

Minds

診療ガイドライン 作成マニュアル

2020 ver. 3.0

編集:Minds診療ガイドライン作成マニュアル編集委員会

発行:公益財団法人 日本医療機能評価機構

お願いとご注意

本マニュアルのタイトル・バージョン管理について

タイトルに含まれる「年」は、「改訂を行った年度」を示しています。また、初版からの通算でバージョンを付記し、原則、小改訂時は小数点、大改訂時は整数を更新します。

本マニュアルは継続的に改訂・修正を行い、最新版を Minds ガイドラインライブラリで公開します。ご利用の際には最新のバージョンであることをご確認ください。

Minds ガイドラインライブラリ
<https://minds.jcqhc.or.jp/>



本マニュアルの引用時の記載例

Minds 診療ガイドライン作成マニュアル編集委員会. Minds 診療ガイドライン作成マニュアル 2020 ver. 3.0. 公益財団法人日本医療機能評価機構 EBM 医療情報部. 2021.

Minds Manual Developing Committee ed. Minds Manual for Guideline Development 2020 ver. 3.0. Tokyo: Japan Council for Quality Health Care, 2021.

(バージョン・発行年は参照するバージョンの情報を反映してください。)

本マニュアルをご活用いただいている皆さまから、内容に関する幅広いコメントを募集しています。コメントは Minds 事務局へメールでお送りください。お寄せいただいたコメントは Minds 診療ガイドライン作成マニュアル編集委員会で協議のうえ、改訂の際に参考にさせていただきます。ただし、全てのコメントの内容を反映できかねることをご了承ください。

連絡先

公益財団法人 日本医療機能評価機構 EBM 医療情報部（Minds 事務局）

minds_gdg_support@jcqhc.or.jp

〒 101-0061 東京都千代田区神田三崎町 1-4-17 東洋ビル

Minds 診療ガイドライン作成マニュアル 2020 ver. 3.0

2021 年 3 月 22 日発行

編集：Minds 診療ガイドライン作成マニュアル編集委員会

発行：公益財団法人 日本医療機能評価機構

© 公益財団法人 日本医療機能評価機構 EBM 医療情報部（Minds 事務局）

無断転載を禁じます。

Minds 診療ガイドライン作成マニュアル編集委員会（五十音順）

委員長

森實 敏夫 日本医療機能評価機構 客員研究主幹
大船中央病院消化器肝臓病センター 非常勤医師

委 員

奥村 晃子	日本医療機能評価機構 EBM 医療情報部 部長
佐藤 康仁	日本医療機能評価機構 客員研究員 東京女子医科大学医学部衛生学公衆衛生学講座 講師
馬場 俊明	日本医療機能評価機構 客員研究員 国立国際医療研究センター国際医療協力局 医師
福岡 敏雄	日本医療機能評価機構 執行理事 (EBM・診療ガイドライン担当) 大原記念倉敷中央医療機構倉敷中央病院 副院長 兼 総合診療科 主任部長 兼 救命救急センター センター長
吉田 雅博	日本医療機能評価機構 客員研究主幹 国際医療福祉大学消化器外科学教室 教授 国際医療福祉大学市川病院 人工透析センター長・一般外科部長

マニュアル作成ワーキンググループ（執筆・作成協力など）（五十音順）

池田 俊也	国際医療福祉大学医学部公衆衛生学 教授
浦久保安輝子	日本医療機能評価機構 EBM 医療情報部 部員
奥村 晃子	日本医療機能評価機構 EBM 医療情報部 部長
河合 富士美	聖路加国際大学学術情報センター マネジャー
清原 康介	大妻女子大学家政学部食物学科公衆衛生学研究室 専任講師
小島原 典子	静岡県立総合病院社会健康医学研究部疫学研究班 上席研究員
後藤 温	公立大学法人横浜市立大学医学群 / データサイエンス研究科ヘルスデータサイエンス専攻 教授
佐藤 康仁	日本医療機能評価機構 客員研究員 東京女子医科大学医学部衛生学公衆衛生学講座 講師
志真 京子	日本医療機能評価機構 EBM 医療情報部 主任
白岩 健	国立保健医療科学院 保健医療経済評価センター 主任研究員
中山 健夫	京都大学大学院医学研究科社会健康医学系専攻健康情報学分野 教授
畠山 洋輔	東邦大学医学部社会医学講座公衆衛生学分野 助教
馬場 俊明	日本医療機能評価機構 客員研究員 国立国際医療研究センター国際医療協力局 医師
福岡 敏雄	日本医療機能評価機構 執行理事 (EBM・診療ガイドライン担当) 大原記念倉敷中央医療機構倉敷中央病院 副院長 兼 総合診療科 主任部長 兼 救命救急センター センター長
森實 敏夫	日本医療機能評価機構 客員研究主幹 大船中央病院消化器肝臓病センター 非常勤医師
吉田 雅博	日本医療機能評価機構 客員研究主幹 国際医療福祉大学消化器外科学教室 教授 国際医療福祉大学市川病院 人工透析センター長・一般外科部長

査読協力者（五十音順）

相原 守夫	相原内科医院 院長
石川 英二	済生会松阪総合病院 内科・腎臓センター長
大田 えりか	聖路加国際大学大学院看護学研究科国際看護学 教授
蓋 若琰	国立社会保障・人口問題研究所社会保障応用分析研究部 第4室長
後藤 励	慶應義塾大学大学院経営管理研究科 准教授
佐伯 晴子	一般社団法人マイインフォームド・コンセント 理事長
古賀 弘志	信州大学医学部皮膚科学教室 講師
豊島 義博	鶴見大学歯学部探索歯学講座
南学 正臣	東京大学大学院医学系研究科腎臓・内分泌内科 教授
南郷 栄秀	独立行政法人地域医療機能推進機構（JCHO）東京城東病院総合診療科
西田 俊彦	株式会社グローバルヘルスコンサルティング・ジャパンコンサルティング部門 コンサルタント
能登 洋	聖路加国際病院内分泌代謝科 部長
福井 次矢	聖路加国際病院 院長
日々澤 肇	東京都医師会 理事 医療法人社団茜遙会日々澤醫院 院長
森 臨太郎	国連人口基金アジア太平洋地域事務所 地域アドバイザー
山口 直人	社会福祉法人恩賜財団済生会済生会保健・医療・福祉総合研究所 研究部門長
山中 寿	医療法人財団順和会山王メディカルセンター 院長
湯浅 秀道	独立行政法人国立病院機構豊橋医療センター歯科口腔外科
渡辺 範雄	コクランジャパン／医療法人社団蘇生会蘇生会総合病院精神科 部長

編集事務局

日本医療機能評価機構 EBM 医療情報部（Minds 事務局）

はじめに

公益財団法人日本医療機能評価機構は、2002年度からEBM普及推進事業（Minds）を開始した。Mindsは2011年度より厚生労働省委託事業として運営されており、本年度で10年目の節目を迎えた。本事業は、質の高い診療ガイドラインの普及を通じて、患者と医療者の意思決定を支援し、医療の質の向上を図ることを目的としている。Mindsは、①診療ガイドライン作成支援、②診療ガイドライン評価選定・公開、③診療ガイドライン活用促進、④患者・市民支援を事業の4つの柱とし、診療ガイドラインデータベース「Mindsガイドラインライブラリ」（<https://minds.jcqhc.or.jp/>）を通じて、誰もが無料で診療ガイドラインや一般向けの解説を閲覧できる環境を用意している。

診療ガイドライン作成マニュアルの作成・改訂は、診療ガイドライン作成支援の最重要課題として継続的に取り組んできた。Mindsでは、診療ガイドライン作成主体に対して多様な作成支援の取り組みを行っているが、全てこの作成マニュアルをベースに実施している。2007年に『Minds 診療ガイドライン作成の手引き 2007』を刊行した後、2011年にInstitute of Medicine (IOM) [現 National Academy of Medicine (NAM)] が新たな診療ガイドラインの定義を発表し、世界的にも診療ガイドライン作成方法が大きく変化した。Mindsは、GRADE (The Grading of Recommendations Assessment, Development and Evaluation) アプローチをはじめ、国際標準の診療ガイドライン作成方法を踏まえた上で、日本の医療に適した診療ガイドライン作成方法を『Minds 診療ガイドライン作成マニュアル 2014』としてまとめ、その要約版である『Minds 診療ガイドライン作成の手引き 2014』を公開・刊行した。続く2017年にはマニュアルの改訂版『Minds 診療ガイドライン作成マニュアル 2017』を公開した。このたび公開する『Minds 診療ガイドライン作成マニュアル 2020 ver. 3.0』の改訂の要点は以下のとおりである。

「1章 診療ガイドライン総論」においてMindsの新しい診療ガイドライン定義を示した。この背景としては、診療ガイドラインが取り扱う内容と診療ガイドラインを取り巻く環境の多様化を鑑み、より実際にかなう新しい定義を提案するに至ったことによる。

「2章 準備」においては、利益相反（conflict of interest : COI）の開示・管理についてより具体的な説明を加え、患者・市民参画の項目を設け、診療ガイドライン作成への患者・市民参画の意義や具体的な方法を説明した。

さらに「5章 医療経済評価」を新たに導入し、診療ガイドライン作成において医療経済評価を組み入れる場合の基本的な考え方や方法について記載した。他章・他項目についても必要なバージョンアップを行った。

本マニュアルの改訂にあたっては、診療ガイドライン作成過程と同様に手順の透明化を重視した。まず、作成支援を担当する専門部会である診療ガイドライン作成支援部会の下、Minds 診療ガイドライン作成マニュアル編集委員会およびマニュアル作成ワーキンググループを編成した。編集委員会を中心に、改訂の企画立案、担当者決定、内容の精査を図り、外部有識者による外部査読とパブリックコメント募集を行い、それらの内容を踏まえて最終調整した。今後も国際動向や国内の診療ガイドラインの現状を考慮しながら継続的に改訂作業を行っていく。

本マニュアルの利用者として想定されるのは、診療ガイドラインに関わる全ての方々である。診療ガイドライン作成にあたっては、作業を進めるガイドとなるようにMindsが提案する方法を、作成過程の流れに沿っ

て解説した。さらに、各作成過程で記載すべき内容と資料をテンプレートとして提示した。診療ガイドラインの利用にあたっては、診療ガイドラインの記述内容や作成過程の全体像を示すことで、ガイドライン推奨内容の理解や実際の利用に生かされるよう配慮した。

なお、Minds の診療ガイドライン評価・選定では、このマニュアルに忠実に従うことを必須としてはいない。作成手順が記録されていること、綿密で体系的な作成方法がとられていること、先入観と偏りを避けるよう工夫されたこと、現場での適用を想定して作成されたこと、などを評価のポイントとしている。診療ガイドラインの作成にあたっては最適な作成過程を選択し、このマニュアルを超えた工夫や、やむなく削除・変更した手順も含めて記録していただければと思う。

2021 年 3 月

公益財団法人 日本医療機能評価機構
執行理事（EBM・診療ガイドライン担当）
福岡 敏雄

目 次

お願いとご注意	ii
Minds 診療ガイドライン作成マニュアル編集委員会／マニュアル作成ワーキンググループ／査読協力者	iii
はじめに	v
第1章 診療ガイドライン総論	1
Key Words	2
1.1 診療ガイドラインとは	3
1.2 診療ガイドライン作成の全体像	3
1.3 診療ガイドライン利活用のための追加文書など	6
1.4 診療ガイドラインに期待される役割	6
文献	8
第2章 準備	9
Key Words	10
2.0 概要	11
2.1 ステップ1：診療ガイドライン統括委員会の設置	12
2.2 ステップ2：診療ガイドライン作成手順およびスケジュールの決定	13
2.3 ステップ3：COI管理方針の決定	14
2.4 ステップ4：診療ガイドライン作成資金の準備	24
2.5 ステップ5：診療ガイドライン作成組織の編成	25
2.6 患者・市民参画	27
テンプレート	34
テンプレート 記入方法	40
文献	46
第3章 スコープ	48
Key Words	49
3.0 概要	51
3.1 ステップ1：疾患トピックの基本的特徴・スコープ作成方針の決定	52
3.2 ステップ2：疾患トピックの基本的特徴の整理	53
3.3 ステップ3：スコープ作成（診療ガイドラインがカバーする内容に関する事項① 重要臨床課題とCQ）	57
3.4 ステップ4：スコープ作成（診療ガイドラインがカバーする内容に関する事項② その他の記載事項）	68
3.5 ステップ5：スコープ作成（システムティックレビューに関する事項）	72
3.6 ステップ6：スコープ作成（推奨作成から公開に向けた最終調整、公開までに関する事項）	74
3.7 ステップ7：スコープの確定	75
テンプレート	77
テンプレート 記入方法	82
テンプレート 記入例	87
文献	94

第4章 システマティックレビュー	95
Key Words		96
4.0 概要		99
4.1 ステップ1：エビデンスの収集		118
4.2 ステップ2：スクリーニング		126
4.3 ステップ3：エビデンスの評価：個々の研究に対する評価（STEP 1）		128
4.4 ステップ4：エビデンス総体の評価（STEP 2）		142
4.5 ステップ5：システムティックレビューレポートの作成		158
4.6 補足資料1：Rとmetaforパッケージを用いたメタアナリシス		160
4.7 補足資料2：診断精度研究のメタアナリシス（RとmadaおよびOpenBUGS）		166
4.8 補足資料3：ネットワークメタアナリシス		168
4.9 補足資料4：システムティックレビューツール		191
テンプレート		192
テンプレート 記入方法		207
テンプレート 記入例		214
文献		224
第5章 医療経済評価	235
Key Words		236
5.0 はじめに		238
5.1 資源利用		239
5.2 概論：診療ガイドラインにおける費用対効果の考慮について		244
5.3 レファレンスケース		247
5.4 費用対効果の評価手順		254
5.5 費用対効果プラン		256
5.6 医療経済評価のシステムティックレビュー		258
5.7 新規の医療経済評価		260
5.8 費用対効果のエビデンスの解釈		263
5.9 補足資料：フィルターの例		264
テンプレート		267
テンプレート 記入例		272
文献		273
第6章 推奨	275
Key Words		276
6.0 概要		278
6.1 ステップ1：推奨作成の具体的方法の決定		279
6.2 ステップ2：推奨文草案の作成		283
6.3 ステップ3：推奨の作成、推奨の強さの判定		288
6.4 ステップ4：解説の執筆		294
6.5 一般向けサマリーの執筆		297
6.6 GRADEアプローチを用いる場合の資料		298

6.7 その他の意思決定に関するフレームワーク	303
テンプレート	304
テンプレート 記入方法	321
テンプレート 記入例	324
文献	327
第7章 公開に向けた最終調整	330
Key Words	331
7.0 概要	332
7.1 ステップ1：診療ガイドライン公開後の対応について協議と決定	333
7.2 ステップ2：作成経過に関する報告事項の作成	335
7.3 ステップ3：診療ガイドライン草案の決定	337
7.4 利用者視点からの再確認	340
7.5 ステップ4：外部評価の実施	341
7.6 診療ガイドライン最終案の決定	346
テンプレート	347
テンプレート 記入方法	353
テンプレート 記入例	359
文献	361
第8章 診療ガイドライン公開後の取り組み	362
Key Words	363
8.0 概要	364
8.1 公開後の組織体制	364
8.2 導入	365
8.3 普及・活用・効果の評価	369
8.4 改訂	370
8.5 公開後の取り組みの決定プロセス	374
テンプレート	375
テンプレート 記入方法	379
テンプレート 記入例	383
文献	384
重要用語集	385
文献一覧	403



各章のテンプレート一式

Minds ガイドラインライブラリよりダウンロードをしてください。

Minds ガイドラインライブラリ <https://minds.jcqhc.or.jp/>

第1章

診療ガイドライン総論

1章 Key Words

◎ 益と害 (benefit and harm)
◎ 益と害のバランス
◎ エビデンス総体 (body of evidence) エビデンスに基づく政策 (evidence-based policy making) 協働意思決定 (shared decision making : SDM)
◎ クリニカルクエスチョン (clinical question : CQ)
◎ 今後の研究 (future research)
◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
◎ システマティックレビューチーム (systematic review team) システムティックレビューレポート (サマリーレポートともいう)
◎ 診療ガイドライン (clinical practice guideline : CPG)
◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう) 診療ガイドライン作成主体
◎ 診療ガイドライン統括委員会
◎ 推奨
◎ スコープ (scope)
◎ クオリティーアンディケータ (quality indicator : QI)
◎ バイアスリスク (risk of bias)
◎ 利益相反 (conflict of interest : COI)
◎：巻末「重要用語集」の収録語

1.1 診療ガイドラインとは

本マニュアルでは、診療ガイドライン（clinical practice guideline : CPG）を以下のように定義する。

Minds 診療ガイドラインの定義

健康に関する重要な課題について、医療利用者と提供者の意思決定を支援するために、システムティックレビューによりエビデンス総体を評価し、益と害のバランスを勘案して、最適と考えられる推奨を提示する文書。

従来、「診療ガイドライン」は、「疾患に対する治療行為」を扱うものと見なされてきたが、今日では、治療に限らず、予防・検診、リハビリテーション、看護介入、保健指導、社会的支援など幅広い内容が扱われるようになってきた。このため診療ガイドラインは「健康に関する重要な課題」を扱うものとした。また、診療ガイドラインの推奨の中には、医療・保健・介護・福祉などの行政担当者など、医療に関わる多くの人が協働しなければ実施できないものもある。これらのこと踏まえて、定義の中の「医療利用者」は、「患者」のみではなく、その家族や一般市民などにも広がる場合もあるだろうし、「医療提供者」も、医師や看護師などの医療系有資格者に限ってはいない。個別には、必要に応じてそれぞれの診療ガイドラインで設定されるべきである。

システムティックレビュー（systematic review : SR、系統的レビューと同義）とエビデンス総体（body of evidence）の取り扱いについては、以降の章で詳述する。

1.2 診療ガイドライン作成の全体像

診療ガイドラインは信頼に足るものでなければならない。そのためには、科学的なエビデンス総体の評価がなされていること、その作成過程では思い込みや偏りを避ける方策がとられていること、さらに治療や検査に伴う害やコスト、負担などを勘案しても意味のあるものであるかが検討されていることが重要である。

診療ガイドラインの作成にあたっては、作成過程での判断の偏りを避けるために、作業を分担し独立して作成することを提案する。以下に、詳細を示す。

1.2.1 三層構造の担当組織

診療ガイドライン作成の過程では、その体制作りが重要である。作業を分担し、独立して作業を行う体制を目指す。作成全体を統括する「診療ガイドライン統括委員会」と、診療ガイドライン作成に不可欠なスコープ（scope）やクリニカルクエスチョン（clinical question : CQ）などを定める「診療ガイドライン作成グループ（guideline development group : GDG、診療ガイドライン作成パネルともいう）」、そしてCQに沿ってシステムティックレビューを作成しエビデンス総体を評価する「システムティックレビューチーム（systematic review team : SR チーム）」の3つの組織からなる体制が望ましい（図1-1 参照）。

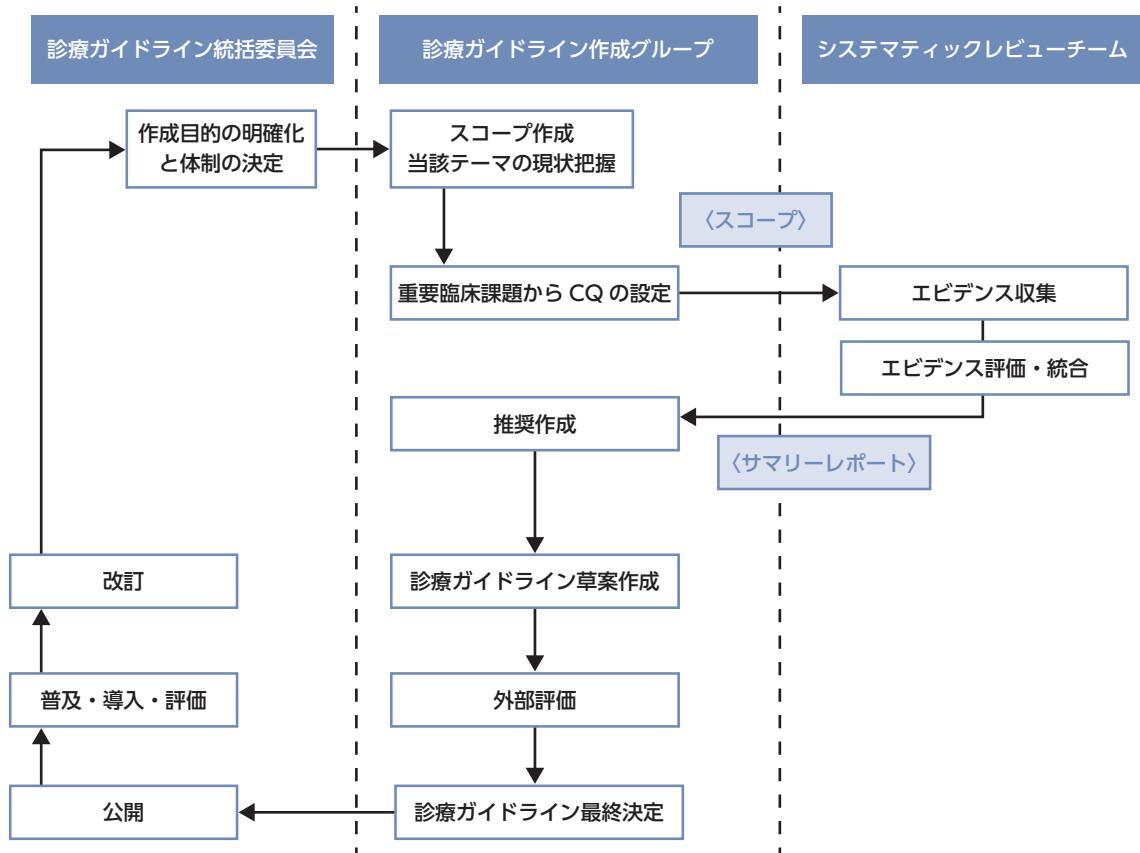


図 1-1 診療ガイドライン作成過程と担当組織

● 診療ガイドライン統括委員会

診療ガイドライン作成全体を制御し、意思決定および予算措置などを行い、診療ガイドライン作成グループの編成・設置を進める。最終的に個々の推奨と診療ガイドライン全体を承認し、公開に向けた作業を進めるのも重要な役割である。診療ガイドライン統括委員会の体制としては、診療ガイドラインを作成する学会などの理事会内や、その委員会として設置される。複数の学会などが協力して作成する場合は、各学会の代表者で構成される協議会的な委員会として設置する。

● 診療ガイドライン作成グループ

診療ガイドラインのスコープを確定し、診療ガイドラインが取り上げるトピック、CQなどを決定する。そして、システムティックレビューチームがまとめたサマリーレポート（システムティックレビューレポートともいう）に基づき、益と害（benefit and harm）のバランスを勘案して推奨を作成し、最終的に診療ガイドラインをまとめる。

● システマティックレビューチーム

診療ガイドライン作成グループが設定したCQに対して、スコープに記載された方法にのっとり、システムティックレビューを実施し、エビデンス総体を評価するための基本情報を提供する。

この3つの組織は役割が異なり構成員に求められる要件も異なる。診療ガイドライン統括委員会は作成主体（学会など）を代表する組織であるのに対して、診療ガイドライン作成グループには、学会員に限らず患者・市民も含めてさまざまな背景を持つ人たちが参加すべきである。また、システムティックレビューチーム構成員には、システムティックレビューが実施できる技能が求められる。

3つの組織の構成員は、一部兼任したり、他の組織と協議するといった対応はあり得るが、作成過程の透明性を確保するために、原則としてそれぞれの作業は独立して進めることが望ましい。

1.2.2 作成過程から先入観と偏りを排除する方法

診療ガイドラインの作成手順は、作成過程全体の透明性を確保するために、詳細に記録し開示することが求められる。それにより、先入観や偏りを避ける手順が一貫して行われたかについての検証が可能になる。

判断の偏りの原因として利益相反 (conflict of interest : COI) がある。診療ガイドラインで言及される医薬品・医療機器に関連する企業の株保有や雇用関係、金銭提供、研究費補助などは経済的COIとして重要視され取り扱いについてさまざまな基準が設けられている。近年このような経済的COI以外のCOI（学術的COIなど）の影響が注目されている。自らが専門とする治療法に好意的な判断を行いがちになったり、所属組織や職業上の立場が判断に影響を与える場合などがある。このため経済的COI以外のCOIも適切に管理・報告することが求められている。これらのCOIの対応については**第2章**で詳述する。

1.2.3 エビデンス総体の評価

エビデンス総体を評価するためには、1つの臨床上の問題（CQ）に対して情報を網羅的に収集し、一定の基準に基づいて選択し、アウトカムごとに評価し集約する必要がある。この、エビデンス総体を構成する臨床研究の論文を検索・収集し、評価・統合する一連のプロセスがシステムティックレビューである。この手順については**第4章**で詳述する。

重要なのは、たとえ信頼性の高い研究デザインであっても、ただ1つの研究結果に過度に依存しないことである。同一のテーマに対するものであっても研究方法（研究デザイン、研究対象、介入方法、アウトカムの測定方法、統計・解析方法など）の相違によって必ずしも同一の結果を示すとは限らない。システムティックレビューを通してエビデンス総体を、不確実性やバイアスリスク（risk of bias）も含め評価することが極めて重要である。

1.2.4 益と害のバランス

介入によってもたらされるものは期待される効果（益）のみではなく、有害な事象（害）も含まれる。推奨決定手順では、このバランスを考慮する。益と害に関する重大なアウトカムを列挙し、システムティックレビューによって幅広くエビデンス総体を評価する。また、患者・市民の不利益として、費用負担の増加や身体的・精神的な負担なども可能な限り考慮することが必要である。

1.3 診療ガイドライン利活用のための追加文書など

診療ガイドライン利用者の現場での利活用を助けるために追加文書を作成することが多なくなった。例えば、医療提供者がその場で参照できるクイックレファレンスガイドや、患者あるいは一般市民に診療ガイドラインのエッセンスを知ってもらうために「一般向けガイドライン解説」を作成することも一般的になっている。

また診断や治療を支援するツールとして、スマートフォンアプリやウェブアプリを提供する診療ガイドラインも見られるようになった。

このような追加文書も診療ガイドラインにひもづけられている必要がある。推奨内容に疑問が生じさらに確認しなければならなくなつた場合などには、基となった診療ガイドラインが参照できなければならない。

1.4 診療ガイドラインに期待される役割

診療ガイドラインは広くわが国の医療全体に貢献するものである。また、医療利用者と医療提供者の協働意思決定（shared decision making : SDM）を支援するためには、医療者のみでなく、患者・家族・市民にも診療ガイドラインの考え方と内容を知つてもらう取り組みが重要となる。

診療ガイドラインの活用が期待される領域として、診療、教育、研究、そして医療政策が考えられる。

● 診療

診療ガイドラインが現場で診療に生かされるためには、医療利用者と医療提供者による協働意思決定の基礎資料として活用しやすいものでなければならない。さらに、現在医療現場では多職種によるチーム医療が行われ、地域では施設を越え医療や介護、福祉を含む地域包括ケアの実現に向かっている。この中で活用される診療ガイドラインであるためには、あらゆる職種、あらゆる施設、あらゆる立場から見て、役に立ち、実際に活用されるものである必要がある。

さらに、実際の利用状況を把握するための測定項目【クオリティーアインディケータ（quality indicator : QI）】を提示する診療ガイドラインや、実際にそれを測定・共有し、改善する活動を行つてゐる診療ガイドライン作成グループや施設、団体も見られるようになった。

これらの測定項目やそれをめぐる活動は、国内全体での医療実践状況の可視化や改善運動の推進にもつながる大きな意義を持つ。

● 教育

診療ガイドラインは医学・医療の学習教材として適している。診療ガイドラインの作成過程は、現場の課題を把握し取り組む手順と重なる。治療の効果ばかりでなく、害や負担などを考慮して治療方針を決めることが実践されている。医療利用者と医療提供者の協働意思決定プロセスを診療ガイドラインから学ぶことができる。また、毎年発表される個々の研究は膨大で、その信頼性や内容も多様である。これらを厳密に評価してアップデートされた診療ガイドラインを活用すれば、効率的な学習が可能になる。

このような診療ガイドラインを活用した学習を、それぞれの職種の学習カリキュラム、専門教育および生涯教育の教材として位置付けられることが望まれる。

● 研究

研究の計画段階で、社会的なニーズを考慮することが求められている。臨床課題に対して推奨を作成する診療ガイドラインの作成過程においては、課題解決のための不足点が明らかになり、ここから今後の研究課題をまとめることができる。近年の診療ガイドラインの中には、そのような不足点を今後の研究（future research）として提示するものが増えてきた。このように診療ガイドライン作成は、次に取り組むべき研究課題を明らかにする手段としても活用されつつある。

● 医療政策

公共政策において、エビデンスに基づく政策（evidence-based policy making）の手法が受け入れられてきた。厳密に作られた診療ガイドラインはエビデンスに基づく医療政策の基盤として活用されるべきものであり、このために診療ガイドラインは、社会における人員・施設・費用などの資源の利用負担や、地域や世代間の公平性、現場での実現可能性などを踏まえたものであることが求められている。

また、診療ガイドラインは医療現場での意思決定を支援するためのものであり、医療政策の視点からも、診療ガイドライン導入による効果や有効性を評価検討する必要性は高まっている。

1章 文献

Institute of Medicine (IOM). (2011) Clinical Practice Guidelines We Can Trust. National Academies Press

第2章 準備

Minds

2章 Key Words

Guidelines International Network (G-I-N)
ICMJE (国際医学雑誌編集者委員会)
◎ Institute of Medicine (IOM) [現 National Academy of Medicine (NAM)]
◎ 益と害 (benefit and harm)
◎ 外部評価
◎ 外部評価委員会
◎ 患者・市民参画 (patient and public involvement : PPI)
◎ 患者・市民の価値観・希望
◎ 協働意思決定 (shared decision making : SDM)
◎ クリニカルクエスチョン (clinical question : CQ)
◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
◎ システマティックレビューチーム (systematic review team : SR チーム)
◎ システマティックレビュー論文
◎ 重要臨床課題 (key clinical issues)
◎ 診療ガイドライン (clinical practice guideline : CPG)
◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
◎ 診療ガイドライン作成事務局
◎ 診療ガイドライン作成主体
◎ 診療ガイドライン統括委員会
◎ 診療ガイドラインの適応 (guideline adaptation)
◎ 推奨
◎ スコープ (scope)
◎ バイアスリスク (risk of bias)
◎ 利益相反 (conflict of interest : COI)

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】

SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】

RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】

PO : post 【第8章 診療ガイドライン公開後の取り組み】

2.0 概要

本章では、診療ガイドライン（clinical practice guideline : CPG）作成前の準備段階で行うべき作業概要および手順について説明する。診療ガイドライン作成開始前の準備作業は、**ステップ1**から**ステップ5**までの5つのステップに分けられる。

○ ステップ

- (1) 診療ガイドライン統括委員会の設置
- (2) 診療ガイドライン作成手順およびスケジュールの決定
- (3) 利益相反 (conflict of interest : COI) 管理方針の決定
- (4) 診療ガイドライン作成資金の準備
- (5) 診療ガイドライン作成組織の編成

COI の管理方針

診療ガイドラインは、信頼性の高い情報源として医療現場を中心に広く社会に認識されることから、作成過程の厳密さやその透明性の担保は非常に重要であり、COIを開示し、COIへの対応策を講じる必要性が極めて高い。

診療ガイドライン作成開始前の準備段階から、診療ガイドライン作成におけるCOIの管理方針をあらかじめ決め、診療ガイドライン作成過程を通じて、その方針にのっとって管理することが求められる。

診療ガイドライン作成への患者・市民参画

・患者・市民参画 (patient and public involvement : PPI) の必要性

診療ガイドラインの目的は、医療現場で医療利用者と医療提供者による協働意思決定 (shared decision making : SDM) を支援することであり、医療者からの視点だけではなく、当該疾患に罹患した経験のある人や、家族、ケア提供者、支援者（患者団体などの担当者）、一般市民の視点を反映することが求められる。診療ガイドラインを医療者のみで作成した場合、患者・市民の価値観や希望、重要視する点などについて見落としたり、配慮したようでも、実際には見誤ったりする可能性がある。診療ガイドラインが対象とする人たちの実情により即した方法で、患者・市民にとって重要なアウトカム（ある介入によって患者・市民に生じる結果のこと）を検討するためには、診療ガイドライン作成過程に患者・市民が参画することが重要である。

・患者・市民参画の方法

- 1) 診療ガイドライン作成グループ (guideline development group : GDG、診療ガイドライン作成パネルともいう) メンバー や外部評価委員に患者・市民を含める。

- 2) スコープ（scope）作成時に患者・市民も含めた協議や文献調査、インタビュー、アンケートなどを行うことで、患者・市民にとっての課題・疑問を収集する。
- 3) 推奨作成時に患者・市民も含めた協議を行い、その他患者・市民の価値観・希望に関する情報などを収集し、推奨の根拠として検討する。
- 4) 診療ガイドラインの草案に対する外部評価の際に、患者・市民からのフィードバックを受け、必要に応じた対応・修正を施す。

2.1 ベーストップ 1：診療ガイドライン統括委員会の設置

診療ガイドライン統括委員会は、作成主体の下で、作成に関わる委員会の設置、予算の決定などの意思決定を担う委員会をいう。単一の学会・研究会の場合、その理事会、常設委員会などが当たることが多い。また、複数の学会・研究会が合同して診療ガイドラインを作成する場合は、各学会・研究会の代表者で構成される「協議会」のような組織が診療ガイドライン統括委員会となることもあり得る。診療ガイドライン統括委員会の役割は、**第1章図 1-1** を参照のこと。

● 手順

- (1) 診療ガイドライン統括委員会メンバー候補の選定
- (2) 診療ガイドライン統括委員会メンバーの決定

2.2 ステップ2：診療ガイドライン作成手順およびスケジュールの決定

診療ガイドラインは、以下の作成手順により作成される。計画にあたっては、全体を通してどのくらいの時間が必要か、各手順にどの程度の時間と費用をかけるかを考慮し、具体的に立案するとよい。

作成主体とは診療ガイドライン作成に責任を持つ学会・研究会などの団体をいう。単一の学会・研究会などの場合もあるが、選定されたテーマ／トピックによっては、関連する複数の学会・研究会などが、準備の段階から合同して作成を進めることがある。例えば、内科系学会と外科系学会、医科系学会と歯科系学会、医科系学会と看護系学会など、多様な専門的視点を得ることで、全過程を通して、偏りの少ない診療ガイドライン作成が実現できる。診療ガイドラインを作成する上で、偏りのない作成組織の編成は最も重要である。特に、診療ガイドライン統括委員長をはじめとした各組織の責任者の選出は、経済的、学術的、その他の利害関係に配慮して慎重に行う。

● 手順

- (1) 作成目的の明確化
- (2) 作成主体の決定
- (3) 事務局・診療ガイドライン作成組織の編成
- (4) スコープ作成
- (5) システマティックレビュー
- (6) 推奨作成
- (7) 診療ガイドライン草案作成
- (8) 外部評価・パブリックコメント募集
- (9) 公開
- (10) 普及・導入・評価
- (11) 改訂

● テンプレート

[PR-1 診療ガイドライン作成手順およびスケジュール]

● テンプレート 記入方法

[PR-1 診療ガイドライン作成手順およびスケジュール 記入方法]

● テンプレート 記入例

(なし)

2.3 ベース 3 : COI 管理方針の決定

COIとは、教育・研究に携わる専門家としての社会的責任と産学連携の活動に伴い生じる利益などが衝突・相反する状態のことをいう。COIは、経済的COIとそれ以外のCOI（学術的COIなど）に大別される。

また、個人的なCOIと同様に組織的なCOIに対する検討および対応も求められる（表2-1参照）。

2.3.1 COIの定義

- National Academy of Medicine [NAM, 旧 Institute of Medicine (IOM)] の定義

COIとは、第一義的な利害（研究の整合性、患者の福利、医学教育の質）に関係した専門家としての判断・行為に対して、二次的な利害（金銭的収益だけではなく、専門家としての達成や、個人的な功績に対する認識や、友人・家族・学生・同僚に対する便宜）が不当に影響を与えてしまうリスクを生み出す状況のセットのことである。

Institute of Medicine (IOM). (2009) Conflict of Interest in Medical Research, Education, and Practice. National Academies Press

独立した観察者が、個人の専門的行動あるいは決断が経済的、アカデミックな昇進、臨床的収入の流れあるいは社会的地位のような個人的獲得利益に動機付けられているかどうかについて適切な疑問を抱くかもしれないが、個人の私的利害とその専門家としての義務の間の相違

Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press

- Guidelines International Network (G-I-N) の定義（2015年）

主要な利害に関する専門的判断あるいは行動が、副次的利害に不当に影響を受けるリスクを生じる一連の状況

〈知的COIの定義〉

開かれた心で科学的疑問にアプローチする個人の能力に影響を与えるかもしれない経済的あるいは知的関係

Schünemann HJ, et al. (2009) An official American Thoracic Society Policy statement: managing conflict of interest in professional societies. Am J Respir Crit Care Med. 180:564-580.

