

はじめて学ぶ やさしい疫学

— 疫学への招待 —

日本疫学会 監修

南江堂

交絡因子

交絡因子とは

調査目標とする因子に、調査対象とする曝露要因以外の原因（または関連要因）が存在し、それが調査対象とする曝露要因と関連しているとき、これらの原因（または関連要因）を交絡因子と呼ぶ。

交絡因子が存在すると、調査対象とする曝露要因と調査目標とする因子とを単純に（交絡因子を無視して）関連づけても、両者の間に存在する真の関連は観察できない。交絡因子は結果に偏り（バイアス, bias）を与える原因の一つである。

たとえば、食塩摂取量と血圧の関連を観察したいとする。年齢が高いほど血圧は高い傾向にあり、年齢が高いほど食塩摂取量も多い傾向にある場合、年齢を無視して食塩摂取量と血圧の関連を観察しても、食塩摂取量と血圧の間に直接の関連があるのか、食塩摂取量が多い人は年齢が高く、年齢が高い人に血圧が高めの人が多いためのかわからない（図1）。

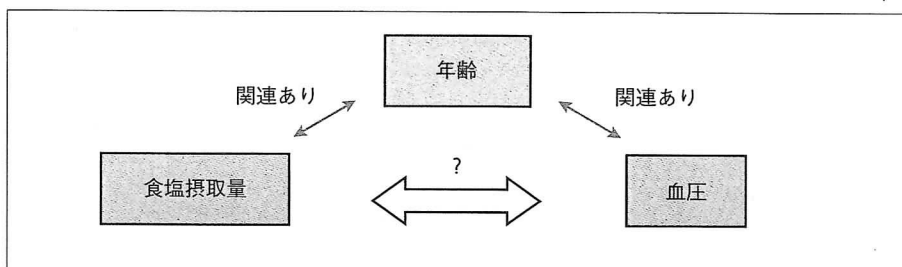


図1 年齢が交絡因子として働く可能性を示すための模式図（食塩摂取量と血圧との関連を検討する場合）

図2の左図は食塩摂取量と血圧との間に正の相関があることを示しているが、同じデータを右図のように年齢階級別(ここでは年齢を三つの階級に分けている)に見直してみると、どの年齢階級でも両者には関連が認められない。この場合、「食塩摂取量と血圧との間には関連がない」が正しい結論である。ここでは年齢が交絡因子となっている。

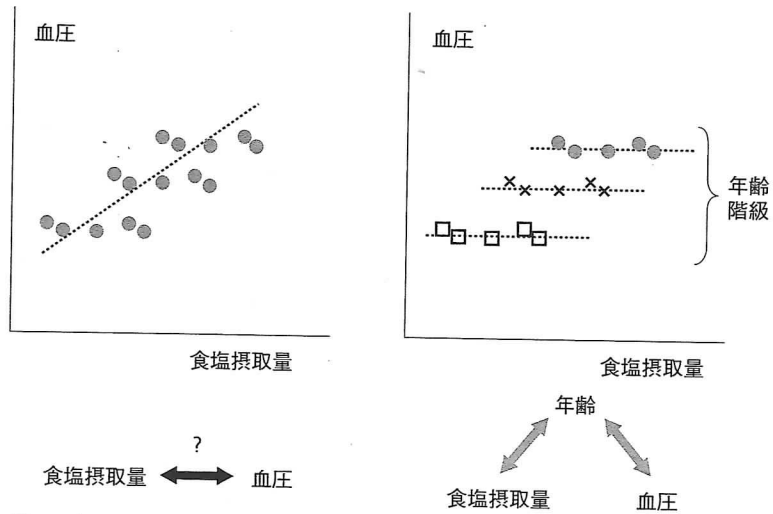


図2 年齢という交絡因子のために食塩摂取量と血圧との間には見かけ上の相関が観察される例
年齢階級ごとに解析すると、関連がないことがわかる。

補足1 他の独立因子が存在する場合

図3の左図は食塩摂取量と血圧との間に意味のある関連はないことを示しているが、同じデータを右図のように年齢階級別(ここでは年齢を三つの階級に分けている)に見直してみると、どの年齢階級でも食塩摂取量と血圧の間に関連が認められる。この場合、「食塩摂取量と血圧との間には関連がある」が正しい結論である。これは、食塩摂取量とは独立に、年齢が血圧と関連しているために、両者とも考慮して検討しなければ、正しい関連がみえてこないという例である。ここでは年齢は食塩摂取量と関連していないため、厳密にいえば交絡因子ではない。

