

過去の去来夢の記事で交絡の問題やその調節について言及されています。観察研究を適切に行うためにこれらの概念は非常に重要です。また、観察研究はこれらの問題の影響を少なからず受けることから、ランダム化比較試験 (randomized controlled trial; RCT) がゴールドスタンダードとも言われます。ですが、そもそもなぜ RCT が大事なのかということを、二回にわたってややフォーマルにとりあつかってみたいと思います。

ここでは、PECO の形で定型化できるリサーチクエスチョンを考えます。このようなリサーチクエスチョンでは、特定の患者集団(P)における、なんらかの要因(E)のアウトカム(O)に与える効果を、適切な比較(C)によって明らかにしようとしています。これは究極的には、なんらかの因果効果(causal effect)を求めようとする「因果推論(causal inference)」の研究デザインと考えられます。なぜこのような研究が重要なのでしょうか。先に結論を述べると、もし要因とアウトカムの間に因果効果が存在するのならば、要因(E)を減らしたり、あるいは、増やしたりすることにより、特定の患者集団(P)におけるアウトカム(O)をより望ましい方向に変化させることが可能と考えられるからです。

「因果関係(causal relationship)がある」というのはどういうことでしょうか。たとえば、頭痛があるあなた (P) が、ある痛み止めを内服(E)したとします。そして、一時間後には痛み(O)がなくなりました。これは痛み止めの内服と頭痛の改善には、因果関係があると言えるでしょうか。商品のコマーシャルなどではよく見掛けるロジックですが、臨床研究の道標[1]では比較群(C)のない不適切な「雨乞い三「た」論法」として紹介されています(図 1; 頭についているものは頭痛を示すと考えてください)。「適切な比較」がないことが問題というわけですが、そもそも「適切な比較」とはなんなのでしょうか。

図 1. 雨乞い三「た」論法

ベースラインの状態

1時間後の未来の状態



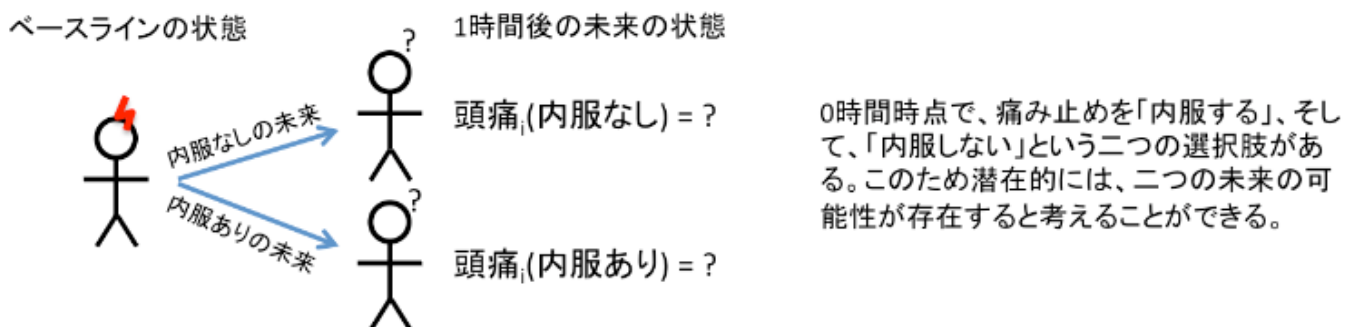
0時間時点で、痛み止めを内服した。
1時間時点で、頭痛はなかった。
痛み止めは効果があった?

まず、今回は「適切な比較」を定義するために、個人レベルの因果効果という概念を「潜在的アウトカム」というフレームワークを用いて定義します。また、この定義方法の結果として、個人レベルの因果効果が測定不可能であることを確認します。次回は、これに続けて集団レベルの因果効果という概念から、なぜ RCT が因果推論のゴールドスタンダードであるのかを確認します。

まずは、因果効果というものを定義する必要があります。ここでは「頭痛_i」を i 番目の個人(ここではあなた)の 1 時間時点での頭痛の有無のアウトカム(0 がなし、1 があり)とします。前述の論法では「0 時間時点で頭痛あり」が「1 時間後時点で頭痛なし」に変化したことをもって、痛み止めの効果を論じようとしています。ある意味では比較(前後比較)を行っているわけですが、因果推論の考え方ではこれを適切な比較とは考えません。適切な比較は、「potential outcome」(定訳はないようですが、以降は「潜在的アウトカム」)[2]という考えかたを用いて定義されます。この考えかたは、Neyman や Fisher らが実験環境(RCT)において先駆し、Rubin(傾向スコア、多重代入法などで知られる)が実験や観察研究を問わず有効な考えかたであることを示しました(数学的表記については解説を参照)。また、ほぼ同じ意味で counterfactuals (反事実)[3]という言葉が使われることもあります。

