



ネットワークメタアナリシスの概要および留意事項
— ISPOR レポート(2011)を参考に —

日本製薬工業協会 医薬品評価委員会

データサイエンス部会

Ver 1.0

2019年11月

目次

略語.....	4
1. はじめに.....	5
2. ネットワークメタアナリシスの概要.....	7
2.1. ネットワークメタアナリシスの意義.....	7
2.2. エビデンスのネットワークの種類.....	7
2.3. エビデンスの統合.....	9
間接比較及び混合比較の妥当性を担保するための仮定.....	9
解析手法並びに効果の指標の選択時の留意点.....	11
統計モデル.....	12
パラメータ推定の方法.....	14
観察研究による比較の利用.....	15
3. ネットワークメタアナリシスの報告のレビュー.....	17
3.1. 内的妥当性と外的妥当性.....	17
内的妥当性.....	17
外的妥当性.....	18
3.2. ネットワークメタアナリシスの報告.....	18
導入.....	20
方法.....	20
結果.....	21
考察.....	24
3.3. 2014年のISPORタスクフォースによる質問項目.....	24
4. 解釈と意思決定.....	27
4.1. 解釈.....	27
4.2. 意思決定.....	27
5. まとめ.....	28
参考文献.....	31
執筆者.....	38

表目次

表 3-1 報告された NMA を評価するときに意思決定を補助するチェックリスト	19
--	----

図目次

図 2-1 ネットワークの種類と分類	8
図 2-2 間接比較における類似性の仮定	10
図 2-3 混合比較における一致性の仮定	11

略語

略語	正式名称及び和名
AIC	Akaike's information criterion. 赤池情報量基準
BIC	Bayesian information criterion. ベイズ情報量基準
CTDS	Clinical trial data sharing
DIC	Deviance information criterion. 逸脱度情報量基準
ECOG	Eastern Cooperative Oncology Group
ISPOR	International Society for Pharmacoeconomics and Outcomes Research. 国際医薬経済・アウトカム研究学会
ITC	Indirect treatment comparison. 間接比較
MCMC	Markov chain Monte Carlo
MTC	Mixed treatment comparison. 混合比較
NMA	Network meta-analysis. ネットワークメタアナリシス
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomized controlled trial. ランダム化比較試験
SUCRA	Surface under the cumulative ranking

1. はじめに

利用可能なエビデンスを集約するための手法としてのメタアナリシスは、医薬品評価などの分野で応用・活用されている。2種の治療を比較することを目的としたメタアナリシス、所謂対比較のメタアナリシスは、例えば同種同効薬剤に対する興味の対象の薬剤の位置づけの確認や、承認申請時の統合データ解析に由来から利用されている。しかし、対比較のメタアナリシスで解析の対象にできるのはあくまで2種の治療であり、興味の治療が3種以上あるときには、これらを同時に評価することはできない。対比較のメタアナリシスを一般化したネットワークメタアナリシス（network meta-analysis, 以下 NMA）は、例えば薬剤の有効性及び安全性などのアウトカムについて、複数の試験を基に複数の薬剤の比較の結果を統合することで、直接的な薬剤間比較（直接比較）のみならず、たとえ直接的な比較の結果が存在しなくとも間接的な薬剤間比較（間接比較）を行うための方法論である。その有用性から、さまざまな疾病領域で本方法論の適用事例が報告されており、その結果を治療ガイドラインの策定や医療政策決定のための情報として活用することが考えられている。

一方、NMA は、その有用性の反面、評価する治療の数が多いことや方法論の複雑さ故にその特性を理解するのに困難を伴うことに加え、結果の妥当性を担保するために満たすべき仮定も存在するなど、NMA を活用するには留意すべき事項がある。この仮定は、例えば、本方法論を適用する治療群が統合するのに十分に類似していること、つまり各試験のデザイン、対象集団、及びアウトカムが十分に類似していること、そして直接比較による結果と間接比較のそれが統合するのに十分に類似していることなどが挙げられる。これらの仮定が十分に満たされない場合には、結果の解釈が困難であり、結果にバイアスが生じている可能性が高いことに注意を払う必要がある。

本報告書では、ISPOR の Indirect Treatment Comparisons Good Research Practices のタスクフォースが 2011 年に公表した NMA に関するレポート[1]を参考に、NMA で得られた結果の解釈において注意を要する点、及び本方法論の利用や結果の活用について説明し、本方法論を応用的観点から解説する。ただし、補足することを目的として 2014 年に公表された同タスクフォースのレポート [2] も参照する。ISPOR とは、International Society for Pharmacoeconomics and Outcomes Research (和名：国際医薬経済・アウトカム研究学会) であり、医薬経済学とアウトカム研究の啓発と普及を推進する国際組織である [3]。なお NMA の適切な評価や解釈に役立つ PRISMA¹のチェックリストが公表されているが、それに関しては既に発表されている文献を参照していただきたい[4, 5]。

第 2 節では NMA の概要として、その意義、ネットワークの種類及び各試験結果の統合方法について説明する。第 3 節では、NMA の結果を利用する意思決定者に向けた、NMA の報告に対する批判的なレビューや解釈を示す。第 4 節では、NMA の結果の解釈及び意思

¹ PRISMA については後述する。

決定を行う上で考慮すべきポイントを述べる。なお、以上の節は ISPOR のタスクフォースのレポート[1, 2]に基づいているが、我々のタスクフォースの見解についてはコラム若しくは脚注として示した。最後に、第 5 章で本報告書をまとめる。

本報告書は統計担当者以外の担当者でも理解できるような内容となっている。しかし、2.3 節などでは統計学の知識を要する箇所もある。

なお、本報告書では NMA の対象とするエビデンスをランダム化比較試験 (randomized controlled trial : 以下, RCT) 由来のものに主眼を置いて議論する。後述するように、観察研究のような非介入試験の比較結果を利用する際には特別な注意を要するためである (2.3 節)。

2. ネットワークメタアナリシスの概要

2.1. ネットワークメタアナリシスの意義

RCT のシステマティックレビューは、エビデンスに基づいた治療選択の手段の一つであり、英国等では保険償還の意思決定にも用いられている。多くのシステマティックレビューでは、複数の類似したデザインの試験結果を定量的に統合し、エビデンスを集約するための手法としてメタアナリシスが利用される [2]。メタアナリシスにおいては 2 種の治療について比較するが、より妥当な意思決定をするためには、全ての関連する治療間の比較が求められ、良くデザインされた RCT によって全ての関連の治療を同時比較することが理想的である。しかしながら、このような RCT を実施することは実質不可能である。その結果として、同時比較されていない複数の RCT のエビデンスに基づく複雑な解釈並びに意思決定を強いられることになる [6, 7, 8, 9]。

関心のある治療同士の直接比較を含む試験が存在しない場合、間接比較がそれらの治療効果の差の有用なエビデンスとなり得る。例えば、2 種の治療が直接比較されていなくとも、共通の比較対照をおいた比較試験が各治療に関して存在する場合、間接比較は共通の比較対照を介した相対的な効果として評価可能である [10, 11, 12, 13]。

間接比較は直接比較のエビデンスが利用不能である場合の拠り所となるが、直接比較と間接比較のいずれのエビデンスも利用可能な場合は、それぞれが全体としてのエビデンスに寄与することを理解することは重要である。即ち、間接比較のエビデンスと直接比較のエビデンスを統合した混合比較の結果は、直接比較のみによる評価よりも強固なエビデンスとなり得る [6]。直接比較のエビデンスが決定的なものであったとしても、混合比較では比較する治療のネットワーク中の情報を相互に関連付け、ネットワークを最大化することにより、混合比較は直接比較のみのエビデンスよりも精度の高い治療効果の推定値をもたらし得る。推測の対象が収集された試験に広がったとみることもできるかもしれない [12]。

ネットワークが直接比較や間接比較、或いはその両方を含む RCT によって構成されるならば、それらは NMA によって統合可能である [14]。対比較のメタアナリシスでは、対象とする全ての比較試験が同じ比較対照に対する同じ治療からなる。即ち、関心のある 2 種の治療を直接比較した RCT で構成される。一方、NMA は対比較のメタアナリシスの概念を複数の治療からなる複数の対比較に拡張した方法論であり、複数の治療間の相対的な治療効果を推定できる。複数の対比較による推定値は効果比較²の目的において有用となる。なお、本レポートでは、効果比較という用語を治療間でのアウトカムの比較を指すものとして利用する [15]。

2.2. エビデンスのネットワークの種類

図 2-1 は複数の治療が比較される際のネットワークのパターンを示している。各節点は

² 原文では comparative effectiveness

治療を表し、2つの節点を結ぶ線は1つ以上のRCTが存在することを意味している。連結されたネットワークに属するどの治療に対しても、相対的な治療効果を推定することができる。