診療ガイドラインに特異的な知的COIは、特異的な推奨に関する個人の判断に不当に影響し得る特異的な見解への執着の可能性を生み出すアカデミックな活動と定義される。

Guyatt G, et al. (2010) The vexing problem of guidelines and conflict of interest: a potential solution. Ann Intern Med. 152:738-741.

- ICMJE (国際医学雑誌編集者委員会) による生物医学雑誌への投稿のための統一規定 (2019年改訂版)

COIは、著者（あるいは著者の所属機関）、査読者または編集者が、自らの意思決定に不適切な影響（バイアス）を与える財政的または個人的な利害関係を有する場合に問題となる。（中略）これらの利害関係は、影響力がごくわずかなものから、意思決定に多大な影響力を持つものまで多様である。（中略）財政的利害関係（雇用、顧問、株式の所有、謝礼金、報酬を受けた専門家証言など）は、最も分かりやすいCOIであり、雑誌、著者、そして科学そのものの信頼性を最も損なうものといえる。COIが存在する可能性の見落としや置き違えを防ぐために、このような情報は論文の一部となっている必要がある。

International Committee of Medical Journal Editors. (2019) Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. (Updated December 2019) <http://www.icmje.org/> (参照 2021/03/03)

- 厚生労働科学研究における利益相反の管理に関する指針（2018年一部改正）

COIとは、具体的には、外部との経済的な利益関係等によって、公的研究で必要とされる公正かつ適正な判断が損なわれる、又は損なわれるのではないかと第三者から懸念が表明されかねない事態をいう。公正かつ適正な判断が妨げられた状態としては、データの改ざん、特定企業の優遇、研究を中止すべきであるのに継続する等の状態が考えられる。

厚生労働省. (2018) 厚生労働科学研究における利益相反の管理に関する指針. <https://www.mhlw.go.jp/stf/seisaku/seisaku-10600000-Daijin-kanboukouseikagakuka/0000152586.pdf> (参照 2019/10/10)

表 2-1 診療ガイドライン作成における COI の取り扱い：開示と管理

	IOM ¹⁾	G-I-N ²⁾	日本医学会 ³⁾
1. 透明性の確保	1.1 診療ガイドライン作成と資金調達の過程は明確、詳細で公衆がアクセスできるようにすべきである。	1. 診療ガイドライン作成者は直接的なCOIあるいは関連のある間接的COIを有する者をメンバーに含まないために可能な限り努力すべきである。	1. CPG策定参加者と学会組織がCOIを開示・公開するだけでなく、CPG策定参加者の資格要件を明確にし、バイアスリスクを回避するためのマネジメントが求められる。
2. COIの管理	2.1 診療ガイドライン作成グループの選任に先立ち、参加を考慮中の個人は作成グループの活動との間でCOIを生じ得る全ての利益と活動を招集者に書面で申告すべきである。 ・申告は現在の、そして計画されている、診療ガイドラインの想定されるスコープに関連する商業的（そこから収入のかなりの部分を得ている業務・サービスも含む）、非商業的、知的、患者・公衆に関する活動の全てを反映すべきである。	2. COIの定義とその管理は、専門性あるいは代表しているステークホルダーに関わらず、診療ガイドライン作成グループの全員に適用され、これはパネルが構成される前に決められるべきである。 3. 診療ガイドライン作成グループは利害申告に標準化されたフォームを用いるべきである。 4. 診療ガイドライン作成グループは全ての直接的な経済的および間接的COIを含む利害を公開すべきで、これらは診療ガイドライン利用者が容易にアクセスできるようにすべきである。 5. 診療ガイドライン作成グループの全てのメンバーはグループの会議のたびにそして定期的に（例えば、診療ガイドライン作成グループにいる間は1年ごとに）利害の変更について申告し更新すべきである。	2. 分科会の長は、常設委員会としてCPG統括委員会を設置する。診断、治療、予防にかかるCPGの新たな策定または改訂を行うためには、CPG策定委員会（ガイドライン作成グループ）、システムティックレビューチーム、外部評価委員（委員会）を設置する。それらの委員会、グループ、チームに参加するすべての個人本人とその配偶者、1親等親族または収入・財産的利益を共有する者がCOI開示の対象となる。

	IOM ¹⁾	G-I-N ²⁾	日本医学会 ³⁾
2. COI の管理	<p>2.2 診療ガイドライン作成グループ内での COI の申告 :</p> <ul style="list-style-type: none"> ・診療ガイドライン作成グループメンバーの全ての COI は、その仕事を開始する前に報告され、これから形成されるであろう作成グループによって議論されるべきである。 ・各パネルメンバーはその COI が診療ガイドライン作成プロセスあるいは個別推奨にどのように影響し得るかを説明すべきである。 	<p>8. 推奨の方向と強さを決める診療ガイドライン作成グループのメンバーは誰も直接的な経済的 COI を有するべきではない。</p> <p>*これらのメンバーは診療ガイドライン作成の場に参加すべきではない。彼らは、推奨の方向と強さに関する議論の場にはいないようにすること。</p>	<p>3. CPG 策定参加候補者の COI 状態が社会的に容認される範囲を超えていると判断される場合、CPG が策定されて、その内容の推奨に関連して、利害を有する当該企業・団体などから、CPG 策定の参加者本人またはその配偶者、1 親等親族または収入・財産的利益を共有する者に対して直接あるいは間接的に大きな経済的利益がもたらされる可能性が想定される場合には、原則として当該候補者を CPG 策定に参加させるべきでない。</p>
	<p>2.3 (株式などの) 処分</p> <ul style="list-style-type: none"> ・診療ガイドライン作成グループメンバーは各自および家族が保有する経済的投资を自ら売却・処分すべきで、その利益が診療ガイドラインの推奨により影響を受け得る企業体のマーケティングあるいはアドバイザリーボードに参加すべきではない。 		
	<p>2.4 適用除外</p> <ul style="list-style-type: none"> ・可能な限り診療ガイドライン作成グループメンバーは COI を有するべきではない。 ・いくつかの状況では、メンバーは診療ガイドラインに関連のあるサービスから収入のかなりの部分を得ている関連のある臨床専門家のような、COI を有するメンバーなしで、その仕事を遂行できないかもしれない。 ・COI を有するメンバーは診療ガイドライン作成グループの少数派に留めるべきである。 ・議長あるいは副議長は COI を有する者であってはならない。 ・資金提供者は診療ガイドライン作成で何らかの役割を担ってはならない。 	<p>6. 診療ガイドライン作成グループの議長は直接的あるいは間接的 COI を有するべきではない。議長の直接的あるいは間接的 COI が避けられない場合、診療ガイドライン作成パネルをリードする COI のない副議長を任命すべきである。</p> <p>*推奨の向きと強さに影響する場合は関連のある COI が存在する。そのような COI を有しない副議長の例として推奨の向きと強さに関係のある利害を持たない方法論専門家がある。</p>	
3. 診療ガイドライン作成グループの構成	<p>3.1 診療ガイドライン作成グループは方法論の専門家、臨床家、そして診療ガイドラインで影響を受けるであろう集団などさまざまな人々から構成され、多くの専門分野のバランスがとれた構成にすべきである。</p>	<p>7. 関連のある COI と特異的な知識あるいは専門性を持つ専門家は個々のトピックの議論への参加を許可し得るが、意見を求められる委員の間で適切なバランスがとられるべきである。</p> <p>*いくつかの設定においては、この役割を満たせる者は診療ガイドライン作成グループの投票するあるいは投票しないメンバーではない専門的アドバイザーと考えてもよい。</p>	<p>4. CPG 策定参加資格に係る基準に基づき策定参加者を委員として選考するが、委員会構成として下記の点に留意する。</p> <ol style="list-style-type: none"> 1) (一定以上の) 金額を有する委員数が委員総数の過半数を超えてはならない。 2) 痠学や統計専門家を含めて、参加者の専門性や領域について多様性を重視した委員構成とし、ある特定の企業や営利団体と利害関係にある委員に偏らないようにする。

	IOM ¹⁾	G-I-N ²⁾	日本医学会 ³⁾
3. 診療ガイドライン作成グループの構成	3.2 患者と市民の参画は現在の患者あるいは疾患経験者、患者支援者あるいは患者・医療利用者団体の代表者を診療ガイドライン作成グループに含むことで【少なくともクリニカルクエスチョン（clinical question : CQ）作成時点と診療ガイドライン草稿のレビューの時点で】強化すべきである。	9. 統括委員会は COI に関連した規則の作成と実行に責任を持つべきである。 * 統括委員会は論争になる問題に対処すべきで、診療ガイドライングループの議長に誰が投票し、誰が投票しないメンバーとするか、そして専門的アドバイザーとして誰を指名するかについてアドバイスすべきである。	5. CPG は医療現場で患者と医療者による意思決定を支援するための情報源として役立つことから、患者や市民団体を代表する参加者も CPG 策定委員として受け入れ、価値観が共有できる委員会構成を目指すべきである。
	3.3 エビデンスの批判的吟味のトレーニングを含む、患者と医療利用者の代表の効果的な参加を増やす戦略が診療ガイドライン作成グループにより採用されるべきである。		
4. 診療ガイドライン・システムティックレビューの交差	4.1 診療ガイドライン作成者は IOM の比較効果研究のシステムティックレビューのスタンダード ⁴⁾ により設定された基準を満たすシステムティックレビューを用いるべきである。		
	4.2 システマティックレビューが特定の診療ガイドラインに情報を与えるために特異的に実施されるときは、診療ガイドライン作成グループとシステムティックレビューチーム（systematic review team:SR チーム）は両者のスコープ、アプローチ、そしてアウトプットについて協働すべきである。		

1) Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press

2) Schünemann HJ, et al. (2015) Guidelines international network: principles for disclosure of interests and management of conflicts in guidelines. Ann Intern Med 163:548-553.

3) 日本医学会利益相反委員会. (2017) 診療ガイドライン策定参加資格基準ガイダンス . http://jams.med.or.jp/guideline/clinical_guidance.pdf (参照 2021/03/03)

4) Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for Systematic Reviews. National Academies Press

2.3.2 COI の種類

COI は、特定の企業／団体との経済的関係、研究費取得などが関係する経済的 COI と、学術的 COI（研究活動や専門性等）などの経済的 COI 以外の COI に大別される（表 2-2 参照）。また、個人的な COI と同様に、診療ガイドライン作成グループメンバーが所属する大学などの教育機関、学会などの学術組織の経済的 COI やその他の COI も診療ガイドライン作成に影響を及ぼす可能性がある。

表 2-2 COI の種類

	経済的 COI	経済的 COI 以外の COI
個人的 COI	<ul style="list-style-type: none"> ・特定の企業／団体から本人、家族への経済的利益の提供 ・研究費取得の利益 ・機器、人材、研究環境の提供、他 	<ul style="list-style-type: none"> ・研究活動 ・個人の専門性・選好 ・昇進・キャリア形成 ・師弟関係などの人間関係、他
組織的 COI	<ul style="list-style-type: none"> ・特定の企業／団体から学会・研究会などの経済的支援 ・学会・研究会の経済的発展、他 	<ul style="list-style-type: none"> ・学会・研究会などが推奨する専門性 ・学会・研究会などの学問的発展 ・利害関係のある他組織との競争関係、他

2.3.3 COI 開示および管理の基本的な考え方と対応の要点

- ・COI の申告および開示の目的は透明性の確保である。
- ・個人的 COI のみならず、組織的 COI についても申告・開示対象とする。
- ・経済的 COI のみならず、経済的 COI 以外の COI についても申告・開示対象とする。
- ・診療ガイドライン利用者がアクセス可能な形での情報開示を行う。
- ・企業名のみの開示ではなく、診療ガイドライン作成グループメンバー（役割を含む）と企業名（金額など）を共に開示する。
- ・COI を有する場合の対応こそが本質的な課題である。
- ・診療ガイドライン作成前に、COI 管理方針・管理方法を統括委員会／学会理事会／学会 COI 管理委員会が決定する。
- ・COI の申告内容に基づいて、診療ガイドライン作成グループの構成を検討する。
→個人メンバーの就任の可否を判断する。
→全ての関連するステークホルダーを含んでいるか、多様性が確保されているかを確認する。
- ・COI 管理方法について診療ガイドライン上に開示する。
- ・COI を有する場合の対応方法を事前に検討し開示する。

● 手順

(1) COI 管理方針 (COI 管理指針, COI 申告書) の決定

1) 個人的 COI への対応

診療ガイドライン作成組織の編成前に、候補者から経済的 COI および学術的 COI などの経済的 COI 以外の COI の自己申告書を診療ガイドライン統括委員会に提出してもらい、作成組織への参加の適否を検討する。診療ガイドラインの内容と関連する企業／団体などから資金提供を受けている候補者は診療ガイドライン作成上のいかなる役割も担わないなどの対応が必要である。特に診療ガイドライン統括委員長、診療ガイドライン作成グループ責任者は、本ガイドラインに関連する COI の視点から、作成の中立性が担保できるか否か、外部からの疑惑の対象にならないかについて、学会理事会などが十分検討を行った上で、適切な人物を選出する必要がある。

学術的 COI などの経済的 COI 以外の COI への対応としては、診療ガイドライン作成グループの構成員が特定の専門領域・関係者に偏らないように配慮する必要がある。

【対応策の例】

- ・COI を有するメンバーをリーダーにしない。
- ・COI を有するメンバーを全メンバーの 2/3 以上にしない。
- ・必要に応じて、COI を有するメンバーの役割制限を設ける。
- ・多様なメンバー構成にする。
- ・推奨作成時には、特定の人物の意向が反映しないような合意形成方法をあらかじめ検討し準備する。

① COI 申告／開示対象

COI の申告／開示対象者は、当該診療ガイドライン作成に携わる全ての個人とする。

② COI 申告／開示対象期間

COI 申告／開示対象期間は、当該診療ガイドライン作成メンバーに就任時からさかのぼって過去 3 年間分とする。なお、就任後についても、申告／開示すべき COI について変更があった場合には、都度速やかに申告する。

③ 個人的・経済的 COI 申告項目 (例)

- ・企業や営利を目的とした団体の役員、顧問職の有無と報酬額
- ・株式の保有と、その株式から得られる利益（最近 1 年間の本株式による利益）
- ・企業や営利を目的とした団体から特許権使用料として支払われた報酬
- ・1 つの企業や営利を目的とした団体より、会議の出席（発表、助言など）に対し支払われた日当、講演料などの報酬
- ・1 つの企業や営利を目的とした団体がパンフレット、座談会記事などの執筆に対して支払った原稿料
- ・1 つの企業や営利を目的とした団体が提供する研究費（産学共同研究、受託研究、治験等）

- ・1つの企業や営利を目的とした団体が提供する奨学（奨励）寄附金
- ・企業などが提供する寄附講座
- ・その他の報酬（研究とは直接に関係しない旅行、贈答品等）

日本医学会利益相反委員会. (2017) 診療ガイドライン策定参加資格基準ガイダンス. http://jams.med.or.jp/guideline/clinical_guidance.pdf (参照 2021/03/03)

日本医学会利益相反委員会. (2020) COI 管理ガイドライン. http://jams.med.or.jp/guideline/coi_guidelines_2020.pdf (参照 2021/03/03)

2) 組織的 COIへの対応

診療ガイドラインの内容に影響を与える可能性のある特定の団体からの寄附金などが診療ガイドラインの作成に影響を及ぼす可能性について、診療ガイドライン統括委員会において十分な議論が必要である。

学術的 COI など経済的 COI 以外の COI への対応としては、診療ガイドラインの内容に関連する可能性のある学会・研究会などが幅広く参加し、多様な専門家・関係者が合同で作成に当たることが極めて重要である。

① COI 申告／開示対象

- ・診療ガイドライン作成主体（学会、研究会など）
- ・当該診療ガイドライン作成に携わる個人の所属組織

② COI 申告／開示対象期間

COI 申告／開示対象期間は、当該診療ガイドライン作成メンバーに就任時からさかのぼって過去 3 年間分とする。なお、就任後についても、申告／開示すべき COI について変更があった場合には、都度速やかに申告する。

③ 組織的・経済的 COI 申告項目（例）

- ・1つの企業や営利を目的とした団体が提供する研究費（産学共同研究、受託研究、治験等）
- ・1つの企業や営利を目的とした団体が提供する奨学（奨励）寄附金

日本医学会利益相反委員会. (2017) 診療ガイドライン策定参加資格基準ガイダンス. http://jams.med.or.jp/guideline/clinical_guidance.pdf (参照 2021/03/03)

日本医学会利益相反委員会. (2020) COI 管理ガイドライン. http://jams.med.or.jp/guideline/coi_guidelines_2020.pdf (参照 2021/03/03)

(2) 診療ガイドライン作成グループメンバーなどの候補者選定**(3) COI 申告／開示依頼・申告／開示内容変更受付**

作成過程で申告／開示すべき COI に変更が生じた場合は、 診療ガイドライン作成グループ委員長に自己申告するように委員全員に周知する。

(4) 役割分担・役割範囲の決定

COI 申告／開示内容に基づき、 診療ガイドライン作成メンバーなどの選定および診療ガイドライン作成過程における役割分担・役割範囲を決定する。

(5) COI に関する情報管理と公開

COI の管理と対応に関する事項は、 診療ガイドライン上に公開すべきである。ただし、 個人別に収集した COI 申告書は秘匿すべき個人情報が含まれている可能性があるので、 診療ガイドライン統括委員会において厳重に管理するとともに、 公開方法についても検討する。COI 関連資料は、 診療ガイドライン公開後、 5 年間を目安に次の改訂時までは保管する。

● テンプレート

【PR-2 COI 申告① 経済的 COI 申告書】

【PR-3 COI 申告② 経済的 COI 以外の COI 申告書】

【PR-4 COI 開示 経済的および経済的 COI 以外の COI 開示サマリーと対応方針記載表】

● テンプレート 記入方法

【PR-2 COI 申告① 経済的 COI 申告書 記入方法】

【PR-3 COI 申告② 経済的 COI 以外の COI 申告書 記入方法】

【PR-4 COI 開示 経済的および経済的 COI 以外の COI 開示サマリーと対応方針記載表 記入方法】

● テンプレート 記入例

(なし)

【COI 管理チェックリスト】

COI 管理フローに基づいた手順項目

(1) 診療ガイドライン統括委員会の設置

- 学会内に診療ガイドライン統括委員会が設置されている。
- 診療ガイドライン統括委員を決めるにあたり、各候補者に COI 申告を依頼した。
- COI 申告内容に基づいて、診療ガイドライン統括委員会の構成員を決定した。
- 診療ガイドライン統括委員長は重大な COI を有していない。
- 診療ガイドライン統括委員長が COI を有している場合、就任の必要性に関する説明がなされている。

(2) 診療ガイドライン統括委員会による組織的 COI の開示

- 診療ガイドライン統括委員会は、学会組織の COI を開示している。

【組織的・経済的 COI の開示項目例】

- ・診療ガイドライン作成資金源
- ・学会等組織運営資金源

【組織的・経済的 COI 以外の COI の開示項目例】

- ・学会等組織活動内容

(3) COI 定義および COI 管理方針の決定

- 診療ガイドライン統括委員会は、COI の定義を定め、公表・周知している。
- 診療ガイドライン統括委員会は、COI 管理方針を定め、公表・周知している。
- 診療ガイドライン統括委員会は、COI 申告フォームを定め、公表・周知している。

【個人的・経済的 COI の開示項目例（本人／家族）】

- ・役職・顧問職
- ・株式
- ・特許権使用料
- ・謝金・講演料
- ・原稿料
- ・研究費
- ・奨学寄附金
- ・寄附講座
- ・その他

【個人的・経済的 COI 以外の COI の開示項目例（本人／家族）】

- ・所属機関
- ・所属学会
- ・所属委員会等
- ・専門分野

(4) 診療ガイドライン作成グループ構成員の候補者選定と COI 申告依頼および構成員の決定

- 診療ガイドライン統括委員会は、診療ガイドライン作成グループ構成員を決めるにあたり、各候補者に COI 申告を依頼した。
- 診療ガイドライン統括委員会は、COI 申告内容に基づいて、診療ガイドライン作成グループ構成員を決定した。
- 診療ガイドライン作成グループ議長（診療ガイドライン作成委員長）は重大な COI を有していない。
- 診療ガイドライン作成グループ議長が COI を有している場合、就任の必要性に関する説明がなされている。
- 診療ガイドライン作成グループ議長が COI を有している場合、COI を有していない共同議長が任命されている。
- メンバー構成員に偏りが生じないよう、多様性を確保する。可能な限り関連する全てのステークホルダーを含めるようにする。

(5) システマティックレビューチームの候補者選定と COI 申告依頼および構成員の決定

- 診療ガイドライン統括委員会は、システムティックレビューチーム構成員を決めるにあたり、各候補者に COI 申告を依頼した。
- 診療ガイドライン統括委員会は、COI 申告内容に基づいて、システムティックレビューチーム構成員を決定した。

(6) 外部評価委員の候補者選定と COI 申告依頼および構成員の決定

- 診療ガイドライン統括委員会は、外部評価委員会構成員を決めるにあたり、各候補者に COI 申告を依頼した。
- 診療ガイドライン統括委員会は、COI 申告内容に基づいて、外部評価委員会構成員を決定した。

(7) COI 申告内容に基づく役割範囲の決定・役割制限の設定

- 診療ガイドライン統括委員会は、COI 申告内容に基づいて、診療ガイドライン作成グループ議長の役割範囲を決定し、必要に応じた役割制限を設けている。
- 診療ガイドライン統括委員会は、COI 申告内容に基づいて、診療ガイドライン作成グループ共同議長の役割範囲を決定し、必要に応じた役割制限を設けている。
- 診療ガイドライン統括委員会は、COI 申告内容に基づいて、診療ガイドライン作成グループ各構成員の役割範囲を決定し、必要に応じた役割制限を設けている。

(8) 推奨作成時における COI 対応

- 推奨作成における合意形成方法については、あらかじめ計画段階で決定している。
- CQ 単位で診療ガイドライン作成グループ構成員の COI の有無を確認している。
- 診療ガイドライン作成グループ議長が COI を有している場合、該当 CQ の推奨作成において、議事進行役（司会）の担当から外れている。

(9) COI に関する情報開示

- 診療ガイドライン作成グループの COI 管理方針および診療ガイドライン作成グループ構成員の COI の有無と状況について、診療ガイドライン上に情報開示がなされている。

(10) COI 申告内容の変更受付および情報更新

- 定期的に診療ガイドライン作成グループ構成員・システムティックレビューチーム構成員・外部評価委員のCOI申告内容の変更の有無を確認し、変更があった場合は、速やかに情報の更新に努め、必要に応じて役割の変更などの対応を行っている。
- 診療ガイドライン作成グループ構成員・システムティックレビューチーム構成員・外部評価委員のCOI申告内容に変更があった場合は、速やかに情報の更新に努め、必要に応じて役割の変更などの対応を行っている。

2.4 ベース4：診療ガイドライン作成資金の準備

診療ガイドライン作成にどの程度の資金が必要か予算を組み、作成資金をどこから拠出するのか記載する。診療ガイドライン作成に要する主な費用には、交通費、会議会場費、文献検索・収集に関わる費用、製本等公開のために必要な費用などがある。公的資金、企業資金だけでなく、学会の資金を使用する場合も記載する。

● 手順

- (1) 作成資金の費用項目を挙げ、予算案を作成する。
- (2) 資金提供者を検討し、確保する。

● テンプレート

【PR-5 診療ガイドライン作成資金】

● テンプレート 記入方法

【PR-5 診療ガイドライン作成資金 記入方法】

● テンプレート 記入例

(なし)

2.5 ベース5：診療ガイドライン作成組織の編成

診療ガイドライン作成組織には、診療ガイドライン統括委員会の下、診療ガイドライン作成事務局、診療ガイドライン作成グループ、システムティックレビューチーム、外部評価委員会が必要である。**ステップ2**であらかじめ決定した作成組織編成方針に従って、組織編成を進めていく。

(1) 診療ガイドライン作成事務局

診療ガイドライン作成の進行管理、診療ガイドライン作成組織のメンバー間の連絡、作成会議の日程調整、会議室の確保、文献収集などの事務作業、診療ガイドライン作成資金の管理などを担当する事務局を設置する。

(2) 診療ガイドライン作成グループ

診療ガイドライン作成グループは、診療ガイドライン作成の企画書とも言うべきスコープを作成し、診療ガイドラインが取り上げるべき問題（CQ）を決定し、システムティックレビューの結果を受けて推奨を作成して診療ガイドライン草案を作成する責任を負う（**第1章図1-1** 参照）。診療ガイドライン作成グループの編成は、診療ガイドライン統括委員会が行う。疾患の特性、どのような診療ガイドラインを作成するかにもよるが、十数名で構成されることが多い。診療ガイドラインの内容に関係する多様な領域の専門家が幅広く含まれるべきである。診療ガイドラインが扱うトピックの専門医を複数の関連学会から利害関係を熟慮した上でバランスよく選出するほか、プライマリケア医、看護師や薬剤師などの医療職、診療ガイドライン作成に関わる専門家（図書館員など医学文献検索専門家、疫学専門家、統計専門家）、医療経済学の専門家、法律家、患者・市民、政策担当者など、性別、経済的およびアカデミックな利害関係に配慮した多様なメンバーで編成することが望ましい。また、**2.3節**で述べたCOIへの配慮が必要である。

(3) システマティックレビューチーム

システムティックレビューチームは、システムティックレビューを担当するグループであり、診療ガイドライン作成グループとは独立したチームとして、診療ガイドライン統括委員会の指示で編成される。システムティックレビューチームのメンバーは、システムティックレビューの方法についてトレーニングを受け、システムティックレビュー作成の十分な経験を持っていることが望ましい。システムティックレビューチームは、各CQに対して、既存のシステムティックレビュー論文の採用、海外の診療ガイドラインの適応（guideline adaptation）の可否を判断して、新たなシステムティックレビューが必要な場合は、それを実施する。わが国では、医師などの医療関係者がシステムティックレビューを行うことが多いと思われるが、文献検索を担当する図書館員など医学文献検索専門家、システムティックレビューの方法論に習熟する疫学専門家・統計専門家などをチームに入れ、システムティックレビューに必要な技術面のサポートを得ることが望まれる。なお、医療経済研究のシステムティックレビューは、①「医療経済評価チーム」がまとめて担当する方法と、②トピックごとに各レビューチームが行う方法のどちらでも可能である。また、診療ガイドライン作成グループメンバーと

同様に COI への配慮が必要である。

(4) 外部評価委員会

外部評価委員は、診療ガイドライン草案を第三者の立場で評価し、改善のための助言を行う。外部評価委員は、診療ガイドライン統括委員会が任命する。**2.3 節**で述べた COI への配慮が必要である。利害関係にある複数の関連医学会の疾患専門医とプライマリケア医、その他の医療職、疫学専門家、経済学者、法律専門家、患者・市民などを指名して評価を受けるほか、テーマによっては、学会のウェブサイトなどに草案を一定期間公開して幅広くパブリックコメントを求める方法も検討が必要になる。

● 手順

- (1) 各組織のメンバー編成方針に従って、メンバー候補者を選定する。
- (2) COI 申告の依頼をする。
- (3) COI 申告内容に基づき、診療ガイドライン作成メンバーなどの選定および診療ガイドライン作成過程における役割分担・役割範囲を決定する。

● テンプレート

【PR-6 診療ガイドライン作成組織】

● テンプレート 記入方法

【PR-6 診療ガイドライン作成組織 記入方法】

● テンプレート 記入例

(なし)

2.6 患者・市民参画

「患者・市民参画（patient and public involvement : PPI）」とは、多岐にわたるヘルスケア分野において、医療専門家以外の疾患経験者や家族、ケア提供者、支援者、一般市民が成果を生み出すプロセスに参画することを指す。診療ガイドライン分野においては、英国を中心に2000年代頃から患者・市民参画の取り組みが活発化し、今では世界中の診療ガイドライン作成組織で、さまざまな形での患者・市民参画が推進されている。

わが国の患者・市民参画は、研究倫理審査、臨床研究、医療・医薬品開発、医療政策の策定などの分野でも進められており、近年臨床研究分野において日本医療研究開発機構（AMED）が患者・市民参画の取り組みを強化している。

診療ガイドライン分野における国内の取り組みは、2004～2006年度厚生労働科学研究『「根拠に基づく診療ガイドライン」の適切な作成・利用・普及に向けた基盤整備に関する研究：患者・医療消費者の参加推進に向けて』、2007～2009年度「診療ガイドラインの新たな可能性と課題：患者・一般国民との情報共有と医療者の生涯学習」（代表：中山健夫）をはじめとして、検討が継続されてきた。それらの成果を取り入れて Minds でも議論を進め、2016年には『「診療ガイドライン作成への患者・市民の参加」の基本的な考え方』を公開し、診療ガイドライン作成プロセスへの患者・市民参画を促進するべく活動を展開している。患者・市民参画は診療ガイドラインの品質基準においても重要視されており、国際的な診療ガイドラインの質評価ツールである AGREE II（AGREE Next Steps Consortium 2009）には、評価項目として「対象集団（患者、一般市民など）の価値観や希望が調べられたこと」が盛り込まれている。また、IOM（現 NAM）が発行した *Clinical Practice Guidelines We Can Trust* (IOM 2011a) でも、信頼に足る診療ガイドラインの基準として、「患者や元患者、ならびに患者支援団体や患者／消費者団体の代表者を作成グループに加える」ことが挙げられている。

2.6.1 患者・市民参画の重要性

患者・市民参画の重要性については、大きく3つの側面から考えることができる。

1点目は、協働意思決定の支援という診療ガイドラインの本質に根差す重要性である。診療ガイドラインは、医療利用者と医療提供者が協働して意思決定を行うことを支援するために、臨床上重要な課題について、最適と思われる治療法等（選択肢）を推奨として示す文書である。科学的に妥当な治療法であっても、それ以外の取り得る選択肢や患者・市民にとって重大な害（副作用、費用、不便、不利益等）がある場合、また患者・市民の価値観・希望が多様な場合など、いずれの治療法を選択するか、患者・市民と医療者が協働して考えるための情報を提供するのも診療ガイドラインの役割である。そのためには、診療ガイドラインが患者・市民のニーズ（知りたいこと）に応えるものである必要があり、その作成過程に患者・市民が参画し、自身の経験や、患者・市民の価値観・希望、重要視する点などについて情報提供を行うこと、または作成グループが患者・市民の価値観・希望について情報収集を行い、吟味することが重要である。

2点目は、より質の高い診療ガイドライン作成に資する手段としての重要性である。診療ガイドライン作成に患者・市民が参画することにより、以下のような点が改善されると考えられる。