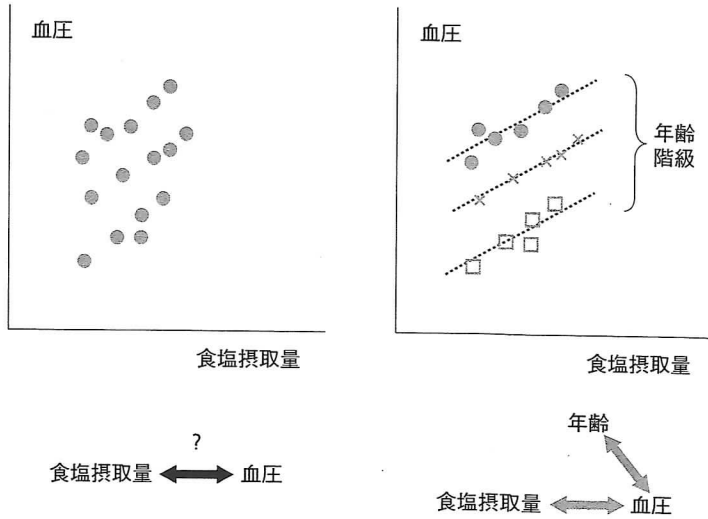
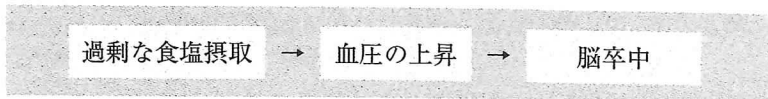


図3 年齢という独立因子のために食塩摂取量と血圧との間で相関が観察されない例
 年齢階級ごとに解析すると、関連があることがわかる。

補足2 原因と結果の間に中間因子が存在する場合

食塩摂取量と脳卒中罹患率との関連を検討したいとする。食塩摂取量と血圧に関連があり、血圧と脳卒中罹患率の間にも関連がある場合、食塩の過剰摂取が高血圧を引き起こし、高血圧が脳卒中を引き起こすと考えると、血圧は中間因子である。この場合、血圧の影響を除去してしまうと、食塩摂取量と脳卒中罹患率との間の関連は正しくは検討できない、この例における血圧は交絡因子ではない。

先の例では、食塩摂取量と年齢との間に関連がみられたとしても、食塩摂取量が年齢の原因とはなりえず(減塩したら若いままでいられる?!)、両者は原因でも結果でもなく、単に相関が観察されたにすぎない。交絡因子か中間因子かの違いは、一方が他方の原因となりうるか否かにある。しかし、実際にはその判別が困難な場合も少なくない。



交絡因子の調整方法

交絡因子を調整する、すなわち、その好まざる影響を可能な限り除去する方法は研究手法によって異なるが、大きく分けて、調査時に行う方法(その計画は調査前に立てるため、事前の処理とも呼ぶ)と、解析時に行う方法(調査後に行うため事後の処理とも呼ぶ)があり、可能な限り事前に行うことが望ましいと考えられている。

事前に行う方法には、① 限定 (restriction)、② 無作為化 (randomization)、③ マッチング (matching) などがある (表1)。横断研究、症例対照研究、コホート研究を含むあらゆる疫学研究で用いられるのが「限定」である。しかし、現実的には十分な数の対象者が得られないという問題をはらんでおり、その実行は困難な場合が多い。性、年齢、居住地域など、限定したい因子が多い場合はさらに困難となる。

表1 交絡因子(例、年齢)を混入させないための調査方法

調整方法	主に用いる研究手法	具体例
限定	あらゆる疫学研究	たとえば、ある年齢階級に限って行う
無作為化	無作為化比較試験	年齢が偏らないように無作為に集団を2群に分ける
マッチング	症例対照研究	それぞれの症例に対して年齢に近い(±2歳未満の)対照を選ぶ

事後の方法(解析方法)には、① 層別解析 (stratified analysis)、② 多変量解析 (multivariate analysis) などがある (表2)。両者は必ずしも区別できるものでなく、②は部分的に①を含んでいる。実際にはたくさんの層に分けて解析を行い、それを数学的に統合して結果を導くため、各層に割り当てられる対象者数が極端に少ない(たとえば、5人未満)と信頼度の高い結果は得られない。たとえば、年齢の影響を調整したうえで要因Xと結果Y(多くの場合、疾病)との関連を検討したいと考えても、年齢の分布が大きく偏っていて、ある年齢階級に属する人数が極端に少ないような場合には、その適用は困難となる。したがって、事後の方法によって交絡因子の影響を調整したい場合には、調整したい因子の分布を調査前にある程度把握しておくことが必要である。さらに、解析に投入されるすべての因子の分布や仮定している関連の仕方など、実際には多数の前提条件が存在するため、調整したい因子を十分に調整することは必ずしも容易ではない。

表2 交絡因子を調整するための解析方法

調整方法	解析手法の例	具体例
	層別解析	層(たとえば年齢階級)に分けて、層ごとに解析を行う
多変量解析*	Mantel-Haenszel 法	2×2分割表によって二つの要因の間の関連(相対危険)を検討する場合に、調整したい因子で層別解析を行い、結果を数学的に統合する方法 すべての層で相対危険は同じであることを前提としている
	重回帰分析	結果に影響を与えていると考えている複数の因子で多項式を作り、多項式によって推定される結果が観察された結果をもっとうまく説明するように、多項式内の係数を決定する方法 結果となる因子がカテゴリーである場合[多くの場合は2値(生存か死亡か、好きか嫌いかなど)である]はロジスティック回帰分析となる

*詳しくは15章を参照されたい。