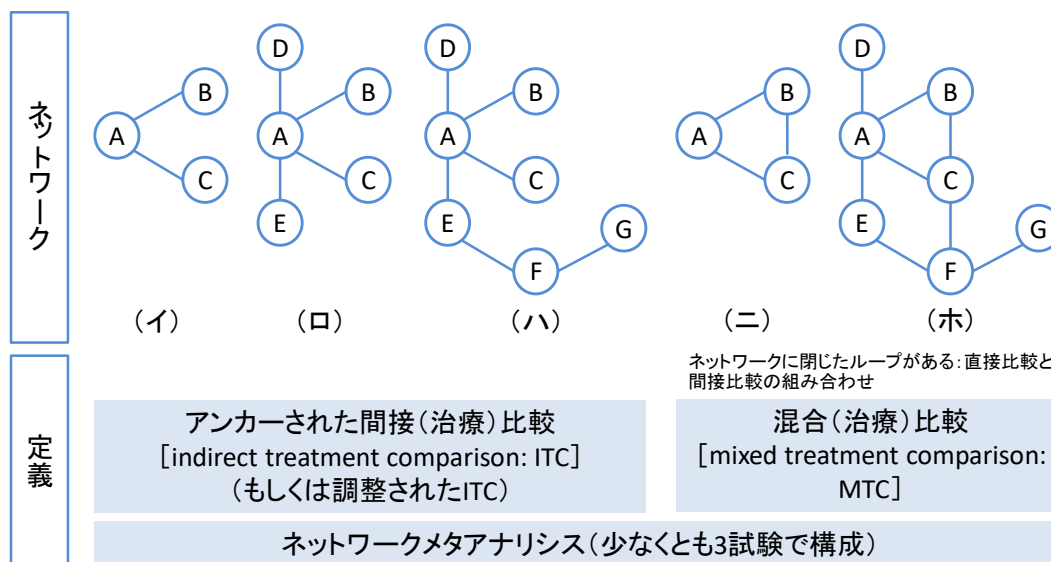


図 2-1 ネットワークの種類と分類

例として、直接比較のRCTが存在しない治療Bと治療Cの比較を考える。図 2-1の(イ)のネットワークでは、治療Bは治療AとのRCTが、治療Cは治療AとのRCTがそれぞれ存在しており、間接比較によって治療Bと治療Cの相対的な治療効果を推定することができる。治療Bと治療Cの間接比較は治療Aが共通の対照治療、つまり“アンカー”として介することにより可能となる³。なお、治療Aは実薬でもプラセボでも構わない。

図 2-1の(ロ)のネットワークでは、全ての治療がネットワーク中で連結しているため(即ち、任意の2治療を繋ぐパスを有している)、共通の対照治療Aをアンカーとして治療B~Eの全ての対比較に対する間接比較が可能である。このような“スター型”ネットワークの事例として、4種の競合治療それぞれのプラセボ対照試験の結果を利用した、骨粗鬆症治療としてのビスホスホネートの比較が挙げられる [16]。

(ハ)のネットワークでは、全ての治療について共通の比較対照治療との直接比較のエビデンスを有していないが、全ての治療が連結している。治療Fと治療Gは、治療Eと治療FのRCT並びに治療Fと治療GのRCTを介して治療A, B, C, D及びEと連結しており、すべての治療間で間接比較が可能である。ただし、長いパスを通じて連結している治療

³ 間接比較が治療Aにより調整されている (adjusted) と表現される文献もある。ただし交絡因子の調整の意味合いはない。

間の間接比較は精度を欠いてしまう [12]。このネットワークの事例としては、全身真菌感染症に関する抗真菌剤の NMA が挙げられる [17]。

(ニ) のネットワークは 1 番目のネットワークと同様に A~C の 3 種の治療から構成されているが、RCT の結果からいずれの対比較も可能である。このネットワークの重要な特徴は閉じたループ (closed loop) であるということであり、各対比較は直接比較と間接比較のエビデンスを有している。例えば、治療 B と治療 C の比較は、これらの治療の RCT に基づく直接比較のエビデンスと治療 A と B 並びに治療 A と C の RCT に基づく間接比較のエビデンスを有しており、他の対比較についても同様である。即ち、このネットワークは混合比較である。3 種の介入を比較した混合比較の事例として、薬物溶出性ステントとベアメタルステントに関する Stettler らの混合比較が挙げられる [18]。

(ホ) のネットワークは治療 A~C による混合比較を含み、さらに治療 A, C, E, F がもうひとつの長いループを構成している。ループを含んでいるネットワークで重要なのは、間接比較が直接比較の相対的な治療効果と一致していることであり、これについては後に詳しく説明する [12, 19, 20]。ループを含む NMA の事例として、Psaty らによる高血圧に対する一次治療の NMA [21]、Cooper らによる心房細動をともなう患者における脳梗塞の予防に関する NMA [22]、Vissers らによる癌による突出痛 (breakthrough cancer pain) に対するオピオイドの NMA [23]、そして Cipriani らによる大うつ病に対する新規の抗うつ薬の NMA [24] が挙げられる。Salanti らは公表された試験のさまざまなネットワーク構造を概観している [25]。

間接比較、混合比較及び NMA は時に交換可能な表現として使われるが、本 ISPOR のタスクフォースのレポートでは「3 種以上の治療が連結したネットワークが 3 つ以上の RCT から構成される場合」を NMA と呼ぶことを提案している。さらに、そのネットワークに少なくとも 1 つの閉じたループを含む場合はその解析は混合比較、ループのないその他全てのネットワークの解析を間接比較と呼ぶことを提案している。

2.3. エビデンスの統合

間接比較及び混合比較の妥当性を担保するための仮定

治療のネットワークとそれを構成する RCT が所与であるとして、解析の目的は個々の RCT の結果を統合して対比較ごとに相対的な治療効果を推定することである。このとき、統合する RCT の試験デザインなどが十分類似しているかどうかは重要な問題である。

対比較のメタアナリシスでは、同じ治療を比較した複数の RCT の結果を統合し、相対的な治療効果 (例えば、ハザード比、オッズ比、相対リスク、ベースラインからの変化量の差など) の統合推定値とばらつきの推定値を得る。ここで、ランダム化は同一試験内での比較

の内的妥当性⁴しか保証しないことを認識することは重要であり、それ故、異なる試験間では試験特性並びに患者⁵特性が異なることがある。もし実際にこれらの特性が試験内の群間でバランスがとれていたとしても、試験間では異なり、かつその特性が相対的な治療効果の修飾因子である場合、統合される試験は異質とみなされる。治療効果の修飾因子としては、患者特性、アウトカム⁶の定義やその測定方法、併用可能な薬剤、及び観察期間といった試験の実施計画上の特性などが例として挙げられる。

同様に、複数の治療の比較を伴う NMA でも、特定の対比較における複数の試験が異質であるという問題は起こり得る。さらに、治療 A と治療 B、治療 A と治療 C の直接比較間でそれぞれを構成する試験の特性に差異があり、かつその差異が相対的な治療効果の修飾因子によるものである場合、治療 A をアンカーとした治療 B と治療 C の間接比較の推定値はバイアスを伴う [11, 26, 20, 27]。この状態は間接比較の類似性 (similarity) の仮定が満たされていないと表現される [26, 20]。

図 2-2 に間接比較における類似性の仮定を図示した。いま、治療 C に対する治療 B の相対的な治療効果の間接比較に基づく推定値を d_{BC} とし、対数オッズ比、対数ハザード比といった正規分布に従う治療効果の差の指標とする。治療 A と治療 B の比較試験と治療 A と治療 C の比較試験について、両試験間で治療効果の修飾因子の分布が類似している場合、 d_{BC} は治療 B に対する治療 A の直接比較に基づく相対的な治療効果 (D_{AB}) と治療 C に対する治療 A の直接比較に基づく相対的な治療効果 (D_{AC}) から⁶、 $d_{BC} = D_{AC} - D_{AB}$ としてバイアスを伴う事なく推定することが可能である⁷。即ち、類似性の仮定では、間接比較から推定される d_{BC} と、3 治療の同時比較試験中での直接比較に基づく相対的な治療効果の推定値が同じ値であることが期待できるとする [12]。

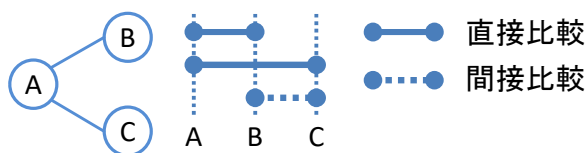


図 2-2 間接比較における類似性の仮定

⁴ 内的妥当性とは、観察された相対的な治療効果が交絡なく治療間の差によるものとし、それにより因果が治療によるものであるとできる程度である [28]。

⁵ 本報告書では、ISPOR のレポートに従って被験者 (subject) を用いず、患者 (patient) で統一した。

⁶ ISPOR のタスクフォースのレポートでは、相対的な治療効果の推定値を間接比較と直接比較の区別なく d としているが、本報告書では区別してそれぞれ d および D とした。

⁷ B と C の比較は間接比較のみなので、直接比較に基づいた推定値を得ることはできない。そこで A と B および A と C の直接比較の推定値で、B と C の比較を間接比較の推定値として得る (類似性の仮定に基づき d_{BC} を $D_{AC} - D_{AB}$ で推定する)。

