

レクチャー

メタアナリシスの進め方

荒木章裕

メタアナリシス (meta-analysis) とは、学術論文の系統的レビュー (systematic review) の統計学的分析に位置付けられており、エビデンスに基づく医療 (Evidence Based Medicine : EBM) の根拠として注目されている。そこには多くの厳格な基準が存在し、円滑な実施のためには基本的な知識が必須となる。本稿では、著者が実際に手掛けた経験を踏まえ、メタアナリシスの基本的な進め方について概説する。

キーワード：メタアナリシス，系統的レビュー，文献検索

I. はじめに

現代の研究においてメタアナリシスは、研究のエビデンスレベルが最も高い位置づけにある¹⁾。そのため本法は厳格な基準に従って実施する必要があり、特に最近では多くの学術学会誌などでも最低必要条件としてPRISMA 声明に準拠することが求められている²⁾。

メタアナリシスとは、過去に報告された複数の研究結果を統合するための統計解析と表現される³⁾。また既存の研究論文をまとめた「系統的レビュー (systematic review)」の一環であり、その中でも統計学的手法を用いた分析方法とも表現される⁴⁾。そこで用いられる論文は適格な基準と方法に沿って収集されるため、研究者は自身の研究課題 (research question) に関する過去の知見を網羅的に把握することができ、さらに科学的根拠に基づく医療 (Evidence Based Medicine ; EBM) の根拠として示すことが可能となる。

先述のとおり、メタアナリシスには厳格な基準が存在するため決して実践が容易なものではないが、少しでもそのハードルを低くすることができるよう、本稿では手掛けるにあたっての基本的なメタアナリシスの手順について解説していく。本稿が実践者のアシストとなれば幸いである。

II. 研究課題 (research question) の設定

基本的なことではあるが、メタアナリシスを実践するにあたっては、解決したい疑問(仮説等)を研究課題として具体的に設定する。例えば、「高齢者の睡眠を促すためにどのような方法が効果的か」ではなく、「高齢者への足浴は睡眠にどの程度効果があるか」などのように、ターゲットを絞り込むことが必要となる。

III. 文献検索

現存する同じテーマの研究論文をすべて検索・収集することは、出版バイアスや英語バイアス、データベースバイアス等の理由から不可能に近いと言える。しかしながら、メタアナリシスで用いる文献は、網羅的かつ偏り

のない検索によって実施されることが求められており、上記のバイアスを可能な限り除去する必要がある。

1. 文献データベース

まず、既存の文献を調べる際に活用されるのが、文献データベースである。データベース一覧と特徴をまとめたものを表1に示す。さらに必要に応じて引用検索、ハンドサーチなどによる検索を実施しながら、場合によっては、専門家に連絡を取る必要がある。ただし、表に挙げたように、日本人であれば日本語と英語を用いた文献検索を行うことが多いと思われるが、それ以外の言語で執筆された報告が隠れている可能性も否定はできない。

2. 検索語

データベース検索ではキーワードを用いて検索を行うことになるが、ここでどれだけ論文が網羅的に収集できるかが鍵となる。例えば、「高齢者への足浴は睡眠にどの程度効果があるか。」という研究課題に対し医中誌Webで検索を行う際、検索語に [高齢者]、[不眠]、[足浴] を用いるだけでは不十分と言える。その理由について以下に記述する。

まず、[高齢者] では広義的すぎる点が挙げられる。検索語からは「65歳以上の成人」というターゲットが連想されるが、検索では自動的に [高血圧] や [バリアフリー住宅]、[後期高齢者医療制度] などの意味合いも含まれてしまう。厳正な検索を行うのであれば、[高齢者看護]、[80歳以上高齢者] などの副項目を活用した検索がベストと言えよう。ただし、余分な論文が抽出されてしまう分には、この後の除外基準で弾かれることになるため、この点については重要な問題とはならない。

次に [不眠] という検索語では <睡眠に対する効果> を網羅できない点が挙げられる。[不眠] は “sleep initiation and maintenance Disorders”、“insomnia, fatal familial”、“night eating syndrome”、“prion proteins” を統制する言葉であり、本来知りたかったはずの “sleep” というターゲットが外れることになる。検索

