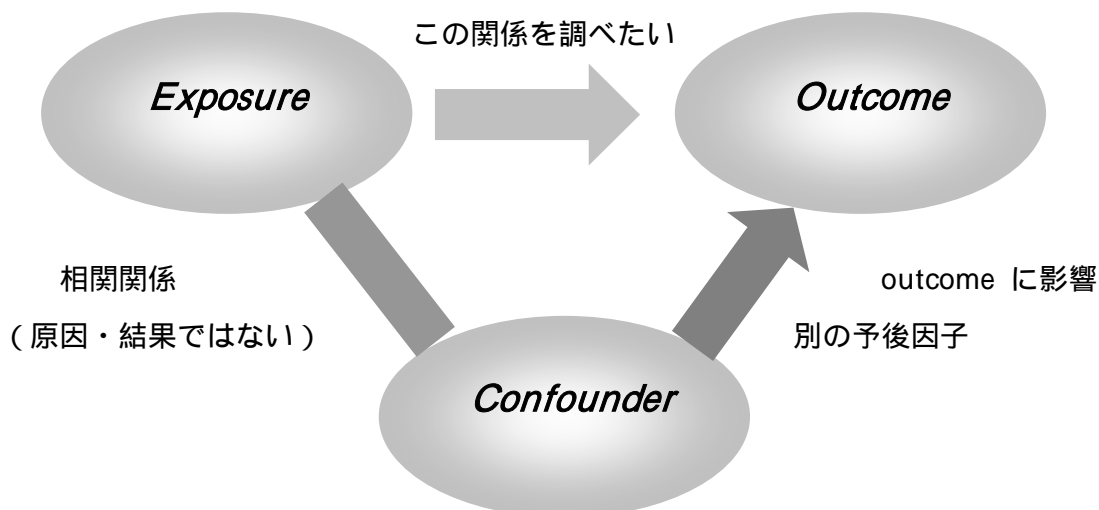


Confounding

Confounder とは？

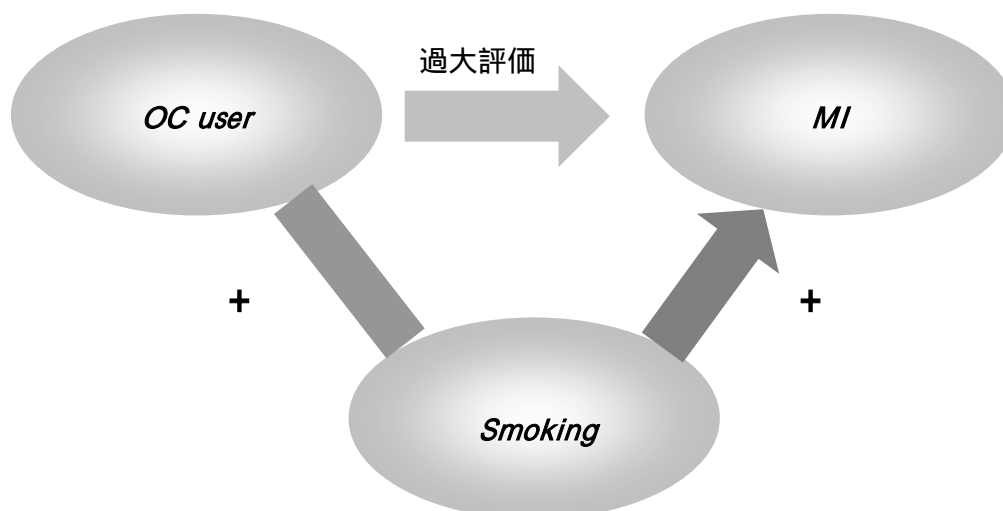
Confounder とは日本語で交絡因子と訳されています。Outcome に対する exposure とは独立した危険因子であり、なおかつ exposure とも相関関係をもつ場合を confounder と呼びます。



例えば「ピザを沢山食べると冠動脈疾患(CHD)になり易いかどうか」調べようと思います。ビールを沢山飲むと CHD になりやすいことは既にわかっているとします(実際そのような報告はないと思いますが)。そしてピザを良く食べる人は、ビールをよく飲む傾向にあると仮定します。本当はピザを食べても CHD 発生には全く影響がなくても、ピザを良く食べる人はビールも良く飲むので CHD の合併が多くなります。そこで誤って「ピザを沢山食べると CHD になりやすい」という結論を導いてしまうかもしれません。この場合はビールは confounder に相当します。