特定の対比較に関して直接比較と間接比較のそれぞれの相対的な治療効果の推定値が統合される際は、間接比較の推定値がバイアスを含まず、かつ直接比較と間接比較の推定値に乖離がないことが重要である [20,25,29,30]。即ち、これらの直接比較と間接比較の間の一致性 (consistency) が示されなければいけない。図 2-3 に一致性の仮定の構成要素を図示した。このネットワークでは、治療 A から C の全ての対比較において直接比較と間接比較のエビデンスが存在する。一致性を保持するためには、 $D_{BC} = D_{AC} - D_{AB}$ が満たされる必要がある⁸ [20,29]。もし、3組の対比較のうち1つ以上で、それらを構成する試験の相対的な治療効果の修飾因子の分布に偏りがある場合は、一致性の仮定は正当化できないだろう。ここで、一致性はエビデンスのループに対してのみ適用される概念であるため、治療 AB 間の対比較と治療 AC 間の対比較で一致性を満たすと言った表現は意味を成さず、全ての対比較⁹で一致性を満たすという表現が適当である。図 2-3 のように閉じたループのネットワーク下で一致性が成立しない簡単な事例を示す。いま標本誤差を無視して、治療 B に対する治療 A の何かしらの反応に関するオッズ比の真値を 0.4 ($OR_{AB}=0.4$)、治療 C に対する治療 A の真値を 0.5 ($OR_{AC}=0.5$) とすると、治療 C に対する治療 B のオッズ比の真値は $OR_{BC}=OR_{AC}/OR_{AB}=1.25$ であると期待される。この時、治療 B と治療 C の直接比較による推定値が 1.25 ではなかった場合は一致性が成立しない¹⁰。

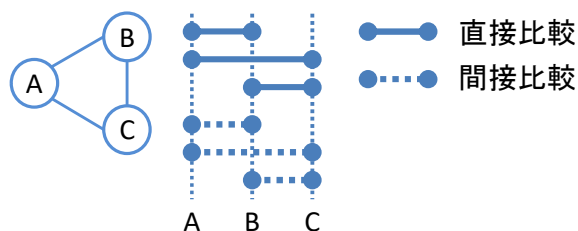


図 2-3 混合比較における一致性の仮定

解析手法並びに効果の指標の選択時の留意点

ネットワーク中の異なる RCT の結果を統合し、全ての対比較についての相対的な治療効果の推定値を得るために必要なことが 2 つある。

⁸ B と C の比較は混合比較であるから、直接比較と間接比較いずれの推定値も得られる。これらの推定値が等しい性質を一致性と呼ぶ (D_{BC} と $D_{AC} - D_{AB}$ は等しいという性質)。

⁹ 全ての対比較とは治療 A, B 及び C のそれぞれの対比較を指している。

¹⁰ ここでは概念の説明の都合上真値が分かっている状況としているが実際的ではない。また、標本誤差を無視しているが、実際の NMA の解析時は常に標本誤差が存在するため、このような点推定値のみに基づいた一致性の評価は本来不適切である。ゆえに「1.25 と大きく異なる場合は一致性が成立しない」とするのが実際的であろう。

第一に各試験内でランダム化により担保される内的妥当性を保持することである。つまり、“break randomization”しないこと、すなわちランダム化状態を崩さないことである [8, 13, 31]。例えば、図 2-1 の左端のネットワークにおいて、治療 A から C がそれぞれ関節リウマチの治療である場合、治療 A と治療 B の RCT で観測された治療 B の反応割合と、治療 A と治療 C の RCT で観測された治療 C の反応割合を単純に比較するのは不適切である。それは、治療 A がプラセボである場合、この方法で比較をするとプラセボ反応と薬剤の有効性を分離することができないためである。ランダム化状態を崩すべきではないもうひとつの理由は、たとえ相対的なリスクが試験間で一致していたとしても、反応の差は異なるベースラインリスクを反映しているかもしれないことである [8, 26]。関心のある被検群のみを比較試験から抽出して各対照群を無視すると、比較結果にバイアスをもたらすため、そのような比較は避けるべきである [11]。従って、各試験でのランダム化の状態を保持するために、オッズ比やハザード比などといった相対的な治療効果で比較しなければならない。

第二に試験間でのランダム化の欠如によるバイアスを最小化する手法を用いることである。これについては 2.3 節で述べるとおりメタ回帰を用いることでバイアスを最小化するアプローチが考えられる。

統計モデル

解析にあたっては、ネットワーク中の任意の直接比較と間接比較の相対的な治療効果の関連性を、統計モデルを用いて定義することができる [12]。A から C の 3 種類の治療からなるネットワーク下で、類似性と一致性の仮定が成立する場合には、相対的な治療効果の推定値は前述のとおり $d_{BC} = D_{AC} - D_{AB}$ と表すことができる。この表記法をあらゆるネットワークにも適用できるように一般化すると、アンカーとした治療を A、興味の対象の治療を k、及び対照治療を b として $\delta_{bk} = \delta_{Ak} - \delta_{Ab}$ と表記できる¹¹。ただし、ここでは類似性と一致性の仮定が成立する条件下で直接比較と間接比較の違いは無視している。ネットワークに依存する形で、k としては治療 B, C, D, E など様々な治療をとり得る。対照治療 b の選択は、結局は全ての対比較を行うので、アンカーとした治療以外であればどれでもいいのだが、アルファベット順に最も早いものから表示するという表示法が一つの方法としてあり得る（ここでは治療 B）。このように表記することで、全ての治療が一つのネットワークとして繋がっている限り、任意の特定の対比較の推定値をアンカーである A に対する相対的な治療効果で表記することができる。この時、 δ_{Ak} (k は A 以外) は“basic parameter”と呼ばれ、利用可能な試験から推定される。A 以外の治療間の対比較のパラメータ（例えば δ_{BC} , δ_{BD} , δ_{CD} など）は“functional parameter”と呼ばれる [29]。このようなアンカーとした治療との対比較が他の全ての治療についてある前提で、K 個の治療と T 組の対比較

¹¹ ISPOR のタスクフォースのレポートでは相対的な治療効果の推定値として d としているが、本報告書では δ とした。本文直後の説明の通り、直接比較の推定値 d と間接比較の推定値 D の違いを無視していることによる。

を含むネットワークの場合、 $K-1$ 個の **basic parameter** と $T-K+1$ 個の **functional parameter** が存在することになる。要約すると、NMA のモデルは、 K 個の治療から定義される複数の対比較を行うために、 $K-1$ 個のパラメータを推定する必要があるという意味で、対比較のメタアナリシスのモデルの拡張であるといえる。このような NMA のモデルはループを含むか否かによらず、いずれのネットワークでも適用される。

NMA は固定効果モデル又は変量効果モデルを仮定して解析される。固定効果モデルでは、特定の対比較の相対的な治療効果に関して試験間の異質性は存在しないことを仮定する [26, 32]。この仮定の下では、特定の対比較において試験間で観測される差異は偶然によるものとみなされる。そこで、試験間で異質性が存在する場合には、変量効果モデルを用いるのが自然であろう。一般に変量効果モデルでは、各試験の相対的な治療効果に関するパラメータが試験間で交換可能 (**exchangeable**) とみなされ、統合した相対的な治療効果を平均値、その標準偏差の大きさに異質性の程度が反映された正規分布からの標本として各試験の相対的な治療効果が表現される¹² [2, 32, 33, 34, 35]。なお、変量効果モデルに基づく NMA では、しばしば異質性を反映する分散の値は全ての対比較において一定であると仮定される [12]。

変量効果モデルでは異質性を明確にモデル化する一方で、異質性そのものを説明していない。そこで、NMA のモデルを治療と共変量の交互作用を含むように拡張することにより、相対的な治療効果の異質性を説明し、かつ共変量の異なる水準での相対的な治療効果を推定することが可能になる¹³。既述のとおり、相対的な治療効果の修飾因子となり得る共変量の分布に試験間差がある場合、NMA の結果はバイアスを伴う [26, 20, 27]。この問題の対処法の一つとして、メタ回帰を用いて試験レベルの共変量を治療と共変量の交互作用とともに考慮することにより類似性や一致性の仮定からの逸脱によるバイアスの影響を軽減できる [20]¹⁴。一方、試験間で分布に差異がある共変量でも、その共変量が相対的な治療効果の修飾因子でない場合には、メタ回帰で当該共変量を考慮する必要はない。

大抵、ネットワークを構成する試験の数は限られているため、メタ回帰による試験レベルの共変量の調整は、疑わしいことがある [26, 36]。加えて、試験単位に要約されたレベルの

¹² 変量効果モデルでは、異質性に対して直接的にモデリングする点で固定効果モデルと異なる。異質性は、試験の特性を表す何かしらの要因によって生じる可能性もある。

¹³ 交互作用のみならず、共変量そのもので調整することも検討される。これは、後述の試験レベルの共変量の調整に相当する。

¹⁴ メタ回帰は異質性に対する興味共変量の説明可能性の程度を調べる方法である [39]。異質性について考えられる原因を理解することは、メタアナリシスで得られた結果の科学的な価値および臨床的妥当性の両方を高めることができるだろう。一方、本文にもある通り、変量効果モデルで異質性の原因となる共変量が何であるかを探ることはできない。

共変量による調整は ecological bias¹⁵を引き起こし、推定される結果の解釈を限定的なものにし得る [36, 37, 38]。これに対し、患者レベルの NMA では、メタ回帰による推定にあたって通常十分な情報量を有しているため、類似性や一致性の仮定からの逸脱によるバイアスの影響を軽減するとともに、試験間での治療効果の差異を探索する機会をもたらす。しかし、ネットワークを構成する全ての RCT について患者レベルのデータを入手するのは非現実的であるため、代替策として、患者レベルのデータが利用可能な試験ではこれを利用し、そうでない場合には試験レベルのデータを利用することで、試験レベルのデータのみを利用する場合よりは相対的な治療効果の推定を改善する¹⁶というアプローチが採用され得る¹⁷。