表1 医学/看護系文献の検索に活用可能なデータベースとその特徴

文献データベース	主な言語	特徴
医中誌 Web	日本語	日本国内の医学・歯学・薬学分野の文献データベース。
CiNii Articles	日本語	日本国内で発行された学協会雑誌および大学研究紀要等に掲載された記事・論文情報のデータベース。
最新看護索引 Web	日本語	日本国内の看護文献データベース。
KAKEN 科学研究補助金データベース	日本語	文部科学省及び日本学術振興会が交付する科学研究費補助金により行われた研究のデータベース。
厚生労働科学研究成果データベース	日本語	厚生労働科学研究費補助金等で実施した研究報告書の概要版及び報告書が閲覧可能なデータベース。
国立国会図書館サーチ	日本語	国会図書館が提供するコンテンツに加え、全国の公共図書館、学術情報機関等のコンテンツが閲覧可能なデータベース。
JDreamIII	日本語	科学技術、医学、薬学関係の文献データベース。
JSTAGE	日本語	科学技術振興機構（JST）が運営する電子ジャーナルサイト、学協会が発行の学会誌・論文誌等が閲覧可能なデータベース。
PubMed	英語	アメリカ国立医学図書館の国立生物工芸情報センターが運営するデータベース。
CINAHL	英語	全米を中心とした看護学・健康関連分野の文献データベース。
Cochrane Library	英語	Cochrane Collaboration（コクラン共同計画）の成果であるシステムティックレビューを中心としたデータベース集。
MEDLINE	英語	世界各国の医学分野の文献データベース。

から除外された論文は、それ以降解析対象には成り得ないため、検索語の選定には細心の注意を払う必要がある。したがって、研究課題の対象を言い当てる言葉を複数選択し、なおかつ可能な限り広義的な言葉を選択する必要がある。（例：睡眠 - 覚醒状態、睡眠不全、睡眠衛生など）

以上のように、検索語は直感に頼らず慎重に選択していく。検索語の選定にあたっては、各種文献データベースでも検索可能なシソーラスや MeSH 用語を活用することが推奨される。

3. 検索式

論文を絞り込んでいく過程においては、検索式の設定も重要となる。また系統的レビューは定期的に更新されていくことが望ましく、過去の報告と新規の報告を含め解析をし直すための手段として、過去のメタアナリシ

スを参考とすることも少なくない。体系的な検索を継続して実施するための再現性という点において、いつ検索を実施したのかという情報に加え、論文中に検索式を記述することが求められる²⁾。

4. スクリーニング

上記の手順により抽出された論文は、予め設定した適格性基準に沿って、解析の対象とするかどうかの選定作業にかけられる。適格性を判断するための包含基準および除外基準とする内容として、①論文が出版された期間、②研究デザイン、③研究対象集団、④介入もしくはリスクファクター、⑤許容可能なコントロール群、⑥研究デザインに関するその他の要件、⑦許容可能なアウトカム⁵⁾の7項目が挙げられている⁵⁾。筆者の経験から特に除外基準を丁寧に設定しておくことを推奨するが、これは解説書等でも“garbage in garbage

out”と示されているように、質が保障されない材料を含めることで、完成品の質が落ちてしまう恐れがあるためである。

このスクリーニング作業は2人以上の査読者によって、2段階に分けて進めていく。1次スクリーニングでは論文タイトルと要旨を対象に判定し、2次スクリーニングでは全文レビューによる判定を行う。各スクリーニングの作業は2人以上の査読者で独立して実施することが望ましく、さらに2人の意見が食い違う場合には合意するまで議論するか、もう1人別の査読者による意見を求める。

5. その他

複数の文献データベースで検索を行う場合、複数のデータベースで同じ論文が抽出されることも少なくない。集計結果を示す際には何件が重複したのかまで報告する必要があるため、各データベースの検索結果を照合する作業も必要となる。

また文献の検索から適格性の判定までの過程で採用された文献と除外された文献は、その件数をフローチャートで報告する。特にスクリーニングの段階では、包含基準にどう合致したか、なぜ除外したのかの理由を記録しておく、文献を除外した理由を説明する必要がある。