前述のシナリオでは、あなたには、0 時間時点で、痛み止めを「内服をする」と「内服をしない」いう二つの治療選択肢があります。このため、潜在的には、二つの治療選択肢に対して、二つのそれぞれ対応した未来があると考えることができます。この二つの潜在的な未来を、「アウトカム(治療選択肢)」という括弧書きでどの選択肢に対応するものかを明確にします。

図 2. 潜在的アウトカムの定義



頭痛_i(内服なし) : 0 時間時点で痛み止めを「内服しない」場合の 1 時間後の未来の頭痛の有無

頭痛_i(内服あり) : 0 時間時点で痛み止めを「内服する」場合の 1 時間後の未来の頭痛の有無

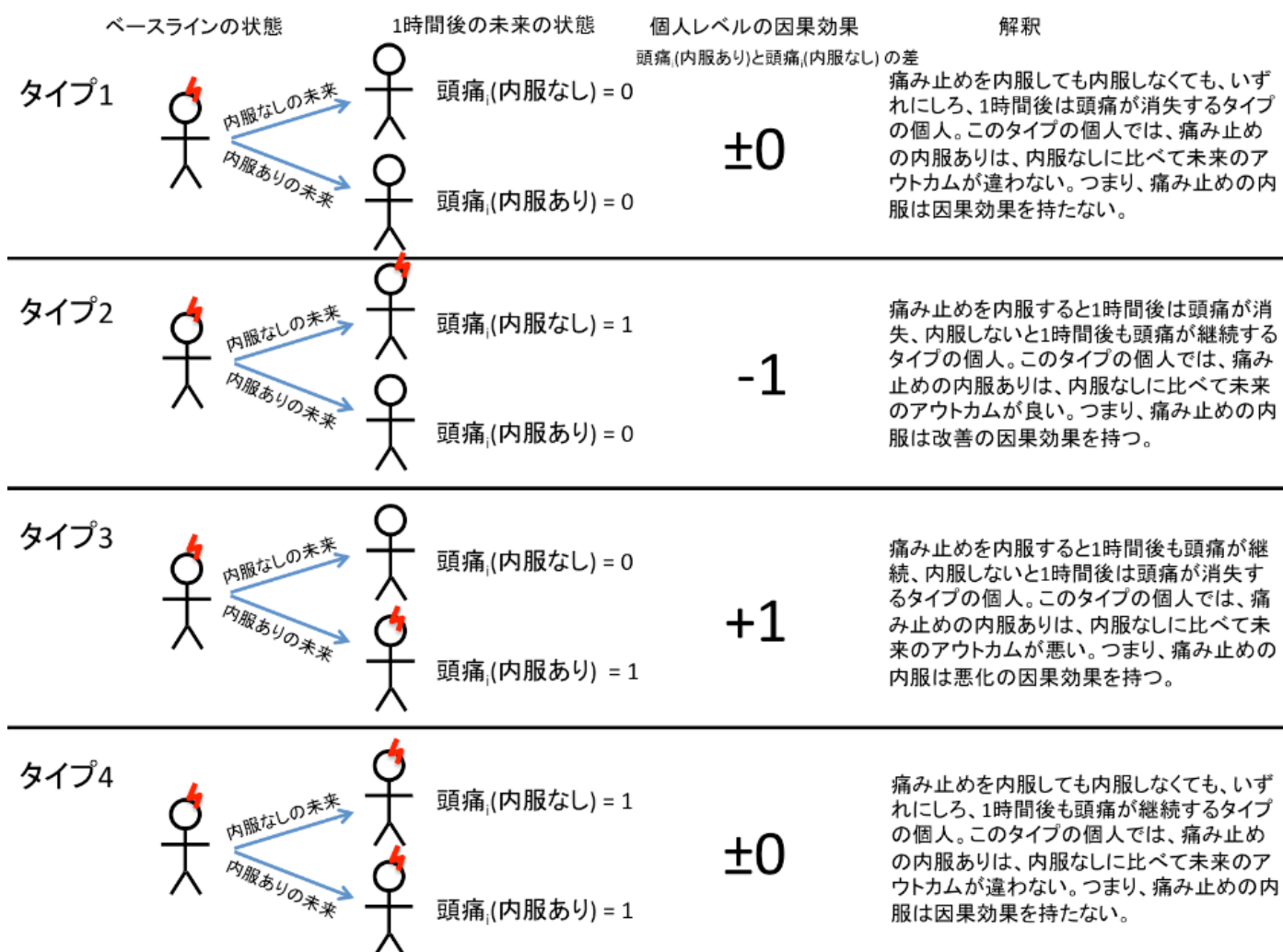
つまり、実際に選択される治療選択肢も、実際に選択されない治療選択肢も含めて、すべての治療選択肢に対して、アウトカムが潜在的に存在すると想定するわけです。臨床的な言葉におきかえると、「0 時間時点で痛み止めを「内服しない」場合の 1 時間後の未来の頭痛の予後」(放置する場合の予後)と、「0 時間時点で痛み止めを「内服する」場合の 1 時間後の未来の頭痛の予後」(介入する場合の予後)と言えます。こう言いかえると、臨床医が意思決定をするにあたって、普段から意識している概念であることがわかります。