変量効果モデルでは試験特異的な相対的な治療効果の異質性が明確にモデル化されるので、固定効果モデルよりも変量効果モデルの方がデータに適合する。同様に、治療と共変量の交互作用項を含めるように固定効果モデル及び変量効果モデルを拡張することでもモデルのデータへの適合度を高めることができる。しかし、特定のデータセットが与えられた下では、推定すべきパラメータが増えるほどパラメータ推定値の不確実性も大きくなる。従って、データに十分適合しかつ交絡によるバイアスを軽減することができ、一方で安定したパラメータ推定値をもたらすモデルを用いることが目標となる。固定効果モデル及び変量効果モデルや、モデルに共変量の交互作用項を含めるかなどのモデルの選択は、競合するモデル間でデータに対するモデル適合度 (goodness-of-fit) の比較によってなされる。その具体的な基準として尤度を用いた赤池情報量基準 (AIC)、ベイズ情報量基準 (BIC)、逸脱度情報量基準 (DIC) などが用いられる [41, 42, 43]。

パラメータ推定の方法

NMA は頻度論の方法とベイズ流の方法のどちらのフレームワークでも実施可能であるが、本レポートではベイズ流の方法に焦点をあてる。ベイズ流の方法では、仮定するモデルのパラメータがとり得る値に対する事前の確信度を反映した事前分布と実際の観察データに基づく尤度を併合し、対応する事後分布を得る [44]。得られるパラメータの事後分布に基づけば、治療 A が治療 B よりも優れた反応を示す確率は x%であるといったように解釈で

¹⁵ 試験間に認められた関連性の方向が、各試験で観察された関連性の方向と反対となる状況を引き起こす可能性のあるバイアスである [40]。メタ回帰や部分集団解析を実施しても、回帰に用いられる共変量や部分集団を構成するのに考慮される変量というのは、あくまで試験レベルの代表値であり、その共変量や変量のアウトカムに対する真の影響を誤って推定してしまうことがある。

¹⁶ 相対的な治療効果の推定を改善するとは、バイアスの影響を軽減することを意味していると考えられる。

¹⁷ 実際に患者レベルのデータを NMA 或いは混合比較でパラメータの推定精度が改善されるという報告が 2012 年に相次いで報告されている [40, 48, 49]。また、臨床試験で得られた被験者レベルの個別被験者データを研究者に共有する取り組み (Clinical Trial Data Sharing : CTDS) もあることから、どのようなデータソースでエビデンスの統合を行うか慎重に検討されるべきである。

きる。観測された結果が設定した事前分布に影響されないように、相対的な治療効果のパラメータに対して無情報事前分布を適用することも可能である。頻度論の方法に対するベイズ流の方法の主たる利点は、事後分布の確率的な解釈によって意思決定を支援するフレームワークをもたらす点である [44, 45, 46]。また、NMA に特異的な利点としては、パラメータの事後分布にもとづいて、競合する治療ごとにその治療が最も優れる確率の計算や、その他の確率的な記述をすることができる点であり [47]、これらは意思決定者（政策決定者、医療専門家や医師など）に直接関連する情報をもたらす。一方で、後述するようにこれらの確率を過剰解釈するリスクがあることにも留意が必要である。ベイズ流の方法の他の利点としては、予測をするのが簡便である点や異なる情報源の不確実性を反映できる点が挙げられる [44, 45]。

観察研究による比較の利用

RCT やそれらの NMA による間接比較の結果が利用できない場合に、ランダム化比較ではない観察研究の比較結果を利用するかどうかは、意思決定者にとって重要な問題である。これまでも述べてきているとおり、試験間で患者特性や試験特性の分布が異なり、かつそれらの特性が相対的な治療効果の修飾因子である場合、これらの試験に基づく間接比較の結果はバイアスを伴う。また、RCT の NMA であってもランダム化の恩恵は試験間には適用されないことを念頭におくことが重要である。即ち、RCT に基づく NMA は観察研究によるエビデンスのような形態をとる。それでも RCT に基づく NMA の方が、観察研究に基づく NMA に比べて交絡によるバイアスは確実に小さくなる。観察研究に基づく NMA では、未測定の変量の分布の差異が治療と結果の両方に影響する場合、結果にバイアスを伴うのに対して、RCT に基づく NMA では、その差異が試験間で生じ、かつその変量が相対的な治療効果の修飾因子である場合に限り結果にバイアスを伴う。観察研究に基づく NMA に比べると、RCT に基づく NMA の方が、恐らくバイアスは生じにくいであろう。故に、RCT に基づく NMA の結果が利用できない場合に観察研究の比較結果を使うべきかどうかを問うことは、意思決定者に示すに十分な内的妥当性（3.1 節参照）を有しているといえる観察研究のエビデンスレベルはどの程度かを問うことと同義である。また、NMA に対して許容できる最低のレベルが RCT に基づくものか観察研究に基づくものかを問うこともいえる。

コラム：観察研究を含めた NMA

NMA の対象となるエビデンスとして RCT に加え観察研究を含めたいような状況、もしくは対象が主に観察研究からのものとなる状況であっても、NMA によるエビデンスを得たいという要求はあるだろう。本文にもある通り、各試験の相対的な治療効果でさえ、未観測の交絡因子によるバイアスが生じている可能性が否定できないなど、NMA の内的妥当性に影響する観察研究由来の特有の問題がある。

その問題に向き合い、その上で NMA を実施するのであれば、その問題によるバイアスのリスクを適切に評価する必要がある（選択バイアス、交絡バイアス、曝露変数やアウトカムの測定に関連したバイアス）。また、観察研究の対比較のメタアナリシスに関するチェックリストは **Stroup** らによって与えられているのでこの観点から各試験を評価するのも、バイアスのリスクの評価の助けとなるであろう [50]。統合の対象となる各試験の評価を行ったうえで、バイアスのリスクを踏まえ、NMA を実施することの妥当性を判断することになる。

RCT と観察研究を混合して NMA を実施する場合には、観察研究のエビデンスを除いて統合した結果とを比較するなどといった感度分析も有用だろう。

3. ネットワークメタアナリシスの報告のレビュー

この節では、NMAの結果を利用する意思決定者に向けた、NMAの報告に対する批判的なレビューや解釈を示す。3.1節で、NMAの内的妥当性、外的妥当性に関する論点を、3.2節で、NMAの報告に対して、適切な評価や解釈に役立つチェックリストを提示する。

3.1. 内的妥当性と外的妥当性

内的妥当性

意思決定者はNMAにより得られた治療間の差がおおよそ真であるとみなせるか、それともバイアスが含まれているかを評価する必要がある。NMAにおける解析の内的妥当性は、1) ネットワークを構成する試験の適切な同定、2) 各RCTの質、3) 類似性や一致性の逸脱による交絡を原因とするバイアスの程度という3つの要素を条件として成立する。

エビデンスネットワーク構成、つまりネットワークを構成するための文献収集に限界があったとしても、関連するRCTの適切な検索及び選択に努めなくてはならない[51, 52]。また、仮に厳密で広範囲な文献検索の方法であったとしても、公表バイアス (publication bias) の程度は検討しなければならない。結果がネガティブな試験又は小規模な試験は公表されにくいことが良く知られており、エビデンスネットワークはそれに依拠して限定的になる[53]。さらに、RCTのネットワークにおいて、特定の比較が他の比較よりも重みが大きい¹⁸と、不釣り合いなネットワークになってしまう[25]。NMAの妥当性は、ネットワークに含まれる各試験が内的妥当性を有することも条件であり、質の低い試験を入れることは問題となり得る。ランダム化はRCTにバイアスがないことを完全に保証しているわけではない[11, 54, 55]。治療群の盲検の欠如は、ランダム化後に試験中止する患者数の群間での不均衡の原因となり得る¹⁹。また、アウトカムの盲検の欠如は治療効果の過大評価をもたらすかもしれない[56]。このように、NMAに含まれる各RCTに対しバイアスの観点から批判的に評価されるべきである。

NMAに含まれる試験間のデザインの類似性も内的妥当性の決定的な要素となる。試験ごとに、患者特性、アウトカムの指標や定義、そして認められている併用治療、追跡期間、試験実施期間のタイムフレームに関して異なるだろう[31]。

2.3節で述べたように、NMAは間接的に比較される治療間で治療効果の修飾因子の分布に相違がある場合、交絡による影響を受ける。統計モデルに治療と共変量の交互作用項を導入し、調整することでその影響を減らすことができるのだが、試験間の類似性を判断でき、潜在的なバイアスを調整できるのは、観測されている試験レベルの共変量 (試験特性そのも

¹⁸ 例えば、極端に例数が多い試験。NMAで目的としている評価項目以外のもので検証するために大きい例数を要する試験、若しくは例数設計を行っていない試験。例数が多いことで分散が小さくなることで重みが大きくなる。

¹⁹ 一般的には被験者、試験担当医師および試験依頼者の判断、行動および評価に影響を及ぼす。

のもしくは患者特性の代表値) によってのみである。このような限界があったとしても、ベースラインリスクの差やプラセボ反応の差は、これらの共変量の試験間の差を反映したものである可能性があるので適切に評価されるべきである²⁰。