IV. 結果の整理・統合

1. データ抽出

一連の過程を経て採用された個々の論文は、予め定めた様式（データ抽出フォーム）に従ってデータを要約していく。データ抽出

フォームに含めることが推奨される項目を表2に示す⁵⁾。メタアナリシスでの結果の統合には、研究対象集団やアウトカム指標などの研究対象と研究方法、実数データや検定結果（ p 値や信頼区間）などの結果に関する情報の2種類が必要となる。この要約作業においても、2名以上の要約者によって独立して実施されることが望ましく、2人の要約内容が異なる場合には、合意するまで議論するか、別の要約者の意見を求める必要がある。

抽出する際に注意したいのは、あくまでも結果で示された内容を抜き出すに留め、それを超える解釈は行わないことである。またスクリーニングに堪えられた論文であっても、様式に含めるべき情報が不足していることもある。そのような場合は、例えば要約者の主観で「採用したい」と思われる論文があったとしても、除外しなければならない。

2. データの統合

メタアナリシスにおけるデータの統合方法にはいくつかの種類があるが、本稿では代表的な2種類（固定効果モデル、ランダム効果モデル）を紹介する。まず、研究間に異質性が認められる場合にはランダム効果モデル（または変量効果モデル：random effect model）が選択され、異質性を認めない場合には固定効果モデル（または母数モデル：fixed effect model）を選択することが一般的である。この2つのモデルにおける統計手法には、ランダム効果モデルでは *DerSimonian-Laird*法、*general variance-based method*があり、固定効果モデルでは *Mantel-Haenzel*法、*Peto*

表2 データ抽出フォームに含めることが推奨される項目

1	適格基準
2	研究デザインの特徴
3	各研究群の参加者の特性と人数
4	介入あるいはリスクファクター
5	以前に設定されたサブグループにおける主たるアウトカム、副次的アウトカム
6	採用した研究の質の評価に必要な要素

note) Stephen B, et al., 2014 より引用⁵⁾

法、*general variance-based method*がある。

異質性（または非一様性；heterogeneity ⇔ 一様性；homogeneity）とは、異なる論文の研究結果の違いやバラツキのことを指す。異なる論文のデータを統合する際には、論文ごとにサンプルサイズや研究集団の特徴、介入内容、アウトカムの種類などが異なるため、単純な合算ではなく「重みづけ」を行った上で併合効果と信頼区間を示す必要がある。異質性の指標には各研究の偏りの分散 τ^2 の割合である I^2 値（%）や、コクランの Q 検定（Cochrane's Q）などが存在する。 I^2 値の範囲の解釈については、コクランハンドブックに記載されているものを参考にするとよい⁶⁾。その他にフォレストプロットを用いて目視で評価する方法もあるが、この手法のみで評価することは避け、上記の統計量と併記することが望ましい。

しかしながら、単純に異質性のみでモデルを決定するという判断は避けた方が良くとされている⁴⁾。その理由は、ランダム効果モデルが「扱った研究以外の研究も含めての母集団であり、統合したデータは母集団からランダムに抜き出されたものである」のに対し、固定効果モデルは「他の集団に一般化する前提がない」という考え方にある。言い換えるならば、明らかにしたい薬剤の効果について、前者は「他の集団に対しても効果があるか」に対し、後者は「既存の研究で示された結果に対して本当に効果があったと言えるか」ということである。ただし、ランダム効果モデルは固定効果モデルよりも信頼区間の幅が広くなり有意差が出にくくなることから、そちらを使用することが推奨されている⁵⁾。逆を言えば、固定効果モデルによって得られた結果は過大評価される可能性がある。いずれにしても、実際にメタアナリシスを実施する際には、使用したモデルとその根拠を示す必要がある。

3. 出版バイアスの評価

出版バイアスは文献検索の際にも気を付ける内容ではあるが、結果の統合でも意識しておく必要がある。出版バイアスとは、有意差

が出なかった研究は報告数が少ない等の可能性のことを指し、出版バイアスが認められるデータを対象にメタアナリシスを実施してしまうと、真の効果とは異なる偏った結果が得られることになる。したがって、結果を示す際には出版バイアスの可能性と併合効果への影響について触れる必要がある。

出版バイアスを評価する主な方法には、小規模研究によって大きな効果が示されていないかを調べる方法、想定される出版バイアスを把握した後に関式を復元することによってエビデンスを再構築する方法、統計学的有意水準に応じた出版バイアスの確立推定、累積メタアナリシスの4つが挙げられている⁷⁾。最も多く用いられるのは、ファンネルプロット（funnel plot）による目視確認（10件以上の研究が含まれている場合に利用が推奨される）であり、この分布が左右非対称になっている場合には出版バイアスが疑われる。しかし目視評価だけでは判断に乏しいこともあるため、さらに Egger の回帰検定等を用いて統計学的に評価することもある。

