モデル	イベントありの人数を 10 で割った数

説明変数の選び方

回帰モデルが検討できる説明変数の数が分かったら、次にどの説明変数を入れるかについて考えましょう。多くの研究者の間では、単変量解析で比較群間に統計的に有意な差が出た背景情報だけを、最終的な多変量回帰モデルに説明変数として投入するという方法が用いられているようですが、このような方法は、必ずしも推奨できるものではありません。例えば、年齢の平均値をアスピリン使用群と非使用群で比べるなどした単変量解析を行い、有意差が確認された結果をみて年齢を多変量解析に説明変数としてモデルに加えるといったことをしてしまうと、今回解析に用いられたデータでしか再現のできない結果になってしまい、将来的に別の研究者が同様の研究を行ったときに同様の結果が得られない可能性が出てきます。すなわち、再現性の乏しい結果が得られてしまう可能性が高くなります。正しい選び方としては、データを一切見ることなく、医学的見地からアウトカムに影響があると思われる因子をその影響の大きなものから順に、回帰モデルに可能な限り入れるという方法が広く推奨されています。

多変量解析は正しく用いることにより、無作為化の行われていない研究において、背景のズレによって起こる交絡の影響を適切に調整できる点で大変優れた方法ですが、上記のような使い方をしてしまうと誤った結果につながり得ますので、十分注意して利用するようにしてください。

冒頭の例における心臓病のハイリスク患者さんにおいて、アスピリンの延命効果を調べた研究では、死亡又は生存が 2 変数として観測され、そのような関心の対象となるイベントの発生有無だけでなく、死亡した患者さんは死亡までの時間、死亡しなかった患者さんではその生存が最終的に確認されるまでの時間も観測されました。そのため、回帰モデルとして比例ハザード回帰モデルが用いられました。背景のズレを無視した解析では死亡に関するハザード比は 1.08 となり、アスピリン使用群のハザード(ハザードについては比例ハザード回帰の単元で詳しく説明しますが、今のところはリスクのようなものと解釈してください)は非使用群に比べて 8%高いという結果になっています。

	ハザード比	P 値
調整なし	1.08	0.5
年齢と性別を調整	0.75	0.02
年齢と性別と CAD を調整	0.57	<0.001
26 の背景因子を全て調整	0.67	0.002

比例ハザード回帰モデルに、アスピリンの投与又は非投与を表す群別因子だけでなく、背景情報である年齢と性別を加えると、(年齢と性別の影響を調整して得られる)ハザード比は 0.75(P=0.02)となり、アスピリンの使用によるハザードの統計的に有意な減少が確認されました。この研究では、無作為化は行われていないため、アスピリン使用群はアスピリン非使用群に比べ、平均年齢が 6 歳高く、男性の占める割合が高く、それらの不均衡が原因で死亡率も高くなってしまっていました。しかし回帰モデルに年齢と性別を説明変数として加えたことで、アスピリン使用群に不利な背景の不均衡の影響を調整でき、アスピリンの効果をより正しく解析できたと考えられます。さらに冠動脈疾患の既往歴はアスピリン使用群で 70%、アスピリン非使用群で 20%と、かなりアスピリン使用群に不利な背景情報のズレが起きていました。そこで冠動脈疾患既往歴に関しても年齢及び性別と同様にモデルに加えると、ハザード比は 0.57 となりました。このように年齢、性別、冠動脈疾患の既往歴を同時にモデルに加えて背景のズレを調整することで、アスピリンの使用の効果がさらに強く出ています。

この研究では、結局 26 個の背景因子とアスピリンの合計 27 個の説明変数が比例ハザード回帰モデルで検討されました。ここで説明変数の数が多すぎないか、上述した目安を参考に確認してみましょう。比例ハザード回帰モデルを用いているので、死亡者数の合計 276 人を 10 で割った数、すなわち、おおよそ 27 個の説明変数を当該モデルに投入できるということになります。したがって、本研究における比例ハザード回帰モデルに基づく多変量解析において、検討した説明変数の個数はこの目安にそっているという意味では理にかなっていると言えます。この帰結として、ここで検討した背景因子については、そのズレによって起こる交絡の問題に対処できているということになります。

また、多変量解析による背景因子の調整は、データが観測された因子を回帰モデルに投入することで行われます。逆にいえば、データが観測できなかった因子やモデルに投入しなかった因子によって交絡が起きている場合には、それらの影響を調整できません。一方、無作為化はデータとして取得されている背景因子も、データとして取得されていない背景因子も全て平均的に比較群間で均衡をとることができるというメリットがあることが知られているため、研究後に多変量解析を行いさえすれば無作為化は必要なくなるというわけではありません。しかし前述したように、無作為化が現実的に不可能な研究や、安全性や有効性が担保されていない治療など、最初から無作為化を行うことが倫理的に難しい状況もあります。このような場合には、まずは観察研究のデータを用いて、交絡の問題に対処できるだけ結果の妥当性を担保できるような統計解析を行って、より科学性の高いエビデンスを得ることが重要です。

本単元は、日本医療研究開発機構(AMED)が実施する研究公正高度化モデル開発支援事業(第一期)の「医系国際誌が規範とする研究の信頼性にかかる倫理教育プログラム」(略称:AMED 支援国際誌プロジェクト、信州大学・大阪市立大学)によって作成された教材です。作成および査読等に参加した専門家の方々の氏名は、[こちら](#)に掲載されています。

参考文献

- [1] Patricia A. Gum, Maran Thamarasan, Junko Watanabe, Eugene H. Blackstone, Michael S. Lauer. Aspirin Use and All-Cause Mortality Among Patients Being Evaluated for Known or Suspected Coronary Artery Disease: A Propensity Analysis. JAMA. 2001; 286(10):1187-1194.

無断転載禁止

この単位に関する国際誌におけるチェックポイントをいくつか紹介します。
(内容は解釈を助けるために一部意訳している部分もあります)

Annals of internal medicine

(<https://annals.org/aim/pages/AuthorInformationStatisticsOnly>)

- Information for Authors - General Statistical Guidance
- 2. Multivariable Analysis

■ 変数選択について

単変量解析によって統計的に有意に比較群間で異なる因子のみを多変量モデルで考慮する方法は妥当ではない。各因子が交絡因子かどうかは、比較群間での統計的な差に関係なく、それが関心のある因子の影響を変化させてしまうかどうか、あるいは他の共変量とともに考慮することで交絡を生じさせるかどうかによってのみ決定される。

■ モデルの構築について

モデルの中における各因子の影響の強さや、それに関する P 値が基準を超えるかどうかによって、各因子を多変量モデルへ含めるか否かを決定するステップワイズ法は利用すべきではない(変数増加法、変数減少法、変数増減法によらず)。P 値などを用いた変数選択方法よりも、臨床における重要性から決定される方が望ましい。ブートストラップ法のようなデータの復元抽出を繰り返しながら変数を自動選択する方法も、結局はステップワイズ法と同じであるため注意が必要である。

著者はどのように多変量モデルを評価したのか、あるいは交互作用などを検討した場合には何故その因子について交互作用を検討したのかについて、その方法と結果を示さなければいけない。

■ 多変量解析の結果報告

時折、一部の交絡因子についてのみ調整した多変量モデルの結果と、全ての交絡因子について調整したモデルの結果とを比較しているような文献が見られる。しかし多くの場合、そのような一部の因子のみを考慮した途中過程の多変量モデルの結果は読者にとって重要ではないことが多く、結果や考察において注目されるものではないため、提示する必要はない。