外的妥当性

NMA の外的妥当性²¹は、ネットワークに含まれる RCT の外的妥当性によって制限されるため、意思決定者は、NMA の結果が興味のある母集団に外挿できるかをレビューする必要がある。治験が外的妥当性のある程度犠牲にした選択された均一な集団となる傾向があることを、念頭に置くことは重要である [57, 58]。患者集団の一定程度の多様性は、それがリアルワールドを反映しているのであれば、効果比較評価や費用対効果評価の意思決定においては歓迎されるだろう。従って、直接比較時の治療効果のある程度の異質性は、それが試験間の治療効果の修飾因子のばらつきよりも大きいのであれば、議論のあるところではあるもののエビデンスネットワークにおける試験間の異質性は外的妥当性を向上させるかもしれない²²。明示的に外的妥当性を評価することは難しいかもしれないが、直接比較に関する分散分析で試験間分散と試験内分散の大きさを比較することで外的妥当性の程度を評価するというアプローチがある²³。

3.2. ネットワークメタアナリシスの報告

意思決定者が結果を解釈するために、NMA の報告に含めるべき項目のチェックリストを表 3-1 に示す。このチェックリスト[1]は、包括的ではないが一般的な指針となることを意図している²⁴。このチェックリストは、レポートの質の評価に焦点をあてており、NMA の内的妥当性や外的妥当性を判断及びスコア化するためのものとしては意図していないこと

²⁰ アウトカム (例: 全生存期間) と関連するようなマーカー [例: ECOG (Eastern Cooperative Oncology Group) パフォーマンスステータス] があり、そのマーカーについてアウトカムが不良となるような値 (例: 2 以上) を取る集団においては、アウトカムの治療間差が小さくなる傾向 (例: 全生存期間のハザード比が 1 に近い値を取る傾向) があるとす。メタアナリシスを実施するにおいては、個人レベルの共変量は得られず、試験レベルの共変量 (例: ECOG パフォーマンスステータスの各投与群の分布) のみが利用可能ではあるものの、マーカーの分布を適切に考慮することで、その分布の違いによる対照群もしくはプラセボ群のアウトカムへの影響や治療間差への影響に、一定程度対応可能であると考えられる。

²¹ 一般化可能性ともいう。NMA で得られた結果を、NMA の対象とした集団より広い集団において、一般性を失わず適用できることを意味する。

²² 交互作用の原因となりうる試験間の治療効果の修飾因子の分布の広がりよりも、直接比較の効果の統合における異質性が相対的に大きいのであれば、治療効果の修飾因子の影響が限定的であると解釈し、外的妥当性を有すると解釈することを意味している。

²³ つまり、内的妥当性と外的妥当性のトレードオフの関係を利用した評価方法である。

²⁴ NMA の結果がこのような視点で読まれることを踏まえると、NMA の報告者はチェックリストの項目が評価できるような情報を提供すべきと考える見方もある。NMA の論文の本文にその情報をすべて記載することは現実的ではないが、付録等に NMA の計画の概要として載せることで実現可能と考えられる。

に留意する必要がある。

表 3-1 報告された NMA を評価するときに意思決定を補助するチェックリスト²⁵

節 (Report section)	チェック項目 (Checklist Item)
導入 (Introduction)	試験実施の根拠や目的が明確に示されているか。(Are the rationale for the study and the study objectives stated clearly?)
方法 (Methods)	<p>方法の節に以下が含まれているか。(Does the methods section include the following?)</p> <ul style="list-style-type: none"> - 組入・除外基準 (Description of eligibility criteria) - 情報源 (Information sources) - 検索方針 (Search strategy) - 試験選択の手順 (Study selection process) - データ抽出 (個々の試験の妥当性や品質の評価) [Data extraction (validity/quality assessment of individual studies)] - アウトカム指標が記載されているか。(Are the outcome measures described?) <hr/> <p>解析方法や統合方法の記載があるか。また、示されている方法には以下が含まれているか。(Is there a description of methods for analysis/synthesis of evidence? Do the methods described include the following?)</p> <ul style="list-style-type: none"> - 解析方法やモデルの記載 (Description of analyses) methods/models - 潜在的なバイアスや不一致性への対応 (Handling of potential bias/inconsistency) - 解析のフレームワーク (Analysis framework) <hr/> <p>感度分析について示されているか。(Are sensitivity analyses presented?)</p>
結果 (Results)	<p>ネットワークに含まれた試験の要約が示されているか。(Do the results include a summary of the studies included in the network of evidence?)</p> <ul style="list-style-type: none"> - 個々の試験のデータ (Individual study data?) - 試験のネットワーク (Network of studies?) <hr/> <p>モデル適合度の評価について記載があるか。また、競合するモデルが比較されているか。(Does the study describe an assessment of model fit? Are competing models being compared?)</p> <hr/> <p>統合したエビデンス (間接比較や直接比較) の結果が明確に示されているか。(Are the results of the evidence synthesis (ITC/MTC) presented clearly?)</p> <hr/> <p>感度分析及びシナリオに基づいた解析 (Sensitivity/scenario analyses)</p>
考察 (Discussion)	<p>考察には以下が含まれているか。(Does the discussion include the following?)</p> <ul style="list-style-type: none"> - 主な結果について記載され、要約されているか。(Description/summary of main findings) - 解析に関する内的妥当性 (Internal validity of analysis)

²⁵ このチェックリストは Jansen ら [1] の Table 1 を簡略して和訳した。

-
- 外的妥当性 (External validity)
 - 対象者に対する結果の意味合い (Implications of results for target audience)
-

導入

本節では、興味の対象の患者集団や競合する治療を示すとともに、意思決定上のリサーチクエスションは何であるかを明確に記載すべきである。NMA で比較している治療を、特定の種類に属する全ての薬剤に限定しているかもしれないが、異なる種類の競合する薬剤や、さらには他の医学的介入を組み入れることも可能である。注目する治療が何であったとしても、選択した根拠が明確に記載されているべきである。

なお、2014 年の ISPOR のタスクフォースのレポート [2] には、下記のような記載がある。

NMA で評価しているアウトカムを示すとともに、それが意思決定者にとって意味のあるもの (例: 心疾患イベントのように患者や医療制度に直接関連するアウトカム) かどうかを示すべきである。また、アウトカムそのものだけでなく、それを評価する時期も記載すべきである。例えば、慢性疾患に対する治療を決める情報として、長い追跡期間の試験を含めた NMA は、短い追跡期間の試験に限定した NMA よりも妥当であるといえる。

また、NMA に含まれている RCT と興味の対象の背景 (状況や環境) 間に相違があるかを示すべきである。例えば、NMA に含まれる試験実施年は、疾患の医療背景が劇的に変わったときや、標準治療が変わったときに興味の対象となり得る。

方法

エビデンスネットワークの構築は、各種データベースの明確な検索ロジックや、事前に定めた試験の選択のための組入及び除外基準からなるシステマティックレビューに従って行うべきである。NMA の実施計画書には明確な検索ロジックや事前に定めた試験の選択の方法を記載することや、アウトカムに関する選択バイアス (outcome selection bias) を避けるために、解析するアウトカムを事前に規定することが望まれる [59]。試験選択後のデータ抽出の段階では厳密なデータ抽出方法を用いるべきであり、二者によるデータ抽出が行われたか、不一致があった場合はどのように解決したか、欠測データはどのように扱われたか明記すべきである。これらの方法は Centre for Reviews and Dissemination Handbook [60] などに詳しく記載されており、PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) ステートメントに従って報告すべきである [61, 62]。

コラム : PRISMA

PRISMA はシステマティックレビュー及びメタアナリシスの報告において、最低限必要とされるエビデンスに基づいた項目のことであり、PRISMA ステートメントには 27 項目のチェックリストと 4 段階のフロー図 (PRISMA ダイアグラム) が含まれている [63, 64]。これらの狙いは、システマティックレビュー及びメタアナリシスの結果を示そうとしている著者に対し、それらの報告の改善を促すことにある。

チェックリストは報告の流れに沿って、タイトル、抄録、はじめに、方法、結果、考察、資金といった観点から 27 項目が設定されている。また、NMA に関するチェックリストも提案されており、ネットワーク構造や一貫性に関する項目が追加されている [4, 5]。

統計解析の節では、アウトカムや評価項目の選択、相対的な効果の推定量、固定効果モデルや変量効果モデルの選択の正当性を含め、使用した統計解析方法の概要を提示すべきである。また、直接比較の結果と間接比較の結果を比較することで一貫性の仮定について考察し、類似性の向上や不一致性を減らすために、試験レベルの共変量を入れたモデルに拡張するかどうかを明示すべきである。ベイズ流の方法で解析を行う場合、パラメータの事前分布の選択及び事前分布に関する感度分析について記載すべきである。

結果

システマティックレビューで同定した試験のリストと NMA の解析対象とした試験のリストを提示すべきである。実際の解析に利用するにあたってデータが不適切 (例: 推定値が欠損) である試験があるときには、これらの 2 つのリストには相違が生じるだろう。試験の選択方法を説明するフロー図²⁶は有用である。システマティックレビューの結果の報告の方法は、PRISMA ステートメントを参照することになるだろう [61]。各試験の主な患者特性及び試験特性に関するリストは表形式で示すべきである。治療効果の修飾因子となり得る試験間の相違があるかどうかを判断することは必須である。例えば、患者の年齢、有病期間、既往歴の違いは治療効果の修飾因子となることがある。また、試験実施地域が、報告されている患者特性では表すことのできない患者集団間の違いを反映しているかもしれない²⁷。

