

# 研究員 の眼

## メタアナリシスの弱点

研究データをたくさん集めれば、正しい分析ができるか？

保険研究部 主席研究員 篠原 拓也

(03)3512-1823 tshino@nli-research.co.jp

社会科学でも自然科学でも、さまざまな分野で、多くの研究者が日々研究を重ねている。注目度の高い分野では、類似した内容の研究を、複数の研究者が手がけていることもまれではない。それぞれの研究者が行った実験や調査から得られた研究データが、論文などの形で公表されている。

そうすると、自らが直接、実験や調査を行うかわりに、こうした公表された研究データを収集して、それらを分析することで、有益な研究につながる場合が出てくる。このように、ある特定のテーマに関する研究報告を網羅的に集めて、それらを評価したり比較したりしながら、分析結果を統合して、そのテーマに関する知見を導き出す研究方法は、「システマティックレビュー」といわれる。

### ◆ システマティックレビューでは複数の文献の差異を分析する

システマティックレビューでは、収集した研究データの量と質が問われる。公衆衛生や医療の分野でみてみよう。量については、医学関連の電子データベースがいくつも立ち上げられている。そして、さまざまな疫学研究に関する文献や論文などが、収集されて、網羅的な検索ができるようになっている。こうした文献や論文では、得られたデータをもとに何らかの結論を導き出すことが一般的に行われている。

システマティックレビューを行う上では、たとえ類似した研究テーマであっても、得られた結論が必ずしも同じになるとは限らない点がポイントといえる。収集した複数の文献や論文のどの部分が一致していて、どこに差異があるのか？ なぜ、そうした差異が生じているのか？ —— といったことを検討するのが、システマティックレビューの真骨頂といえる。

### ◆ 複数の文献のデータを統合して解析する「メタアナリシス」も行われている

ただし、それぞれの研究結果が異なり、結論が大きく食い違う場合もある。こうした場合、各文献の結論を定性的に評価するだけでは、分析結果の統合ができないことがある。

また、複数の文献のうち、どの研究成果が他のものに比べて有用なのか、を捉えにくい場合もある。

そこで、それぞれの文献で行われている同一のテーマに関するデータを、統計学的に統合して解析していく場合がある。これは、「メタアナリシス」と呼ばれる。

肺がんと喫煙の関係をテーマに、メタアナリシスをみてみよう。(以下は、本稿のために筆者が作成した架空の事例)

肺がんと喫煙の関係についての研究 (架空の事例)

研究	調査時期	調査対象 (いずれも成人男性)	研究方法	結果(リスク指標(オッズ比)の区間推定)	0 1 2 3			
					0	1	2	3
A	1994-96年	V市在住の非喫煙者1,000人、喫煙1,000人	前向き	1.5 (1.4-1.6)				
B	2001-05年	W病院の患者200人と、健常者200人	後ろ向き	1.6 (1.2-1.9)				
C	2009-11年	X県在住の非喫煙者300人、喫煙者200人	前向き	1.4 (1.0-1.8)				
D	2010-12年	Y病院の患者30人と、健常者30人	後ろ向き	2.0 (0.8-3.0)				
E	2019年	Z病院の患者50人と、健常者50人	後ろ向き	0.2 (0.1-0.5)				

※ 区間推定は95%区間。なお、前向き研究は、非喫煙者と喫煙者を取りだして、以降の肺がん罹患状況をみていくもの。後ろ向き研究は、肺がん患者と健康な人を取りだして、それぞれの過去の喫煙状況をみていくもの。一般に、リスク指標として、前向き研究ではリスク比(=喫煙者の発症割合÷非喫煙者の発症割合)、後ろ向き研究ではオッズ比(=喫煙者の発症者/喫煙の未発症者)÷(非喫煙者の発症者/非喫煙の未発症者)が用いられる。この表では、各研究の比較のために、前向き研究もオッズ比で表示している。(通常、オッズ比を示す際は、例えば0.5と2が、1から等距離に位置するようにするために対数目盛で表示するが、この表ではわかりやすさを優先して対数目盛を用いていない。)リスク指標の数値が高いほど、肺がんと喫煙の関係が強いことがうかがわれる。(前向き研究や後ろ向き研究についての詳細は、「治験の概要-臨床試験の現状(前編)」篠原拓也(基礎研レポート, ニッセイ基礎研究所, 2021年7月29日)をご参照いただきたい。)