- 1) 患者・市民にとって重要であっても、医療者が見落としてしまう対処すべき課題・疑問やCQのアウトカムを拾い上げることができる。
- 2) 当該診療ガイドラインにおいて課題となる治療法が、患者・市民に実際にどのような影響を与えるかについて知り、益と害（副作用、費用、不便、不利益等）の推定をより具体的に行うことができる。
- 3) 推奨を作成する際、科学的なエビデンスを補完・補強するもの、または疑義を呈するものとして、患者・市民の見解を反映できる。
- 4) 推奨文などが、分かりやすく、患者・市民を尊重した表現で作成されているか、検証することができる。
- 5) 診療ガイドラインの普及と活用（患者・市民向け解説含む）について、示唆が得られる。

最後に重要な点として、診療ガイドラインへの患者・市民参画が、当該診療ガイドラインの社会的信用の基盤となることが挙げられる。患者・市民は、いわゆる医療利用者（医療消費者）として、あるいは医療保険制度における納税者として、医療に社会的な関わりを持つ存在でもある。こうした人々のニーズや意見が診療ガイドラインに反映されることは、広く社会からも重要視されるといえる。診療ガイドラインが、患者・市民を含む社会の多様なステークホルダーによって、民主的で透明性のあるプロセスを経て作成されたということが、社会への説明責任を果たし、広く信頼できる情報として受け入れられることにつながる。

前述のとおり、近年では診療ガイドラインが扱うテーマなどによって、家族、ケア提供者、支援者、一般市民といった患者以外の人々の価値観・希望を反映することも重要だと考えられている。ただし、こうした人々の意見が患者本人の利益を損ねないよう、慎重に取り扱うことが求められるという声もあり、診療ガイドラインが考慮すべき価値観・希望の取り入れ方については今後さらなる議論が必要であると考えられる。

2.6.2 患者・市民参画についてのポリシー

前述のとおり、診療ガイドラインにおける患者・市民参画は、診療ガイドラインの果たすべき役割やその質の向上、社会的責任を果たすという観点から重要性を増している。作成プロセスにおいて患者・市民の価値観・希望などについて情報収集し、そのことを診療ガイドライン中に明記することで、当該診療ガイドラインの信頼性や透明性、公平性を内外に示すことができる。具体的には、診療ガイドラインの中の「作成方法」に関する箇所や「解説」などに、患者・市民の価値観・希望などをどのような形で推奨に取り入れたか、または検討したかということを明記することが望ましい。

また、診療ガイドライン作成グループのメンバーや外部評価委員として患者・市民が参加した場合には、プライバシーに配慮した上で、作成委員や外部評価委員、協力者などのリストの中に明記することも重要である。

2.6.3 患者・市民参画の具体的な方法

患者・市民参画は、診療ガイドライン作成プロセスの以下の段階において重要とされる。

(1) スコープ作成

患者・市民にとって重要な課題・疑問について情報収集を行い、医療者にとっての課題・疑問などと合わせて検討し、重要臨床課題（key clinical issues）として何を取り上げるか決定する。

(2) CQ の作成

アウトカムを検討する際、患者・市民にとって重要なアウトカムを見落とさないよう、情報収集を行う、もしくは患者・市民委員と協議する。

(3) システマティックレビュー

患者・市民の価値観・希望、重要視する点、選好やニーズなどをテーマとした研究などについて情報収集を行う。

(4) 推奨作成

患者・市民にとっての益と害の推定や、最終的に患者・市民が重要視するアウトカムについて患者・市民から情報収集を行う、もしくは患者・市民委員と協議する。推奨の強さの決定においては、患者・市民の価値観・希望は重要な検討要素となる。

(5) 外部評価・パブリックコメント

外部評価委員として患者・市民に参画してもらい、診療ガイドラインの草案に対して患者・市民の視点で、必要な情報が盛り込まれているか、推奨は適切か、患者・市民にとって分かりやすい表現で記載されているかなどについてフィードバックを得る。

(6) 活用・普及

診療ガイドラインの活用や普及に関する取り組みにおいて、患者・市民と協働する。患者・市民向け解説やデシジョンエイド（意思決定ガイド）などの患者・市民向けの情報について助言を受ける。

2.6.4 患者・市民参画の手法

患者・市民参画の手法には、委員として作成グループに参加する以外にもさまざまな方法があり（表2-3参照）、診療ガイドライン作成の全ての段階で取り入れることが推奨されるが、作成グループの人的、財政的リソースや方針により、柔軟に検討することが重要である。

表 2-3 患者・市民参画の手法

手法		手順	対象者の要件	募集	その他
作成グループ委員としての参加		<ul style="list-style-type: none"> スコープの作成から推奨作成まで、全ての過程を通して患者・市民に委員として診療ガイドライン作成グループに参加してもらい、意見を取り入れながら診療ガイドラインを作り上げる。 スコープの作成や推奨作成など、患者・市民の意見が必要な段階で作成グループの議論に参加してもらうことも検討できる。 	<ul style="list-style-type: none"> 対象者を決めて作成グループへの参加を依頼する。 会議の開催が決またら通知する。 議論にあたって必要な資料を渡し、質問などを受け付ける。 作成グループの会議が開催される際に、出席してもらう。 <p>(例)</p> <ul style="list-style-type: none"> 自分自身が患者である／家族である／ケア提供者である／支援者である／医療者ではない一般市民である。 患者・市民団体などの活動をとりまとめた経験がある。 多くの患者・市民の代表として意見を述べる意思がある。また述べる能力がある。 会議や準備、作業に必要な時間が割ける。 COIを申告することに同意する。守秘義務を守れる（作成グループ内に主治医がないなど）。 	<ul style="list-style-type: none"> 公募 推薦（患者団体などからの推薦、作成グループ委員長や委員の推薦） 	<ul style="list-style-type: none"> COIや守秘義務に留意する。 作業や会議参加が発生した場合は、謝礼や交通費を支払うことが望ましい。
外部評価委員としての参加		<ul style="list-style-type: none"> スコープや推奨、診療ガイドラインの草案が出来上がった段階で、作成グループ委員ではない外部の患者・市民の評価を受ける。 			
情報収集	インタビュー	<ul style="list-style-type: none"> 1対1の個別インタビューや、複数人が参加するフォーカスグループインタビューなどを実施する。 患者・市民にとっての課題・疑問を収集し、スコープや推奨の作成に活用する。 	<ul style="list-style-type: none"> インタビューの対象者や質問項目を検討する。 インタビューへの参加を依頼する。 インタビューの実施日・場所などを決め出席してもらう。 インタビュー内容を記録する。 結果を分析する。 可能な範囲で診療ガイドライン策定結果を報告する。 <p>(例)</p> <ul style="list-style-type: none"> 自分自身が患者である／家族である／ケア提供者である／支援者である／医療者ではない一般市民である。 作業に必要な時間が割ける。 必要に応じてCOIを申告することに同意する。守秘義務を守れる。 		
	アンケート	<ul style="list-style-type: none"> 郵送や電話、ウェブなどで質問紙調査を実施する。 患者・市民にとっての課題・疑問を収集し、スコープや推奨の作成に活用する。 	<ul style="list-style-type: none"> 調査対象、調査期間、調査方法、調査票などを確定する。 郵送による質問紙調査やネット調査などを実施する。 結果を分析する。 可能な範囲で診療ガイドライン策定結果を報告する。 		
	文献調査	<ul style="list-style-type: none"> 患者・市民の価値観・希望、重要視する点、選好やニーズなどを調査した研究論文やその他の文献から患者・市民にとっての課題・疑問を収集し、スコープや推奨の作成に活用する。 	<ul style="list-style-type: none"> 検索用語を検討し、研究論文や文献を検索する。 必要に応じて結果を統合する。 	—	—
	パブリックコメント	<ul style="list-style-type: none"> スコープや推奨、診療ガイドラインの草案が完成した時点で草案を公開し、広く意見を募集する。 寄せられた意見について回答するとともに、スコープや推奨、診療ガイドラインの草案について再考する材料とする。 	<ul style="list-style-type: none"> 学会のウェブサイトやメディア媒体などで草案を公開し、パブリックコメント募集の案内を行う。 メールや入力フォームなどで意見を提出してもらう。 寄せられた意見をとりまとめ、回答を作成する。 意見と回答を公表する。 草案を修正する。 	—	—

2.6.5 作成委員としての参加の場合

① 患者・市民委員の人選

対象は、当該疾患に罹患した経験のある人、家族、ケア提供者、支援者（患者団体などの担当者）、一般市民などである。人数については、2名以上とすることが望ましい。2名以上が参加することで、個別的な経験によって意見が偏った方向に進むことを避けることができる。またグループ内で孤立せず、安心して発言できる状況を作り出すことにもつながる。

患者・市民委員の候補を募る方法については、公募、関連する患者団体などからの推薦や、作成グループの委員からの推薦などがある。

② 作成委員としての関わり方

患者・市民を作成グループの委員（患者・市民委員）として迎え入れる際、原則としてスコープ作成から推奨作成、普及・活用までの全ての段階に参画してもらうことが望ましい。作成委員として全てのプロセスに参加する以外に、患者・市民が外部評価委員として草案について意見を述べる、またはプロセスの中でも患者・市民の意見が特に重要とされる段階で情報提供を受けるなどの関わり方も考えられる。作成グループは、患者・市民にどのような形で、作成プロセスのどの段階に関わってもらうかを準備段階でよく検討しておく必要がある。

③ COIに関する考え方

作成グループに参加する全ての委員からCOIを申告してもらうことを推奨する。患者・市民委員についても同様であり、事前にCOIについて説明しておくとよい。特に患者の場合、個人的にはCOIがなくとも、所属する患者団体が特定の組織から支援を受けているケースもあり、注意が必要となる。

また、経済的COIに当てはまらない患者・市民特有のCOIとして、主治医や、関係している医療者が診療ガイドライン作成グループに含まれる可能性があることが挙げられる。その場合も事前に申告してもらい、委員として診療ガイドライン作成グループに参加することが適切かどうか判断する必要がある。ただし、診療ガイドライン作成グループに関係する医療者がいるからといって、必ずしも委員を外れる必要はなく、その点を考慮して、委員長などが対応をよく検討することが求められる。

④ 作成グループの委員長および委員に求められること

患者・市民は当該疾患の当事者としての経験を持つエキスパートであり、診療ガイドライン作成における重要なステークホルダーである。一方で、患者・市民委員は、診療ガイドライン作成に関する知識などについて、サポートを必要とする場合もある。委員長や委員には、知識や技能の面で生じる障壁をできる限り減らすよう努め、患者・市民委員がより意見を表明しやすくなるよう、適切なサポートを提供することが求められる。

まずは、作成プロセスの開始前に、委員長を含めた委員全員が、患者・市民参画の意義を共有し、診療ガイドライン作成に関わる患者・市民と医療者は対等な立場であることを確認しておくことが望ましい。患者・市民委員がグループのメンバーとして尊重されていることを感じられるような、包摂的な環境を作ることに留意し、発言しやすい雰囲気作りや働きかけを常に行っていくことが重要である。ただし、患者・市民委員の全ての意見を取り入れることは必須ではなく、患者・市民の価値観・希望などを推定する際の1つの検討材料として、十分に吟味することが最も重要である。当該診療ガイドライン作成において反映しない場合であっても、今後の参考・課題として記録に残すことに意味がある。

〈患者・市民委員への配慮の例〉

- ・患者・市民委員が診療ガイドライン作成グループに参加する前の段階あるいは初期の段階で、疾患に関する基礎知識や、どのような意見が求められているか、不安や懸念がないかどうか、などについて、話し合いの場を持つ。
- ・発言の機会を意識的に作るよう努める。
- ・議論の中で、随時論点を整理し、患者・市民委員が分からぬことや知らない専門用語があれば説明する。
- ・患者・市民委員の「沈黙」は、必ずしも「意見がない」「賛成している」わけではないことに留意し、患者・市民委員に丁寧に問い合わせるよう努める。
- ・立場の違いを強調するような呼称になるべく使わない（例：患者さんに「様」をつけない、○○先生や○○教授ではなく○○さんという呼称に統一する）。
- ・患者・市民委員が、孤立した席ではなく、委員長の近くや全体の輪の中心に座れるよう配慮するなど、安心して発言できる環境を整える工夫を行う。

⑤ 患者・市民委員としての参加のための準備——患者・市民へのサポート

診療ガイドライン作成グループへの効果的な患者・市民参画を実現するためには、患者・市民委員が診療ガイドラインの基礎知識や大勢の中で意見を表明するためのコミュニケーションスキルを学ぶ研修など、一定の準備作業が必要と考えられる。また、過去に作成グループに委員として参加したことのある経験者と話をする機会を設けることや、作成プロセスが進む中で困りごとがあつた際に相談できる担当者を決めておくことなども、患者・市民委員が自分自身の考えを十分にグループ内で述べるために重要である。

2.6.6 患者・市民の価値観・希望などに関するさまざまな情報収集

スコープ作成における患者・市民由来の課題・疑問の抽出や、推奨作成のためのエビデンスの探索の一環として、さまざまな方法で患者・市民から情報を収集することが望ましい。作成グループへ患者・市民を委員として迎えることが難しい場合でも、こうした方法で患者・市民の視点についての情報収集を行い、診療ガイドラインに反映していくことも重要である。また、やり方によっては少ない人的、財政的リソースで実施できるものもあり、作成グループの状況に応じて、柔軟に組み合わせていくことができる。

① インタビュー

対面や電話などを通じて、患者・市民にインタビューを行い、情報を収集する。1対1のインタビューの他、複数名の患者・市民が参加するフォーカスグループインタビューを行うことも考えられる。

② アンケート

郵送やウェブなどで経由で質問紙調査を行い、情報を収集する。関連する患者団体などの協力を得て、団体会員を対象にアンケートを実施することも検討できる。倫理審査委員会で承認を受けて、研究として論文化できれば、当該の診療ガイドライン作成だけでなく、他の同様の取り組みを行う組織の資料としても役立てることができる。

③ 文献検索

疾患や特定の治療に関する患者・市民の課題や疑問、価値観・希望、重要視する点、選好やニーズ、また心理・社会的な問題などを扱っている論文を収集し（質的研究含む）、評価・精査を行い、必要に応じて結果を統合するなどして、スコープや推奨の作成に活用する。

④ パブリックコメント

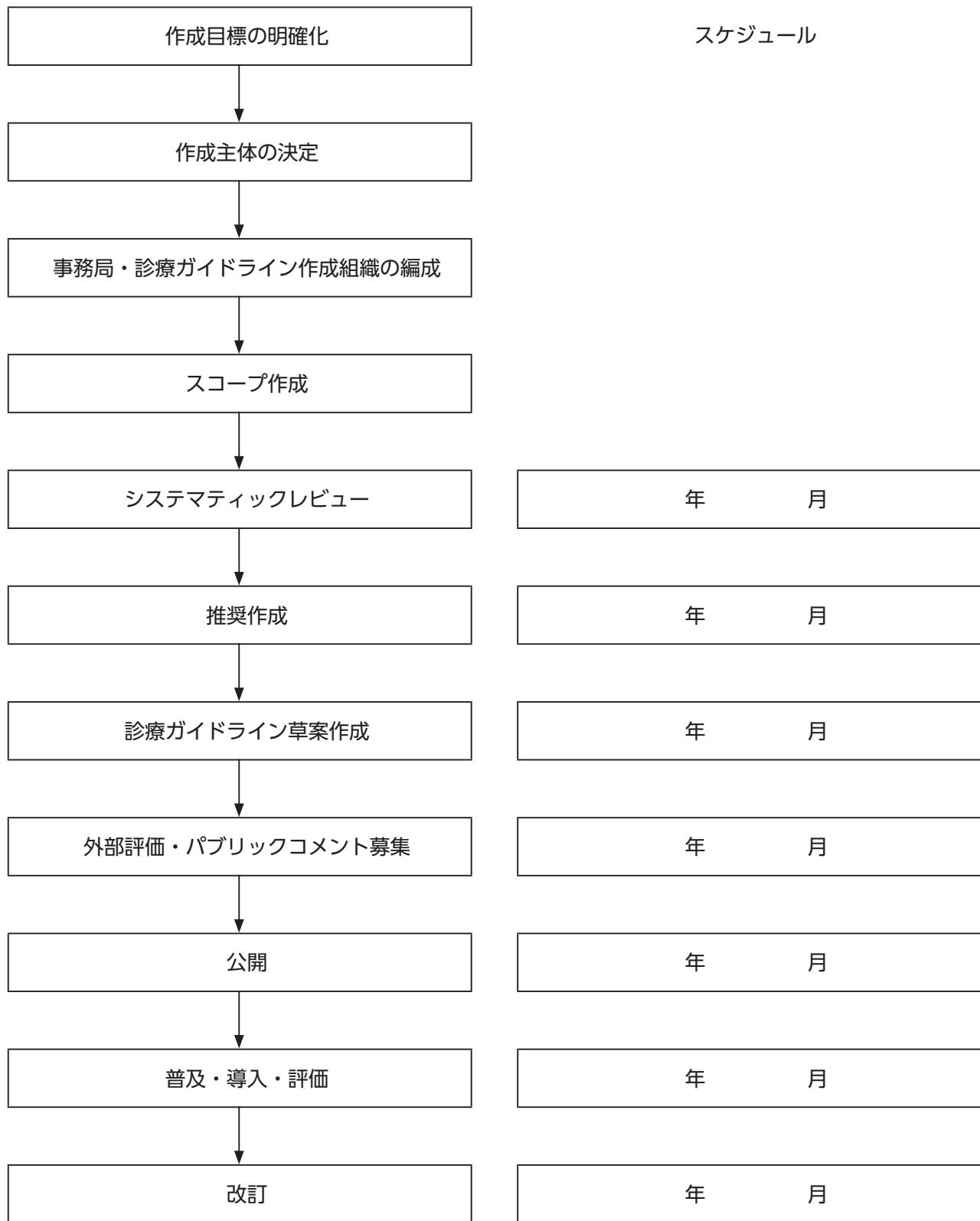
スコープや推奨、診療ガイドラインの草案が完成した時点で、一般市民から広く意見を募集する。パブリックコメントでは、より多くのさまざまな立場の人から意見を募り、集まった意見に対しては、それぞれ回答を作成し、公表することが望ましい。

⑤ その他

診療ガイドラインが取り扱うテーマについて、患者・市民から医療者や医療機関、患者団体などに寄せられた問い合わせ内容などから、患者・市民が疑問に感じやすい点や困りごとなどを把握し、どのようなCQを立てるべきか、スコープ作成の段階などで参考にすることもできる。

2章 テンプレート

【PR-1 診療ガイドライン作成手順およびスケジュール】



【PR-2 COI 申告① 経済的 COI 申告書】

診療ガイドライン名	
氏 名	
所 属	

上記の診療ガイドラインに関連する____年____月____日から____年____月____日の期間の企業・組織・団体との経済的関係について以下の通り申告する。

関連項目	申告基準	COI の有無	本人／家族	時期	企業・組織・団体名	備考
役員・顧問職	万円以上/年	有・無	本人 家族			
株式	万円以上/年	有・無	本人 家族			
特許権使用料	万円以上/年	有・無	本人 家族			
謝金・講演料	万円以上/年	有・無	本人			
原稿料	万円以上/年	有・無	本人			
研究費（受託・共同研究費）	万円以上/年	有・無	本人			
奨学（奨励）寄付金	万円以上/年	有・無	本人			
寄附講座	所属の有無	有・無	本人			
その他 ()	万円以上/年	有・無	本人 家族			

申告日： _____ 年 _____ 月 _____ 日

署名： _____ 印 _____

【PR-3 COI 申告② 経済的 COI 以外の COI 申告書】

経済的 COI 以外の COI の種類	COI の内容

【PR-4 COI 開示 経済的および経済的 COI 以外の COI 開示サマリーと対応方針記載表】

COI 開示サマリー	
経済的 COI 開示サマリー	
上記への対応方針	
経済的 COI 以外の COI 開示サマリー	
上記への対応方針	

【PR-5 診療ガイドライン作成資金】

費用項目	予算	資金提供者	備考

【PR-6 診療ガイドライン作成組織】

(1) 診療ガイドライン 作成主体	学会・研究会名			
	関連・協力学会名			
	関連・協力学会名			
	関連・協力学会名			

(2) 診療ガイドライン 統括委員会	代表	氏名	所属機関／専門分野	所属学会	作成上の役割

(3) 診療ガイドライン 作成事務局	代表	氏名	所属機関／専門分野	所属学会	作成上の役割

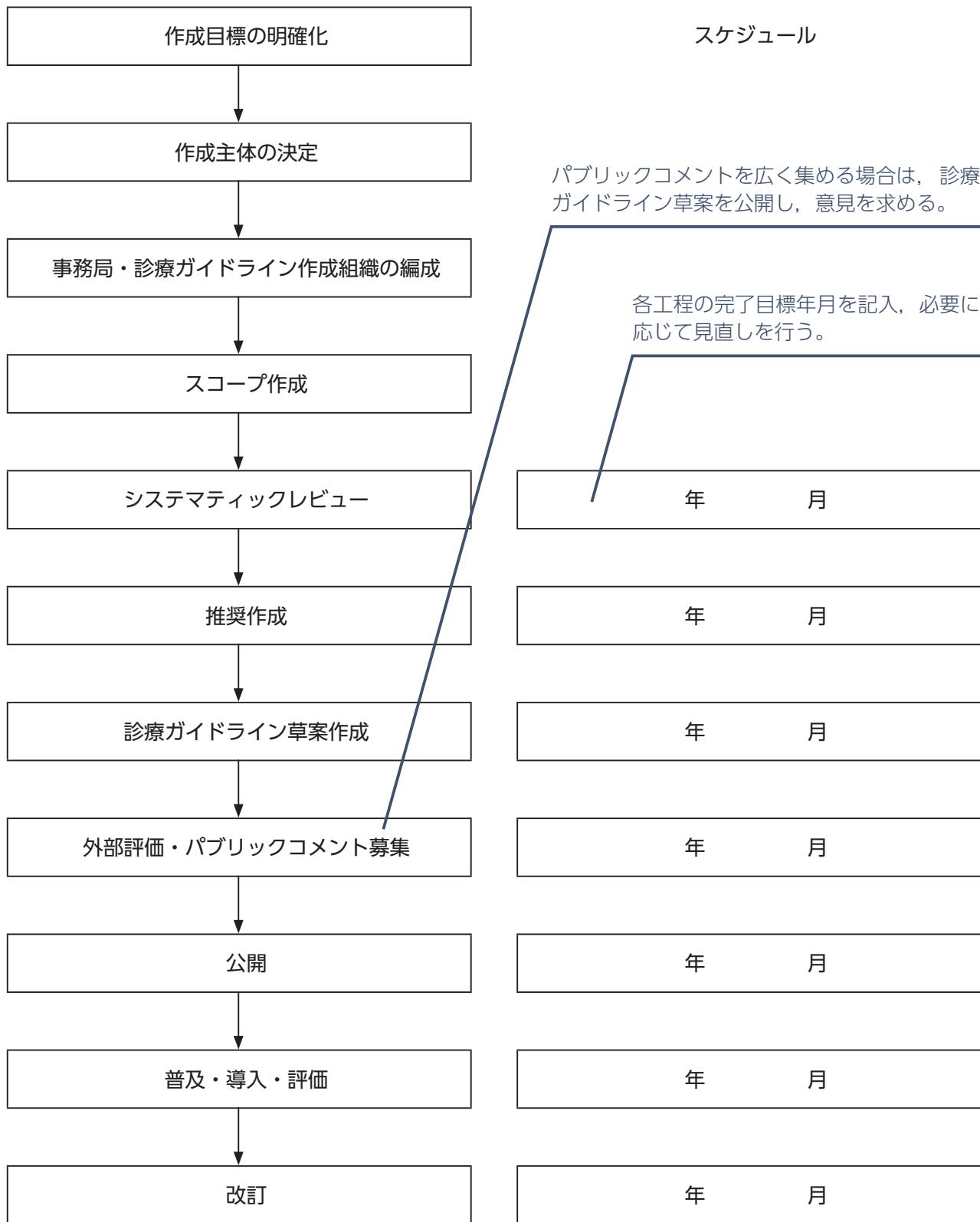
(4) 診療ガイドライン 作成グループ	代表	氏名	所属機関／専門分野	所属学会	作成上の役割

(5) システムティック レビューチーム	代表	氏名	所属機関／専門分野	所属学会	作成上の役割

(6) 外部評価委員会	代表	氏名	所属機関／専門分野	所属学会

2章 テンプレート 記入方法

【PR-1 診療ガイドライン作成手順およびスケジュール 記入方法】



【PR-2 COI 申告① 経済的 COI 申告書 記入方法】

診療ガイドライン名	
氏 名	
所 属	

上記の診療ガイドラインに関連する____年____月____日から____年____月____日の期間の企業・組織・団体との経済的関係について以下の通り申告する。

初回は、直近3年分について記載する。

関連項目	申告基準	COI の有無	本人 ／ 家族	時期	企業・組織・団体名	備考
役員・顧問職	万円以上/年	有・無	本人 家族			COI がある場合、 関係が生じた時期 について記載する。
株式	万円以上/年	有・無	本人 家族			
特許権使用料	万円以上/年	有・無	本人 家族			
謝金・講演料	万円以上/年	有・無	本人			
原稿料	万円以上/年	有・無	本人			
研究費(受託・共同研究費)	万円以上/年	有・無	本人			
奨学(奨励)寄付金	万円以上/年	有・無	本人			・実際に関係があったのが 本人か家族か記載する。 ・家族は一般に2親等以内
寄附講座	所属の有無	有・無	本人			
その他()	万円以上/年	有・無	本人 家族			

申告日： _____ 年 _____ 月 _____ 日

署名： _____ 印 _____

[PR-3 COI 申告② 経済的 COI 以外の COI 申告書 記入方法]

経済的 COI 以外の COI の種類	COI の内容

経済的 COI 以外の COI について、種別・内容を記入する。
分類が難しい場合は、内容のみを記載してもよい。

【PR-4 COI 開示 経済的および経済的 COI 以外の COI 開示サマリーと対応方針記載表 記入方法】

COI 開示サマリー	
経済的 COI 開示サマリー	対象となるトピックスに関連する COI をまとめて公開してもよい。公開の基準は明記するが、公開方法は作成組織で決定する。
上記への対応方針	
経済的 COI 以外の COI 開示サマリー	
上記への対応方針	

【PR-5 診療ガイドライン作成資金 記入方法】

診療ガイドライン作成に必要な費用項目を挙げ、予算および資金提供者も含めて記載する。

【PR-6 診療ガイドライン作成組織 記入方法】

(1) 診療ガイドライン 作成主体	学会・研究会名	対象となるテーマ／トピックスに関連する複数の 関連・協力学会で作成する場合に記載する。			
	関連・協力学会名				
	関連・協力学会名				
	関連・協力学会名				
(2) 診療ガイドライン 統括委員会	代表	氏名	所属機関／専門分野	所属学会	作成上の役割
(3) 診療ガイドライン 作成事務局	代表	氏名	所属機関／専門分野	所属学会	作成上の役割
(4) 診療ガイドライン 作成グループ	代表	氏名	所属機関／専門分野	所属学会	作成上の役割
(5) システムティック レビューチーム	代表	氏名	所属機関／専門分野	所属学会	作成上の役割
(6) 外部評価委員会	代表	氏名	所属機関／専門分野	所属学会	

2章 文献

AGREE Next Steps Consortium. (2009) The AGREE II Instrument [Electronic version]. <http://www.agreertrust.org> (参照 2013/02/28)

AGREE Next Steps Consortium. (2017) The AGREE II Instrument [Electronic version]. <http://www.agreertrust.org> (参照 2019/10/10)

Guidelines International Network (G-I-N). (2015) G-I-N Public Toolkit: Patient and Public Involvement in Guidelines. <https://g-i-n.net/document-store/working-groups-documents/g-i-n-public/toolkit/toolkit-2015> (参照 2021/02/24)

Guyatt G, et al. (2010) The vexing problem of guidelines and conflict of interest: a potential solution. Ann Intern Med 152:738-741. doi: <10.7326/0003-4819-152-11-201006010-00254>. PMID: <20479011>

International Committee of Medical Journal Editors. (2019) Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals (Updated December 2019) <http://www.icmje.org/> (参照 2021/03/03)

Institute of Medicine (IOM). (2009) Conflict of Interest in Medical Research, Education, and Practice. National Academies Press

Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press

Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for Systematic Reviews. National Academies Press

Melissa J, et al. (2017) Framework for enhancing clinical practice guidelines through continuous patient engagement. Health Expect 20:3-10. doi: <10.1111/hex.12467>. PMID: <27115476>

Minds 患者・市民支援部会. (2017) 「診療ガイドライン作成への患者・市民の参加」の基本的な考え方. <https://minds.jcqhc.or.jp/docs/minds/guideline/pdf/Proposal3.pdf> (参照 2021/02/24)

National Institute for Health and Care Excellence (NICE). (2015) NICE's approach to public involvement in guidance and standards: a practical guide. <https://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/Public-involvement-programme/PIP-process-guide-apr-2015.pdf> (参照 2021/02/24)

Schünemann HJ, et al. (2009) An official American Thoracic Society Policy statement: managing conflict of interest in professional societies. Am J Respir Crit Care Med 180:564-580. doi: <10.1164/rccm.200901-0126ST>. PMID: <19734351>

Schünemann HJ, et al. (2015) Guidelines international network: principles for disclosure of interests and management of conflicts in guidelines. Ann Intern Med 163:548-553. doi: <10.7326/M14-1885>. PMID: <26436619>

Scottish Intercollegiate Guidelines Network (SIGN). (2019) SIGN 100: A handbook for patient and carer representatives. <https://www.sign.ac.uk/assets/sign100.pdf> (参照 2021/02/24)

浅見万里子 他. (2007) 診療ガイドライン作成過程への患者・支援者参画のためのガイドライン “Patient Involvement Guidelines” (略称：PIGL) . <http://www.kanjyakai.net/manage/wp-content/themes/kanjakai/docs/guideline01.pdf> (参照 2021/02/24)

厚生労働省. (2018) 厚生労働科学研究における利益相反の管理に関する指針 . <https://www.mhlw.go.jp/stf/seisaku/jouhou-10600000-Daijinkanboukouseikagakuka/0000152586.pdf> (参照 2019/10/10)

全国医学部長病院長会議. (2018) 医学系研究機関における組織 COI 管理ガイドンス . https://www.ajmc.jp/pdf/20190425_01.pdf (参照 2021/03/03)

日本医学会利益相反委員会. (2017) 診療ガイドライン策定参加資格基準ガイドンス . http://jams.med.or.jp/guideline/clinical_guidance.pdf (参照 2021/03/03)

日本医学会利益相反委員会. (2020) COI 管理ガイドンス . http://jams.med.or.jp/guideline/coi_guidelines_2020.pdf (参照 2021/03/03)

第3章 スコープ

Minds

3章 Key Words

- ◎ Delphi (デルファイ) 法
- ◎ EtD フレームワーク (evidence to decisions framework)
- ◎ GPS (good practice statement)
- ◎ GRADE (The Grading of Recommendations Assessment, Development and Evaluation)
 - GRADE ワーキンググループ (GRADE Working Group)
- ◎ individual perspective (個人の立場)
- ◎ International Guideline Library
- ◎ MEDLINE
- ◎ National Guideline Clearinghouse (NGC)
- ◎ National Institute for Health and Care Excellence (NICE)
- ◎ PICO
- ◎ population perspective (集団の立場)
- ◎ PROSPERO (International Prospective Register of Systematic Reviews)
- ◎ PubMed
- ◎ SoF 表 (summary of findings table)
- ◎ The Cochrane Library
- ◎ アウトカムの重要性
- ◎ 医中誌 Web
- ◎ 益と害 (benefit and harm)
- ◎ 益と害のバランス
- ◎ エビデンス総体 (body of evidence)
- ◎ エビデンス総体の質的統合
- ◎ エビデンス総体の総括 (overall evidence)
- ◎ エビデンス総体の量的統合
- ◎ エビデンスの確実性
- ◎ 介入研究 (interventional study)
- ◎ 外部評価
- ◎ 観察研究 (observational study)
- ◎ 患者アウトカム
- ◎ クリニカルクエスチョン (clinical question : CQ)
- ◎ 最終調整 (finalization)

◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
◎ システマティックレビューチーム (systematic review team : SR チーム)
◎ システマティックレビュー論文
◎ 重要臨床課題 (key clinical issues)
◎ 診療アルゴリズム
◎ 診療ガイドライン (clinical practice guideline : CPG)
◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
◎ 診療ガイドラインデータベース
◎ 診療ガイドライン統括委員会
◎ 推奨
◎ スコーピングサーチ
◎ スコープ (scope)
◎ 前景疑問 (foreground question)
◎ バイアスリスク (risk of bias)
◎ 背景疑問 (background question)
◎ 非直接性 (indirectness)
◎ 非ランダム化比較試験 (non-randomized controlled trial : non-RCT) フォーマルな合意形成 (formal consensus formation)
◎ メタアナリシス (meta-analysis, メタ解析と同義)
◎ ランダム化比較試験 (randomized controlled trial : RCT)

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】

SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】

RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】

PO : post 【第8章 診療ガイドライン公開後の取り組み】

3.0 概要

スコープ（scope）は、診療ガイドライン（clinical practice guideline : CPG）の作成にあたり、診療ガイドラインがカバーする内容に関する事項、システムティックレビュー（systematic review : SR、系統的レビューと同義）に関する事項、推奨作成から公開に向けた最終調整（finalization）、公開に関する事項などを明確にするために作成される文書であり、診療ガイドライン作成の企画書ともいえる文書である。

スコープの作成に先立ち、診療ガイドラインで取り上げる疾患トピック^{*1}の基本的特徴を整理する。疾患トピックの基本的特徴には、疾患トピックの臨床的特徴や疫学的特徴、現在の診療の全体的な流れが含まれる。また、現在行われている診療の全体的な流れは、診療アルゴリズムという形で図示することが望ましい。これらを基にして、作成する診療ガイドラインで取り上げるべき臨床上の課題が何であるかを明確にし、スコープを作成する。