なお、2014 年の ISPOR のタスクフォースのレポート [2] には下記の記載がある。

試験間で未知又は測定されていない治療効果の修飾因子の分布に不均衡が生じるリスクが常にあることを認識しておくことが重要である。従って、観察されているすべての治療効果の修飾因子の分布が試験間で均衡の取れている場合でも交絡バイア

²⁶ 所謂、PRISMA ダイアグラム。

²⁷ 地域そのものが NMA の解析対象の評価項目に関する試験間の異質性の直接的な要因となることは通常考えにくく、他の特性が影響していると考えるのが妥当である。つまり地域が見かけ上の交絡因子として働いていることになる。しかし、そのような特性を全試験で収集しているとは限らないため、代替的に地域という特性で表現することを意図している。

スのリスクが常に存在する。

どの患者特性或いは試験特性が治療効果の修飾因子であるかを各試験の結果を得た後に決定することで、不一致性や間接比較におけるバイアスの原因について誤った結論を下してしまう可能性がある [65, 66]。従って、事前に得られている知見や各試験で報告されたサブグループ解析の結果をもとに、各試験間の結果を比較する前に治療効果の修飾因子として可能性のあるものを一覧として示し、次に一覧にした試験特性及び患者特性を試験間で比較し、試験間で特性の分布に不均衡がないか検討する必要がある。

エビデンスネットワークの図示は有用で、解析の透明性を向上させるだろう。

個々の試験で得られた点推定値とそれに対応する不確実性の指標は、表形式で各試験の各治療群に対して示すべきである。これはネットワークの理解を促し、試験間で共通の対照治療（若しくはプラセボ）のアウトカムを比較することで、試験間の相違を評価するのに役立つ。RCTの相対的な効果をフォレストプロットのような図や表で提示することも、NMAの統合結果と各試験の結果間の比較することができる点で有用である。

NMAの結果が示される節では、複数の統計モデルを適用した場合には、モデルを選択した根拠を示すべきである。最低限の記述として、共通の対照治療又はアンカーに対する相対的な効果（例えば、オッズ比、ハザード比、平均値の差）の推定値や95%信頼区間或いは信用区間を記述すべきである。NMAの報告の質を高めるためには、全ての対比較の結果を示すことを推奨する。

モデル解析で用いたものとは異なる尺度の相対的な治療効果の推定値を報告することは、意思決定者にとって有益な場合もある。オッズ比でNMAの結果を報告しているところを、相対リスク、リスク差、NNTで置き換えてしまうのが一例である。これらの推定値は、対照治療における推定された反応確率から得ることができる。ベイズ流の方法の統計解析ソフトウェアであるWinBUGSを用いた解析では、異なるスケールの相対的な治療効果の推定値の導出が容易である [67]。

ベイズ流の方法で解析する場合、相対的な効果の推定値の不確実性（例：95%信用区間の幅）は、推定値に基づいてなされる決定の不確実性の程度と言い換えることができる²⁸。例えば、共通のアンカーに対する各治療のオッズ比とその95%信用区間は、各治療が他の全ての治療よりも有効である確率を算出することを可能とする²⁹。

²⁸ 仮に信用区間で推定値の不確実性を測るのであれば、信用区間の幅が広い場合、たとえ分布の中央値等の代表値においてより良い推定値を与えているとしても、その代表値の確実性は低くなる。

²⁹ ベイズ流の推測がシミュレーションベース（Markov chain Monte Carlo methods）：

ここまで、ベイズ流の方法の使用による重要な利点を示したが、各治療に対する順位の“広がり”の情報は重要であり、1位の確率のみを提示しているときは注意すべきである。例えば、試験数が少なく、結果的に広い信用区間になる治療は、最善及び最悪となってしまう確率がそれぞれ50%になってしまうこともある。従って、全ての治療に対して効果の順位の期待値を算出することが有用である。

相対的な治療効果の推定値（オッズ比、相対リスクなど）に加え、二値応答のアウトカムに関する絶対的な確率（オッズ、リスクなど）を示すのも有用であろう。この確率の算出にはアンカーとした治療の確率が必要である。推定には臨床試験データか他のソースに基づくことになるが、推定方法は明確に記述すべきであり、感度分析は必要だろう。

コラム：NMAにおける指標としての順位及びSUCRA

ベイズ流の方法論を用いたNMAでは、興味のアウトカムに関し各治療が最善の治療となる確率を推定することができ（下表）、さらにはその確率の高さに順位を付けることができる。順位を付ける方法としてsurface under the cumulative ranking (SUCRA)が知られている[68]。この方法は、横軸をベイズ流の推定でのシミュレーション1回ごとの順位、縦軸はシミュレーション回数に対する横軸の順位及びそれよりも良い順位となる割合を表している。SUCRAはこれらに基づいて作図された曲線の下側面積で定義され、この値の大きい順に順位が付与される（下図ではSUCRAの値が最大であるBが1位となる。）。曲線の下側面積は、順位の累積確率の平均値若しくは曲面下面積の割合として解釈できる。順位は非常にわかりやすい指標であり、NMAの結果として望まれることがある。しかし、単純な指標なだけにその解釈には注意を要することも理解したうえで、SUCRAの値と併記するなどして評価を行うべきである。

表 仮想例：各薬剤の各順位となる確率（括弧内は累積確率）とSUCRA

Rank <i>b</i>	Treatment <i>j</i>			
	A	B	C	D
1	0.25 (0.25)	0.50 (0.50)	0.25 (0.25)	0 (0)
2	0.25 (0.50)	0.25 (0.75)	0.50 (0.75)	0 (0)
3	0.25 (0.75)	0.25 (1)	0.25 (1)	0.25 (0.25)
4	0.25 (1)	0 (1)	0 (1)	0.75 (1)
SUCRA	50%	75%	67%	8%

MCMC) で実施されることがあることに由来する。

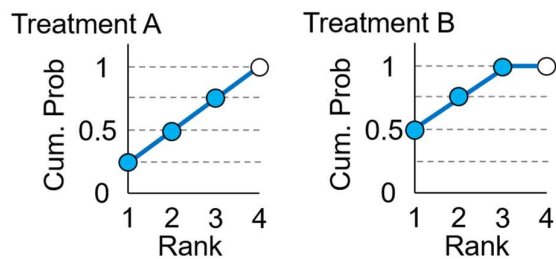


図 仮想例：順位の累積確率

考察

内的妥当性及び外的妥当性の観点から結果の批判的な評価を示すべきである。NMA 論文の著者らは、類似性と一致性を仮定することができるかどうかの十分な検討をすべきである。この検討をする節では、当該 NMA の結果がメタアナリシス（対比較のメタアナリシス及び NMA）や他のエビデンス（観察研究など）に基づく結果に沿ったもの（一致したもの）か説明すべきである [69]。さらに、生物学的、臨床的な観点から、比較した治療間の観察された相違を説明することを推奨する。結果の適切性とは別に、実臨床での意思決定に関する結論の妥当性についても説明すべきである。

なお、2014 年の ISPOR のタスクフォースのレポート [2] には、利益相反について下記の記載がある。

NMA 論文の著者（又は著者の所属する組織や雇用主）が、著者の意思決定、研究又は当該論文にバイアスを生じ得る金銭的又は個人的な利害関係をもつ場合、利益相反の可能性がある。

この問題に対処するためには、すべての利益相反が明示され、投稿した論文は査読を受けている必要がある。また、各著者の役割を開示するために、貢献した内容が明確に記載されている必要がある。

3.3. 2014 年の ISPOR タスクフォースによる質問項目

2014 年の ISPOR のタスクフォースのレポート [2] では、間接比較および NMA の妥当性と信頼性を評価するための質問を提供している。妥当性に関する質問は、意思決定者の興味の対象の設定で NMA が行われているか判断するために設けられている。信頼性に関する質問は、NMA の結果が意思決定者に対して根拠のある回答を与えるかを判断するために、設けられている。

- 妥当性

- 母集団は妥当か（意思決定者にとって興味の対象の母集団であるか）

- NMA に含まれるべき介入が評価されているか
- NMA に含まれるべきアウトカムが評価されているか
- NMA に含めた試験の設定や状況は受け入れ可能か

- **信頼性**

- **間接比較及びNMA に用いられたエビデンス**

- ◇ NMA の実施者は関連する全ての RCT を同定し含めようとしたか
- ◇ 興味の治療に対する試験が連結した RCT のネットワークを構成しているか
- ◇ 質の低い試験が含まれているのは明らかか。それによりバイアスをもたらしているのは明らかか
- ◇ アウトカムの選択的な報告によってバイアスは含まれていそうか
- ◇ ネットワークの異なる治療比較の間に治療効果の修飾因子(治療効果に影響のある患者のベースライン特性や試験の特性)の分布に系統的な違いがあるか。もしあれば、それについて個々の試験の結果を比較する前に検討されているか。

- **解析**

- ◇ 試験内のランダム化を保つ統計的方法は用いられたか (naive comparisons ではない)
- ◇ 興味の対象の対比較において直接比較と間接比較の両方が可能な場合 (閉じたループである場合), 治療効果の一致性は評価又は議論されたか
- ◇ 一致性がいえるなら, 直接比較および間接比較のエビデンスを共に NMA に含めたか
- ◇ 一致性に抵触する又は異なる比較間で治療効果の修飾因子の分布に不均衡があるとき, 実施者は解析でこのバイアスを最小にしようとしたか
- ◇ 変量効果又は固定効果モデルの使用に対して妥当な論拠はあるか
- ◇ 変量効果モデルが用いられた場合, 異質性についての仮定は検討又は議論されたか
- ◇ 異質性の可能性がある場合, 部分集団解析又は予め特定された共変量でのメタ回帰が実施されたか