4. サブグループ解析と感度分析

メタアナリシスで扱った研究間の報告に異質性が認められ、さらにその原因が集団や介入、アウトカムにあると考えられる場合、異質性の原因と成り得る項目ごとにサブグループ解析を実施することが望ましい。これにより、異質性がなぜ生じたかを説明することが可能となり、より真の効果に近い結果が得られる。一方で、多くの項目でサブグループ解析を実施すると、多重検定のため有意差が出やすくなり、誤った結果が得られる可能性がある。また本解析では有意な結果が得られず、サブグループ解析で有意な結果が出てしまうことで、後者を選択的に報告されてしまう「サブグループ解析バイアス」というものも存在するので注意したい⁸⁾。メタアナリシスを実施する際には、あらかじめ実施する予定のサブグループ解析項目とその理由を明確にしておくことや、異質性の要因ごとに異なる推定値を提示することが望ましい。

また、「外れ値」に該当するような研究結

果があるのであれば、それを加えた場合と加えない場合との併合結果の違いを示す感度分析が有用となる。メタアナリシスに加えるかどうか曖昧な基準にあるデータに対しては、結果の確かさを得るためにも実施することが望ましい。

V. 最後に

本稿は筆者が経験したメタアナリシスの実践から投稿までの経験を基に作成したものである。その際の査読では特に方法の箇所について丁寧な記載を求められたこともあり、説明が必須と思われる箇所については重点的に説明を加えている。分析にあたっては無料のソフトウェアも多数あり、知識があれば実践すること自体は容易と言える。また本稿はあくまでも基礎的な知識を記述したものであるが、インターネットを通してメタアナリシスに関するさらに多くの知識が容易に入手でき、手掛けるための敷居は低くなっていると思われる。これから手掛けようとする方は、本稿を参考にしながら、学会誌等への投稿基準として提示されることが多くなったPRISMA 声明に沿って整理し、決して諦めることなく完遂していただきたい。

また冒頭にも述べたとおり、メタアナリシスは系統的レビューの統計学的な分析方法という位置づけであり、データの抽出の項までは系統的レビューでも通用する内容である。これから研究を始める方や、文献レビューを手掛けようとしている方にも活用してもらえると幸いである。

引用・参考文献

- 1) 卓興鋼, 梅垣 敬三, 渡邊 昌. 【健康機能食品を科学する】国内における健康機能食品の現状と動向. アンチ・エイジング医学. 4 (1), 21-29, 2008
- 2) 卓興鋼, 吉田 佳督, 大森 豊緑. エビデンスに基づく医療 (EBM) の実践ガイドライン システマティックレビューおよびメタアナリシスのための優先的報告項目 (PRISMA 声明). 情報管理. 54 (5), 254-266, 2011
- 3) 丹後 俊郎. 新版メタ・アナリシス入門 エビデンスの統合をめざす統計手法. p1. 朝倉書店, 2020
- 4) 相原 守夫. Anonymous 診療ガイドラインのための GRADE システム (3 版). 34-37, 中外医学社, 2018
- 5) Stephen B Hulley, Steven R. Cummings, Warren S. Browner, Deborah G. Grady, Thomas B. Newman (2014). 木原 雅子, 木原 正博 (2014). 医学的研究のデザイン 研究の質を高める疫学的アプローチ (4 版), 228-237. メディカル・サイエンス・インターナショナル, 2014
- 6) Jonathan J. Deeks, Julian PT. Higgins, Douglas G. Altman. Cochrane Handbook for Systematic Reviews of Interventions Version 6.1. 2021 (閲覧日: Mar 9., 2020.)
- 7) 相原 守夫. 診療ガイドラインのための GRADE システム. 68-70. 中外医学社, 2018
- 8) 丹後 俊郎. 新版メタ・アナリシス入門 エビデンスの統合をめざす統計手法. 55-62, 朝倉書店, 2020

連絡先：荒木章裕
茨城県土浦市真鍋 6 丁目 8-33
TEL：029-826-6622
Mail：a-araki@tius.ac.jp