Confounder の方向性と程度

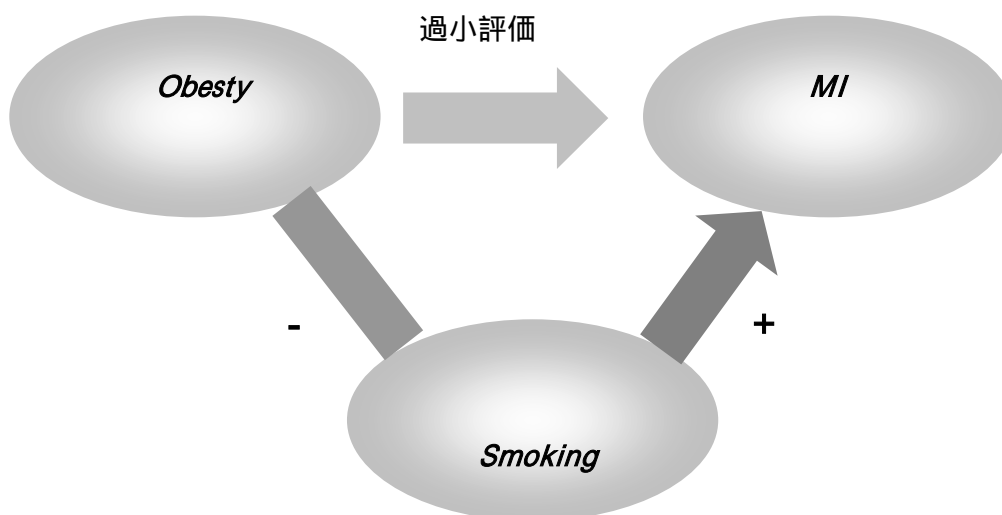
もう1つ例を挙げましょう。経口避妊薬（OC）を使用していると心筋梗塞（MI）が増えることがわかっています(OR = 2.06 (1.27-3.33))。しかしながら OC 使用者では喫煙率も高く、喫煙は MI の危険因子でもあります。喫煙は exposure{OC 使用}と outcome (MI 発症) 両方に関与していますので、confounder と言えます。



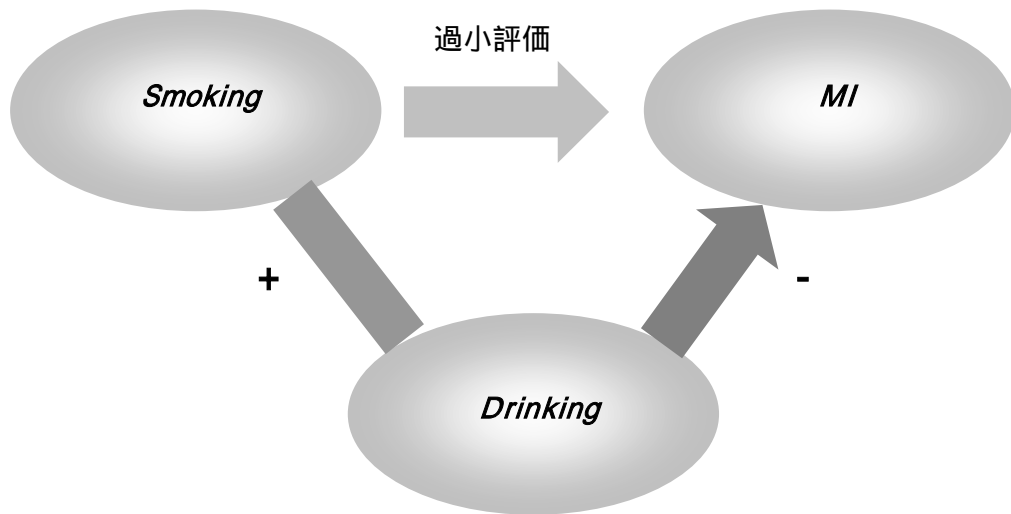
さらに OC 使用者が増えれば喫煙者も増えるので両者の間には陽性の関係があります。また喫煙は MI 発症を助長しますので、ここにも陽性の関係が存在します。このような場合、confounder を調整(adjusting)しないで crude data だけで判断すると過大評価をしてしまうことになります。OC 使用は MI 発症率を 2 倍に押し上げるとしますと、confounder を調節後は例えば 1.5 まで低下するといった具合です。一方、喫煙が MI を 3 倍に押し上げるとしますと、OC 使用は MI 発症に何ら関係がないということもありえます。しかし、喫煙によって MI が 1.5 倍にしか上昇しないとすると、OC 使用による MI の 2 倍は全て説明がつかず(喫煙以外 confounder が無いと仮定して)、よって OC 使用は少なくともある程度 MI 発症に寄与しているといえます。OC 使用者間での喫煙率が 20%、非使用者間の喫煙率が 7%であったとすると、喫煙率は 2.86 倍となります。もし OC 使用者で MI 発生が 2 倍にしかならないのであれば、むしろ OC 使用と喫煙の方が強く関係しており、OC 使用による MI 発生増加は全て喫煙で説明がつく可能性もあります。一方、OC 使用により MI 発症が 5 倍にまで押し上げられるとすれば、confounder の影響を差し引いても OC 使用は MI を引き起こすことでしょう。それで