- **報告の質と透明性**

- ◇ エビデンスネットワークの視覚的な提示又は表形式での提示においては, 直接比較ごとに RCT の数に関する情報を与えられたか
- ◇ 個々の試験結果は報告されているか
- ◇ 直接比較の結果は, 間接比較又は NMA の結果と別々に報告されているか

- ◇ NMA の対象となった全ての治療の組み合わせの結果は不確実性の指標とともに報告されているか
- ◇ 治療の順位はアウトカムによる報告された治療効果とその不確実性を提供しているか
- ◇ 治療効果において重要な患者特性の効果は報告されているか

- 解釈
 - ◇ 結論は公平でバランスが取れているか

- 利益相反
 - ◇ 利益相反はあったか。あるなら、対策を講じたか

4. 解釈と意思決定

本節では NMA の結果の解釈及び意思決定を行う上で考慮すべきポイントを述べる。

4.1. 解釈

第一に、NMA に用いる試験数が少ない場合、試験間の類似性が成り立たないことで生じるバイアスを、メタ回帰などの統計解析手法で調整するには限界があることはよくある。試験間の類似性が成り立っていないことを理由に NMA の結果を直ちに無視するのではなく、間接比較におけるバイアスの方向について仮説を立て、バイアスのない間接比較とどれだけ異なり得るのか、バイアスにより誤った結論や意思決定を導き出してしまうのかどうかを検討すべきである。

第二に、有効性の評価項目のみに基づいて“最良”な治療を特定することはできない。NMA での有効性の結果は観察研究の結果や安全性及び利便性といった有効性以外の特徴も踏まえて解釈しなければならない。総合的な解釈を得るため複数の評価項目に対する NMA の結果をどのようにまとめるか検討する必要がある。1つのアプローチとして、相対的な重要性に応じて各評価項目に重みを付け、各治療が“最良”となる確率を計算する手法が提案されている [70]。

4.2. 意思決定

2.3 節で述べたように、NMA では RCT に基づくエビデンスを必要とするが、意思決定をする状況のうち多くの場合 RCT によるエビデンスを利用できないかもしれない。意思決定のためのエビデンスとしての RCT、RCT に基づく NMA、観察研究（例：コホート研究）の適切な利用に関する論争は今後も続きそうである。直接比較の RCT によるエビデンスが欠如した状況下では、「より質の低いエビデンスを用いることによる潜在的なリスクを把握して」、それと同時に「そのリスクを取ることが可能である場合」に限り、観察研究のエビデンスを使用できると考える。直接比較の RCT や RCT に基づく間接比較のエビデンスが創出されるのを待つと判断した場合、意思決定者は“新しい”治療よりも“古い”治療を選ぶであろうことを、その潜在的な社会的影響（societal implication）とともに認識しなければならない³⁰。論争の結論が何であっても、どんなタイプのエビデンスが意思決定に利用されたのかを明確にしてかつ透明性を担保すること、そして意思決定の限界と影響を評価することで意思決定の質は向上するだろう。

コラム：諸刃の刃 NMA

NMA は強力なツールであるが、適用には満たすべき仮定があることはこれまでも述べ

³⁰ 古い治療を選択することは、既存のエビデンスをもとに治療を決定する保守的なアプローチを意味すると考えられる。潜在的な社会的影響とは、新しいエビデンスが出てくるまでは既存の治療の中で最良のものを決定することで、真に有用性の高い新しい治療が用いられないことを意味すると考えられる。

てきた。どのような解析であっても、データを入力し、解析プログラムを実行すれば何かしらの結果が得られる。それは NMA でも同じである。そして、その結果が自分にとって望ましい結果であれば、そのまま利用したくなるかもしれない。しかし、NMA から得られた結果が批判に耐えうるためには、慎重な検討が必要である。

Benjamin らは、NMA を実施するには対比較のメタアナリシスよりも時間が掛かることから、手を付ける前にどれくらいのリソースが必要となるか見積もることを推奨している [71]。

また、試験計画、解析計画、モニタリングの方法などを含む臨床試験の方法論は、時代と共に発展をしていて、類似性と一致性からの逸脱に影響を与えるような要素として考慮せねばならないだろう。昨今、National Research Council は欠測データに対する解析方法に強い影響を与えたと考えられており [72]、ICH E9(R1) step 2 により “estimand” の考えが臨床試験において導入されたことから [73]、NMA の対象となる臨床試験でこれまでとは異なる解析が実施されるようになっているかもしれない。また、中間解析やアダプティブデザインによる試験結果をどのように統合すれば良いかについては未だ研究領域でもある [74, 75]。

ここで繰り返して強調したいことは、NMA は強力なツールであるが、適用には満たすべき仮定が多くあり、NMA を実施するにはリソースが必要であるということ。そして、安易に NMA を利用し、その結果を誤用し、批判に耐えられないということを守るためにも、慎重な対応が必要であるということである。

5. まとめ

本報告書をまとめる。

ネットワークメタアナリシスの結果の妥当性を担保するために満たすべき仮定

NMA の有用性の反面、評価する治療の数が多いことや方法論の複雑さからその特性を理解するのに困難を伴うことに加え、結果の妥当性を担保するために満たすべき仮定がある。2.3 節で紹介した類似性と一致性である。

メタアナリシス全般にいえることだが、各試験における患者の組み入れ基準やアウトカムといった試験実施計画が完全に同一ということではなく、検証不可能な仮定に基づく結果しか得ることができない。それ故、対象を広く取ることができる NMA では、試験間の類似性の逸脱の可能性がより大きくなる。また、一致性は間接比較を有する NMA に特徴的な性質である。

これらの仮定が十分に満たされないと考えられる場合には、NMA の適用結果にバイアスが生じている可能性があることに注意を払って結果を解釈する必要がある。これらの仮定の逸脱について、その影響の程度を感度分析やメタ回帰で定量的に評価し考察することは、NMA の質を向上させる助けになるであろう。たとえ感度分析の結果、NMA の主要解析の結果に影響を与えることになったとしても、その結果を真摯に吟味すべきである。

企業における既存のネットワークメタアナリシスの結果の活用と新規実施

NMA を製薬企業の観点からみると、治療選択をサポートするための情報提供、医療経済性評価、自社製品のポジショニング戦略などで NMA の活用が期待される。これは、自社で有しているデータは、主に自社の新薬の承認取得のための臨床試験のデータに限られ、プラセボ若しくは標準治療と比較した試験のみである場合が多いという背景に基づいている。NMA により有効性及び安全性に関して同種同効薬剤に対する位置づけを総合的かつ定量的に考察できる。

製薬企業に限ったことではないが NMA によるエビデンスを得るには 2 つ方法がある。ひとつは既存の NMA に基づく公表論文などを参照しその結果を活用する方法、もうひとつは自ら NMA を用いた研究を計画し結果を得る方法である。

NMA の公表論文を参照する場合には、一般的にいわれている選択バイアスと同様なことが生じうる。つまり、自分たちに都合のいい結果が載っている論文のみを信じたくなくなってしまうことを意味する。これは、参照した論文の結果のみを追い求めていることから陥ってしまうのであるが、たとえいい結果であったとしても、それが自分の目的に沿ったものであるか、そして適切に実施されているかを客観的に吟味せねばならない。このような意味で、NMA の結果を引用する者にリテラシーが問われる。客観的な吟味の助けとなるものとしての NMA の妥当性や信憑性の評価のためのチェックリストは既に示している (表 3-1)。

一方、自ら NMA を実施する場合には、基となる論文の公表バイアスが懸念されることはよく知られているが、NMA の実施方法の妥当性も重要な問題である。この問題に対応するために、前述のチェックリスト (表 3-1) や PRISMA の extension statement [4, 5] を実施者として活用する方法もある。計画している若しくは実施した計画の公表は、妥当性を示す直接的なアプローチではないが、透明性を確保する意味で良い (公表の例として [76] の supplementary appendix)。特に製薬企業が自ら実施する際にはより透明性が求められることを考慮するとこのアプローチは重要である [76]。

最後に

NMA は正しく適用することができれば、直接比較の RCT を実施しなくても興味の治療の全体の中のポジショニングを知るという究極的とも言える要求を満たし得る。前述した

ように NMA は対比較のメタアナリシスと比べて結果の妥当性を担保するためにより多くの仮定を要することから、その実施や結果の解釈には十分な注意を要する。たとえ目下の NMA の結果が自分たちにとってポジティブな結果であったとしても、立ち止まってその結果を批判的に見るべきであるし、臨床担当者や統計担当者の専門知識を総動員しなければその正しい解釈はできないだろう。

参考文献

1. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri JC. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* 2011; 14(4): 417-28
2. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect Treatment Comparison/Network Meta-Analysis Study Questionnaire to Assess Relevance and Credibility to Inform Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report. *Value Health* 2014; 17(2): 157-73.
3. 国際医薬経済・アウトカム研究学会（ISPOR）日本部会 <<https://ispor-jp.org/index.html>> 2019年10月17日アクセス
4. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Annals of Internal Medicine* 2015; 162: 777-84.
5. 大西良佳, 渡辺範雄. PRISMA Extension for Network Meta-analysis (PROSMA-NMA) : メタアナリシスおよびシステマティックレビューの手法の展開に対応するためのツール (PRISMA 拡張版). *薬理と治療* ; 46 (1), 2018.
6. Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331: 897-900.
7. Ioannidis JPA. Indirect comparisons: the mesh and mess of clinical trials. *Lancet* 2006; 368: 1470-2.
8. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008; 26: 753-67.
9. Wells GA, Sultan SA, Chen L, et al. Indirect Evidence: Indirect Treatment Comparisons in Met-Analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.
10. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50: 683-91.