では、個人レベルの因果効果を定義しましょう。潜在的アウトカムのフレームワークでは、0 時間時点での痛み止めの「内服あり」と「内服なし」を比較した、1 時間後の頭痛に対する因果効果を、対応する二つの潜在的アウトカムの比較(ここでは差)として定義します。

個人レベルの因果効果(差): 頭痛_i(内服あり) - 頭痛_i(内服なし)

つまり、各個人について「0 時間時点で痛み止めを「内服する」場合の 1 時間後の未来の頭痛の有無」と「0 時間時点で痛み止めを「内服しない」場合の 1 時間後の未来の頭痛の有無」の差を見ています。この例ではある個人における「介入する場合の予後」と「放置する場合の予後」の差と言えるでしょう。二つの未来(二種類の潜在的アウトカム)のそれぞれが、0(頭痛なし)か 1(頭痛あり)かの値をとりますので、図のように 4 種類のタイプの個人がいることがわかります。

図 3. 潜在的アウトカムを用いた因果効果の定義



さて、個人レベルでの因果効果というものが定義されたわけですが、これは残念ながら測定不可能です。潜在的アウトカムという言葉が示唆する通り、現実には起ったアウトカムとは限らないからです。実際に選択された治療選択肢に対応する潜在的アウトカム、例えば、実際には痛み止めを内服したあなたの、「痛み止めを内服した場合の 1 時間時点での頭痛の有無」は、実際に起ったアウトカムであり測定可能です(厳密には SUTVA の用語解説参照)。前述の図を使えば、図 1 で実際に観測された 1 時間後のアウトカム(内服ありの未来)は、図 2 の潜在的アウトカムのうちの下のもの(内服ありの未来)に一致します。

しかし、実際に選択されなかった治療選択肢に対応する潜在的アウトカム、例えば、実際には痛み止めを内服したあなたの、「痛み止めを内服しなかった場合の一時間時点での頭痛の有無」は、概念として想定することは問題ないですが、この現実世界に実際に測定できるかたちでは存在しません。前述の図を使えば、図 1 では「内服なしの未来」は実際に内服を行ったため観測されておらず、よって図 2 の潜在的アウトカムのうち上のもの(内服なしの未来)については全くデータがありません。

このもう一方の潜在的アウトカムを測定するためには、全く同じあなたという個人が、全く同じ場所、全く同じ時間、全く同じ頭痛、および、全く同じ全身状態で痛み止めを「内服しなかった」状態を観察する必要があります。これは、あなたの内服状況が違う以外はすべてが同じパラレルワールドに入りこむ、あるいは、タイムマシンで全く同じ時間にもどって痛み止めを内服しないという選択肢をとらせる、といった、やりたくても現実にはできない SF めいた研究デザインが必要です。これを因果推論の根本問題と言います。臨床研究の文脈で解釈すると、「一例報告では行われた治療の因果効果を測定できない」ということです。臨床的には、この一症例で「実際行われた治療を行わなかった場合の予後」、つまり、「治療なしの場合の潜在的アウトカム」を臨床経験からおぎなうことで、インフォーマルに解釈をします。しかし、これは検証不可能な仮定(「治療なしの場合の潜在的アウトカム」の best guess)を導入することで解釈(測定ではない)をしているわけです。

では、因果推論というものはまったくの机上の空論で役に立たないのでしょうか。実は、この根本問題に対する解決法は、興味の対象を個人レベルの因果効果から、集団レベルの因果効果に移すことで解決されます。次回はこの集団レベルの因果効果と RCT の関係を考えてみます。