本マニュアルでは、一連のプロセスを明示するため、下記のボックス内のテンプレートを提案する。

【SC-1 疾患トピックの基本的特徴】

1. 疾患トピックの臨床的特徴
2. 疾患トピックの疫学的特徴
3. 疾患トピックの診療の全体的な流れ

【SC-2 診療アルゴリズム】

診療アルゴリズム（図）

【SC-3 スコープ】

1. 診療ガイドラインがカバーする内容に関する事項
2. システマティックレビューに関する事項
3. 推奨作成から公開に向けた最終調整、公開までに関する事項

【SC-4 CQ の設定】

○ ステップ

診療ガイドラインの作成は、主に診療ガイドライン作成グループ（guideline development group : GDG、診療ガイドライン作成パネルともいう）とシステムティックレビューチーム（systematic review team）によって進められる。診療ガイドライン作成の開始にあたって、疾患トピックの基本的特徴を整理してスコープを作成するのは、診療ガイドライン作成グループの役割である。スコープの草案を作成した段階で、外部評価委員による外部評価を実施する場合もある。また、診療ガイドラインの利用が想定される医療者、患者や、診療ガイドラインによって影響を受ける行政、団体など、全ての利害関係者の意見を聞くパブリックコメントのプロセスを経て確定することもある。スコープ作成の過程で変更の必要が生じた場合は、変更の理由、変更の承認プロセスを明記して変更を加え、日付とバージョンを明記して改訂版であることが分かるよ

* 1 疾患名そのものもあるが、特定の病態、または診断、治療、予防などのテーマに特化したトピックを取り上げることもあり得るので、「疾患トピック」とした。

うにすることが望ましい。診療ガイドライン作成グループの討議によってスコープが確定すると、それを基にしてシステムティックレビューのプロセスに進むため、原則としてスコープは確定後に内容を変更することはできない。ただし、スコープで作成したクリニカルクエスチョン（clinical question : CQ）は、システムティックレビューのプロセスの中で、システムティックレビューチームと診療ガイドライン作成グループとのやり取りを通じて変更される可能性がある。これについては後述する（**3.3 節**参照）。

医療経済性を当該診療ガイドラインで考慮するかどうかについては、スコープ作成の段階で決定する（**3.4 節**および**3.6 節**参照）。考慮しない場合は、その旨を理由とともに記載する。医療経済性を考慮する場合、**第5章**を参考に評価を実施する。

3.1 ステップ1：疾患トピックの基本的特徴・スコープ作成方針の決定

本マニュアルでは、疾患トピックの基本的特徴やスコープで記述すべき標準的な項目をテンプレートに示しているが、個々の診療ガイドラインの事情によって、項目の追加あるいは省略は可能である。どのような記載項目を採用するかは、診療ガイドライン作成グループの最初の会議において、診療ガイドライン統括委員会の決定事項を確認した上で方針を決定する。

● 手順

(1) 診療ガイドライン統括委員会が決定した事項の確認

診療ガイドライン統括委員会が決定した事項を確認する。統括委員会からの口頭による説明、文書による提示などの方法がとられる。

(2) 診療ガイドライン作成プロセスの全体像の確認

第1章を参照し、作成プロセスの全体像を確認する。患者・市民を代表する立場の委員など、診療ガイドライン作成に関して予備的な知識を持たない委員が参加することを前提に、委員全体で全体像を確認することが望ましい。

(3) 疾患トピックの基本的特徴およびスコープで記載する事項の決定

本マニュアルが提案するテンプレートをベースにしつつ、記載項目の追加あるいは省略について決定する。

3.2 ベースとなる疾患トピックの整理

診療ガイドラインが取り上げる疾患トピックの基本的特徴を、臨床的特徴、疫学的特徴、診療の全体的な流れについて整理する。診療の全体的な流れについては、「診療アルゴリズム」として図示することを推奨する。これらの事項を整理することにより、診療の現状を把握し、診療ガイドラインが取り上げるべき臨床上の課題を重要臨床課題（key clinical issues、[ステップ3 参照](#)）として明確にすることが可能となる。これらは診療ガイドライン作成の「背景」に当たる部分であり、診療ガイドライン本体の冒頭に掲載する。

診療ガイドラインに、推奨の理解のために必要な知識についての背景疑問（[ステップ3 参照](#)）を設定し、その解説を含めることを計画する場合には、疾患トピックに関する記述の重要性が高くなる。その場合には、スコープ作成だけに必要な知識の範囲を超えると考えられ、十分な情報の収集とまとめが重要となるので、スコープの疾患トピックにはその短縮版を用いることも検討する。

● 手順

(1) 臨床的特徴、疫学的特徴の検討

臨床的特徴、疫学的特徴は、教科書のように網羅的な解説は必要なく、診療ガイドラインの推奨を理解するために必要な事項を取り上げる。本マニュアルでは作業の効率化を図るために、診療ガイドライン作成グループ会議の前に草案を作成することを前提としている。診療ガイドライン作成グループ会議では、草案を基に検討を進める。また、この部分の作成はスコープ作成の後半で行ってもよい。

臨床的特徴としては、病態生理、臨床分類、歴史的事項などが主な記載事項となる。疫学的特徴としては、罹患率、死亡率、受療率、生存期間などの現状、経年変化、地域特性などが主な記載事項となる。

(2) 診療の全体的な流れの検討

「診療の全体的な流れ」とは、診療ガイドラインが取り上げる疾患トピックにおける診療の全体像を、疾患の臨床分類と、分類された各群の患者への治療の選択という形式でまとめたものである。

事前に作成した草案を基に、診療ガイドライン作成グループ会議で検討する。草案作成にあたっては、新しいエビデンスを把握し、確実に反映させるために、[表3-1](#)に示すさまざまな方法がとられる。委員による提案は重要な情報であるが、患者代表などの委員構成による偏りには注意を要する。スコーピングサーチは偏りのない情報が得られる利点があり、最も重要な情報源となる。ただし、検索法に関する技術を要する（具体的方法は次項を参照）。学会内外からの意見募集も診療現場のニーズを取り上げができる利点がある。ただし、時間がかかるために、時間に余裕がない場合には、スコープ草案作成後のコメント募集として実施する方法もある。

表 3-1 診療の全体的な流れを把握するための情報収集の方法

委員による提案	診療ガイドライン作成グループ委員の専門知識に基づいて提案を出し合う。
スコーピングサーチ	文献検索による予備的な情報収集を実施する。海外で公表された診療ガイドライン、システムティックレビュー論文 (Cochrane など)、RCT などが検索の対象となる。
意見募集	学会員あるいは学会外から広く意見を募集する。

(3) スコーピングサーチの方法

スコーピングサーチとは、スコープを作成する際に新しいエビデンスを把握し、確実に反映させるために、偏りのない文献検索による予備的な情報収集を実施する方法である。国内・海外で公表された診療ガイドライン、システムティックレビュー論文 (Cochrane など)、ランダム化比較試験 (randomized controlled trial : RCT) などが検索の対象となる。この段階では網羅的に検索する必要はなく、疾患トピックや介入方法に関するキーワードに Publication Type を掛け合わせるなどして、それぞれの文献数や出版状況を把握する。スコーピングサーチの結果、RCT でのシステムティックレビューが難しい場合は観察研究 (observational study) の検索と、それを用いたシステムティックレビューをシステムティックレビューチームに指示することになる。

1) 先行する診療ガイドラインの検索

国内・海外の先行する診療ガイドラインの検索は診療ガイドラインデータベースを用いる。**表 3-2** に示すように、ECRI Guidelines Trust[®]、National Institute for Health and Care Excellence (NICE) Evidence Search などが対象となる。国内の網羅的なデータベースとしては「東邦大学・医中誌診療ガイドライン情報データベース」がある。他の作成団体による先行する診療ガイドラインが公表されている場合は十分に吟味し、整合性についてあらかじめ検討する必要がある。なお、質の高い選定された診療ガイドラインは、Minds ガイドラインライブラリに収載されている。

さらに、PubMed で疾患名に Practice Guideline [Publication Type] を掛け合わせて検索する。

表 3-2 主な診療ガイドライン検索のためのデータベース

データベース名	組織名	国名	URL	閲覧要件
ECRI Guidelines Trust [®]	ECRI Institute	米国	https://guidelines.ecri.org/	要登録
NICE Evidence Search	National Institute for Health and Care Excellence (NICE)	英国	https://www.evidence.nhs.uk/	なし
International Guideline Library	Guidelines International Network (G-I-N)	国際組織	http://www.g-i-n.net/	要契約
東邦大学・医中誌診療ガイドライン情報データベース	東邦大学医学メディアセンター・医学中央雑誌刊行会 (共同運営)	日本	https://guideline.jamas.or.jp/guideline/	なし
Minds ガイドラインライブラリ	(公財) 日本医療機能評価機構	日本	https://minds.jcqhc.or.jp/	なし

2) システマティックレビューの検索

既存のシステムティックレビューを探す場合、PubMed など文献データベースや The Cochrane Library の The Cochrane Database of Systematic Reviews を用いる。

PubMed でシステムティックレビューを探す場合、Systematic Review [Publication Type] を用いる方法と systematic [sb] というフィルターを用いる方法がある。いずれもメタアナリシス (meta-analysis, メタ解析と同義) は含まれていないので、必要な場合はさらに Meta-Analysis [Publication Type] を追加する。

また、システムティックレビューの国際的な登録データベースである PROSPERO^{*2} には進行中のシステムティックレビュー情報も含まれているので、必ず検索し、確認する。

(4) 診療アルゴリズムの検討

綿密なスコーピングサーチの結果などに基づき、診療の全体的な流れが把握できたら、診療アルゴリズムという形式で図示するのが効果的である。最も単純な診療アルゴリズムの例を図 3-1 に示す。ここでは、診療ガイドラインの対象疾患 X は、ステージ 1、ステージ 2 に分類される。ステージ 1 に対する治療としては、手術単独療法が確立している。一方、ステージ 2 に対する現状の治療法としては、手術単独療法と手術 + 放射線療法の 2 種類が治療選択肢 (alternative care options) として考えられる、という例である。このように、診療アルゴリズムを作成することにより、診療ガイドラインで取り上げるべき臨床上の課題（重要臨床課題）は何かを明確にすることができる。例えば、この例における疾患 X のステージ 2 のように、複数の選択肢が存在し、これらの益と害 (benefit and harm) のバランスが不明瞭な場合、重要臨床課題の候補となる。

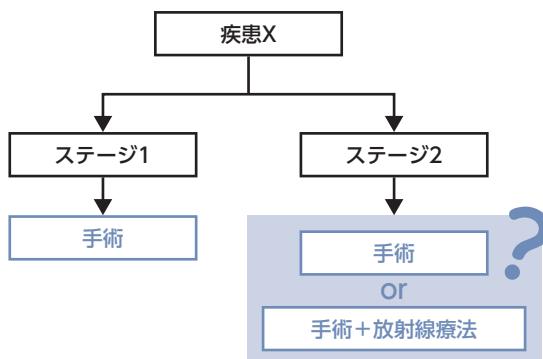


図 3-1 診療アルゴリズムの例

* 2 PROSPERO (<https://www.crd.york.ac.uk/PROSPERO>)

● テンプレート

[SC-1 疾患トピックの基本的特徴]

[SC-2 診療アルゴリズム]

● テンプレート 記入方法

[SC-1 疾患トピックの基本的特徴 記入方法]

[SC-2 診療アルゴリズム 記入方法]

● テンプレート 記入例

[SC-1 疾患トピックの基本的特徴 記入例]

[SC-2 診療アルゴリズム 記入例]

3.3

ステップ3:スコープ作成(診療ガイドラインがカバーする内容に関する事項① 重要臨床課題とCQ)

疾患トピックの基本的特徴が整理できたら、それに基づいてスコープの本体を作成する。本マニュアルでは、スコープの記載事項として、「診療ガイドラインがカバーする内容に関する事項」、「システムティックレビューに関する事項」、「推奨作成から最終調整、公開までに関する事項」の3項目に分けて記載することを提案している。

まず、「診療ガイドラインがカバーする内容に関する事項」を検討する。診療ガイドラインがカバーする内容に関する事項の項目には、「タイトル」、「目的」、「トピック」、「診療ガイドラインがカバーする視点」、「想定される利用者、利用施設」、「既存の診療ガイドラインとの関係」、「診療ガイドラインがカバーする範囲」、「重要臨床課題」、「CQリスト」などが含まれる。

これらの中でも、重要臨床課題とCQの設定は、作成した診療ガイドラインの出来不出来を左右する極めて重要なポイントである。このステップでは、重要臨床課題を決定し、それを基にしてCQを設定する手順について解説する。

診療の全体的な流れや診療アルゴリズムの検討結果に基づいて、診療ガイドラインが取り上げる実臨床における問題を重要臨床課題として設定する。重要臨床課題としては、現時点の診療プロセスにおいて複数の選択肢が存在し、これらの益と害のバランスが不明瞭な臨床上の課題を取り上げる。どのような臨床上の問題点を重要臨床課題として取り上げるかを明確にするためにも、綿密なスコーピングサーチを行った上で診療の全体的な流れを把握し、診療アルゴリズムを整理することが望ましい。

スコープで取り上げるべき重要臨床課題が決定したら、それを基にしてCQを設定する。ここでは、CQの定義を「診療ガイドラインで取り上げることが決まった重要臨床課題に基づいて、疑問の構成要素を抽出したもの」とする。CQを明確に設定することは、この後で行うエビデンスの収集や推奨作成のために非常に重要である。CQの構成要素は、エビデンス検索を行う際の検索キーワード設定の基礎となる（**第4章4.1節**参照）。また、構成要素を明瞭に設定することはエビデンスの非直接性（indirectness）の評価の際にも役立つ（**第4章4.3節**および**4.4節**参照）。さらに、CQ設定時点での介入の益と害に関するアウトカムを明確に定め、患者にとっての重要性を点数化しておくことは、推奨を作成する際に必要となる（**第6章6.2節**および**6.3節**参照）。

● 手順

重要臨床課題を決定し、それを基にしてCQを設定する手順の一例を以下に示す（図3-2参照）。まず重要な臨床課題を設定し、それを基にCQの構成要素〔PICO（P: Patients, Problem, Population, I: Interventions, C: Comparisons, Controls, Comparators, O: Outcomes）〕を抽出する。次に、抽出したO（アウトカム）の相対的重要性を評価し、最後に抽出した構成要素を用いてCQを1文で表現する。本項ではこの流れに沿って、重要臨床課題からCQを設定する手順を解説する。ただし、作成する診療ガイドラインの最終的な推奨の提示方法によっては必ずしもCQを文として表現する必要はなく、このプロセスは省略することも可能である。

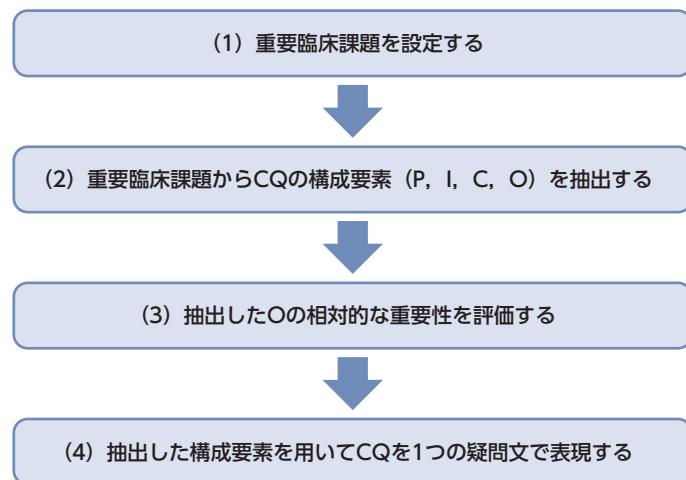


図3-2 CQ設定の手順

(1) 重要臨床課題の設定

診療の全体的な流れおよび診療アルゴリズムの検討結果に基づき、重要臨床課題を決定する。図3-3は、実臨床の場で生じるさまざまな疑問を診療ガイドラインの中でどのように扱うかを分類したものである。臨床的な疑問は、背景疑問（background question）と前景疑問（foreground question）に大別される（Straus SE 2010）。

まず、疾患の罹患率、症状、発症経過など、疾患トピックの背景となるような情報に関する疑問を背景疑問という。これらの臨床的・疫学的特徴に関する疑問は、重要臨床課題として取り上げる必要はなく、疾患トピックの基本的特徴の中に記載するだけでよい。

実臨床の場において医療行為を選択する意思決定に関する疑問を前景疑問という。前景疑問の中で、現時点の診療プロセスにおいて複数の選択肢が存在し、これらの益と害のバランスが不明瞭な臨床上の課題など、推奨を出すことで患者アウトカムの改善が期待されるものを重要臨床課題として取り上げる。診療ガイドラインで取り上げる重要臨床課題は、研究のための疑問ではなく、臨床上の意思決定において推奨を示す必要がある問題であることに留意されたい。このような重要臨床課題を基にCQを設定し、システムティックレビューを行い、エビデンス総体（body of evidence）の評価によって推奨を作成する。

ただし、利用者の知識や経験によって何を重要な課題と考えるかが変わるのは当然であり、医学的

にすでに明確な結論が出ているものであっても重要臨床課題に含まれることがあり得る。前述の診療アルゴリズム（図3-1参照）のように、疾患Xのステージ1に対する治療としては手術単独療法がすでに確立しており、他の選択肢を考慮する余地がないような場合など、システムティックレビューを行うまでもなく自明な疑問がこれに当たる。このような疑問に対する回答は、GPS（good practice statement）として、推奨とは明確に分けて提示することもある。GPSは「診療上の重要度の高い医療行為について、新たにシステムティックレビューを行わなくとも、明確な理論的根拠や大きな正味の益があると診療ガイドライン作成グループが判断した医療行為を提示するもの」と定義される（Guyatt GH 2015, Guyatt GH 2016, WHO 2014）。重要臨床課題を基にして作成したCQに対する回答をGPSとして提示するかどうかは慎重に判断する必要があり、表3-3に示した4項目を全て満たす場合にのみGPSとすべきとされる（Guyatt GH 2015, Guyatt GH 2016）。GPSの多用・誤用の懸念が示されており、その適用には細心の注意が必要である。

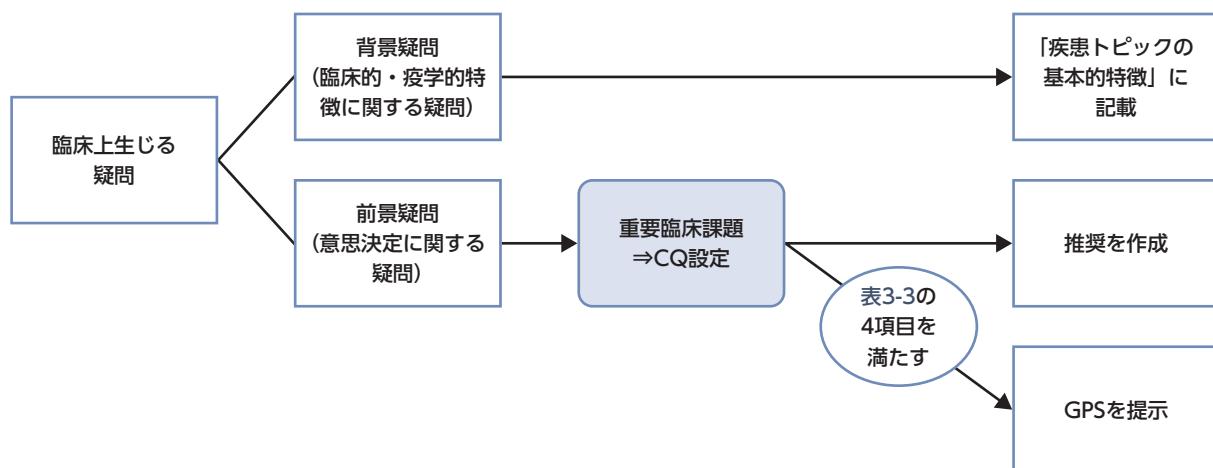


図3-3 疑問の分類と診療ガイドライン作成における扱い

表3-3 CQに対する回答をGPSとして提示するかどうか検討するための要件

どのような回答にも当てはまる要件
(i) 提示が明確であり、実行可能である。
GPSに特有の要件
(ii) 実臨床の場において真に必要なメッセージとなる。
(iii) 関連する全てのアウトカムと起こり得る結果を考慮した上で、GPSを導入することが広範な有益性をもたらす。
(iv) エビデンスを収集して要約するのは診療ガイドライン作成パネルの限られた時間と労力の無駄遣いである（機会損失が大きい）。
(v) 間接的証拠を結び付ける十分に裏付けされた明白な理論的根拠がある。
(ii)～(v)の全てに当てはまる回答がGPSになり得る。

出典：Guyatt GH, et al. (2016) Guideline panels should seldom make good practice statements: guidance from the GRADE working group. J Clin Epidemiol 80:3-7. より作成

1つの重要臨床課題から設定されるCQの数は、1つのこともあるが、複数の場合もある（図3-4参照）。最終的なCQ数の絞り込みについては（6）項に後述するが、この段階で重要臨床課題がある程度数を絞っておくのが望ましい。日常診療において医療行為を選択する意思決定の場面は数多くあるが、その全てについてシステムティックレビューを実施し、推奨を作成するのは非常に多くの労力を要するため、現実的ではない。そのため、診療ガイドラインでは実臨床における問題に対する回答を導き出したいものを厳選して重要臨床課題として取り上げることが望まれる。1つの診療ガイドラインで扱うことが可能な重要臨床課題の数は診療ガイドライン作成グループやシステムティックレビューチームの人数、経験やスキル、作成にかけられる資金や期間によって異なる。それぞれの重要臨床課題について、作成したい診療ガイドラインにとってどれほど重要であるか検討し、実際に取り上げるものリストに列挙する。そのためにも、綿密なスコーピングサーチを行った上で現状の診療の全体的な流れを整理し、優先的に重要臨床課題として取り上げるものを取り捨てる必要がある。

取り上げる重要臨床課題が決まったら、CQを設定して推奨を出すことで患者アウトカムの改善が期待される理由をスコープの中に記述する。また、結果的に取り上げないことになった重要臨床課題についても、同様にその理由を記述しておくことが望ましい。

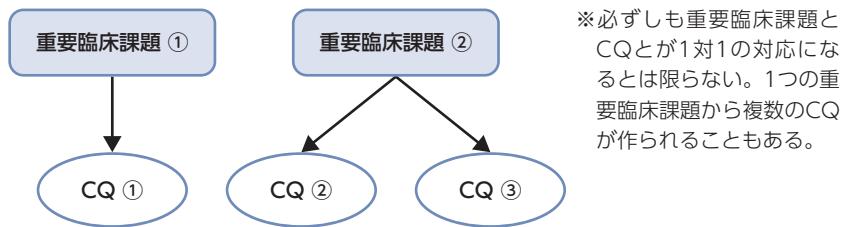


図3-4 CQ作成のイメージ

(2) CQの構成要素の抽出

重要臨床課題が決定したら、重要臨床課題ごとにCQの構成要素を抽出する。CQの構成要素として一般的に用いられているのは、PICOと呼ばれる形式である。ここでは、重要臨床課題からCQの要素であるPICOを詳細に抽出して、テンプレート【SC-4】に記載する方法を紹介する。

1) Pの設定

P (Patients, Problem, Population)とは、介入を受ける対象のことである。ここでは、年齢や性別などの患者特性や症状・病態だけでなく、地理的要件などの要素も考慮する。すなわち、Pとは、介入方法の選択が行われる「状況」そのものを指す。Pの設定で考慮すべきポイントを以下に示す。

- ・介入の対象となる患者特性（性別や年齢など）を明確にする。
- ・疾患や病態、症状などを詳細に設定する。
- ・特定の地理的要件などがあればここに加える。

- ・Pが広義にわたる場合には、必要に応じてCQを複数に分けることも検討する。例えば、対象とする患者の年齢によって介入の選択肢が異なる状況がある場合には、年齢層別にCQを設定することもある。

2) I／Cの設定

I (Interventions) とは、設定したPに対して行うことを推奨するかどうか検討する介入の選択肢である。C (Comparisons, Controls, Comparators) は、Iと比較検討したい介入である。2つの介入を比較する際にはどちらをIとしてどちらをCとするべきか判断しかねることもあると思われるが、この後システムティックレビューチームがシステムティックレビューを行う際に悩まないように、ここではどちらかをI、どちらかをCとして決めておくことを提案する。また、3つ以上の介入を同列に検討したい場合も、システムティックレビューで比較するIとCの組み合わせを列举しておくとよい（例えば、①I：介入その1・C：介入その2、②I：介入その1・C：介入その3、③I：介入その2・C：介入その3）。I／Cの設定で考慮すべきポイントを以下に示す。

- ・Pに対して行うことを推奨するかどうか検討したい介入の選択肢をリストアップする。
- ・介入の期間やアウトカムの評価時期などの時間的要素も考慮する。
- ・「無治療（プラセボ）」との比較を検討する場合、「無治療」も選択肢の中に入れ、選択肢の候補が1つだけにならないように注意する。
- ・どちらかをI、どちらかをCと決めておく。

3) Oの設定

O (Outcomes) とはI／Cで設定した介入を行った結果として起こり得るアウトカム事象（転帰事象）のことである。Oの設定で考慮すべきポイントを以下に示す。

- ・実臨床に基づき（practice-driven）、どの介入が最も推奨されるか判断するための基準となるアウトカムを網羅的にリストアップする。そのアウトカムに対する介入の効果を知ることによって、推奨が変わる可能性があるものが重要なアウトカムである。
- ・「益」のアウトカム（Iを行う方がCを行うより患者にとって望ましい結果になることが期待されるアウトカム）、「害」のアウトカム（Iを行う方がCを行うより患者にとって望ましくない結果になると考えられるアウトカム）の両方のアウトカムを取り上げる（Schünemann H 2013）。
- ・個人に重複して生起するアウトカムの取り扱いには注意が必要である。例えば、「死亡」というアウトカムと、手技の有害事象としての「脳内出血」というアウトカムを取り上げる場合には、両者に強い関連があることが考えられる。このような場合には、合成アウトカムを用いる、どちらかだけのアウトカムを用いる、または両方とも取り上げるなどの選択肢があるが、いずれにしても十分な検討が必要である。
- ・可能な限り、「代替アウトカム」ではなく、「患者にとって真に重要なアウトカム」を取り上げる。代替アウトカムとは、検査値の変化など、臨床医が重視するかもしれない代理、代替、生理的アウトカムである。患者にとって重要なアウトカムとは、生死や症状の変化など、患者自身

が重視するであろう直接的なアウトカムである。後述のテンプレート【SC-4】記入例の虫垂炎の例を取り上げると、虫垂炎の再発や入院期間などは患者にとって重要なアウトカムであるが、炎症反応（CRP値）は代理アウトカムであり、患者にとっての重要性は高くないといえる。

- ・ここでは、コストはアウトカムに含めない（コストについては第5章や第6章を参照）。
- ・アウトカム設定に、コアアウトカムセットの COMET database^{*3} を参考とする方法もある（NICE 2014）。コアアウトカムセットは、ある疾患トピックに関する臨床試験において、最低でも測定・報告されるべき標準的なアウトカムのセットである（Williamson PR 2012）。

(3) 抽出した〇の相対的な重要性の評価

以上のプロセスで検討するアウトカムが全てリストアップされたら、それぞれのアウトカムの重要性を評価して点数を付与し、必要に応じて数を絞り込む。

1) アウトカムの重要性の点数化

それぞれのアウトカムが「介入を受ける患者にとってどの程度重要と考えられるか」を評価する。点数は1～9点とし、得点が高いほどそのアウトカムは患者にとって重要性が高いとする。点数の判定は、診療ガイドライン作成グループの経験や既存の研究結果に関する予備知識などに基づいて主観的かつ相対的に行う。アウトカムの評価には患者の視点を取り入れることが望ましい。ここで付与した点数が1～3点であれば「重要性は高くない」、4～6点であれば「重要」、7～9点であれば「重大」として分類（図3-5参照）し、実際にシステムティックレビューに含むアウトカムは「重要」なものと「重大」なものを採用する（Schünemann H 2013）。なお、ここで決定するアウトカムの重要性の点数はシステムティックレビューの対象とするアウトカムを決定するためのものであり、この後の推奨決定までのプロセスにおいて複数回再評価し、変更する可能性がある。推奨作成時に患者・市民の価値観・希望によってアウトカムの重み付けをする場合には、CQ設定時のアウトカムの重要度はそのまま用いない。エビデンス総体の総括（overall evidence）の確実性の決定や推奨決定時の評価基準としてどのように重要性を決めるかは、各章の記載を参照のこと（第4章、第6章参照）。



図3-5 アウトカムの重要性の点数と分類

出典：Schünemann H, et al. (2013) 3. Selecting and rating the importance of outcomes. GRADE Handbook.
Permission by the editors. <https://gdt.gradepro.org/app/handbook/handbook.html> (参照 2021/01/22)
より作成

* 3 COMET database (<https://www.comet-initiative.org/>)

2) 点数化の一例

アウトカムの点数付与の具体例として、腎不全で高リン酸血症の患者に対するリン酸低下薬投与の有用性について検討したケースを紹介する。アウトカムの候補として、「死亡」、「骨折」、「軟部組織の石灰化による疼痛」、「疼痛を伴わない腹部膨満」の4つを取り上げたとする。診療ガイドライン作成グループ内で検討した結果、患者にとって極めて重大なアウトカムであると考えられる「死亡」には9点を付与し、「骨折」には7点を付与した。患者にとって重要だが重大とまではいえないと考えられる「軟部組織の石灰化による疼痛」には5点を付与した。また、患者にとっては重要な「疼痛を伴わない腹部膨満」には2点を付与した。ただし、これらはあくまで一例にすぎず、同じアウトカムであっても、検討する患者の状態や対象とする疾患によってアウトカムの重要性も異なってくるため、点数が変わり得ることに留意する。

3) Delphi（デルファイ）法によるアウトカムの重要性の点数決定の合意形成

各アウトカムの重要性の点数は、専門家や患者代表で構成された診療ガイドライン作成グループ内で合意を形成して決定する。一般に、グループ内の意見を集約する方法として、会議や審議会、パネルディスカッションでの意見交換が用いられることが多い。しかし、これらの方法はグループにおける権威者や発言力のある者など、特定の人物の意見が結果に過度の影響を及ぼす可能性がある。このような影響力を極力排除する手段のひとつとして、Delphi法がある（Dalkey N 1969）。Delphi法は、多数の専門家や個人にアンケート調査を行い、その結果を回答者にフィードバックし、さらにアンケート調査を繰り返すことにより全体の回答や意見を絞っていく方法である。ここでは、この手法を用いて各アウトカムの重要性の点数の合意形成を図るプロセスの例を紹介する（図3-6 参照）。

はじめに合意形成のルールを決めておく。例えば、「アウトカムの重要性の点数の範囲が3点以内に全員の回答が集約されるまで投票を繰り返す。ただし投票回数は最大5回までとする。採用する点数は最終投票結果の中央値とする。」などである。合意形成のルール設定ができたら、中立的な立場のファシリテーターを1人決め、各メンバーは投票用紙にそれぞれのアウトカムの重要性の点数と、その点数を付けた理由を記入して匿名で投票する。ファシリテーターはメンバーの投票結果を集約し、結果を各人にフィードバックする。各メンバーはフィードバックされた結果を参考にして点数を再検討し、再度投票する。この投票とフィードバック、点数の再検討というプロセスを、事前に設定した合意形成のルールに従って繰り返し行う。リストアップされた全てのアウトカムについての合意形成が終了するまでこの作業を行う。

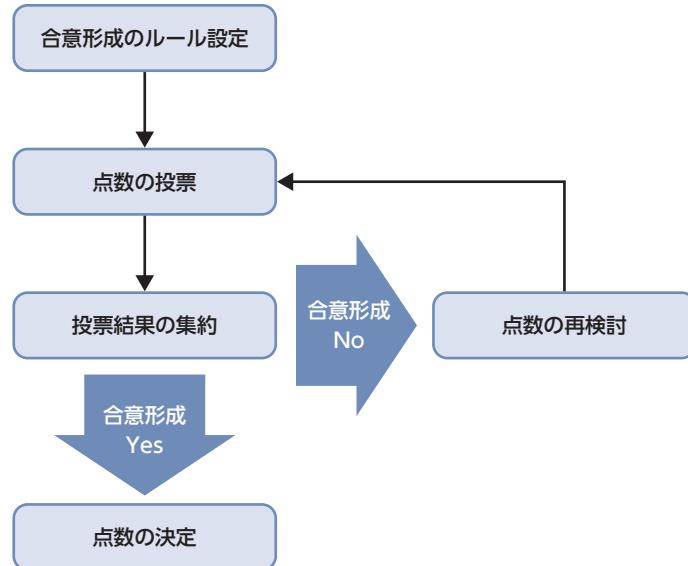


図 3-6 Delphi 法によるアウトカムの重要性の点数の合意形成フローの例

4) 採用するアウトカムの絞り込み

「重大」 または「重要」 に分類されたアウトカムの数が多い場合は、 必要に応じてアウトカムの絞り込みを行う。採用するアウトカムの数はシステムティックレビューを行うメンバーの経験やスキル、 診療ガイドライン作成にかけられる時間を考慮して決定する。ひとつの目安として、「重大」 または「重要」 に分類されたアウトカムのうち、 重要性の得点が高いものから最大7個程度を上限の目安として採用する。採用しなかったアウトカムについては、 その理由を記述しておくことが望ましい。

(4) 抽出した構成要素を用いた CQ の表現

以上のプロセスで抽出した構成要素 (P, I, C) を用いて CQ を 1 つの疑問文で表現する。ただし、 作成する診療ガイドラインの最終的な推奨の提示方法によっては、 必ずしも CQ を文で表現する必要はない。例えば、 CQ と推奨とを併記して Q & A のような形式で提示するのではなく、 章ごと・項目ごとに推奨一覧をまとめて提示するというような場合には、 CQ を文で表現したものを作成する必要はない。ただし、 その場合でも CQ の構成要素 (PICO と O の重要性) は明示しておく必要がある (本マニュアルで提案する **テンプレート【SC-4】** の CQ の構成要素の各項目を提示する)。