11. Song F, Altman DG, Glenny A, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003; 326: 472.
12. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23: 3105–24.
13. Song F, Loke YK, Walsh T, et al. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.
14. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002; 21: 2313–24.
15. Working Group on Relative Effectiveness The Pharmaceutical Forum. <http://ec.europa.eu/pharmaforum/effectiveness_en.htm> 2019年10月17日アクセス.
16. Jansen JP, Bergman GJ, Huels J, Olson M. Prevention of vertebral fractures in osteoporosis: mixed treatment comparison of bisphosphonate therapies. *Curr Med Res Opin* 2009; 25: 1861–8.
17. Mills EJ, Perri D, Cooper C, et al. Antifungal treatment for invasive *Candida* infections: a mixed treatment comparison meta-analysis. *Ann Clin Microbiol Antimicrob* 2009; 8: 23.
18. Stettler C, Wandel S, Allemann S, et al. Outcomes associated with drug eluting and bare-metal stents: a collaborative network meta-analysis. *Lancet* 2007; 370: 937–48.
19. Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Stat Med* 2003; 22: 2995–3016.
20. Cooper NJ, Sutton AJ, Morris D, et al. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: application to stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Stat Med* 2009; 28: 1861–81.
21. Psaty BM, Lumley T, Furberg CD, et al. Outcomes associated with various antihypertensive therapies used as first-line agents: a network meta-analysis. *JAMA*; 289: 2534–44.
22. Cooper NJ, Sutton AJ, Lu G. Mixed comparison of stroke prevention treatments in

- individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006; 166: 1269–75.
23. Vissers D, Stam W, Nolte T, et al. Efficacy of intranasal fentanyl spray versus other opioids for breakthrough pain in cancer. *Curr Med Res Opin* 2010; 26: 1037–45.
 24. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple treatments meta-analysis. *Lancet* 2009; 373: 746–58.
 25. Salanti G, Kavvoura FK, Ioannidis JPA. Exploring the geometry of treatment networks. *Ann Intern Med* 2008; 148: 544–53.
 26. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health* 2008; 11: 956–64.
 27. Coory M, Jordan S. Frequency of treatment-effect modification affecting indirect comparisons: a systematic review. *Pharmacoeconomics* 2010; 28: 723–32.
 28. Sedgwick P. External and internal validity in clinical trials. *BMJ* 2012; 344: e1004.
 29. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006; 101: 447–59.
 30. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; 29: 932–44.
 31. Glenny AM. Statistical methods for indirect treatment comparisons. *Health Technol Assess* 2005; 9: 17–49.
 32. Borenstein M, Hedges LV, Higgins JPT, Rothstein H. *Introduction to Meta-Analysis*. Chichester, England: John Wiley & Sons, Ltd., 2009.
 33. Skene AM, Wakefield JC. Hierarchical models for multicentre binary response studies. *Stat Med* 1990; 9: 919–29.
 34. Gelman AB, Carlin JS, Stern HS, Rubin DB. *Bayesian Data Analysis*. Boca Raton, FL: Chapman and Hall–CRC; 1995.
 35. Cappelleri JC, Ioannidis JPA, Lau, J. Meta-analysis of therapeutic trials. In: Chow S-C, ed., *Encyclopedia of Biopharmaceutical Statistics* (3rd ed.), Revised and Expanded. New York, NY: Informa Healthcare, 2010.

36. Berlin JA, Santanna J, Schmid CH, et al. Individual patient-versus group-level data meta-regression for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002; 21: 371–87.
37. Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. *Int J Epidemiol* 1989; 18: 269–74.
38. Lambert PC, Sutton AJ, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol* 2002; 55: 86–94.
39. Thompson SG, Sharp SJ. Explaining heterogeneity in meta - analysis: a comparison of methods. *Stat Med* 1999; 18(20): 2693–708.
40. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Cochrane, 2019. <www.training.cochrane.org/handbook> 2019 年 10 月 17 日アクセス
41. McCullagh P, Nelder J. *Generalized Linear Models*, Second Edition. Chapman & Hall/CRC, 1989.
42. Dempster AP. The direct use of likelihood for significance testing. *Stat Comput* 1997; 7: 247–52.
43. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc (Series B)* 2002; 64: 583–639.
44. Sutton AJ, Abrams KR, Jones DR, et al. *Methods for Meta-Analysis in Medical Research*. London, UK: Wiley, 2000.
45. Spiegelhalter D, Abrams K, Myles J. *Bayesian Approaches to Clinical Trials and Health Care Evaluation*. Chichester, UK: John Wiley & Sons, 2004.
46. Luce BR, Claxton K. Redefining the analytical approach to pharmacoeconomics. *Health Econ* 1999; 8: 187–9.
47. Goodman, SN. Towards evidence based medical statistics: 1. The P value fallacy. *Ann Intern Med* 1999; 120: 995–1004.
48. Jansen JP. Network meta-analysis of individual and aggregate level data. *Res Synth Meth* 2012; 3: 177–90.

49. Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. *Stat Med* 2012; 31: 3516–36.
50. Stroup DF, Berlin JA, Morton SC, et al, Meta-analysis of Observational Studies in Epidemiology. A Proposal for Reporting. *JAMA*. 2000; 283(15): 2008-12. doi: 10.1001/jama.283.15.2008.
51. Hawkins N, Scott DA, Woods BS, Thather N. No study left behind: a network meta-analysis in non-small-cell lung cancer demonstrating the importance of considering all relevant data. *Value Health* 2009; 12: 996–1003.
52. Hawkins N, Scott DA, Wood BS. How far do you go? Efficient searching for indirect evidence. *Med Decis Making* 2009; 29: 273–81.
53. Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17: 279–301.
54. Cappelleri JC, Ioannidis JP, Schmid CH, et al. Large trials vs meta-analysis of smaller trials: how do their results compare? *JAMA* 1996; 276: 1332–8.
55. Abel U, Koch A. The role of randomization in clinical studies: myths and beliefs. *J Clin Epidemiol* 1999; 52: 487–97.
56. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273: 408–12.
57. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967; 20: 637–48.
58. Rothwell PM. Factors that can affect the external validity of randomized controlled trials. *PLOS Clin Trials* 2006; 1: e9.
59. Williamson PR, Outcome selection bias in meta-analysis *Stat Methods Med Res* 2005; 14: 515–24.
60. Centre for Reviews and Dissemination. Systematic Reviews. <<https://www.york.ac.uk/crd/guidance/>> 2019年10月17日アクセス
61. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA Statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009; 151: W65–94.

62. Moher D, Liberati A, Tetzlaff J, Altman DG, and the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Ann Intern Med* 2009; 151: 264–9.
63. PRISMA <<http://www.prisma-statement.org/>> 2019 年 10 月 17 日アクセス
64. Moher D, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7): e1000097. doi:10.1371/journal.pmed.1000097.
65. Jansen JP, Schmid CH, Salanti G. Direct acyclic graphs can help understand bias in indirect and mixed treatment comparisons. *J Clin Epidemiol* 2012; 65: 798–807.
66. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002; 21: 1559–73.
67. Spiegelhalter D, Thomas A, Best N, Lunn D. WinBUGS User Manual: Version 1.4. MRC Biostatistics Unit: Cambridge, 2003.
68. Salanti G, Ades AE, Ioannides JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011; 64: 163–71.
69. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004; 328: 1490.
70. Jansen JP, Bergman GJD, Huels J, Olson M. The efficacy of bisphosphonates in the prevention of vertebral, hip, and nonvertebral non-hip fractures in osteoporosis: a network meta-analysis. *Semin Arthritis Rheum* 2011; 40: 275–84.
71. Benjamin R et al. Network Meta-Analysis: An Introduction for Clinicians. *Intern Emerg Med*. 2017; 12(1): 103–111
72. National Research Council. The prevention and treatment of missing data in clinical trials. 2010. National Academy Press.
73. ICH HARMONISED GUIDELINE: Estimands and Sensitivity Analysis in Clinical Trials E9(R1) Current Step 2 version dated 16 June 2017.
74. FDA draft guidance: Adaptive Designs for Clinical Trials of Drugs and Biologics Guidance for Industry 2018.
75. REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY

CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN. 2007.

76. Leucht S, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*, 2013. 382, 951-62.

執筆者

日本製薬工業協会 医薬品評価委員会 データサイエンス部会

- 浅川 誉 中外製薬株式会社
- 飯森 孝行 元杏林製薬株式会社
- 岩間 康弘 Meiji Seika ファルマ株式会社
- 兼清 道雄 エーザイ株式会社
- 阪本 亘 日本ベーリンガーインゲルハイム株式会社
- 菅波 秀規 興和株式会社 *
- 野口 英久 ヤンセン ファーマ株式会社
- 藤井 陽介 ファイザーR&D 合同会社
- 眞野 博貴 帝人ファーマ株式会社

* はネットワークメタアナリシスタスクフォース推進委員，それ以外はタスクフォースメンバー。