5つの研究のうち、A、B、Cの3つは、平均的な結果となっている。研究Aは、区間推定の範囲は狭く、結果に対する信頼度は高い。ただ、少し気になるのは調査時期が1990年代とやや古い点だ。

研究Bは、5年間にわたる研究の結果で、区間推定の範囲は5つの研究の平均的なものとなっている。研究Cも、区間推定の範囲が平均的となっている。いずれも、調査対象の人数が数百人規模であり、このメタアナリシスで中央に位置する結果と見られる。

研究Dは、リスク指標が高く、肺がんと喫煙の関係性を肯定する結果を示している。ただし、その区間推定の範囲は広いので注意を要する。これは、この研究の調査対象の人数が少なく、結果の変動幅が大きいためとみられる。

研究Eは、調査時期が比較的最近となっている。リスク指標は低く、肺がんと喫煙の関係があまりみられないことを示唆している。研究Eは、研究の期間が1年と短いことが、結果に何らかの影響を及ぼしている可能性もある。

このように、複数の研究を相互に比較して、それぞれの研究の統合可能性を検討する。場合によっては、統合可能性を判定するために、統計的検定が行われることもある。そして、統合可能と判断された研究は、加重平均法などを用いて結果を統合して結論を導いていく。

### ◆ メタアナリシスの大きな弱点

実は、このメタアナリシスには、大きな弱点がある。読者の皆さんは、もうお気づきだろうか？

それは、いずれも、研究が完了して文献として公表されたものを分析している点だ。メタアナリシスの対象は、作業が完了して成果が公表された研究結果に限られる。つまり、「研究が完了した」という一種のフィルターを通過したものばかりになってしまうわけだ。

先ほどの肺がんと喫煙の関係をテーマとした研究に戻って考えてみよう。

表で比較されていた5つの研究の他に、たとえば研究Fとして、実施はされたものの、結果が未公表のものがあるかもしれない。未公表となっている原因は、「喫煙者よりも非喫煙者のほうが、肺がん罹患しやすい」などと、通常予想と異なる結果が示されたために、その理由について詳細な分析を進める必要があることなどが考えられる。こうした研究Fの結果は、詳細な分析が完了すれば、いずれ公表されるかもしれない。

また、研究Gとして、研究自体を打ち切ったケースがあるかもしれない。打ち切りの理由は、非喫煙者と申告していた調査対象の何人かが、実はタバコを吸っていたことが研究の途中で判明したためだ。こうした研究Gの(打ち切りまでに得られた途中の)結果は、今後も、たぶん公表されることはないだろう。

このように、現時点では結果が未公表だがいずれ公表されるかもしれない研究や、中止となつてしまい結果が公表されない研究は、メタアナリシスの対象にはなりえない。このため、研究が完了して

公表された研究結果を網羅的に収集しても、こうした未公表の研究は抜け落ちてしまう。つまり、メタアナリシスでは、手がけられたすべての研究を把握することはできないこととなる。

#### ◆ 報じられていない情報の中に重要なものがあるかも

これは、なにもメタアナリシスに限られた話ではない。たとえば、どんなメディアでも、伝えられた情報には、伝えようとする人の意思が込められている。どの情報を伝えて、どの情報を伝えないでおくかは、伝える人の考えによって決まってくる。

ニュースなどでは、もしかすると、報じられていない情報の中に、実は重要なものが眠っているかもしれない。視聴者が報じられた情報だけを鵜呑みにして、考えを進めてしまうと、重大な検討漏れを起こしてしまう恐れがある。

会社の会議などでも、同じようなことが起こりうる。議論をする際のベースとなる資料には、情報が網羅的に記載されていないかもしれない。資料を作成した人が、何らかの理由であえて一部の情報を示さなかったり、示せなかったりすることも考えられる。

いま得られている情報の裏に、何か隠されている情報はないだろうか？—— 物事を考える際、たまには、そのように疑ってみることも必要と思われるが、いかがだろうか。

(参考文献)

「新版 医学への統計学」丹後俊郎(著) / 古川俊之(監修) (朝倉書店, 統計ライブラリー, 1993年)

「統計学のセンス—デザインする視点・データを見る目」丹後俊郎(著) (朝倉書店, 医学統計学シリーズ 1, 1998年)

「宇宙怪人しまりす 医療統計を学ぶ」佐藤俊哉(著) (岩波書店, 岩波科学ライブラリー 114, 2005年)

「宇宙怪人しまりす 医療統計を学ぶ 検定の巻」佐藤俊哉(著) (岩波書店, 岩波科学ライブラリー 194, 2012年)