CQ を文で表現する際に考慮すべきポイントは、 以下のとおりである。

- ・ 1 つのセンテンスとする。
- ・ ? で終わる疑問文形式とする。
- ・ 「～推奨されるか」、「～有用か」といった表現を使用する。
- ・ I, C は可能であれば全て列挙する。
- ・ O は入れる必要はない。

特に、 文末の疑問表現には注意する必要がある。CQ の疑問に対応する回答として推奨文が作成されるが、 例えば「～は有効か？」 という疑問表現の CQ を作成した場合、 その回答は「～は有効である」

または「～は有効でない」、「～は有効とはいえない」などになる。しかし、有効性が認められたとしても、副作用やコストを考慮すれば推奨できないという場合もあり、これらの回答は診療ガイドラインの推奨文としては不明瞭である。そのため、CQは「～推奨されるか？」などの言い回しで締め、対応する推奨文がI／Cで設定した治療を行った方がよいのか、行わない方がよいのか明確になるように配慮することが望ましい。「推奨されるか？」という表現の他には、「有用か？」などの表現が用いられることがある。また、CQで検討するIやCの数が非常に多く、1つの文の中に全てを書き入れると冗長になってしまう場合には、重要なものに絞るなどとして簡潔に表現することもある。ただし、そのような場合にも、CQの構成要素が明瞭に分かるように、そのリストを別途明示しておくことが必要である。

2つの介入を比較検討するCQを1つの文で表現する例として、「Pに対して、IとCのどちらを用いることが推奨されるか？」などが考えられる。なお、1つのCQに対する回答として作成される推奨の数は、1のことともあれば、複数のこともある。

(5) CQのチェック

以上のプロセスによりCQを作成したら、そのCQを基にしてシステムティックレビューを行うことができそうか、また推奨を決定するまでのプロセスを滞りなく完遂できそうかどうかを入念に検討する必要がある。再検討すべきCQの例として、以下のようなものが挙げられる。

・特定の治療・介入の実施が前提となっているCQ

「どのような症例に×××を投与するべきか」、「△△△が有効な患者の特徴は」、「○○○の適応は？」というように、特定の治療・介入を実施することを前提としたCQは、診療ガイドラインにおけるCQには適さない。診療ガイドラインでは、特定の臨床状況に対する治療・介入方法として複数の選択肢が存在し、これらの益と害のバランスが不明瞭な状況に対して推奨を出す。そのため、特定の臨床状況に対して治療・介入を選択する形にCQを構成しなおす必要がある。

・介入の選択肢が1つだけになっているCQ

上記の点とも関連するが、他の選択肢がないものに対して推奨を出す必要はない。そのため、無治療・無介入と比較する場合でも、Iを介入、Cを無治療・無介入と明示しておくことが重要である。

・「益」のアウトカムだけが設定されているCQ

推奨を決定する際には、検討する治療・介入の益と害のバランスとを考慮する。そのため、設定するアウトカムには、「益」だけではなく「害」のアウトカムも含まれている必要がある。

・構成要素が漠然としているCQ

PICOそれぞれがあいまいな表現（測定できないもの）になっていると、システムティックレビューの実施が難しくなるため、明瞭に提示するべきである。例えば、「重篤な副作用」というアウトカムを設定したとしたら、それが具体的に何を指すのかを明示しておく必要がある。

・全く同じ臨床状況に対する介入の CQ が複数立てられている場合

同じ臨床状況に対する治療・介入の選択に関して、複数の CQ が設定されることは好ましくない。例えば、「疾患 X のステージ 2 に対して手術を行うことが推奨されるか？」という CQ と、「疾患 X のステージ 2 に対して薬物療法を行うことが推奨されるか？」という CQ を別々に設定することは避けるべきである。当該の診療ガイドラインの利用者にとってみれば、結局のところ疾患 X のステージ 2 には手術と薬物療法のどちらの方がより推奨されるのかが分からなくなってしまうからである。この場合は、「疾患 X のステージ 2 に対して手術と薬物療法のどちらを行うことが推奨されるか？」というような CQ に統合して推奨を出すのが望ましい。

(6) CQ の絞り込みおよび変更

以上のプロセスにより、全ての重要臨床課題について CQ を作成し終えたら、実際に診療ガイドラインで取り上げる CQ の絞り込みを行う。重要臨床課題の絞り込みでも述べたが、CQ を基にしたシステムティックレビューの実施や推奨の作成には多くの労力を要するため、CQ もそれらの実現可能性を考慮して数を絞り込む必要がある。

実際に 1 つの診療ガイドラインで扱うことが可能な CQ の数は、診療ガイドライン作成グループやシステムティックレビューチームの人数、経験やスキル、作成にかけられる資金や期間によって異なる。それぞれの CQ について、作成したい診療ガイドラインにとってどれほど重要であるか検討し、実際に取り上げる CQ を選択し、スコープの CQ リストに列挙する。

また、スコープ確定後、スコープで設定した CQ を基にシステムティックレビューチームがシステムティックレビュー作成を行うが、エビデンスを検索・レビューするプロセスの中で、システムティックレビューチームの提案により CQ を変更することもある。その際には、システムティックレビューチームと診療ガイドライン作成グループとが協議して CQ を見直し、最終的にシステムティックレビューを実施する CQ を確定する。そのプロセスの中で、アウトカムの重要性評価を再度行う場合もある。通常、重大または重要なアウトカムの選定はスコープ作成段階で決定しておくことが原則である。しかし、当初は考慮していなかったものの、エビデンスのレビューを通じて明らかになった重要なアウトカム（例えば深刻な有害事象など）があった場合などは、そのアウトカムをリストに追加し、アウトカムの重要性を再評価する。また、必要に応じて採用するアウトカムの絞り込みも再度行う。

(7) 分析的枠組みの利用

重要臨床課題から CQ を定式化する際には、分析的枠組み（analytic framework）を用いて視覚的に重要臨床課題と PICO の臨床的位置付けを図示するのが効果的である。分析的枠組みは診療アルゴリズムとは異なり、臨床課題を解析するための枠組みである（Harris RP 2001, AHRQ 2014）。CQ の構成要素（PICO）を具体的なものにするアプローチであるとともに、この後行うシステムティックレビューの基盤として利用でき、また診療ガイドライン利用者が課題の捉え方の適切さを確認するのに役立つ。

図 3-7 は治療・介入を簡略化した分析的枠組みの例である。左側に患者集団（すなわち P）、その右上に治療・介入（I と C）を置く。治療・介入の副作用や有害事象は橢円の中に示し、治療・介入から波線でつなぐ。治療・介入のアウトカム（O）として、中間アウトカム・代替アウトカムは角丸四角形の中

に示し、真に重大なアウトカムは長方形の中に示す。①はその治療・介入が真に重大なアウトカムを改善するかどうかという包括的な疑問を表す。②は治療・介入が中間アウトカム・代替アウトカムを改善するかどうかという疑問、③は中間アウトカム・代替アウトカムの改善が真に重大なアウトカムの改善と関連するかどうかという疑問、④は治療・介入がどのような害をもたらすかという疑問を表す。

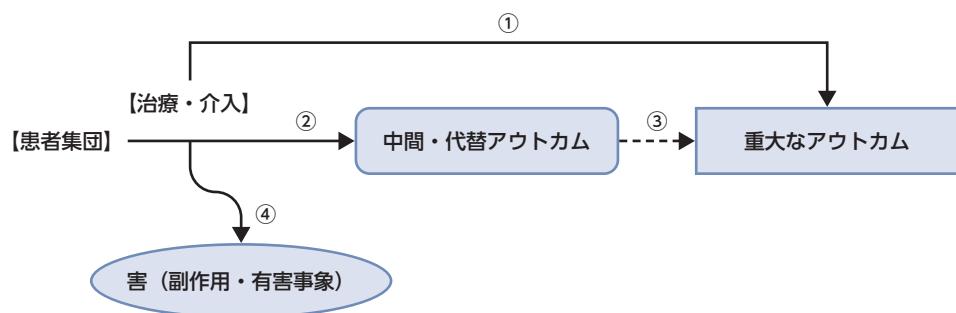


図 3-7 治療・介入を簡略化した分析的枠組みの例

出典：Chapter 4 Appendix A. In:Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov. p.83 より作成

● テンプレート

【SC-3 スコープ】

【SC-4 CQ の設定】

● テンプレート 記入方法

【SC-3 スコープ 記入方法】

【SC-4 CQ の設定 記入方法】

● テンプレート 記入例

【SC-3 スコープ 記入例】

【SC-4 CQ の設定 記入例】

3.4

ステップ4:スコープ作成(診療ガイドラインがカバーする内容に関する事項② その他の記載事項)

このステップでは、スコープの「診療ガイドラインがカバーする内容に関する事項」の残りの記載項目である、タイトル、目的、トピック、診療ガイドラインがカバーする視点、想定される利用者、利用施設、既存の診療ガイドラインとの関係、診療ガイドラインがカバーする範囲について検討する。

● 手順

(1) タイトルの検討

タイトルには、診療ガイドラインのタイトルを記載する。改訂版の場合は、改訂版であることがタイトルから分かることが望ましい。正式タイトルと同時に簡略タイトルを用いる場合は、それも記載する。

(2) 目的の検討

目的には、診療ガイドラインが全体として取り組む患者アウトカムの改善を記載する。

(3) トピックの検討

トピックには、診療ガイドラインが取り上げる疾患トピック・テーマを簡潔に記述する。ある疾患全体を取り上げる場合は疾患名を記載する。疾患全体ではなく、例えば、治療のみを取り上げる場合は、「小児急性中耳炎の治療」のような組み合わせで、取り上げるトピックを明記する。

(4) 診療ガイドラインがカバーする視点の検討

診療ガイドラインがカバーする立場 (perspective) には、患者個別の立場から推奨を考える場合 (individual perspective : 個人の立場) と、集団としての医療費などを考慮して推奨を考える場合 (population perspective : 集団の立場) が考えられる。どちらの立場で推奨を作成するかを記述する。

Individual perspective, population perspective の考慮の仕方を、以下に述べる。

Perspective は診療ガイドライン作成者の立場によって異なり、作成過程の多くの作業に影響を与える。特に、individual perspective, population perspective の違いは推奨作成に大きな影響を与え、同じ CQ に対する推奨が perspective によって異なるものになることがある (Alonso-Coello P 2016a)。EtD フレームワーク (evidence to decisions framework) の用語解説によると、perspective とは、「推奨または決断のための視点で、含める要素 (例えば、費用、効果、容認性、公平性) を決定する」とされている。

推奨を作成する組織が通常パネルが採用する perspective を決める (Alonso-Coello P 2016b) という考えが一般的に受け入れられている。例えば、特定の疾患を取り扱う学会や班会議などが作成主体の場合は、選択をする個々の患者と臨床家の意思決定を支援する展望を持って、individual perspective を採用することになる。

一方、英国NICEのような、国レベルの診療ガイドライン作成者であれば、その国の医療費の最適な使用を確実にする義務があるため、政府あるいは保健省のperspectiveを採用することになる。わが国でも、地方自治体、あるいは国レベルの検診やワクチン接種などに関しては、集団の立場（population perspective）が採用され得る。

また、同じ診療ガイドラインの中で、CQによって違うperspectiveが採用される場合もあり得る。例えば、医療費が特に高額な場合のみ、population perspectiveで推奨を作成するような場合である。

いずれのperspectiveを採用するかがEtDフレームワークの基準項目の選択に影響を与えることについては、GRADEワーキンググループ（GRADE Working Group）が発表しており（Alonso-Coello P 2016b）、各診療ガイドライン作成グループはこれを参考に基準項目を選択することができる（表3-4参照）。

表3-4 Individual perspectiveとpopulation perspectiveのEtDフレームワーク基準項目の選択の違い

基準項目	individual perspective	population perspective
1. 問題	個人の立場から見て問題は優先事項か？	集団の立場から見て問題は優先事項か？
2. 望ましい効果	望ましい効果の大きさはどれくらいか？	
3. 望ましくない効果	望ましくない効果の大きさはどれくらいか？	
4. エビデンスの確実性	効果のエビデンスの全体の確実性はどれくらいか？	
5. 値値観	主要なアウトカムに対する人々の価値観に重要な不確実性あるいはばらつきがあるか？	
6. 効果のバランス	望ましい効果と望ましくない効果のバランスは介入あるいは対照を支持するか？	
7. 必要資源量	介入の費用対効果（正味の望ましい効果に対する自己負担分）は介入あるいは対照を支持するか？	必要資源量（コスト）はどれくらいの大きさか？
8. 必要資源量に関するエビデンスの確実性		必要資源量（コスト）はどれくらいの大きさか、コストのエビデンスはどれくらいの確実か？
9. 費用対効果		介入の費用対効果のエビデンスの確実性はどれくらいか？
10. 公平性	(適用される機会は少ない)	健康の公平性に対する影響はどれくらいか？
11. 容認性	介入は患者、介護者、医療提供者に受け入れ可能か？	介入は全てのステークホルダーに受け入れ可能か？
12. 実行可能性	介入は患者、介護者、医療提供者にとって実行可能か？	介入は実行可能か？

出典：Alonso-Coello P, et al. (2016) GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ 353:i2089. より作成

特に、7. 必要資源量、8. 必要資源量に関するエビデンスの確実性、9. 費用対効果の基準項目が、2つのperspectiveで異なる。Individual perspectiveでは「介入の費用対効果（正味の望ましい効果に対する自己負担分）は介入あるいは対照を支持するか？」とされており、自己負担分のみを取り扱えば十分とされている。これら3つの基準項目を全て取り上げる必要はなく、必要資源量だけを取り上げ、そこで自己負担分を検討すればよいが、自己負担分だけでなく、保険や公的助成でカバーされる費用を含めた総医療費についても提示することが望ましい。また、individual perspectiveであっても、これら3つの基準項目を全て取り上げ、**第5章**で述べる費用効果分析の結果を用いる場合もある。

また、11. 容認性と12. 実行可能性の基準項目については、対象者をどこまで含めて判断するかという点が異なっている。

なお、1つのCQに対して両方の観点で推奨が作成される場合もあるが、最終的にどちらを採用するかは診療ガイドライン作成グループの決定による。

(5) 想定される利用者、利用施設の検討

想定される利用者および利用施設には、診療ガイドラインの活用が想定される医療者や施設などを明記する。医療者としては、医師／看護師／保健師／理学療法士などを記載する。施設としては、病院／診療所／介護施設／検診施設などを記載する。また、医療のステージとして、一次医療（プライマリケア）／二次医療／三次医療の中で限定される場合には明記する。

(6) 既存の診療ガイドラインとの関係の検討

既存の診療ガイドラインとの関係には、改訂版の場合に、どの診療ガイドラインの改訂版であるかを明示する。また、改訂版ではないが、関連する診療ガイドラインがある場合は、関係を明示する。例えば、診断ガイドラインがすでに作成されていて、今回作成する診療ガイドラインが同一疾患の治療ガイドラインである場合、その関連を明記する。

(7) 診療ガイドラインがカバーする範囲の検討

診療ガイドラインがカバーする範囲には、診療ガイドラインが取り上げるトピックについて、より詳細な定義を行い、記載する。診療ガイドラインがカバーする範囲とカバーしない範囲を整理して記述することが望ましい。例えば、トピックとして、「高血圧症」を取り上げる場合、「本診療ガイドラインがカバーするのは、本態性高血圧症である。二次性高血圧、妊娠高血圧は除外する。本診療ガイドラインが対象とする年齢範囲は成人であり、小児期の高血圧は対象としない。」などと明確に範囲を記述する。

● テンプレート

[SC-3 スコープ]

● テンプレート 記入方法

[SC-3 スコープ 記入方法]

● テンプレート 記入例

[SC-3 スコープ 記入例]

3.5 ベンチマーク5：スコープ作成（システムティックレビューに関する事項）

次に、「システムティックレビューに関する事項」を検討する。診療ガイドライン作成プロセスの全体像は**第1章図1-1**で示したとおりであり、診療ガイドライン作成グループが決定したスコープに従って、システムティックレビューチームがシステムティックレビュー作成を実施し、その後、システムティックレビューの結果を基に診療ガイドライン作成グループが推奨作成と診療ガイドラインの最終調整を進める。

システムティックレビューは診療ガイドライン作成グループとは独立したシステムティックレビューチームによって実施されることを想定しているため、スコープで決めるシステムティックレビューの基本方針は、診療ガイドライン作成グループからシステムティックレビューチームへの「指示書」の性格を持つ。ただし、システムティックレビューに関する基本方針は、スコープの中で診療ガイドライン作成グループが決定する場合もあるが、スコープでの決定は最小限として、システムティックレビューチームがシステムティックレビュープロトコルとして決定する場合もあり得る。また、診療ガイドライン作成グループとシステムティックレビューチームが協働して作成する方法も考えられる。

システムティックレビューに関してスコープで決定すべき事項は、実施スケジュール、エビデンスの検索、エビデンスの選出、エビデンスの評価と統合の方法などに関する基本方針である。

● 手順

(1) 実施スケジュールの検討

システムティックレビュー作成の所要日数は、エビデンスの検索とエビデンスの選出に要する日数、エビデンスの評価と統合に要する日数によって決まる。エビデンスの評価と統合に要する日数は、システムティックレビューチームのレビュアーハウス数と、CQ数、そして、CQごとに評価対象となるアウトカム数の総和で決まる。CQ数が多いと、それだけシステムティックレビューチームへの負荷が大きくなり、所要時間も長くなる。

(2) エビデンス検索の方法の検討

エビデンスタイプとして、①個別研究論文〔RCT、非ランダム化比較試験（non-randomized controlled trial：non-RCT）、観察研究など〕、②システムティックレビュー論文、③既存の診療ガイドラインの中で、どれを検索対象とするかを決定し、検索の優先順位を明確にする。また、個別研究論文については、RCT、non-RCT、観察研究などの研究タイプの中で、検索対象に含める研究タイプを決定する。

検索式を決定する際の基本方針を決定する。また、検索外で文献の追加を行う場合は、検索に偏りが生じない具体的方法について記載する。

検索の対象とするデータベースをエビデンスタイプごとに決定する。

検索対象期間は、データベースごとに、検索範囲となる期間を年月日で明示する。

(3) エビデンス選出の方法の検討

既存の診療ガイドライン、システムティックレビュー論文、RCT、non-RCT、観察研究をどのように選出し統合するか、優先順位と採用条件の基本方針を決定する。

採用条件については、診療ガイドライン、システムティックレビュー論文、RCT、non-RCT、観察研究のそれぞれについて基本的な考え方を決定する。

(4) エビデンスの評価と統合の方法の検討

エビデンス総体の評価方法を決定する。また、エビデンス総体の示す強さの表現方法を決定する。

エビデンス総体の質的統合と量的統合に対する基本的な考え方を決定する。

● テンプレート

[SC-3 スコープ]

● テンプレート 記入方法

[SC-3 スコープ 記入方法]

● テンプレート 記入例

[SC-3 スコープ 記入例]

3.6

ステップ6：スコープ作成（推奨作成から公開に向けた最終調整、公開までに関する事項）

最後に、「推奨作成から公開に向けた最終調整、公開までに関する事項」について検討する。この事項には、推奨作成方法、診療ガイドラインの最終調整の方法、外部評価の方法、公開の予定などが含まれる。

(1) 推奨作成方法の検討

推奨の作成は、診療ガイドライン作成グループが行う。コンセンサス形成のプロセスについて、オープンな討議による方法、投票による方法、フォーマルな合意形成（formal consensus formation）による方法などの中で採用する具体的方法を決定する。また、推奨決定の際に考慮する因子（エビデンスの強さ、益と害のバランス、患者・市民の価値観・希望の多様性、経済的視点など）を決定する。決断分析を行い、結果を推奨決定に反映させる場合には、その具体的方法を決定する。医療経済性を当該診療ガイドラインで考慮するかどうかについては、スコープ作成の段階で決定する。医療経済性を考慮する場合、**第5章**を参考に評価を実施する。考慮しない場合は、その旨を理由とともに決定する。推奨を文章で表現する際に準拠するルール、推奨の強さを表現する基準をあらかじめ決めておく。

推奨作成方法の詳細については、**第6章**を参照のこと。

(2) 公開に向けた最終調整の方法の検討

完成した診療ガイドラインの草案を最終決定するための手続きを決定する。外部評価の実施、パブリックコメントの募集などが含まれる。

(3) 外部評価の具体的方法の検討

外部評価として実施する方法を決定する。例えば、外部評価委員による評価、パブリックコメントの実施、一部の医療施設での試行などを決定する。

(4) 公開の予定の検討

公開の予定期日、公開の方法などを決定する。

● テンプレート

[SC-3 スコープ]

● テンプレート 記入方法

[SC-3 スコープ 記入方法]

● テンプレート 記入例

[SC-3 スコープ 記入例]

3.7 ステップ7：スコープの確定

ステップ6まで作成されたスコープ草案を確定する。確定にあたって、外部評価委員による外部評価を実施する場合もある。また、スコープをウェブサイトなどで公開して、広くパブリックコメントを募集する方法がとられることもある。以上のプロセスでスコープの内容を変更した場合には、変更の履歴、理由を明示することが望ましい。

以上の手順での意見を反映させた上で、診療ガイドライン作成グループがスコープを確定する。確定後には、原則としてスコープの内容を変更することはできない。ただし、スコープで設定したCQについては、システムティックレビューのプロセスの中でシステムティックレビューチームと診療ガイドライン作成グループとのやり取りを通じて適宜変更する可能性がある。

【参考 スコープ作成の進め方の一例】

第1回会議 診療ガイドライン作成プロセスの全体像の確認

スコープで決定すべき事項の確認

疾患トピックの基本的特徴、重要臨床課題とCQ設定の進め方の決定

(疾患トピックの基本的特徴／重要臨床課題／CQの検討案作成)

第2回会議 疾患トピックの基本的特徴／重要臨床課題／CQの検討・決定

(決定に至らない場合は会議を追加する)

(システムティックレビューに関する検討案作成)

第3回会議 システムティックレビューの検討・決定

(推奨作成、最終調整、公開など、残された事項の検討案作成)

第4回会議 推奨作成、最終調整、公開など、残された事項の検討・決定

スコープ草案の決定と公開

(学会内外からの質問・コメントの受付)

第5回会議 スコープ最終決定

今後の進め方についての確認

第1回会議

最初に、診療ガイドライン統括委員会が診療ガイドライン作成に関して決定した事項を確認する。続いて、診療ガイドライン作成プロセスの全体像（本マニュアルでは**第1章**に記載されている事項）を委員全員で確認する。

次に、審議事項として、スコープで記載すべき事項を決定する。本マニュアルが提案するスコープの項目は標準的な例であり、個々の診療ガイドラインの特徴に応じて項目の追加あるいは省略はあり得る。

さらに、疾患トピックの基本的特徴の整理の進め方、スコープ作成で最も重要な事項のひとつである重要臨床課題とCQ決定に向けた作業の進め方を審議し、決定する。

第2回会議で疾患トピックの基本的特徴、重要臨床課題、CQを決定することを前提として、スコーピングサーチ、意見募集を行うべきかどうかを決定する。実施する情報収集の作業を実施して、第2回会議までに検討案を作成する。

第2回会議

疾患トピックの基本的特徴を踏まえて、重要臨床課題とCQの検討を行い決定する。(決定に至らない場合は会議を追加する。)

疾患トピックの基本的特徴、重要臨床課題、CQが決定したら、第3回会議の審議事項であるシステムティックレビューに関する審議事項を取り上げて、第3回会議までに実施する準備作業を決定する。

第3回会議

システムティックレビューに関する事項を検討し決定する。この会議には、システムティックレビューチームの中核メンバーが出席して協議に参加することもある。

システムティックレビューに関する事項が決定したら、次に、第4回会議の審議事項である推奨作成、診療ガイドラインの最終調整、そして、公開に関する事項について、第4回会議までに実施する準備作業を決定する。

第4回会議

推奨作成、最終調整、公開など、残された事項を検討し決定する。これまでの決定事項をまとめて、スコープ草案を決定する。スコープ草案は公開して、内外からの質問・コメントを受け付ける期間を設けることが望ましい。

第5回会議

内外からのコメントを受けて、必要な修正を行い、スコープを最終決定する。そして、今後の活動方針について審議・決定して終了する。

3章 テンプレート

【SC-1 疾患トピックの基本的特徴】

臨床的特徴
疫学的特徴
診療の全体的な流れ

【SC-2 診療アルゴリズム】

診療アルゴリズム（図）

【SC-3 スコープ】

バージョン :	
作成日 :	
作成者 :	
1. 診療ガイドラインがカバーする内容に関する事項	
(1) タイトル	
(2) 目的	
(3) トピック	
(4) 診療ガイドラインがカバーする視点	
(5) 想定される利用者、利用施設	
(6) 既存の診療ガイドラインとの関係	
(7) 診療ガイドラインがカバーする範囲	
(8) 重要臨床課題	<p>重要臨床課題 1 :</p> <p>重要臨床課題 2 :</p> <p>重要臨床課題 3 :</p>

(9) クリニカルクエスチョン (CQ) リスト	CQ1 : CQ2 : CQ3 :
2. システマティックレビューに関する事項	
(1) 実施スケジュール	
(2) エビデンスの検索	
(3) 文献の選択基準, 除外基準	
(4) エビデンスの評価と統合の方法	
3. 推奨作成から公開に向けた最終調整, 公開までに関する事項	
(1) 推奨作成の基本方針	
(2) 最終調整	
(3) 外部評価の具体的方法	
(4) 公開の予定	

【SC-4 CQ の設定】

スコープで取り上げた重要臨床課題 (key clinical issues)				
CQ の構成要素				
P (Patients, Problem, Population)				
性別	(指定なし ・ 男性 ・ 女性)			
年齢	(指定なし ・ _____)			
疾患・病態				
地理的要件				
その他				
I (Interventions)	C (Comparisons, Controls, Comparators)			
O (Outcomes) のリスト				
	Outcome の内容	益か害か	重要度	採用可否
O ₁		(益 ・ 害)	_____点	
O ₂		(益 ・ 害)	_____点	
O ₃		(益 ・ 害)	_____点	
O ₄		(益 ・ 害)	_____点	
O ₅		(益 ・ 害)	_____点	
O ₆		(益 ・ 害)	_____点	
O ₇		(益 ・ 害)	_____点	
O ₈		(益 ・ 害)	_____点	
O ₉		(益 ・ 害)	_____点	
O ₁₀		(益 ・ 害)	_____点	
作成した CQ				

3章 テンプレート 記入方法

[SC-1 疾患トピックの基本的特徴 記入方法]

臨床的特徴

診療ガイドラインの推奨を理解し、診療で活用するために必要な臨床的事項を記載する。病態生理、臨床分類、歴史的事項など。

疫学的特徴

診療ガイドラインの推奨を理解し、診療で活用するために必要な疫学的事項を記載する。罹患率、死亡率、受療率、生存期間などの現状、経年変化、地域特性など。

診療の全体的な流れ

実地で行われている診療の全体的な流れについて、解説する。

例えば、当該疾患がステージ1、ステージ2に臨床分類される場合、以下のようになる。

ステージ1：手術療法が標準的治療として確立されている。

ステージ2：手術療法、手術+放射線療法が並立しており、CQとしての評価が必要（本診療ガイドラインで推奨作成）。

【SC-2 診療アルゴリズム 記入方法】

診療アルゴリズム（図）

解説の理解を助けるために、診療アルゴリズムの図を追加することが望ましい。前述の例でのステージ2のように、スコープ作成の段階ではCQとして推奨作成の対象となっている場合は、スコープの段階では「CQ（CQ番号）として推奨作成」と記述しておき、推奨作成が完了した後に、推奨内容を追記して診療アルゴリズムとして完成する（**第6章**）。

【SC-3 スコープ 記入方法】

バージョン： 作成日： 作成者：	
1. 診療ガイドラインがカバーする内容に関する事項	
(1) タイトル	診療ガイドラインのタイトルを記載する。改訂版の場合は、改訂版であることがタイトルから分かることが望ましい。
(2) 目的	診療ガイドラインが全体として取り組む患者アウトカムの改善を目的として記載する。
(3) トピック	診療ガイドラインが取り上げるテーマを記述する。疾患全体を取り上げる場合は「疾患名」、疾患全体ではなく、例えば、治療のみを取り上げる場合は、「小児急性中耳炎の治療」のような組み合わせで、取り上げるトピックを明記する。
(4) 診療ガイドラインがカバーする視点	患者個別の立場から推奨を考える場合 (individual perspective : 個人の立場) と、集団としての医療費などを考慮して推奨を考える場合 (population perspective : 集団の立場) のどちらの立場で推奨を作成するかを記述する。
(5) 想定される利用者、利用施設	診療ガイドラインの活用が想定される施設、医療者などを明記する。
(6) 既存の診療ガイドラインとの関係	改訂版の場合は、どの診療ガイドラインの改訂版であるかを明示する。また、改訂版ではないが、関連する診療ガイドラインがある場合は、関係を明示する。
(7) 診療ガイドラインがカバーする範囲	診療ガイドラインが取り上げるトピックについて、より詳細な定義を行う。診療ガイドラインがカバーする範囲とカバーしない範囲を整理して記述する。
(8) 重要臨床課題	診療ガイドラインが取り上げる課題を重要臨床課題として記述する。重要臨床課題を記述するにあたって、予備的な文献検索を実施して、課題の概略を把握する。取り上げた重要臨床課題に基づいてCQを設定して推奨を出すことで、患者アウトカムの改善が期待される理由も記載する。また、結果的に取り上げないことになった課題についても、同様にその理由を記述しておくことが望ましい。 重要臨床課題1： 重要臨床課題2： 重要臨床課題3：

(9) クリニカルクエスチョン (CQ) リスト	<p>「重要臨床課題」で取り上げられた課題を基に、患者にとって重要なアウトカムを改善するために必要な問題を CQ として設定する。スコープでは、設定された CQ をリストアップする。</p> <p>CQ1 : CQ2 : CQ3 :</p>
2. システマティックレビューに関する事項	
(1) 実施スケジュール	システムティックレビューの実施スケジュールを記述する。
(2) エビデンスの検索	エビデンスタイプとして、①個別研究論文 (RCT, non-RCT, 観察研究など), ②システムティックレビュー論文, ③既存の診療ガイドラインの中で、どれを検索対象とするかを記載し、検索の優先順位を明確にする。また、個別研究論文については、RCT, non-RCT, 観察研究などの研究タイプの中で、検索対象に含める研究タイプを記載する。検索式を決定する際の基本方針を記載する。検索の対象とするデータベースをエビデンスタイプごとに記載する。検索対象期間は、データベースごとに、検索範囲となる期間を年月日で記載する。
(3) 文献の選択基準、除外基準	既存の診療ガイドライン、システムティックレビュー論文、RCT、non-RCT、観察研究をどのように選出し統合するか、優先順位と採用条件の基本方針を記載する。
(4) エビデンスの評価と統合の方法	選出された論文からなるエビデンスの評価方法を記載する。また、エビデンス総体の示す強さの表現方法を記載する。エビデンス総体の質的統合と量的統合に対する基本的な考え方を記載する。
3. 推奨作成から公開に向けた最終調整、公開までに関する事項	
(1) 推奨作成の基本方針	コンセンサス形成の具体的方法、推奨作成の際に考慮する因子、推奨を文章で表現する際に準拠するルール、推奨の強さを表現する基準をあらかじめ決めて記述する。医療経済性を考慮する場合、具体的な方法を記載する。
(2) 最終調整	完成した診療ガイドラインの草案を最終決定するための手続きを記載する。
(3) 外部評価の具体的方法	外部評価として実施する方法を記載する。例えば、外部評価委員による評価、パブリックコメントの実施、一部の医療施設での試行など。
(4) 公開の予定	公開の予定期日、公開の方法などを記載する。

【SC-4 CQ の設定 記入方法】

スコープで取り上げた重要臨床課題 (key clinical issues)							
<p>スコープで取り上げることに決まった重要臨床課題 (key clinical issues) の 1 つを書き入れる。</p>							
CQ の構成要素							
P (Patients, Problem, Population)							
性別	(指定なし	・	男性	・	女性)
年齢	(指定なし	・)
疾患・病態							
地理的要件	介入方法の選択が行われる「状況」を 詳細に設定し、書き入れる。						
その他							
I (Interventions)			C (Comparisons, Controls, Comparators)				
			<p>介入方法の選択肢として検討したいものを I と C に分けて記載する。</p>				
O (Outcomes) のリスト							
	Outcome の内容	益か害か	重要度	採用可否			
O ₁		(益 ・ 害)	_____点				
O ₂		<ul style="list-style-type: none"> 介入を行った結果として起こり得るアウトカム事象のうち、患者にとって重要と考えられるものを網羅的に列挙する。ここでは 10 個まで記入できるようにしてあるが、足りない場合は行数を増やして使用する。 					
O ₃		<ul style="list-style-type: none"> Iを行う方が C を行うより患者にとって望ましい結果になることが期待されるアウトカムは「益」に、望ましくない結果になると考えられるアウトカムは「害」に○をつける。両方のアウトカムを少なくとも 1 つずつ取り上げること。 					
O ₄		<ul style="list-style-type: none"> 合議の上で各アウトカムの重要度の点数を決定して書き入れ、システムティックレビューに含めるかどうかを○、×で採用可否に表記する。 					
O ₅							
O ₆							
O ₇							
O ₈							
O ₉							
O ₁₀							
作成した CQ							
<p>上で抽出した PICO を用いて CQ を 1 文で表現し、記載する。</p>							

3章 テンプレート 記入例

【SC-1 疾患トピックの基本的特徴 記入例】

※これは成人における急性虫垂炎の診断と治療をテーマとした記載例であり、実際の診療ガイドラインではないことにご注意ください。

臨床的特徴

【病態】

虫垂炎は、虫垂の内腔で細菌が増殖し、虫垂に炎症が起った状態であり、細菌繁殖の原因は、便や糞石の貯留によるものが多い。虫垂が壊死すると膿汁や腸液などが漏れ出し、腹膜炎・敗血症となる。