引用文献

1. 福原俊一. 臨床研究の道標(みちしるべ)―7つのステップで学ぶ研究デザイン. 東京: 特定非営利活動法人 健康医療評価研究機構; 2013.
2. Imbens GW, Rubin DB. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. 1 edition. New York: Cambridge University Press; 2015.
3. Hernan MA, Robins JM. Causal Inference [Internet]. Chapman & Hall/CRC; 2016. Available: <http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>

潜在的アウトカムの概念は、古典的な統計学の枠組みの外側にある概念です。このため数学的表記も古典的な統計学で使われるものを拡張して用いる必要があります。因果推論には学派がいろいろあり、ほぼ同じことを意味するのにいろいろな表記方法があります。ここでは、おそらく一番読みやすい Rubin ら [2] の括弧書きの潜在的アウトカム記法を用います。

まずは、古典的な統計学の枠組みの変数を定義します。

Z_i : 0 時間時点で実際に選択された治療選択肢。なしなら 0、ありなら 1。

Y_i : 実際に観測された 1 時間後の頭痛の有無。なしなら 0、ありなら 1。

次に、潜在的アウトカムの括弧書きの記法を定義します。

$Y_i(z)$: 治療選択肢が仮に z であった場合に起る 1 時間後の未来の頭痛の有無。なしなら 0、ありなら 1。

ここでは実際に選択された治療選択肢と関係なく、治療選択肢の変数が仮に z という値であった場合に起こる未来のはなしをしているわけです。治療の選択肢はここでは、痛み止めの「内服なし」と「内服あり」の二種類がありますので、各個人に対して二種類の潜在的アウトカムが定義されます。

$Y_i(0)$: 治療選択肢が仮に「内服なし」であった場合に起る 1 時間後の未来の頭痛の有無。

$Y_i(1)$: 治療選択肢が仮に「内服あり」であった場合に起る 1 時間後の未来の頭痛の有無。

これらを用いて、個人レベルでの因果効果を、治療選択肢が仮に「内服あり」であった場合に起る 1 時間後の未来の頭痛の有無と、治療選択肢が仮に「内服なし」であった場合に起る 1 時間後の未来の頭痛の有無、その二つの潜在的アウトカムの差として定義します。

個人レベルの因果効果: $Y_i(1) - Y_i(0)$

それぞれの項は 0 か 1 ですので、-1、0、+1 の値を取ります。それぞれの解釈は図 2 を参考にしてください。

本文中で実際に選択された治療選択肢に対応する潜在的アウトカムは測定可能と仮定しました。つまり、下記が成り立つと仮定しました。

$$Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0)$$

左側の Y_i はこの個人 i において 実際に測定された一時間時点での頭痛の有無です。式の右側はこれを二つの潜在的アウトカムに結びつけています。痛み止めの有無(A_i)は 0, 1 で定義されているので、右側は $Z_i = 1$ であれば $Y_i(1)$ 、 $Z_i = 0$ であれば $Y_i(0)$ になります。これを stable unit treatment value assumption (SUTVA)[2]といいます。一見、あたりまえに思えますが、これが明確に定義されるためには二つの要件を満たす必要があります。

ひとつは、治療の選択肢が明確に定義されている必要があります。今回のシナリオでは、厳密には他の鎮痛薬を内服せずにアセトアミノフェン 500mg を内服する($Z_i = 1$)、アセトアミノフェンも含めて一切の鎮痛薬を内服しない($Z_i = 0$)というように、それぞれの選択肢がプロトコールとして定義されている必要があります。「なにか痛み止めを内服する」というような表現では、選択肢が一意に決まっておらず(どの薬剤か? 内服量は? など)、このため、対応する潜在的アウトカムも一意に決まりません。ただ、臨床的な意義が考えにくい些細な治療のバリエーション(アセトアミノフェン 500mg を、西を向いて内服するか、東を向いて内服するかなど)は無視して構いません[3]。

もう一つは、non-interference[2]という概念になります。これは他の個人が受けた治療が、あなたの潜在的アウトカムに影響を与えないという仮定です。これの反例としては感染症疫学領域の herd immunity がわかりやすいでしょう。つまり、あなた個人がワクチンのある感染症に罹患するかどうかの潜在的アウトカムは、あなた個人のワクチン接種状況という個人レベルの暴露要因だけでなく、コミュニティ内でどれくらいワクチン接種が行われているかという他人の暴露要因にも依存しています。感染症関連の研究や教育介入の研究などでは特に重要になり、研究デザインに工夫が必要になります。