は、OC 使用が MI を 5 倍にする、喫煙は MI を 2 倍にする、OC 使用者の喫煙率は非 OC 使用者に対して 2.86 倍だとします。Confounder の寄与率はせいぜい 2 倍（2.86 倍と 2 倍の小さい方）です。よって（他に confounder が存在しないと仮定して）OC 使用が MI 増加に寄与している可能性は十分あると考えられます。

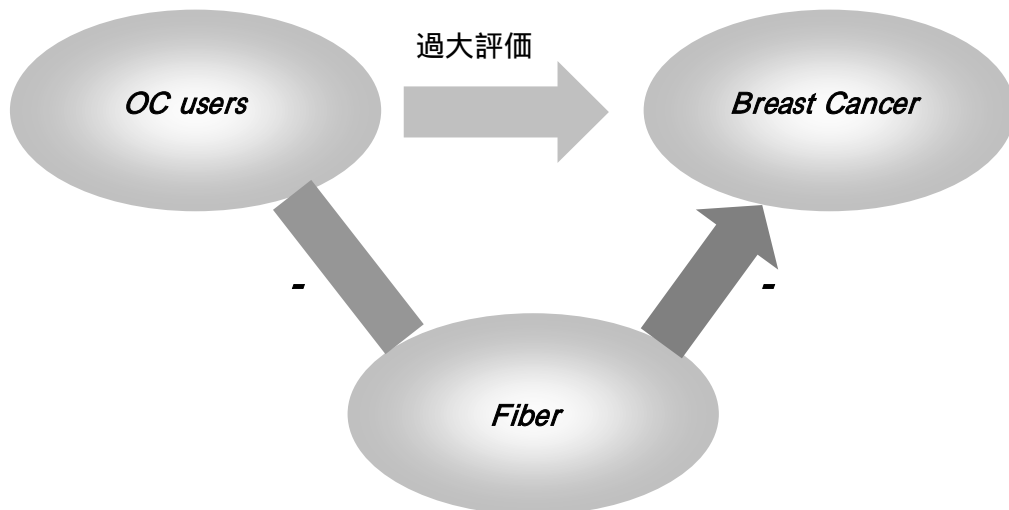
OC 使用でなく肥満だったらどうでしょう。そしてやせている人の方がより喫煙率が高いとします。肥満が多くなればなるほど、喫煙率は下がるので、今度は陰性の関係となります。喫煙と MI の関係は同じ陽性です。両者はマイナスとプラスを掛け合わせてマイナスとなるため confounder を調整しない状態で結果を評価すると過小評価してしまいます。つまり本当は 2 なのに 1.5 倍のリスクと言った具合です。



喫煙が exposure でアルコールが他の危険因子であるとしましょう。喫煙はアルコールを好む傾向があるとすると陽性の関係です。一方適度のアルコールは冠動脈疾患をかえって抑制するとします。よってアルコールは exposure/outcome の両者に関連していますから confounder です。その方向性はプラスとマイナスを掛け合わせてマイナスに働くので crude data は adjusted data より低い値となります。



OC 使用と乳癌の関連について調べています。OC 使用者は子供が少ない傾向にあり、子供の数が増えると乳癌の発症は減ることがわかっています。やはり exposure に associate するもう1つの抑制ファクターであるため confounder です。マイナスとマイナスが働くとプラスとなり、confounder を adjust しないと crude data を overestimate してしまいます。



Cohort study において飲酒と冠動脈疾患発生について調査しています。可能性のある confounder として食物繊維を選定しました。飲酒は冠動脈発生を抑制する傾向にありましたが、そして食物繊維の摂取は飲酒と逆相関し(negative association)、冠動脈発生に対しても抑制的に働くことがわかりました。Negative x negative で結局食物繊維に関

してadjustしないとアルコール摂取が増えると食物繊維摂取減少を介して冠動脈疾患発生が増加してしまいます。つまりcrude dataはadjusted dataより大きな値をとることになります。具体的には $RR_{\text{crude}} = 0.7$ だとすると $RR_{\text{adjusted}}=0.5$ といったデータになるかもしれません。

以上みてきたようにexposureとconfounder, confounderとoutcomeの関係によりexposureとoutcomeの関係は過大評価overestimateあるいは過小評価underestimateされます。またその程度(magnitude)はconfounderのexposureとoutcomeとの関係によって規定されます。すなわちexposureとconfounderの間に強い関係が存在してもconfounderがoutcomeに極軽度しか影響しなければ大したことはありません。逆も真なりです。両者が強く関係していればconfounderの影響は当然大きくなります。またconfounderの存在は95%CIやp-valueをも狂わせ、おかしい結果解釈につながります。狭義ではconfounderとbiasは区別されますが、広義ではconfounderもbiasに含まれます。

Confounderを除去するためには

Confounderは危険因子であり生物学的特徴として内在するものなので関係自体を変えることはできません。しかしながら暴露とconfounderの関係を調整することはできます。

- (1) Matching
- (2) Restriction
- (3) Randomization

Confounderの概念だけを解説しました。実際の調整法については後述します。