【診断】

上腹部全体の不快感や痛みで発覚することが多く、その後、右下腹部に痛みが集中していくのが典型例である。これらの症状に加え、腹部診察による腹膜炎症状の診断、血液検査による白血球の増加やCRP上昇などで診断が行われてきた。

〈診断の歴史的事項〉

1980年代以降、腹部CTと超音波検査が臨床で用いられるようになった。腫大した虫垂が確認できれば虫垂炎と診断できるとされ、急速に虫垂炎診断への利用が広がった。

しかし、診断の精度は向上するが、両検査は必須なのか、どのような場合にどちらが有用なのかに関する明確な情報は提示されていない。

【治療】

診断精度が低い時代は、虫垂炎に対する治療は外科手術が基本であった。これは、様子を見ることで重症化する可能性があるので、まずは手術することで安全性を確保・最優先したいという理由からである。

〈治療の歴史的事項〉

1) 前述のように、超音波、CTなどの画像診断が進化し、また有効な抗菌薬も開発されたことから、腹膜炎に至っていない虫垂炎には、手術のみではなく、抗菌薬治療も行われるようになった。いわゆる薬で「散らす」という方法である。

2) さらに、1990年代から腹腔鏡下手術が広まり、虫垂炎の外科手術を行う際に、この腹腔鏡下手術が行われるようになり、傷が小さく、社会復帰が早まるという利点が報告された。

しかし、手術に対する抗菌薬治療の有用性、開腹手術に対する腹腔鏡下手術の有用性についての明確な情報は提示されていない。

疫学的特徴

男女差なし。10～20代の発症が多い。

予後は良好であるが、重症化し、腹膜炎・敗血症を併発すると死亡する場合もある。

診療の全体的な流れ

1. 診察：

病歴聴取（腹痛症状や部位の変化）、腹部診察など



2. 検査：

血液検査、超音波検査、腹部CT検査、腹部レントゲン検査

検査に関する検討点

- ・全員に全ての検査が必要か？
- ・妊娠している可能性がある場合どうするか？

↓

3. 重症度評価：

腹膜炎を併発しているか？
他に併存疾患があるか（心臓、脳、がん治療中など）
そのほか、何に注意する必要があるか？

↓

4. 治療

腹膜炎合併例では、現在の日本では、通常外科手術が行われている。

- ・開腹手術か腹腔鏡下手術か

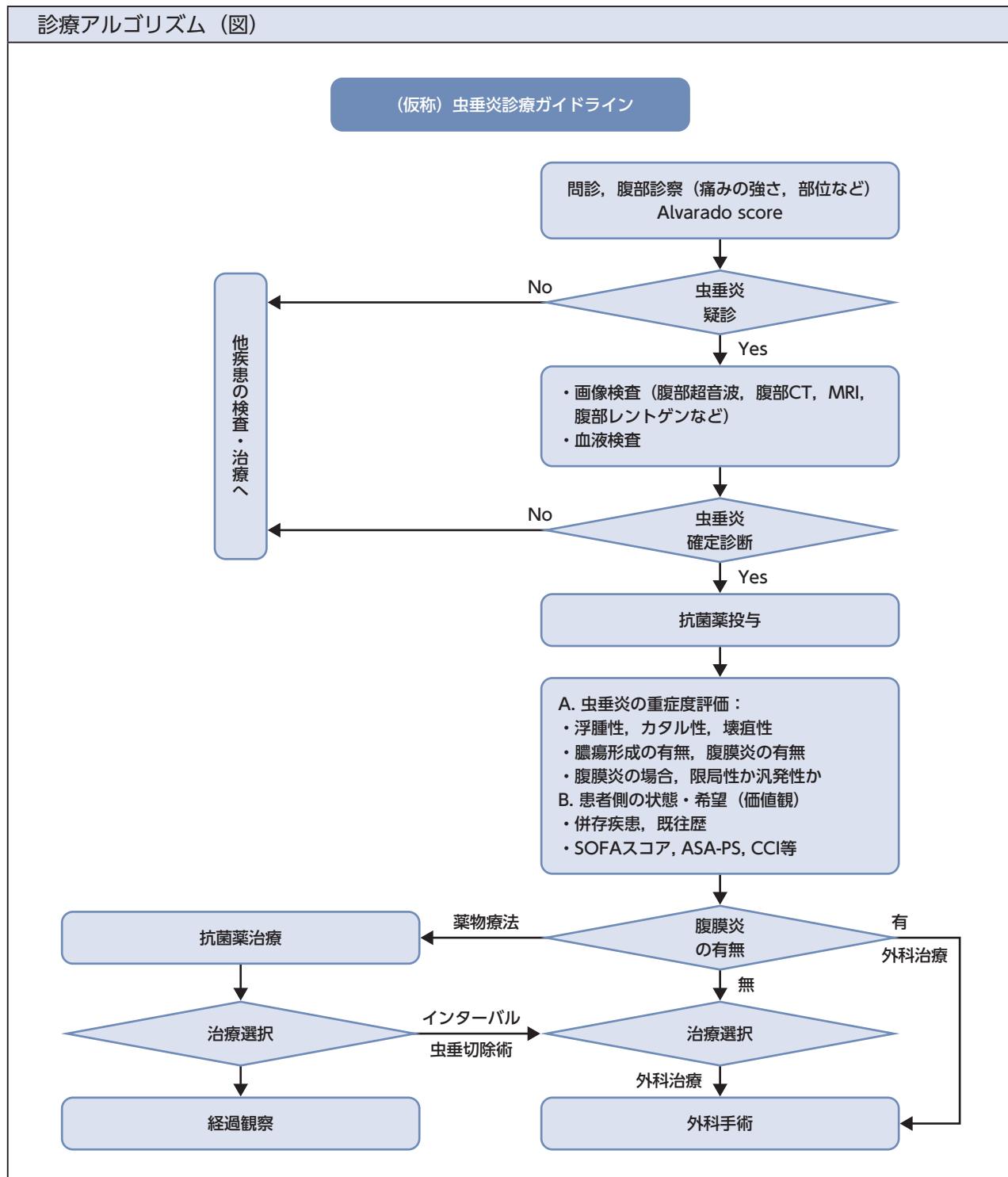
腹膜炎を合併していない場合

- ・抗菌薬か外科治療か
- ・外科治療の場合、開腹手術か腹腔鏡下手術か
- ・患者の併存疾患で治療方法を変えるべきか
- ・患者の希望はどのように反映されるか

↓

5. 食事開始と退院、社会復帰

【SC-2 診療アルゴリズム 記入例】



【SC-3 スコープ 記入例】

<p>バージョン：1.x</p> <p>作成日：●年●月●日作成（○年○月○日最終改訂）</p> <p>作成者：○○グループ（個人名××××、△△△△、□□□□…）</p>	
<p>1. 診療ガイドラインがカバーする内容に関する事項</p>	
(1) タイトル	成人における急性虫垂炎の診断と治療 (簡略タイトル：虫垂炎)
(2) 目的	以下のアウトカムを改善することを目的とする <ul style="list-style-type: none"> ● 診断（偽陽性）に伴う不必要的手術と、診断（偽陰性）による手術遅れと合併症 ● 治療後1年以内の再発 ● 治療に伴う主要な合併症（腹膜炎、腹膜瘻着による腸閉塞、出血など。創感染含まず） ● 入院期間
(3) トピック	成人における急性虫垂炎の診断と治療
(4) 診療ガイドラインがカバーする視点	本診療ガイドラインは、individual perspective（個人視点）で作成する。
(5) 想定される利用者、利用施設	適用が想定される臨床現場 <ul style="list-style-type: none"> ● 一次医療（プライマリケア） ● 二次医療（救急を含む）
(6) 既存の診療ガイドラインとの関係	過去5年間で、PubMed、医学中央雑誌、Amazonインターネット版を検索した範囲では、急性虫垂炎に関する診療ガイドラインとの記載がある文献は、収集できなかった。合併症を有する腹腔内感染症に対する抗菌薬治療に関して、2010年に Diagnosis and Management of Complicated Intra-abdominal Infection in Adults and Children: Guidelines by the Surgical Infection Society and the Infectious Diseases Society of America の報告があるのみである。
(7) 診療ガイドラインがカバーする範囲	本診療ガイドラインがカバーする範囲 <ul style="list-style-type: none"> ● 虫垂炎と診断された成人（18歳以上）および・高齢者（上限なし） ● 穿孔、膿瘍などの合併症を伴わない虫垂炎 本診療ガイドラインがカバーしない範囲 <ul style="list-style-type: none"> ● 小児、若年者（18歳未満） ● 穿孔、膿瘍などの合併症を伴った虫垂炎患者 ● 抗菌薬アレルギーのある患者 ● ステロイド治療や抗血栓療法を受けている患者 ● 炎症性腸疾患の既往がある患者 ● 妊娠している可能性のある女性
(8) 重要臨床課題	<ul style="list-style-type: none"> ● 重要臨床課題1： 成人の急性虫垂炎に対する画像診断として、腹部超音波と腹部CTのいずれを選択すべきかについて明らかにする必要がある。 ● 重要臨床課題2：

	<p>診断、検査に基づいて急性虫垂炎である可能性が高まった段階で、治療の方針について決定する必要がある。現在、急性虫垂炎の治療法についてのゴールドスタンダードは外科治療とされている。外科治療には開腹手術と腹腔鏡下手術があるが、程度の差はあるものの虫垂を切除するために必然的に身体への侵襲を伴うことになる。そのため、急性虫垂炎の治療については、身体への侵襲のない抗菌薬投与も選択肢のひとつとして考慮されている。どちらの治療法が推奨されるか明確になれば、臨床決断の大きな助けとなることが期待される。</p> <ul style="list-style-type: none"> ● 重要臨床課題 3 : <p>成人の急性虫垂炎に対して外科的虫垂切除術を行う場合、開腹手術（脊椎麻酔）と腹腔鏡下手術（全身麻酔）のいずれを選択すべきかについて明らかにする必要がある。</p>
(9) クリニカルクエスチョン (CQ) リスト	<ul style="list-style-type: none"> ● 重要臨床課題 1 の CQ CQ1：成人の急性虫垂炎に対する画像診断として、腹部超音波は、腹部 CT に比べ推奨できるか？ ● 重要臨床課題 2 の CQ CQ2：成人の急性虫垂炎において抗菌薬投与による保存的治療は、外科的虫垂切除術と比べ推奨できるか？ ● 重要臨床課題 3 の CQ CQ3：成人の急性虫垂炎に対する外科治療として、腹腔鏡下手術（全身麻酔）は、開腹手術（脊椎麻酔）に比べ推奨できるか？

2. システマティックレビューに関する事項

(1) 実施スケジュール	<ul style="list-style-type: none"> ● 文献検索に 1 カ月 ● 文献の選出に 2 カ月 ● エビデンス総体の評価と統合に 2 カ月
(2) エビデンスの検索	<p>(1) エビデンスタイプ：</p> <ul style="list-style-type: none"> ● 既存の診療ガイドライン、SR / MA 論文、個別研究論文を、この順番の優先順位で検索する。優先順位の高いエビデンスタイプで十分なエビデンスが見い出された場合は、そこで検索を終了してエビデンスの評価と統合に進む。 ● 個別研究論文としては、ランダム化比較試験 (RCT)，非ランダム化比較試験 (non-RCT)，観察研究を検索の対象とする。 <p>(2) データベース：</p> <ul style="list-style-type: none"> ● 個別研究論文については、MEDLINE、医中誌 Web、CENTRAL ● SR / MA 論文については、MEDLINE、医中誌 Web、CDSR ● 既存の診療ガイドラインについては、Guidelines International Network の International Guideline Library、米国 AHRQ の National Guideline Clearinghouse (NGC)、医中誌 Web、Amazon <p>(3) 検索の基本方針：</p> <p>介入の検索に際しては、PICO フォーマットを用いる。P と I の組み合わせが基本で、</p>

	<p>ときに C も特定する。O については特定しない。</p> <p>(4) 検索対象期間 :</p> <p>全てのデータベースについて、2015年11月末まで The Cochrane Library は、2015 issue 11 まで</p>
(3) 文献の選択基準、除外基準	<ul style="list-style-type: none"> ● 採用条件を満たす診療ガイドライン、システムティックレビュー論文が存在する場合は、それを第一優先とする。検索日が1年以上経過している診療ガイドラインやシステムティックレビュー論文は、新しいRCTが出ていないか確認する。 ● 採用条件を満たす診療ガイドライン、システムティックレビュー論文がない場合は、個別研究論文を対象として <i>de novo</i> でシステムティックレビューを実施する。または検索されていない年数だけ検索する。 ● <i>de novo</i> システマティックレビューでは、採用条件を満たす RCT を優先して実施する。 ● 採用条件を満たす RCT がない場合には non-RCT（介入研究）、または観察研究（対照群があるもの）を対象とする。 ● 採用条件を満たす観察研究がない場合は、システムティックレビューは実施しない。
(4) エビデンスの評価と統合の方法	<ul style="list-style-type: none"> ● 個々の研究のバイアスリスク評価には Cochrane の評価ツールを利用し、エビデンス総体の評価には GRADE アプローチの方法に基づく。 ● 効果指標の統合は、質的な統合を基本とし、適切な場合は量的な統合も実施する。
3. 推奨作成から公開に向けた最終調整、公開までに関する事項	
(1) 推奨作成の基本方針	<ul style="list-style-type: none"> ● 推奨の決定は、診療ガイドライン作成グループのパネル会議に基づく。意見の一致をみない場合には、投票を行って決定する。 ● 推奨の決定は、エビデンスの評価と統合で作成された資料を参考に、「アウトカム全体にわたる総括的なエビデンスの確実性」、「望ましい効果と望ましくない効果のバランス」、「患者・市民の価値観と希望」、「資源の利用（コスト）」などを考慮して行う。具体的には、システムティックレビューによって作成された評価シートや SoF 表（summary of findings table）などを参考に、EtD フレームワークを用いて、推奨とその強さを決定する。 ● 医療経済性を考慮する具体的方法については別紙参照のこと。
(2) 最終調整	<ul style="list-style-type: none"> ● 追加すべき事項（活用方法、評価方法など）を記載し、草案を作成する。 ● 草案に対して、外部評価、およびパブリックコメントを募集する。 ● 上記評価を参考にして、診療ガイドラインを最終調整する。
(3) 外部評価の具体的方法	<ul style="list-style-type: none"> ● 外部評価委員が個別にコメントを提出する。 ● 診療ガイドライン作成グループは、各コメントに対して診療ガイドラインの内容を変更する必要性を討議して、対応を決定する。 ● パブリックコメントに対しても同様に、診療ガイドライン作成グループは、各コメントに対して診療ガイドラインの内容を変更する必要性を討議して、対応を決定する。
(4) 公開の予定	<ul style="list-style-type: none"> ● 外部評価、パブリックコメントへの対応が終了したら、診療ガイドライン統括委員会が公開の最終決定をする。 ● 公開の方法は、診療ガイドライン作成グループと診療ガイドライン統括委員会が協議の上決定する。

【SC-4 CQ の設定 記入例】(実際にはCQの数だけ作成する)

スコープで取り上げた重要臨床課題 (key clinical issues)				
重要臨床課題 2 : 診断、検査に基づいて急性虫垂炎である可能性が高まった段階で、治療の方針について決定する必要がある。現在、急性虫垂炎の治療法についてのゴールドスタンダードは外科治療とされている。外科治療には開腹手術と腹腔鏡下手術とがあるが、程度の差はあるものの虫垂を切除するために必然的に身体への侵襲を伴うことになる。そのため、急性虫垂炎の治療については、身体への侵襲のない抗菌薬投与も選択肢のひとつとして考慮されている。どちらの治療法が推奨されるかについて明確になれば、臨床決断の大きな助けとなることが期待される。				
CQ の構成要素				
P (Patients, Problem, Population)				
性別	(指定なし)	・	男性	・ 女性)
年齢	(指定なし)	・	18 歳以上の成人)
疾患・病態	急性虫垂炎で穿孔、膿瘍形成などの合併症を伴わないもの			
地理的要件	医療体制の確立した地域			
その他	特になし			
I (Interventions)	C (Comparisons, Controls, Comparators)			
抗菌薬投与による保存的治療	外科的虫垂切除 (開腹術および腹腔鏡下手術は問わない)			
O (Outcomes) のリスト				
	Outcome の内容	益か害か	重要度	採用可否
O ₁	虫垂炎の再発 (1年間)	(益 ・ 害)	9点	○
O ₂	虫垂切除術 (治療開始1ヵ月以内)	(益 ・ 害)	8点	○
O ₃	主要な合併症 (1年間 : 外科的治療、内視鏡的治療、IVR治療をするような合併症、臓器不全、死亡、深部感染、腹壁瘢痕ヘルニア、癒着性腸管閉塞)	(益 ・ 害)	8点	○
O ₄	その他の合併症 (1年間 : 表在性創傷感染、下痢、腹部不快感など)	(益 ・ 害)	5点	○
O ₅	入院期間	(益 ・ 害)	7点	○
O ₆	病休期間	(益 ・ 害)	6点	○
O ₇	炎症反応 (CRP値)	(益 ・ 害)	3点	×
O ₈		(益 ・ 害)	_____点	
O ₉		(益 ・ 害)	_____点	
O ₁₀		(益 ・ 害)	_____点	
作成したCQ				
成人の急性虫垂炎において抗菌薬投与による保存的治療は、外科的虫垂切除術と比べ推奨できるか？				

3章 文献

- Agency for Healthcare Research and Quality (AHRQ). (2014) Methods Guide for Effectiveness and Comparative Effectiveness Reviews. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/cer-methods-guide_overview.pdf (参照 2021/01/22)
- Alonso-Coello P, et al. (2016a) GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 353:i2016. doi: [10.1136/bmj.i2016](https://doi.org/10.1136/bmj.i2016). PMID: [27353417](#)
- Alonso-Coello P, et al. (2016b) GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ 353:i2089. doi: [10.1136/bmj.i2089](https://doi.org/10.1136/bmj.i2089). PMID: [27365494](#)
- Dalkey N, et al. (1969) The Delphi Method, III: Use of Self Ratings to Improve Group Estimates. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2006/RM6115.pdf (参照 2021/01/22)
- GRADE Working Group. (2015) Evidence to Decision (EtD) Frameworks Guidance. https://s3.amazonaws.com/ietd_pdf/EtD+guidance+updated+2015+05+19.pdf (参照 2021/01/22)
- Guyatt GH, et al. (2015) Guideline panels should not GRADE good practice statements. J Clin Epidemiol 68:597-600. doi: [10.1016/j.jclinepi.2014.12.011](https://doi.org/10.1016/j.jclinepi.2014.12.011). PMID: [25660962](#)
- Guyatt GH, et al. (2016) Guideline panels should seldom make good practice statements: guidance from the GRADE working group. J Clin Epidemiol 80:3-7. doi: [10.1016/j.jclinepi.2016.07.006](https://doi.org/10.1016/j.jclinepi.2016.07.006). PMID: [27452192](#)
- Harris RP, et al. (2001) Current methods of the US preventive services task force: a review of the process. Am J Prev Med 20:21-35. doi: [10.1016/s0749-3797\(01\)00261-6](https://doi.org/10.1016/s0749-3797(01)00261-6). PMID: [11306229](#)
- National Institute for Health and Care Excellence (NICE). (2014) Developing NICE guidelines: the manual. <https://www.nice.org.uk/process/pmg20/chapter/developing-review-questions-and-planning-the-evidence-review> (参照 2021/01/22)
- Schünemann H, et al. (2013) 3. Selecting and rating the importance of outcomes. GRADE Handbook. <https://gdt.gradepro.org/app/handbook/handbook.html> (参照 2021/01/22)
- Straus SE, et al. (2010) Evidence-Based Medicine: How to Practice and Teach It. Fourth ed. Churchill Livingstone
- World Health Organization (WHO). (2014) 14. Strong recommendations when the evidence is low quality. WHO handbook for guideline development. Second ed. World Health Organization.
- Williamson PR, et al. (2012) Developing core outcome sets for clinical trials: issues to consider. Trials 13:132. doi: [10.1186/1745-6215-13-132](https://doi.org/10.1186/1745-6215-13-132). PMID: [22867278](#)

第4章

システムティックレビュー

Minds

4章 Key Words

- AGREE II (The Appraisal of Guidelines for Research & Evaluation II)
- AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews 2)
- CINAHL (Cumulative Index to Nursing and Allied Health Literature)
- Embase
- GRADE (The Grading of Recommendations Assessment, Development and Evaluation)
- GRADEpro GDT
- I^2 統計値
- Institute of Medicine (IOM)
- IOM の SR の基準 (ガイドライン評価における)
- IOM の基準 (ガイドライン評価における)
- ITT (intention-to-treat) 解析
- JMEDPlus
- MEDLINE
- MeSH (Medical Subject Headings)
- NNT (number needed to treat)
- PICO
- PRESS (Peer Review of Electronic Search Strategies)
- PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)
- PROSPERO (International Prospective Register of Systematic Reviews)
- PsycINFO
- PubMed
- Q 統計値
- Rayyan
- RevMan (Review Manager)
- ROBINS-I (Risk of Bias In Non-randomized Studies for Interventions)
- SoF 表 (summary of findings table)
- The Cochrane Library
- アウトカムの重要性
- アブストラクト
- 医中誌 Web

◎ 益と害 (benefit and harm)
◎ 益と害のバランス
◎ エビデンスの確実性
◎ エビデンスの確実性（強さ）
◎ エビデンスの確実性（質）
◎ エビデンス総体 (body of evidence)
◎ エビデンス総体の総括 (overall evidence)
◎ オッズ比 (odds ratio : OR)
◎ 介入研究 (interventional study)
◎ 観察研究 (observational study)
◎ 感度分析 (sensitivity analysis)
◎ クリニカルクエスチョン (clinical question : CQ)
◎ グレイリテラチャー (grey literature)
◎ 傾向スコア分析
◎ 検出バイアス (detection bias)
◎ 交絡因子 (confounding factor)
◎ 固定効果モデル (メタアナリシスにおける)
◎ コホート研究 (cohort study)
◎ 今後の研究 (future research)
◎ コンシールメント (concealment)
◎ サンプルサイズ (sample size)
◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
◎ システマティックレビューチーム (systematic review team : SR チーム)
◎ システマティックレビューの評価 システムティックレビューレポート (サマリーレポートともいう)
◎ 実行バイアス (performance bias)
◎ 重要臨床課題 (key clinical issues)
◎ 出版 (報告) バイアス (publication bias, reporting bias)
◎ 症例減少バイアス (attrition bias)
◎ 症例集積 (case series)
◎ 症例対照研究 (case-control study)
◎ 症例報告 (case reports)
◎ 診断精度 (diagnostic test accuracy : DTA) 研究
◎ 診療ガイドライン (clinical practice guideline : CPG)

◎ 診療ガイドラインのためのシステムティックレビュー
◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
◎ 推奨
◎ スコープ (scope)
◎ 選択バイアス (selection bias)
◎ 早期試験中止バイアス
◎ 多基準決断分析 (multi-criteria decision analysis : MCDA)
◎ ネットワークメタアナリシス (network meta-analysis : NMA)
◎ バイアスリスク (risk of bias)
◎ ハザード比 (hazard ratio : HR)
◎ 非一貫性 (inconsistency)
◎ 非直接性 (indirectness)
評価シート (evidence profile)
◎ 標準化平均値差 (standardized mean difference : SMD)
◎ 非ランダム化比較試験 (non-randomized controlled trial : non-RCT)
◎ ファンネルプロット (funnel plot)
◎ フォレストプロット (forest plot)
◎ 不正確性 (imprecision)
◎ 米国医療研究・品質庁 (Agency for Healthcare Research and Quality : AHRQ)
◎ メタアナリシス (meta-analysis, メタ解析と同義)
◎ 盲検化 (blinding)
◎ 用量-反応勾配 (dose-dependent gradient)
◎ ランダム化 (randomization)
◎ ランダム化比較試験 (randomized controlled trial : RCT)
◎ ランダム効果モデル (メタアナリシスにおける)
◎ 利益相反 (conflict of interest : COI)
◎ リスク差 (risk difference : RD)
◎ リスク比 (risk ratio : RR)

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】

SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】

RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】

PO : post 【第8章 診療ガイドライン公開後の取り組み】

4.0 概要

診療ガイドライン（clinical practice guideline : CPG）は推奨を作成するために、存在する全てのエビデンスを科学的に活用することが望まれる。そのため、システムティックレビュー（systematic review : SR、系統的レビューと同義）が行われる。システムティックレビューでは包括的に研究の検索を行い、一定の基準で選定された研究論文を一定の基準で評価し、アウトカムごとにエビデンス総体（body of evidence）のエビデンスの強さを決定することが望ましい。1つのクリニカルクエスチョン（clinical question : CQ）には複数のアウトカムが設定されるので、アウトカムごとに選定された研究集合ができる、それらがさらに研究デザインごとにまとめられることになる。

なお、エビデンスに対する評価について、質、レベル、限界、強さ、確信度、確実性などの言葉が同じ意味で用いられてきている。本マニュアルでは、広義のエビデンスに対する評価を「エビデンスの確実性」とし、推奨を考慮しない段階で行われるシステムティックレビューでは効果指標の確実性に対する確信という意味で「エビデンスの確実性（質）」、もしくは、「エビデンスの質」という言葉を用い、推奨作成の段階では推奨を支持する強さに対する確信という意味で「エビデンスの確実性（強さ）」、もしくは、「（エビデンス総体の）エビデンスの強さ」という言葉を用いる。この分類は、GRADE アプローチにおいて、システムティックレビューのエビデンスの確実性と推奨作成におけるエビデンスの確実性では異なる定義を用いることが明確化されている（Hultcrantz M 2017, Turner RM 2009）ことにも対応している。

診療ガイドライン作成のためのシステムティックレビューは、科学的作業であるとともに医療へ直接影響を与えるものであることを認識し、科学性、包括性、透明性、不偏性の確保に努めることが望ましい。

システムティックレビューにより、それぞれのアウトカムに対する効果の大きさとその不確実性が、効果指標の統合値と信頼区間として示される。それらが、益と害（benefit and harm）の評価の科学的根拠となる。

包括性を確保するため、データベース検索の際の検索式を工夫するとともに、必要に応じて検索外での文献追加も実行することが望ましい。透明性を高めるため、用いたデータベース、検索方法、検索式を記録する。可能であれば、論文タイトル、アブストラクトによる一次スクリーニング、全文に基づく二次スクリーニングをあらかじめ定めた採用基準、除外基準に基づいて行い、採用論文のリストを作成するとともに、不採用論文のリストも作成することが望ましい。

また、可能であれば、アウトカムごとに作成された研究集合の個別論文、およびそれらを統合したエビデンス総体のエビデンス評価を一定の基準で行う。定性的な評価と定量的な評価の両者に基づき最終的にエビデンスの強さを評価し、最後にアウトカム横断的にエビデンス総体の総括（overall evidence）の評価を行い、推奨作成を行うことが望ましい。

エビデンス評価の基準については、国際的に共通の認識があるので（AHRQ 2012, Berkman ND 2014, Higgins JP 2011a, Higgins JPT 2011b, Deeks JJ web, Schünemann H 2013），それらに準じて行なうことが望ましい。

4.0.1 システマティックレビューの定義

システムティックレビューはさまざまに定義付けが行われているのが現状である。「システムティックレビュー」という言葉が「メタアナリシス」と同義語で用いられていることもある。しかし、実際にはメタアナリシス（meta-analysis、メタ解析と同義）を伴わないシステムティックレビューも存在し、システムティックレビューではないメタアナリシスもあり得るので、完全な同義語ではない。

Institute of Medicine (IOM) はシステムティックレビューを以下のように定義している。「特定の問題に絞って、類似した、しかし別々の研究の知見を見つけ出し、選択し、評価し、まとめるために、明確で計画された科学的方法を用いる科学的研究。別々の研究からの結果の定量的統合（メタアナリシス）を含むことも含まないこともある」(IOM 2011a: 21, IOM 2011b: 23)。診療ガイドライン作成のためのメタアナリシスとは、「研究標本から特定の集団に対する推定と適用を可能とすることを企図して、類似した研究の結果を定量的に結合するために統計学的な方法を用いるシステムティックレビュー」である(IOM 2011a: 21)。いずれの定義も方法論的側面に着目した定義となっている。また、統合（synthesis）とは、「綿密な情報の照合、結合およびシステムティックレビューの結果のまとめのこと」と定義されている(IOM 2011a: 173)。

方法論的な面からの定義として IOM の定義に賛同するものであるが、ゴールに基づく基準^{*1}として考えると、システムティックレビューと呼べるための条件としては、1. 参照した研究に漏れがない、2. 採択された研究に偏りがない、3. 中立の立場で一定の基準に基づき各研究を評価：①アウトカムに及ぼす効果の大きさ、②効果の確実性、4. 結論に評価の結果が反映されている、を提案する(図 4-1 参照)。

実際の作業の面から定義すると、システムティックレビューとは、「CQ に対して、研究を網羅的に調査し、研究デザインごとに同質の研究をまとめ、バイアスを評価しながら分析・統合を行うこと」である。

Cochrane review をはじめとし、数多くのシステムティックレビューあるいはメタアナリシスが発表されている。現在、システムティックレビューを計画時点で登録するウェブサイト PROSPERO^{*2}が運用されているので、診療ガイドラインのための CQ に基づくシステムティックレビューも、論文として発表を計画する場合には、個々に PROSPERO に登録することを検討すべきである。

システムティックレビューの事前登録をするときには、システムティックレビューのプロトコルが必要である。登録をしないときでも、スコープ（scope）に記載されたシステムティックレビューの方法に加えて、プロトコルを作成することが望ましい。

* 1 システマティックレビューの要件を目標としてどれだけ達成しているかという観点から設定する基準である。いわゆるチェックリストによって研究を評価する際に、各項目が合致するかどうかを見ていく評価基準は分類評価基準（criteria-based criteria）と呼ばれる。これに対して、理想的な状態を目標として想定して、それをどれだけ達成できたかを見ていく評価基準は目標評価基準（goal-based criteria）と呼ぶ。

* 2 PROSPERO (International Prospective Register of Systematic Reviews) <http://www.crd.york.ac.uk/PROSPERO>

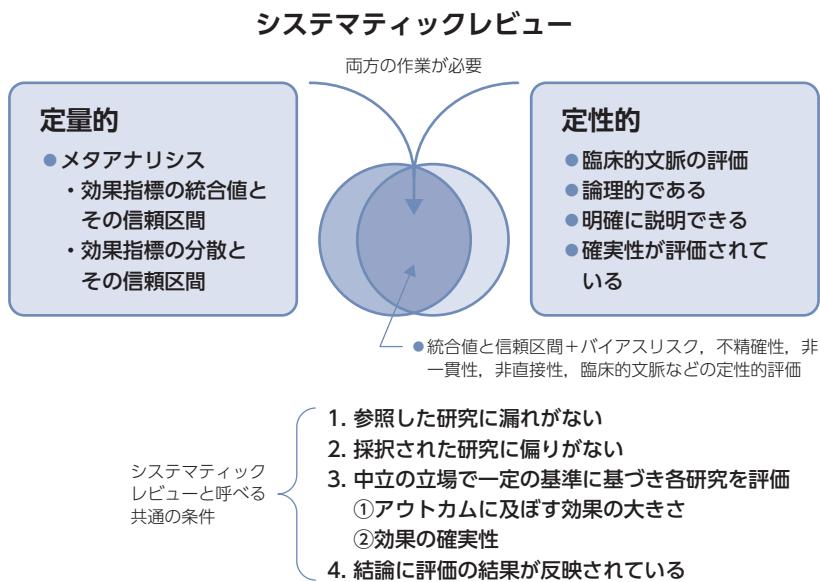


図 4-1 定量的システムティックレビューと定性的システムティックレビュー

4.0.2 定性的システムティックレビューおよび定量的システムティックレビュー

(1) 定性的システムティックレビュー

採択された研究や除外された研究の数、対象の特性と数、比較と介入、バイアスリスク（risk of bias）の評価などを記述し、深い理解を与えるために定性的にまとめたものをいう。定性的な評価の結果を効果指標の値の調整に用いる手法が開発されているが一般的ではないので、定性的システムティックレビューの結果は最終的にエビデンスの強さの判定に反映させる。

参考として、IOM の定性的統合の主要な目的と本マニュアルの考える定性的システムティックレビューにおける実際の作業について表 4-1 に示す。IOM は採択された研究や除外された研究の数、サンプルサイズ（sample size）、比較された介入、バイアスリスクの評価などの記述に留まらず、介入がどのように役立つか、誰のためか、どのような状況で用いられるかなどについてより深い理解を与えるものとして定性的システムティックレビューを位置付けている^{*3}。

* 3 定量的統合あるいは定性的統合という場合は、必ずしもシステムティックという概念を含んでいない。得られる全てのエビデンスを統合する場合は、定量的システムティックレビューと定量的統合、また、定性的システムティックレビューと定性的統合は同じものになる。システムティックレビューで行われるメタアナリシスと単なるメタアナリシスも同様の関係にあり、メタアナリシスは一部の研究だけを定量的に統合する場合にも適用される。

表 4-1 定性的システムティックレビューの主要な目的

目的	該当する実際の作業
1. 臨床的展望の中で利用者に方向付けを与える。	臨床課題あるいは CQ を臨床的文脈において位置付ける。
2. 研究中に参加者に実際に起きたことを記述する。	害、脱落の頻度と理由も評価する。
3. エビデンス総体を論評する。	バイアスリスク、非直接性 (indirectness)、不正確性 (imprecision)、非一貫性 (inconsistency)、などを評価する。
4. なぜ結果が異なるかを説明する個々の研究デザインおよび実行の相違点を明らかにする。	非一貫性、研究間の異質性を評価する。
5. 個々の研究のデザインおよび実行が実世界の臨床的状況との関連にどのように影響するかを述べる。	非直接性を評価する。
6. エビデンスの全般的まとめとセッティングと患者集団に基づくサブグループ解析の結果を取り込む。	非直接性を評価する。
7. 研究が不十分あるいは結果が異なる患者集団に対する注意を促す。	非直接性を評価する。
8. メタアナリシスの結果の堅牢さを解釈し評価する。	感度分析 (sensitivity analysis) の結果、出版 (報告) バイアス (publication bias, reporting bias) を評価する。
9. システマティックレビューの結果が医療業界や社会の通念と異なるかを記述する。	専門医、一般医、医療者、患者・市民代表の評価を取り込む。

(2) 定量的システムティックレビュー（メタアナリシス）

定量的システムティックレビューでも、いわゆるメタアナリシスの前にバイアスの評価など定性的な評価を行い、定量的に統合できるか、異質性 (heterogeneity) を検討する必要がある。メタアナリシスとは、効果指標の値が統計学的に統合され、統合値と信頼区間を計算し、定量的統合を行うことである。

メタアナリシスの原理を簡単に解説すると、いくつかの集団のサンプル数と平均値が分かっているが、個々のサンプルの測定値が分からぬ場合に、全体の平均値を求めることに類似している。分散 σ^2 の正規分布に従う母集団からの n 個のランダムサンプルの分散は、中心極限定理より σ^2/n である。したがって、このランダムサンプルの分散の逆数 n/σ^2 はサンプル数 n に比例することになり、分散の逆数で各集団の平均値を重み付けして算出した平均値は全体の平均値となる、というのが原理である。

(3) 定性的システムティックレビューと定量的システムティックレビューの関係

システムティックレビューとメタアナリシスを同じものと見なすと、複数の類似した研究の効果指標の値を統計学的手法で統合することがシステムティックレビューであるという考えが生じてしまう。個別の研究のサンプルサイズはさまざまで研究の実行の厳密さもさまざまなので、バイアスリスクなどによる研究の質は研究により異なり、得られた結果の確実性はさまざまである。各研究の質をなんらかのチェックリストで評価し、その結果をスコア化し効果指標の値の重

み付けに用いるメタアナリシスが提案されたが、その後その手法の問題点が指摘され、そのような研究の質による効果指標の調整は、現在では用いないことが推奨されている（Jüni P 1999）。一方、バイアスリスクの各項目を定量的に評価し、それを効果指標の値の調整に用いる方法が提案され（Turner RM 2009）、正しく機能することが示されているが、高度なスキルが要求されるため、一般化するに至っていない。

このような状況で、診療ガイドライン作成のためのシステムティックレビューでは、効果指標の値をメタアナリシスの手法で統合し、エビデンス総体の定性的評価の結果によって、エビデンスの確実性の評価をする方法がとられている。例えば、生存をアウトカムとしてランダム化比較試験（randomized controlled trial：RCT）のメタアナリシスによりハザード比（hazard ratio：HR）0.5、95%信頼区間0.41～0.61という結果が得られていて、効果が大きいと見なされる場合でも、各研究のバイアスリスクが高く、研究間の非一貫性も高く、非直接性も高いと判定されれば、エビデンス総体の強さはA「強い」ではなく、B「中」あるいはC「弱い」、場合によってはD「非常に弱い」と判定する。これは、もし定性的評価を効果指標の統合値と信頼区間に反映させることができれば、真のハザード比が0.6で信頼区間が0.38～0.94であると推定するということに相当する。すなわち、効果がより小さめで、確実性が低いと見なすことになる。

複数の研究をエビデンス総体としてまとめる場合に、研究デザイン、対象、介入、対照、アウトカム（PICO）や効果指標の類似性が十分な場合には、定量的システムティックレビューが可能である。しかしそうでない場合には、個々の研究が効果指標の値を提示していても、定量的システムティックレビューができない場合もある。また、定量的な効果指標の値が得られない研究の場合もあり得る。これら全ての研究をエビデンス総体としてまとめ、確実性を評価し、最終的にエビデンスの確実性（強さ）を評価することが望ましい。

定性的システムティックレビューは定量的システムティックレビューと並行して行われるものと、定量的システムティックレビューすなわちメタアナリシスが適用できない複数の研究に適用されるものとがある。前者では、バイアスリスク、不精確性、非一貫性、非直接性、出版（報告）バイアス、臨床的文脈などの評価を行うことが定性的システムティックレビューに該当する。一方、後者では対象となる研究で、研究デザインが異なったり、PICOの違いがあつたりするため、バイアスリスク（Guyatt GH 2011c）、不精確性（Guyatt GH 2011e）、非一貫性（Guyatt GH 2011f）、非直接性（Guyatt GH 2011g）、出版（報告）バイアス（Guyatt GH 2011d）の評価に加え、臨床的文脈を明確にし、論理的で明確な説明をし、確実性を評価することが重要なとなる。

例えば、RCTが1件、症例対照研究（case-control study）が1件しかないような場合でも、それぞれの効果指標の評価と、定性的な評価は可能であり、それらをまとめて結論を導き出すことはシステムティックレビューと呼ぶことができる。

4.0.3 診療ガイドラインのためのシステムティックレビューと通常のシステムティックレビュー

Cochrane review をはじめとし、数多くのシステムティックレビューあるいはメタアナリシスが発表されている。これらのシステムティックレビューの多くは診療ガイドライン作成を目的として行われたものではない。これら通常のシステムティックレビューと診療ガイドラインのためのシステムティックレビューにはいくつかの相違点がある（表4-2参照）。通常のシステムティックレビューはエビデンスの確実性（質）を明らかにすることを目的としており、診療ガイドライン作成のためのシステムティックレビューは推奨と関連したエビデンスの確実性（強さ）を明らかにすることを目的としている。

表4-2 診療ガイドラインのためのシステムティックレビューと通常のシステムティックレビュー

項目	診療ガイドライン作成のためのシステムティックレビュー	通常のシステムティックレビュー
目的	益と害の判定のために効果の大きさとエビデンスの強さを明らかにする。	効果の大きさ（効果指標の統合値と信頼区間）と確実性を明らかにする。
研究の対象者	属性が限定的でより特異的なものから広範囲まで。	同一疾患で属性が広範囲。
アウトカム	複数の益のアウトカムと害のアウトカムに対する効果が評価される。	1～3つ程度の益の主要アウトカムに対する効果の評価が中心になることが多い。
害のアウトカム	・益のアウトカムと同じように重要性が評価される。 ・必要な場合観察研究も対象とされ、包括的に解析される。	・副次的に扱われる可能性がある。 ・RCTの統合では頻度の低い害は解析されない可能性が高い。
統合対象研究のデザイン	異なる研究デザインを含む。	1つの研究デザインだけを取り扱っても成立する。
非直接性	PICOの各項目についてCQとの類似性が評価される。	統合される研究間の類似性が評価される。
研究の質の評価	同じ研究でもアウトカムごとにあるいは非直接性により異なる可能性がある。	研究ごとに一意的に決めることが可能。
文献の管理	1つの文献が複数のアウトカム、あるいはCQに関連し、複雑になりやすい。	比較的単純。
文献の選定	アウトカムごとに行うと同じ文献を何回もチェックしなければならない。	アウトカムが1つであれば比較的単純。
研究の集合	1つの介入でアウトカムごとに複数の集合が必要。	1つの介入で1つの集合がある。

これらの相違点で特に注意すべき点は、診療ガイドラインのためのシステムティックレビューでは益と害のアウトカムの両方が重要視されること、同じ研究でもアウトカムが異なると質の評価が異なること、同じ研究が異なるアウトカムに対して適用される可能性があるため文献管理が複雑になることである。

4.0.4 既存のシステムティックレビューの利用

診療ガイドライン作成を目的としたシステムティックレビューは、エビデンス総体の強さを評価し推奨を決定することを目的としている。特に注意すべき点は、明確なCQに対する網羅的文献検索(McDonagh M 2013)と文献採用基準が求められること、益と害のアウトカムの両方が重要視されること、同じ研究でもアウトカムが異なると質の評価が異なること、同じ研究が異なるアウトカムに対して適用される可能性があるため文献管理を一元的に行う工夫が必要となることである。

すでに同じCQに対応するシステムティックレビューが出版されている場合には、それを利用することができる場合もある。図4-2にその際の方針を示す(White CM 2009)。すでに同じCQに対応する診療ガイドラインが出版されている場合には、診療ガイドラインについてはAGREE II(The Appraisal of Guidelines for Research & Evaluation II), AMSTAR 2(A Measurement Tool to Assess Systematic Reviews 2)(Shea BJ 2017)[AMSTAR(Shea BJ 2007)の改訂版]などのツールを用いて評価し、質の高いものを選択し、その中に含まれるシステムティックレビューの利用については、上記の既存のシステムティックレビューの場合と同じ方針を採用する。

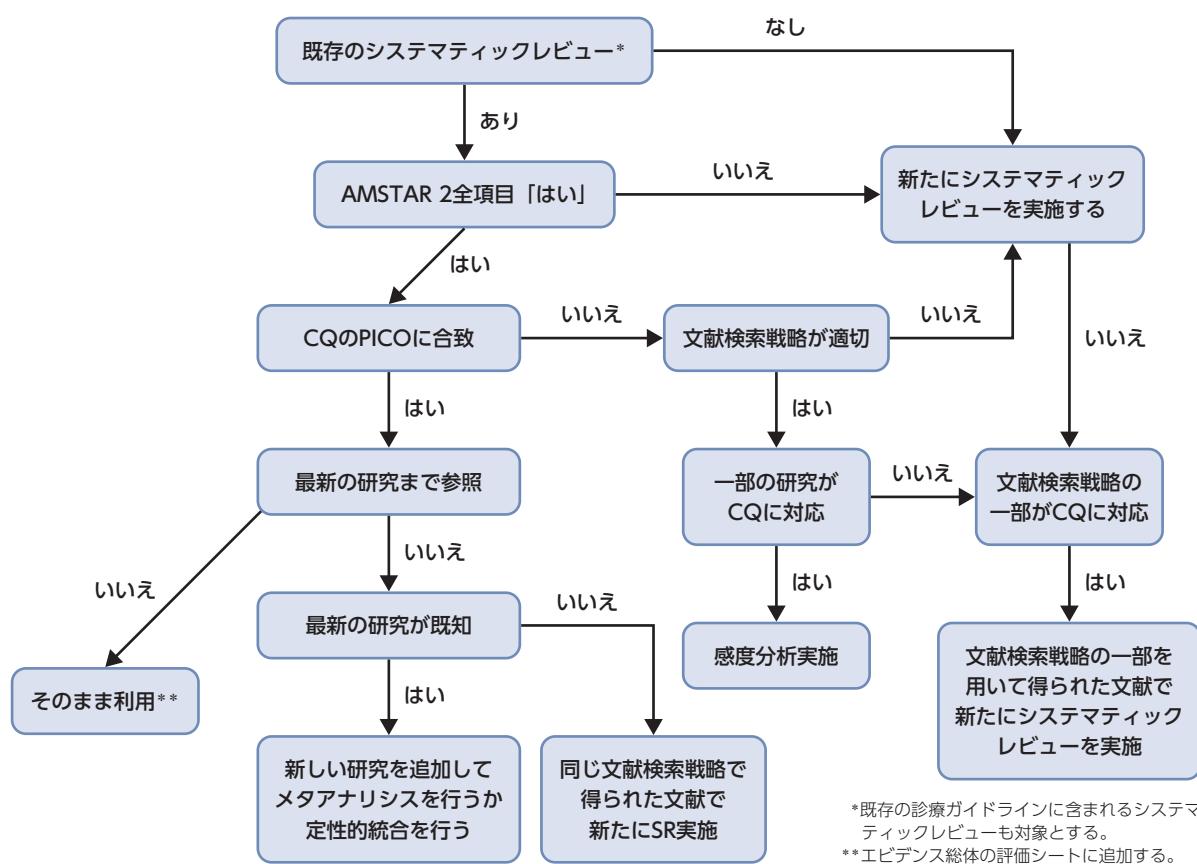


図4-2 既存のシステムティックレビューを利用する場合の方針

出典：White CM, et al. (2009) Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. In: Agency for Healthcare Research and Quality (AHRQ). Methods Guide for Comparative Effectiveness Reviews. By courtesy of Dr. White https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/methods-guidance-de-novo-processes_methods.pdf (参照2021/02/24) より作成

- ① 統合結果をそのまま利用する。
- ② 論文で採択されている元の研究データを再評価し、元の研究から CQ に適合するものを選択してシステムティックレビューを行う。
- ③ 同じ文献検索戦略を用いて得られた文献で新たにシステムティックレビューを行う。
- ④ 新しい研究を追加してメタアナリシスを行うか定性的な統合を行う。
- ⑤ 文献検索戦略の一部を用いて得られた文献で新たにシステムティックレビューを行う。

どの方針を用いるかは、そのシステムティックレビューそのものの質 [これは、以下の表 4-3 などの評価で行うものであり、エビデンスの確実性（質）や研究の質（バイアスリスク）ではない] が高いか低いか、また、そのシステムティックレビュー出版後の新たな研究がどれくらいあるか、それらの結果を含めると結果が変わる可能性があるなどと評価して決める。

表 4-3 AMSTAR 2：ヘルスケア介入のランダム化または非ランダム化研究あるいは両方を含むシステムティックレビューのための厳密な評価ツール

項目 1 レビューの際、リサーチ・クエスチョンと組入れ基準に PICO の要素を含めたか？

Yes の場合： <input type="checkbox"/> 参加者 <input type="checkbox"/> 介入 <input type="checkbox"/> コントロール群 <input type="checkbox"/> アウトカム	選択項目（推奨） <input type="checkbox"/> フォローアップ期間	<input type="checkbox"/> Yes <input type="checkbox"/> No
---	--	---

項目 2 レビューの報告に、レビュー実施前にその方法を決定したことを明示し、重大な計画変更があれば正当な理由を示したか？

部分的な Yes の場合： 著者は以下のすべてを含むプロトコールや基準があることを明示する： <input type="checkbox"/> レビュー・クエスチョン <input type="checkbox"/> 検索戦略 <input type="checkbox"/> 組入れ／除外基準 <input type="checkbox"/> バイアスリスク評価	Yes： 部分的な Yes の場合に加え、プロトコールが事前に登録され、また特定されるべき： <input type="checkbox"/> 必要に応じて、メタ分析／併合の計画、および <input type="checkbox"/> 異質性の原因を探る計画 <input type="checkbox"/> プロトコールからの何らかの逸脱に対する妥当性	<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No
---	--	---

項目 3 レビュー著者は、レビューに組入れる際の研究デザインの選択基準について説明したか？

Yes の場合、そのレビューは以下の 1 つを満たすべきである： <input type="checkbox"/> RCT のみを組入れたことの説明 <input type="checkbox"/> または、NRSI *のみを組入れたことの説明 <input type="checkbox"/> または、RCT と NRSI の両方を組入れたことの説明	<input type="checkbox"/> Yes <input type="checkbox"/> No
---	---

*NRSI : 非ランダム化ヘルスケア介入研究 (Non-Randomised Studies of healthcare Interventions)

項目4 レビュー著者は、網羅的な文献検索の手法を用いたか？

部分的な Yes の場合（以下のすべて）：	Yes の場合、次も必要（以下のすべて）：
<input type="checkbox"/> 少なくとも 2 つのデータベースで検索した（リサーチ・クエスチョンに関して）	<input type="checkbox"/> 検索した参考文献リスト／組入れた研究の文献一覧
<input type="checkbox"/> キーワードと検索戦略を示した	<input type="checkbox"/> 検索した試験／研究登録データベース
<input type="checkbox"/> 出版形態の制限の正当性を示した（例：言語）	<input type="checkbox"/> 検索で組入れた／相談したその分野の専門家
	<input type="checkbox"/> 関連する場合は、検索した灰色文献
	<input type="checkbox"/> レビュー完成の 24 カ月以内に検索を実施した

項目5 レビュー著者は、複数人で研究選択を行ったか？

Yes の場合、以下のいずれか 1 つを含む：	
<input type="checkbox"/> 少なくとも 2 人のレビュー者が、基準を満たす研究の選択について独立して合意し、組入れる研究について合意形成した	<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> または、2 人のレビュー者が基準を満たす研究のサンプルを選択し、良好な一致度（少なくとも 80%）に至り、残りは 1 人のレビューによって選択された	

項目6 レビュー著者は、複数人でデータ抽出を行ったか？

Yes の場合、以下のいずれか 1 つを含む：	
<input type="checkbox"/> 少なくとも 2 人のレビュー者が、組入れられた研究から抽出するデータについて合意形成した	<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> または、2 人のレビュー者が基準を満たす研究のサンプルからデータを抽出し、良好な一致度（少なくとも 80%）に至り、残りは 1 人のレビューによって抽出された	

項目7 レビュー著者は、除外した研究のリストを提示し、その除外が正当である理由を示したか？

部分的な Yes の場合：	Yes の場合、次も含まれなければならない：	
<input type="checkbox"/> 除外されたがフルペーパーとして十分に読み込まれたすべての潜在的に関連する研究のリストを提供した	<input type="checkbox"/> 潜在的に関連する研究についてレビューから除外したことの正当性を示した	<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No

項目8 レビュー著者は、組入れた研究を十分詳細に記述したか？

部分的な Yes の場合（以下のすべて）：	Yes の場合、次も必要（以下のすべて）：	
<input type="checkbox"/> 参加者の記載	<input type="checkbox"/> 詳細な参加者の記載	<input type="checkbox"/> Yes
<input type="checkbox"/> 介入の記載	<input type="checkbox"/> 詳細な介入の記載（関連する投与量も含む）	<input type="checkbox"/> Partial Yes
<input type="checkbox"/> 比較対照の記載	<input type="checkbox"/> 詳細な比較対照の記載（関連する投与量も含む）	<input type="checkbox"/> No
<input type="checkbox"/> アウトカムの記載	<input type="checkbox"/> 研究の場所の記載	
<input type="checkbox"/> 研究デザインの記載	<input type="checkbox"/> フォローアップ期間の記載	

項目9 レビュー著者は、レビューに組入れた各研究のバイアスリスク評価に際し、満足のいく手法を用いたか？

RCT	
部分的な Yes の場合、以下の状態のバイアスリスクが評価されていなければならぬ：	Yes の場合、以下の状態のバイアスリスクも評価されていなければならぬ：
<input type="checkbox"/> 隠蔽されていない割振り、および <input type="checkbox"/> アウトカムを評価する際の、参加者と評価者のブラインディングの欠如（すべての死亡原因のような客観的なアウトカムの場合には不要）	<input type="checkbox"/> 本当にランダムではない割振り手順、および <input type="checkbox"/> ある特定のアウトカムについて多重測定や多重解析から報告された結果の選択

NRSI

部分的な Yes の場合、以下の状態のバイアスリスクが評価されなければならない：

- 交絡因子、および
- 選択バイアス

Yes の場合、以下のバイアスリスクも評価されなければならない：

- 曲線とアウトカムを確認するために使用された方法、および
- ある特定のアウトカムについて多重測定や多重解析から報告された結果の選択

- Yes
- Partial Yes
- No
- Includes only NRSI

項目 10 レビュー著者は、レビューに組入れた研究に対する資金源を報告したか？**Yes の場合：**

- レビューに組入れられた個々の研究の資金源が報告されなければならない（注：レビュー者がこの情報を探したが、著者によって報告されていない場合は、そのように報告する）

- Yes
- No

項目 11 メタ分析が行われた場合、レビュー著者は、結果の統計的併合において適正な方法を用いたか？**RCT****Yes の場合：**

- 著者はメタ分析でデータ併合を正しく行った
- そして、研究結果の併合に適した重みづけ法を用い、存在する場合には異質性を調整した
- そして、異質性の原因について調べた

- Yes
- No
- No meta-analysis conducted

NRSI**Yes の場合：**

- 著者はメタ分析でデータ併合を正しく行った
- そして、研究結果の併合に適した重みづけ法を用い、存在する場合には異質性を調整した
- そして、生データを併合するのではなく、交絡因子を調整した NRSI からの効果量推定値を統計的に併合した、または調整された効果量推定値が利用できなかった場合に生データを併合することの正当性を示した
- そして、RCT と NRSI の両方をレビューに組入れた場合、それぞれに分けて効果量を報告した

- Yes
- No
- No meta-analysis conducted

項目 12 メタ分析が行われた場合、レビュー著者は、各研究のバイアスリスクがメタ分析や他の統合結果に与える影響について評価したか？**Yes の場合：**

- 低バイアスリスクの RCT のみを用いた
- または、プールされた効果量がさまざまなバイアスリスクでの RCT または NRSI に基づく場合、要約効果量におけるバイアスリスクの影響を調べる分析を行った

- Yes
- No
- No meta-analysis conducted

項目 13 レビュー著者は、レビュー結果の解釈や考察をする際にバイアスリスクを考慮したか？**Yes の場合：**

- 低バイアスリスクの RCT のみを用いた
- または、中から高バイアスリスクの RCT や NRSI をレビューに組入れた場合、結果におけるバイアスリスクの影響を考察した

- Yes
- No

項目 14 レビュー著者は、レビュー結果にみられるいかなる異質性に対しても、満足のいく説明や考察を行ったか？**Yes の場合：**

- 結果において有意な異質性はなかった
- または、異質性があった場合、結果における異質性の原因を調べ、レビュー結果に与える異質性の影響を考察した

- Yes
- No

項目 15 量的な統合がなされた場合、レビュー著者は出版バイアス（小規模研究バイアス）について十分な検討、およびレビュー結果に与える影響の可能性を考察したか？

Yes の場合：

- 出版バイアスに対するグラフィック的または統計的検定を行い、出版バイアスの影響の程度を考察した

- Yes
- No
- No meta-analysis conducted

項目 16 レビュー著者は、レビュー実施に際して受領した財源を含め、考えられる利益相反について報告したか？

Yes の場合：

- レビュー著者は利益相反がないことを報告した、または
- レビュー著者は資金源を記述し、潜在的な利益相反にどのように対処したかを記載した

- Yes
- No

出典：上岡洋晴 他. (2018) 「AMSTAR 2：ヘルスケア介入のランダム化または非ランダム化研究あるいは両方を含むシステムティック・レビューのための厳密な評価ツール」の紹介と解説. 薬理と治療 46:111-113.

(1) ペア比較の研究のシステムティックレビュー

2つの介入を直接比較する RCT の結果を統合するシステムティックレビューであり、用いられるメタアナリシスの手法については、米国医療研究・品質庁（Agency for Healthcare Research and Quality : AHRQ）の見解によれば（Morton SC 2018）、ランダム効果モデルを用い、効果指標としては、リスク比（risk ratio : RR）あるいはリスク差（risk difference : RD）を用い、イベント率が1%以下の場合はPeto法を用いること、ゼロイベントが1アームの場合は+0.5などで補正する（この場合はMantel-Haenszel法は用いない）ことが推奨されている。

RDは医療者に理解されやすく決断分析に適しているが、RRに比べるとばらつきが大きくなりやすく、イベント率が低い場合やベースラインリスクのばらつきが大きい場合は検出力が低くなることに注意する必要がある。GRADEアプローチでも、非一貫性の評価の際にはRR、生存分析の場合はハザード比によるフォレストプロット（forest plot）を参照することが勧められている。

ペア比較のメタアナリシスについては、**4.4.4 項（2）** を参照されたい。

(2) ネットワークメタアナリシス

ネットワークメタアナリシス（network meta-analysis : NMA）またはmulti-treatment meta-analysisあるいはmixed treatment comparisons（MTC）（Dias S 2010）は、3つ以上の介入・要因曝露の効果を直接的な比較（head-to-head comparison）だけでなく、間接的な比較も含めてペアでの比較や順位付けを可能にするメタアナリシスである（Dias S 2018）。前提としては、介入の効果を変化させる因子が研究間で差がないこと、すなわち移行性（transitivity）が成立することが必要である。移行性が十分成立する場合、直接比較の情報に間接比較の情報が加味され、より堅牢な結果が得られると考えられている。

例えば、**図4-3**の例1の場合のように治療法がA、B、Cと3種類あり、AとBを直接比較したRCTと、AとCを直接比較したRCTがあり、BとCの比較も行って、いずれの治療効果が最も高いといえるかを知りたいような場合に用いることができる。この例の場合、効果推定値

のパラメータ δ について A と B の差 δ^{AB} , A と C の差 δ^{AC} , B と C の差 δ^{BC} の関係は, $\delta^{AB} + \delta^{BC} = \delta^{AC}$ が成立し, 移項すると $\delta^{BC} = \delta^{AC} - \delta^{AB}$ となり, 治療 B と C を間接比較した場合の効果推定値が得られる。分散 V については, $V^{BC} = V^{AB} + V^{AC}$ が成立する。

図 4-3 の例 2 の場合には, 治療 B と C は直接比較と間接比較の両方が可能となる。この場合は, 直接比較 δ_d^{BC} と間接比較 δ_i^{BC} をそれぞれの分散の逆数で重み付けした平均値を算出する。直接比較の分散 V_d^{BC} , 間接比較の分散 V_i^{BC} となり, 前者は B と C の直接比較の研究データから, 後者は上記のように A と B, A と C の直接比較の研究データから算出された分散から算出される。

ネットワークメタアナリシスでは比較される介入の数が多くなるに従い, 比較されるペアの数が増加する。 S 個の介入を比較する場合, 最大 $S(S-1)/2$ の比較が必要となる。

また, ネットワークメタアナリシスでは, 効果の大きさに順位をつけ, それぞれの治療法が最善, すなわち 1 位になる確率を算出することができる。

ネットワークメタアナリシスに関する執筆ガイダンスとして, PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) 声明の拡張版が発表されており (Hutton B 2015), 可能な限り順守することが望ましい。

R では特定の解析を目的とした「パッケージ」と呼ばれるプログラムが世界中の開発者により提供されている。頻度論派の手法による解析用 R パッケージとして, netmeta パッケージ (Rücker G 2019), ベイジアンの手法による R のパッケージとして, gemtc パッケージ (van Valkenhoef G 2016) がある。gemtc は JAGS (just another gibbs sampler) (Plummer M 2017) を, rjags (Plummer M 2019) を介して動作させるプログラムである。ベイジアンの手法によるネットワークメタアナリシスの多くは OpenBUGS を用いて行われており, そのためのコードが数多く発表されている (Dias S web, Salanti G web) (R については **4.4.4 項 (2)** を参照)。

ネットワークメタアナリシスの解析結果としてネットワークグラフ, フォレストプロット, 順位グラフ (rankogram), ノード分割モデルプロット (node-splitting model plot) (van Valkenhoef G 2016), SUCRA プロット (surface under the cumulative ranking curves plot) (Salanti G 2011) などが得られる。

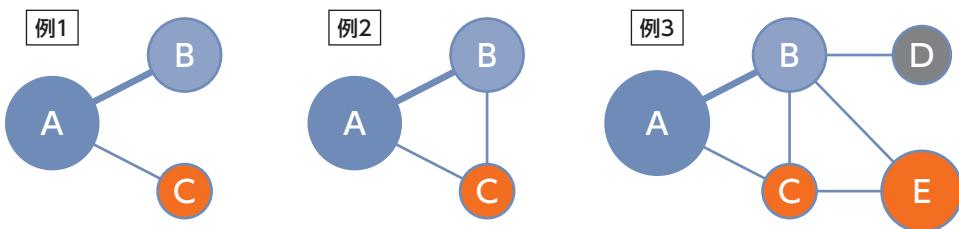


図 4-3 ネットワークグラフ。例 1 では治療 B と治療 C の間接比較が可能になる。例 2 では治療 B と治療 C の直接比較と間接比較も可能になる。例 3 はさらに複雑なモデルである。

ネットワークグラフは各ノードが1つの介入を示し、ノードの大きさは症例数を反映する。ノードを結ぶ直線は直接比較が可能なペアを示し、直線の太さは研究数を反映する。ネットワークグラフが示す概念はネットワークジオメトリ（network geometry）であり、それは介入比較の構造を示す。比較される介入、それぞれの介入にランダム割り付けされた症例数、特定の介入や比較が選好されたあるいは忌避されたかを見て取ることができる。ネットワークはさまざまな形をとり、結合線が少ない場合は、直接比較が少ないと表し、そのようなネットワークメタアナリシスは信頼性が低いと見なされる。ネットワークジオメトリの定性的な記述が必要とされ、ネットワークグラフに付随されるべきであるとされている（Mills EJ 2013）。確立したネットワークジオメトリの見方はまだないので、それぞれの解釈については可能な限り付随する説明を用意することが望ましい。

ペア比較の結果で、直接比較の場合は各研究間の異質性の評価は通常のメタアナリシスと同様で、統計学的異質性（statistical heterogeneity）はQ統計値、 I^2 統計値などで検定されるが、その限界もよく知られているところであり（Gagnier JJ 2012）、臨床的異質性と研究デザインや研究実施の異質性も評価する必要がある。Mills EJ ら（2013）は後者を概念的異質性（conceptual heterogeneity）という言葉で呼んでいる。

ネットワークメタアナリシスの場合は、さらに直接比較の結果と間接比較の結果を比べて不一致がないかを検討し、エビデンスの確実性を評価する必要がある。これは、間接比較も可能なクローズドループ内で行われる。不一致がある場合、インコヒレンス（incoherence）と呼ばれる。具体的には、直接比較と間接比較の結果の統合値と95%信頼区間を一覧表にまとめるなどして、一つひとつ検討することになる。gemtcでは（van Valkenhoef G 2016）、ノード分割モデルの方法で自動的に直接比較と間接比較の統合値と信頼区間のグラフを作成でき、このような作業に用いることができる。

ネットワークメタアナリシスエビデンスに関する従来のGRADEワーキンググループ（GRADE Working Group）の提言では、次のように述べられている。

- ① ネットワーク内のそれぞれのペア比較のエビデンスの確実性は別々に評価する。
- ② 直接比較と間接比較を両方考慮する（Puhan MA 2014）。
- ③ ネットワークメタアナリシスの効果推定値のレーティングには不精確性を考慮する必要はない。
- ④ 直接比較のエビデンスが間接比較のエビデンスよりも確実性が高く、ネットワーク推定値への直接比較の貢献が少なくとも間接比較と同じ程度であれば、間接比較のエビデンス評価は必要ない。
- ⑤ ペア比較のレベルでのインコヒレンスの評価においては、ネットワーク全体のインコヒレンスの統計学的検定を信頼すべきでない。
- ⑥ 直接比較と間接比較のエビデンスにインコヒレンスがある場合、それぞれの推定値のエビデンスの確実性により、どちらのエビデンスを信じるべきかを決める。

したがって、それぞれのペアの比較について、通常のメタアナリシスの場合と同様の評価ドメイン・項目を評価し、エビデンスの確実性を評価する必要がある。

また、順位については、1つの介入が1位になる確率が圧倒的に高いことはまれであり、不確実性を伴っていることを前提に解釈すべきである。

また、GRADE ワーキンググループは図 4-4 のようなインコヒレンスの評価の手順について提言している (Brignardello-Petersen R 2019)。優越的な推定値がなく、バイアスあるいは非直接性が認められる場合はレートダウンし、非移行性のみの場合はレートダウンしない。バイアスと非直接性の両方が影響している場合は、2 レベルのレートダウンを考慮する。

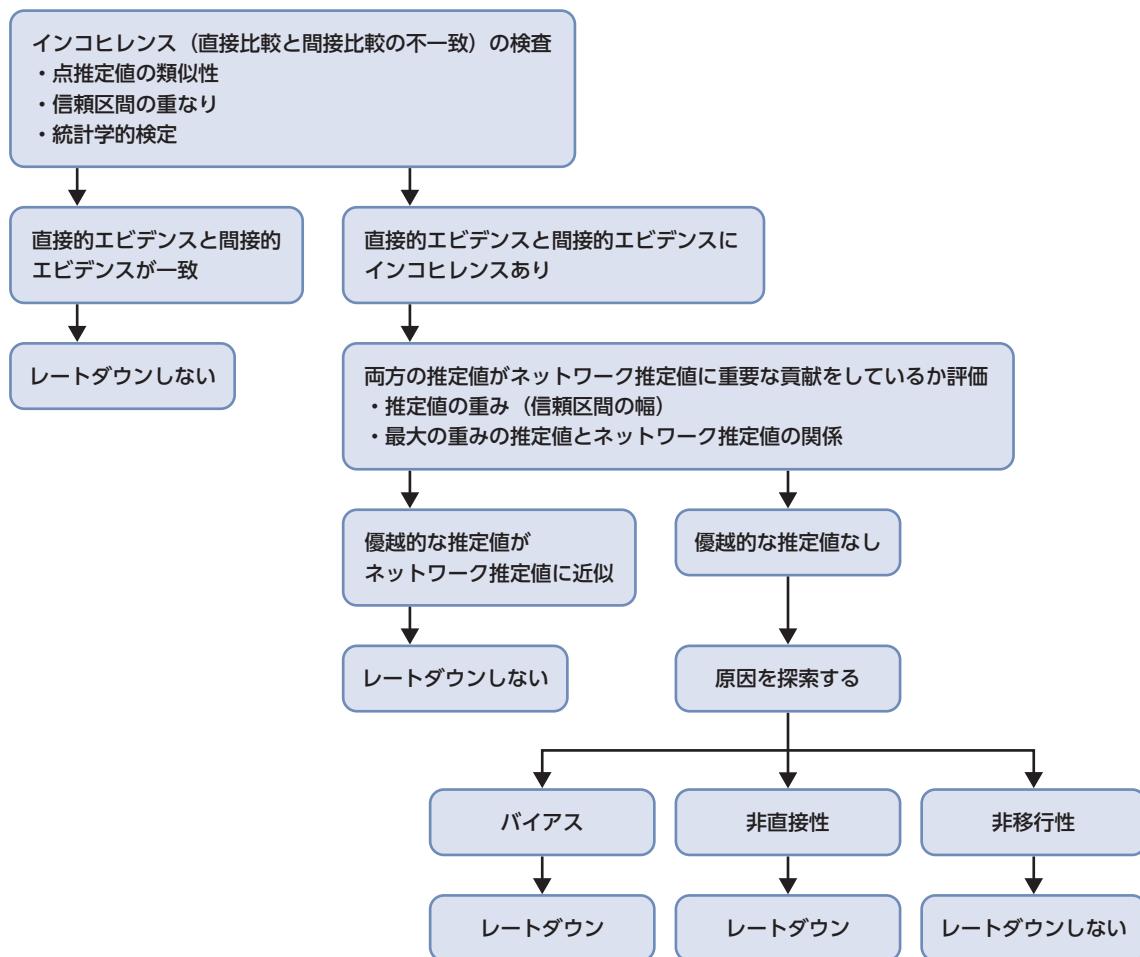


図 4-4 ネットワークメタアナリシスのインコヒレンス評価手順 (GRADE)

出典：Reprinted from Journal of Clinical Epidemiology, Vol 108, Romina Brignardello-Petersen et al., GRADE approach to rate the certainty from a network meta-analysis: addressing incoherence, p.77-p.85, Copyright (2018), with permission from Elsevier より作成

また GRADE ワーキンググループからは、ネットワークメタアナリシスに対応した結果のまとめ表、SoF 表 (summary of findings table) のフォーマットが発表されている (Yepes-Nuñez JJ 2019)。

(3) 診断精度研究のシステムティックレビュー

診断法の臨床的アウトカムへの直接的な効果を測定することは困難なことが多く、癌のスクリーニングなどにほとんど限定されている。そのような研究、特にRCTが行われている場合には、通常の治療的介入の場合と同じ手法でシステムティックレビューが可能となる。CQはPICO形式で設定可能であり、研究の評価、エビデンス総体の評価も通常のRCTの場合と同じように行うことができる。益と害の評価についても、同じ手法が適用できる。

診断法に関する研究の多くは診断能を明らかにする目的で行われており、診断精度(diagnostic test accuracy: DTA)研究と呼ばれている。その目的は、診断法の感度・特異度を明らかにすることであり、何らかの参照基準で診断標的の有無を決定した対象者において診断法を実施し、診断標的を有する者における陽性率(感度)ならびに診断標的を有さない者における陰性率(特異度)を明らかにする。

診断精度研究のシステムティックレビューではCQの設定も異なってくる。WHOの診療ガイドライン作成に関するハンドブック(WHO Handbook for Guideline Development)では、PIRT〔参加者(Participants), インデックス検査(Index test(s)), 参照基準(Reference standard), 標的状態(Target condition)〕形式でCQを設定することを述べている(WHO 2019)。AHRQは臨床的なアウトカムの場合と同じように、PICOTS(Patient population, Intervention, Comparator, Outcomes, Timing, Setting)形式でCQ設定を用いることを提唱しており、Outcomesに患者健康アウトカムまたは中間アウトカムを設定し、中間アウトカムに(診断法の)技術的仕様、精度、決断や治療へのインパクトを設定することを提案している(Chang SM 2012)。

Mindsでは特別寄稿(森實敏夫 2015)で述べているように、CQの形式はPICOを適用することを提言している。診断精度研究の場合は、Oに診断標的を設定し、診断標的を検出する方法として参照基準を併記すること、インデックス診断法が1つの場合は、Iにそれを設定し、Cは空欄とすること、直接的な比較で2つの診断法を比較する場合は、Cに直接比較される診断法を設定することを提言している。また、参照基準が不完全な場合や、研究により異なる参照基準が用いられている場合の方法についても述べている。

メタアナリシスの手法としては、検査閾値を考慮したReitsmaまたはRutter-Gatsonisの二変量モデル、あるいはそれに該当する方法を用い、HSROC(hierarchical summary receiver operating characteristic)解析を用いることが一般的である。頻度論派の手法によるRパッケージであるmadaパッケージ、あるいはベイジアンの手法によるOpenBUGSを用いた解析のコードから、感度・特異度のフォレストプロットおよび統合値とSROC曲線が得られる。

診断精度研究のネットワークメタアナリシスについてはPRISMA声明が発表されており(McInnes MDF 2018)、今後活用が一般化する可能性がある。

感度・特異度の統合値から、一定の事前確率に対して、真陽性(true positive: TP)、偽陰性(false negative: FN)、偽陽性(false positive: FP)、真陰性(true negative: TN)の割合を算出することができる。これら4つに分類された対象者は、その後、一定の治療的介入を受け、その治療による益と害を受けることになる。その診断法を実施した場合を実施しない場合と比較し

て、あるいは別の診断法を実施した場合と比較して、益と害に差があるかどうかを検討する作業が必要になるが、診断法実施後の転帰について実際のデータが不十分な場合が多く、この作業は困難なことが多いと考えられる。また、事前確率が異なると TP, FN, FP, TN の割合も異なるので、益と害の大きさとバランスも異なってくる。さらに、診断法実施後の臨床的アウトカムの生起には時間が関与し、マルコフモデル（Markov model）、マイクロシミュレーション（microsimulation）（Knudsen AB 2016）などの決断モデルが必要になると考えられる。

したがって、診断精度研究のシステムティックレビューは感度・特異度の統合値をメタアナリシスによって明らかにすることが中心にならざるを得ない。

GRADE ワーキンググループは、診断精度とその後の治療を含めた転帰のリンクエビデンスの確実性を評価する必要性およびその際のエビデンス総体の総括の確実性としては一番低いレベルを採用することを提案しているが（Schünemann HJ 2019b），益と害の評価には今後、決断モデルの適用が必要であると述べている。

また、診断法実施後の治療の益とコスト（害と負担）を効用値として設定可能な場合は、決定木を作成して（Sox HC 2013），無治療診断閾値（診断法を実施せず治療もしない場合の効用値が、診断法を実施して陽性なら治療する場合の効用値を上回る疾患確率），治療診断閾値（診断法を実施せず治療する場合の効用値が、診断法を実施して陽性なら治療する場合の効用値を上回る疾患確率）を明らかにすることが可能である。無治療診断閾値と治療診断閾値の間の疾患確率は、診断法を実施して陽性なら治療をする場合の効用値が高い範囲となるので、これらの情報を提供することは、有用であろう。

診断精度研究のバイアスリスクは QUADAS-2（Whiting PF 2011）に基づき、以下のドメインおよび評価項目（括弧で示したもの）を評価する。

- ・選択バイアス（臨床に即したランダム選択）
- ・インデックス検査（盲検化）
- ・参照基準（盲検化、不完全な参考基準）
- ・症例減少バイアス（不完全な検査実施）
- ・フローとタイミング（同時期に実施）
- ・その他（データ欠損など）

個別研究について、これらは「高（-2）」、「中／疑い（-1）」、「低（0）」の3段階で評価する。非直接性はCQの対象、インデックス検査、参考基準、アウトカムと各研究のこれら項目との一致性を評価する。また、上記のように3段階評価を行う。

エビデンス総体については介入研究（interventional study）の場合と同様であり、「高」からスタートして各ドメインに問題があればダウングレードの方向で検討する。

4.0.5 エビデンス総体の評価

(1) エビデンス総体

ある臨床上の問題（CQ）に対して収集し得た全ての研究報告を、アウトカムごと、研究デザインごとに評価し、その結果をまとめたものをエビデンス総体と呼ぶ。

このアウトカムは、スコープ作成時に、1つのCQに対して、「患者にとって重要、重大な各アウトカム」として列挙されたものである。

(2) エビデンス総体の総括

エビデンス総体をアウトカム横断的に統合した全体を、「エビデンス総体の総括」と呼ぶ。診療ガイドライン作成過程において、CQに対する推奨を提示する場合（**第6章**参照）に、エビデンス総体の総括としてのエビデンスの確実性（強さ）を1つだけ決定する（Guyatt G 2013a）。

論理的には、推奨への影響が最も大きいアウトカムに対するエビデンス総体のエビデンスの確実性（強さ）を、エビデンス総体の総括の確実性（強さ）として設定すべきである。例えば、益と害のバランスが、益に傾き、益の大きさが十分であると判断した場合、その益のアウトカムに対するエビデンス総体のエビデンスの確実性（強さ）を、総括のエビデンスの確実性（強さ）として設定することが妥当である。

なお、GRADEアプローチでは、推奨にとって重大なアウトカムに対するエビデンス総体の確実性のうち一番低いものに決定する。

(3) 研究デザインの評価とエビデンスレベル

このように、エビデンスに関する評価は、エビデンス総体に関する評価、および、後述するように複数のアウトカムに対するエビデンス総体を見渡して、エビデンス総体の全体に関する評価を行うことで、推奨に関するエビデンスの強さについての評価とする。

「研究報告単位の研究デザインのみに基づいたエビデンスレベル付け」は、行わなくてもよい。

(4) エビデンス総体の評価

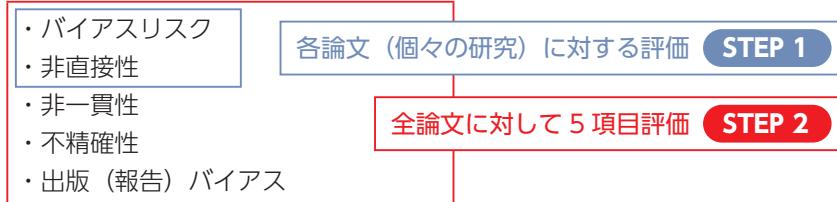
診療ガイドラインの作成においてはシステムティックレビューを実施し、エビデンスの評価、その強さと限界について明確に提示することが望ましい。なお、エビデンスの強さは研究デザインのみで決定されるのではなく、報告内容を詳細に評価、検討し、場合によってはさらにメタアナリシスを行って評価することが望ましい。

エビデンス評価はアウトカムごとに行い、エビデンス総体の強さと限界を明確に記載することが望ましい（**表4-4**参照）。

表 4-4 エビデンスの確実性の評価と判定方法の具体的な作業

- (1) CQに対し、アウトカムを列挙し、重大、重要なアウトカムを採用
 (2) 採用された一つひとつのアウトカムに対し以下の作業を行う
 1) 研究デザインによる分類（RCT群、観察研究群としてまとめる）
 2) RCTに対するエビデンスの確実性の評価

エビデンスの確実性の評価を下げる項目



- 3) 観察研究に対するエビデンスの確実性の評価

エビデンスの確実性の評価を上げる項目

- ・介入による大きな効果 (large effect)
- ・用量 - 反応勾配 (dose-dependent gradient)
- ・可能性のある交絡因子による効果の減弱 (plausible confounder)

注) 各項目の重み付けは CQ によって異なる

出典 : Schünemann H, et al. (2013) GRADE Handbook. Permission by the editors. <http://gdt.guidelinedevelopment.org/app/handbook/handbook.html> (参照 2021/02/24) より作成

4.0.6 エビデンスの強さの考え方

診療ガイドラインにおけるエビデンスの強さは、その治療効果推定値に対する確信が、ある特定の推奨を支持する上で、どの程度適切かを反映している (Balshem H 2011)。その決定の際に、複数の益と害のアウトカムに対するそれぞれのエビデンスの確実性を臨床的文脈全体の中で相互の判断への影響を考慮すべきであることが、GRADE ワーキンググループから提言されており (Hultcrantz M 2017)，システムティックレビューの段階のエビデンス総体の確実性、すなわちエビデンスの確実性（質）と、推奨作成段階のエビデンス総体の確実性、すなわちエビデンスの確実性（強さ）が異なる場合があり得る。システムティックレビュ一段階のエビデンス総体の確実性はエビデンスの質の評価として確定しておき、推奨作成の時点で評価が変わる場合は、その説明を記録することが望ましい。

エビデンスの確実性は、個々の研究報告の研究デザインのみに基づいて決定されるのではなく、①個々の報告内容を詳細に評価検討した上で、②場合によってはさらに統合解析を行って、アウトカムごとにまとめられた研究報告の「総体」としてのエビデンスが決定される。

システムティックレビュ一段階でのエビデンス総体のエビデンスの確実性の基準については、診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)において採用する基準をあらかじめ決めておくことが望ましい。表 4-5 に本マニュアルで推奨するエビデンスの確実性の例を示す。これは GRADE アプローチと同じ分類である。

表 4-5 エビデンス総体のエビデンスの確実性（質）

A (高) :	効果の推定値に強く確信がある
B (中) :	効果の推定値に中程度の確信がある
C (低) :	効果の推定値に対する確信は限定的である
D (とても低い) :	効果の推定値がほとんど確信できない

○ ステップ

- (1) エビデンスの収集
- (2) スクリーニング
- (3) エビデンスの評価：個々の報告評価 **STEP 1**
- (4) エビデンス総体の評価 **STEP 2**
- (5) システマティックレビューレポートの作成

4.1 ステップ1：エビデンスの収集

4.1.1 文献検索データベース

包括的な検索を行うため、複数のデータベース検索を行う（表4-6 参照）。MEDLINE/PubMed^{*4}、The Cochrane Library^{*5}、医中誌 Web の3つのデータベースは必ず検索する。医学領域のEmbase^{*6}やJMEDPlus、社会・心理学的研究のPsycINFO[®]、看護研究のCINAHL（Cumulative Index to Nursing and Allied Health Literature）など、診療ガイドラインのテーマにより重要なデータベースも併せて検索する。必要と考えられる場合は、PROSPERO、臨床試験登録、会議録、医薬品副作用情報なども検索することを検討する。

データベースによって検索式の記述法は異なるが、多くの場合互いに変換することも可能である。

表4-6 主な文献検索データベース

重要度	データベース名	URL	閲覧要件
必須	PubMed	https://pubmed.ncbi.nlm.nih.gov/	なし
	MEDLINE	https://ovidsp.dc2.ovid.com/ovid-a/ovidweb.cgi など	要契約
	The Cochrane Library	https://www.cochranelibrary.com/	検索はフリーだが検索結果の保存は要契約
	医中誌 Web	https://search.jamas.or.jp/	要契約
重要	Embase	https://www.embase.com/	要契約
	PsycINFO [®]	https://www.apa.org/pubs/databases/psycinfo	要契約
	CINAHL	http://web.b.ebscohost.com/	要契約
	JMEDPlus	https://jdream3.com/	要契約
	Open Gray	http://www.opengrey.eu/	なし
	Clinical Trials.gov	https://www.clinicaltrials.gov/	なし
	World Health Organization (WHO)	https://apps.who.int/trialsearch/	なし
	UMIN 臨床試験登録システム	https://www.umin.ac.jp/ctr/index-j.htm	なし
	国立保健医療科学院	https://rctportal.niph.go.jp/	なし

*4 MEDLINEにはPubMed以外にOvidなどの商用データベースもある。

*5 The Cochrane LibraryにはCDSR (The Cochrane Database of Systematic Review)、CENTRAL (Cochrane Central Register of Controlled Trials)などのデータベースがある。例えばCENTRALだけを検索した場合はCENTRALと、指定していない場合はThe Cochrane Libraryと記載する。

*6 RCTに関しては、The Cochrane LibraryのCENTRALから検索可能だがタイムラグがある。チーム内で利用環境が整っていれば検索対象とする。

4.1.2 文献検索戦略

システムティックレビューチーム (systematic review team : SR チーム) は、CQ ごとにキーワード、シソーラス (MeSH など) を組み合わせた検索式を 2 名 (1 名は図書館員など医学文献検索専門家などであることが望ましい) が確認して最適な検索式を作成し、データベースごとに検索式、検索期間、検索日を記録する。検索者が 1 名しか得られない場合は検索式の査読を受けることも検討する。システムティックレビューの検索式査読のためのガイドライン PRESS (Peer Review of Electronic Search Strategies) が開発されており (McGowan J 2016), チェックリストが日本語に訳されている (河合富士美 2017)。

適切な検索を行うためには、診療ガイドライン作成グループと検索担当者との間での重要臨床課題 (key clinical issues) と CQ に対する理解の共有が重要となるため、両者の共同作業が必要である。いつ共同作業が必要かは、次のようにいくつかの時点に分けて考えるとよい。

- ① 検索語句 (キーワード) を選出する時点
- ② 検索式を作成する時点
- ③ 検索実施後、検索結果が得られた時点 (例えば、文献数が多すぎたり少なすぎたりした場合)
- ④ 文献選定後に漏れやデータベースの追加などで検索式の見直しが必要になった時点

それぞれの時点で独立作業と、照合とピアレビュー・修正の共同作業を行うことがあり得る。また、複数の検索式が用いられ、文献検索と文献の選定作業が同時進行するような場合には、すでに選定作業の終了した文献集合を除くなど、選定作業に重複が生じないような工夫が必要である。

通常のデータベース検索では不十分と考えられる場合は、引用文献、会議録などのグレイリテラチャ (grey literature)、教科書の参照など、検索外の追加を検討する。その場合も手段と件数を記録しておく。ただし、検索外の追加については委員会での十分な検討が必要である。

全ての検索式、文献選択の経過は **テンプレート【SR-1】** に記録し、さらに **テンプレート【SR-2】** に示すように、CQ ごとに PRISMA 声明のフローダイアグラムを改変した文献検索フローチャートに記載する。

(1) 検索の進め方

システムティックレビュー、臨床研究は、PICO の P, I, C を用いて系統的に検索するが、害と不利益を含んだ幅広いアウトカムを拾うため、検索式に O は含めない。プロトコルにあらかじめ記載されている場合にはさらに研究デザインでの絞り込みを行う。**テンプレート【SR-1】** を用いた益のアウトカムに関する検索例を以下に示す。

【PubMed検索例1 〈益の検索〉】

タイトル： Stroke and rt-PA
CQ： 発症6時間以内の80歳未満の急性脳塞栓患者（patients）に対して、発症3時間以降のrt-PA投与（intervention）は推奨されるか？
データベース： PubMed
日付： 2019/08/25
検索者： NK／FK

#	検索式	文献数
1	"acute ischemic stroke" [TIAB]	12946
2	"acute ischaemic stroke" [TIAB]	1874
3	"Brain Ischemia" [Mesh]	103113
4	"Stroke" [Mesh]	124695
5	#1 OR #2 OR #3 OR #4	175665
6	"Tissue Plasminogen Activator" [Mesh]	17046
7	rt-pa [TIAB]	2554
8	alteplase [TW]	1559
9	#6 OR #7 OR #8	19457
10	#5 AND #9	5914
11	<u>#10 AND (Meta-Analysis [PT] OR systematic [SB])</u>	<u>144</u>
12	<u>#10 AND "Randomized Controlled Trial" [PT]</u>	<u>328</u>

【PubMed検索例2〈益の検索〉】

#	検索式	文献数
1	"Cerebrovascular Disorders" [Mesh:NoExp]	45882
2	"Brain Ischemia" [Mesh]	103113
3	"Carotid Artery Diseases" [Mesh:NoExp] OR "Carotid Artery Thrombosis" [Mesh]	24050
4	"Stroke" [Mesh:NoExp] OR "Brain Infarction" [Mesh]	124325
5	"Hypoxia-Ischemia, Brain" [Mesh]	5319
6	"Cerebral Arterial Diseases" [Mesh:NoExp] OR "Intracranial Arterial Diseases" [Mesh:NoExp]	2921
7	"Intracranial Embolism and Thrombosis" [Mesh:NoExp]	8680
8	stroke* [TW] OR apoplex* [TW] OR "cerebral vascular" [TW] OR cerebrovasc* [TW] OR cva [TW] OR "transient ischemic attack*" [TW] OR "transient ischaemic attack*" [TW] OR tia [TW]	393220
9	brain [TW] OR cerebr* [TW] OR cerebell* [TW] OR vertebrobasil* [TW] OR hemispher* [TW] OR intracran* [TW] OR intracerebral [TW] OR infratentorial [TW] OR supratentorial [TW] OR "middle cerebral" [TW] OR mca* [TW] OR "anterior circulation" [TW]	1851347
10	ischemi* [TW] OR ischaemi* [TW] OR infarct* [TW] OR thrombos* [TW] OR emboli* [TW] OR occlus* [TW] OR hypoxi* [TW]	1172113
11	#9 AND #10	238225
12	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #11	531896
13	"Tissue Plasminogen Activator" [Mesh]	18149
14	"PLAT protein, human" [Supplementary Concept]	149
15	plasminogen [TW] OR plasmin [TW] OR tPA [TW] OR t-PA [TW] OR rtPA [TW] OR rt-PA [TW]	78965
16	alteplase [TW]	1890
17	#13 OR #14 OR #15 OR #16	79409
18	#12 AND #17	11379
19	<u>#18 AND (Meta-Analysis [PT] OR systematic [SB])</u>	<u>245</u>
20	("Randomized Controlled Trial" [PT] OR "Controlled Clinical Trial" [PT] OR randomized [TIAB] OR placebo [TIAB] OR "Clinical Trials as Topic" [Mesh: noexp] OR randomly [TIAB] OR trial [TI]) NOT (Animals [MH] NOT Humans [MH])	1144087
21	<u>#18 AND #20</u>	<u>2007</u>

(2) PubMedにおける研究デザインの絞り込み方法

1) システマティックレビュー

初めに、先行するシステムティックレビューを検索する。最新で質の高いものが見つかった場合は個別研究の検索が省略できる、または掲載されている検索式を利用して続く検索を行い文献を追加することができる可能性がある。具体的には、systematic [SB] というサブセットを検索結果に掛け合わせる。ただし、別に Meta-Analysis [PT] という Publication Type もあるので、両者を OR で結ぶとより包括性が高まる。なお、Systematic Review [PT] という Publication Type も 2019 年に新設されている。他にも BMJ^{*7} や CADTH^{*8} などのフィルターがある。

2) RCT

スコーピングサーチの結果、RCT でシステムティックレビューを行うことが予定されている場合は検索結果を RCT に絞る。方法としては、PubMed の "Randomized Controlled Trial" [PT]、またはその上位概念である Controlled Clinical Trial [PT] を掛け合わせる。さらに広く拾う場合は Cochrane が用いているフィルター^{*9}を掛け合わせることを検討する。

Cochrane には PubMed 検索用の RCT フィルター 2 種類が用意されている。

感度最大化バージョン (2008 revision), PubMed フォーマット

randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab]

感度・正確度最大化バージョン (2008 revision), PubMed フォーマット

randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR placebo [tiab] OR clinical trials as topic [mesh: noexp] OR randomly [tiab] OR trial [ti] NOT(animals [mh] NOT humans [mh])

Cochrane の検索戦略の考え方は以下のとおりである。

① 検索式は 3 つの検索語句のセットからなる。

- ・対象（健康状態を表す語句）P
- ・介入を表す語句 I
- ・研究デザイン

* 7 BMJ (<https://bestpractice.bmj.com/info/us/toolkit/learn-ebm/study-design-search-filters/>) ※ Ovid 用

* 8 CADTH (<https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#syst>)

* 9 <https://work.cochrane.org/pubmed>

- ② 検索戦略の開発は、すでに引き出された文献集合に基づいて、検索語句が修正・改変される繰り返しプロセスである。
- ③ 検索式の修正と再検索を繰り返していくと追加で見つかる文献の数は減少していき、それ以上の繰り返しのリターンが労力に見合わない点に達する。
- ④ 多すぎる異なる検索コンセプトは避ける。しかし、それぞれのコンセプトの中で広範な同義語と関連語句（フリーテキストと制御された語彙の用語の両方を含む）を OR で結合して用いる。
- ⑤ 異なるコンセプトを AND で結合する。
- ⑥ RCT のための感度最大化検索フィルターを最初に使用し、多すぎたら感度・正確度最大化フィルターを用いる。
- ⑦ 新しい文献の検索にはインデックスされていない文献用に別の検索を行う（可能な場合）。

(3) 害の検索の進め方

観察期間が短いため RCT に含まれる害のアウトカムでは不十分なことが多い。積極的に害についての研究を探す場合は、研究デザインを RCT ではなくコホート研究（cohort study）などの観察研究（observational study）とし、“adverse effects”, “chemically induced”, “complications”などのサブヘーディングを掛け合わせる方法がある。害に関する検索フィルター^{*10}はいくつか開発されている^{*11}。

(4) 費用対効果検索の進め方

医療経済評価研究の検索には、近年の研究から MEDLINE (PubMed), Embase, HTA Database^{*12} の 3 つのデータベース (Arber M 2017) が必要とされ、そのうち PubMed, HTA Database は無料で利用できる。より完全な医療経済評価研究を含むデータベースの一覧リストは SuRe Info^{*13} から入手可能である。費用対効果の検索をする場合は、可能な限り医療経済の専門家に相談する。医療経済評価の文献検索、システムティックレビュー、新規の医療経済評価の方法についての詳細は第5章を参照のこと。

また、費用対効果を推奨作成において考慮することを決定している場合、治療効果、診断精度などの文献検索の際に、医療経済評価研究の包含基準に当たるものが見つかる可能性もあるため、医療経済評価研究のシステムティックレビューの包含基準を考慮したスクリーニングを行うことが望ましい。

* 10 害の検索フィルター紹介サイト：ISSG (Search Filters Resource: Filters to Identify Studies of Adverse Effects) <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/adverse-events-filters>

* 11 なお、益の検索の結果、十分な害に関するエビデンスが収集される場合は、害の検索は省略可能である。

* 12 CADTH Evidence Driven (<https://www.cadth.ca/resources/hta-database-canadian-search-interface>)

* 13 HTAi Vortal SuRe Info (<http://vortal.h tai.org/index.php?q=home>)

(5) 診断精度研究の検索

検索式は原則として、「インデックス検査（OR 参照基準）AND 標的診断」から構成され、さらに必要に応じて「AND 検索フィルター」が組み合わされる。

それぞれのデータベースごとに多数の診断用検索フィルターが提案されているが、システムティックレビュー作成にはどれも用いるべきでないとの Cochrane review も報告されている。しかし、検索結果の論文数が非常に多く、NNR (number needed to read) が非常に大きい場合、フィルターの感度と特異度を認識し、漏れ（必要な文献の見逃し）のリスクと作業効率のバランスを考慮した上で、いずれかのフィルターの使用を検討してもよい。

また、すでに診断精度に関するシステムティックレビュー／メタアナリシスが発表されている場合には、非直接性の評価を含めた CQ への適合度、AMSTAR による妥当性評価結果に基づいて採用を考える。

PubMed の Clinical Queries では、診断に関する検索のためのフィルター（Sensitive/Broad）が用いられている。その感度は 98%、特異度は 74% と報告されている（Haynes RB 2004）。

```
(sensitivity* [Title/Abstract] OR sensitivity and specificity [MeSH Terms] OR diagnosis* [Title/Abstract]
OR diagnosis [MeSH:noexp] OR diagnostic* [MeSH:noexp] OR diagnosis,differential [MeSH:noexp]
OR diagnosis [Subheading:noexp])
```

(6) 比較効果研究の検索

PubMed Special Queries: Directory of Topic-Specific PubMed Queries には比較効果研究 (comparative effectiveness research : CER) の検索式が用意されている。

Research Category として、

- Randomized controlled trials
- Observational Studies (cohort, administrative data, registries, and electronic health records)
- Systematic Reviews, Simulations, Models

Selected Topics として、

- Health Disparities
- Costs and Cost Analysis
- Comparative Effectiveness Research as Subject

などがある。これらのいずれか 1 つあるいは全てを検索語句と組み合わせて検索ができる。

(7) 価値観、意向、効用、QOL 研究の検索

患者・市民が診療ガイドライン作成に関わる方法のひとつとして、患者・市民の価値観や希望について調査された文献のレビューがある。この検索は主に推奨作成で必要となっており、通常のシステムティックレビューとは別に行う。

主なキーワードとしては Patient Preference [MeSH] があるが、この MeSH は 2010 年に採用されている。また、最近では患者・市民の価値観や希望に関する検索フィルター (van Hoorn R 2016, Selva A 2017) もいくつか提案されている。

健康効用 (health utility) や QOL について検索する場合の主なキーワードは Quality of Life [MeSH] であるが、広く探す場合はその上位の Life [MeSH] を用いるか、フィルターの利用を検討する。

4.1.3 文献の再検索

検索時期から 1 年以上経過して診療ガイドラインがまだ完成していない場合は新しい文献が出ていないか確認のため、再検索を行う。未発表の大規模な研究が把握された場合、または重要な研究が見つかった場合は、レビューをアップデートすることを検討する。しかし時間的に難しい場合は追加情報として文献情報を記載しておく。

● 手順

- (1) 検索式の作成
- (2) 検索の実行

● テンプレート

- [SR-1 データベース検索結果]
[SR-2 文献検索フローチャート] (PRISMA 声明を改変)

● テンプレート 記入方法

- [SR-1 データベース検索結果 記入方法]
[SR-2 文献検索フローチャート 記入方法]

● テンプレート 記入例

- [SR-1 データベース検索結果 記入例]
[SR-2 文献検索フローチャート 記入例]

4.2 ベースライン調査

4.2.1 一次スクリーニング

原則としてシステムティックレビューチームの2名が独立して一次スクリーニングを行う。一次スクリーニングでは、タイトル、アブストラクトからCQに合っていないもの、採用基準に合致しないものを除外する。例えば、明らかに臨床研究、疫学研究のデザインでないもの（Letter、系統的でない総説）、医薬品の臨床試験のうちphase Iに相当するもの、法令、官庁からの通知なども、採用基準に合致しない場合は除外する。なお、ここではアウトカムは考慮しない。抄録で判断できないものは原則として残す。

2名の結果を照合し、二次スクリーニング用データセットを作成し、文献を収集する。

4.2.2 二次スクリーニング

原則としてシステムティックレビューチームの2名が独立して全文を読み、二次スクリーニングを行う。選択基準に合った論文を選び、2名の結果を照合するが、意見が異なる場合は第三者の意見を取り入れ、採用論文を決定する。研究の二次スクリーニング後の一覧表をテンプレート【SR-3】としてまとめることとする。

4.2.3 文献集合の作成

エビデンスの統合は研究デザインごとにまとめられたアウトカムについて行われ、さらにそれらがまとめられてエビデンス総体を形成する。テンプレート【SR-3】にまとめられた文献集合をアウトカムごと、研究デザインごとに並べ替え、エビデンス総体の評価の準備をする。

4.2.4 文献集合の管理

電子的に収集した文献をインターネット上のクラウドなどで共有する、あるいは大量に印刷して配布する行為は、著作権侵害に当たる可能性があるので十分注意する。文献は、文献コードによって一元的に管理することが望ましい。文献管理の方法はいくつかあるが、異なるCQで何度も同じ文献が出てくる可能性があることから、著者名、発行年で本文中に挿入し、CQごとにまとめて引用文献を記載するハーバード方式（例：Smith J 2013）を推奨する。引用文献は、CQごとに採用論文、不採用論文、その他の引用論文に分けて筆頭著者のアルファベット順に列挙する。以下に本文中の記載例を示す。ハーバード方式では、同年に公表された論文は登場順にアルファベット（a, b, c, ……）を付して区別するが、診療ガイドライン作成のための引用には年内の発行順にこだわる必要はない。

- ① 1件のRCT (Smith J 2013a) では、……
- ② 米国の大規模研究 (Smith J 2013b) では、……
- ③ Smith A 2013 らの報告によれば、……

①, ②は同一人物の 2013 年の研究, ③は別の人物の研究であることが分かる。

また、不採用論文についてはその理由を記録に残すことが重要である。テンプレート【SR-4】を用いてこれらを記録する。

● 手順

- (1) 一次スクリーニング
- (2) 二次スクリーニング
- (3) 文献の管理方針を決定する
- (4) 方針に従って管理する
- (5) 引用文献を記載する

● テンプレート

【SR-3 二次スクリーニング後の一覧表】

【SR-4 引用文献リスト】

● テンプレート 記入方法

【SR-3 二次スクリーニング後の一覧表 記入方法】

【SR-4 引用文献リスト 記入方法】

● テンプレート 記入例

【SR-3 二次スクリーニング後の一覧表 記入例】

【SR-4 引用文献リスト 記入例】

4.3 ベビデンスの評価：個々の研究に対する評価（STEP 1）

エビデンスを評価し、強さと限界については明確に提示する。エビデンスの質は研究デザインのみで決定されるのではなく、報告内容を詳細に評価、検討し、場合によってはさらにメタアナリシスを行って評価することが望ましい。エビデンス評価は、アウトカムごとに行い、エビデンス総体の質あるいは限界を明確に記載することを心掛ける。

4.3.1 論文の選出・選択

系統的検索によって収集された文献（4.1節、4.2節参照）を基に、以下の論文選出・選択作業を行う。

① 診療ガイドラインやシステムティックレビューが抽出された場合

- 現在作成中のCQと比較して、P（対象は同一か）、I（介入は同一か）、C（対照は同一か）、O（アウトカムは同一か）などの詳細な設定を詳しく評価し、さらにCQに関係する全てのアウトカムが診療ガイドラインやシステムティックレビューに含まれているかを確認する。
- 診療ガイドラインやシステムティックレビューが検索対象とした期間を確認し、同じ検索式を用いてアップデートする。その後に新たなRCTなどが出ていている場合は、これを含めて再度評価する。場合によってはメタアナリシスを行う。

② システマティックレビューが複数抽出された場合

- 上記の「診療ガイドラインやシステムティックレビューが抽出された場合」と同様の評価を行い、場合によってはメタアナリシスを行う。
- 益と害の評価に十分な情報があると判定される場合は、観察研究の評価が省略できる場合がある。

③ 除外論文の取り扱い

抽出された論文のうち、表面的には適格基準を満たしているように見えても、さらに調べてみると満たしていないかった研究は、その論文名と除外理由をリストにしておく。

④ 益の評価における観察研究の取り扱い

益の評価でシステムティックレビューや十分なRCTがある場合は、観察研究の評価が省略できる場合がある。

⑤ 害の評価における観察研究の取り扱い

- ・RCT が複数抽出されて、観察期間が十分長い場合は、観察研究の評価を省略できる場合があるが、特に害の評価においては RCT のみでは限界があることを認識すべきである。これは、RCT の研究対象症例数が益の視点から決定されているからであり、害については、観察研究も評価対象とすることが望ましい。
- ・薬剤副作用情報、添付文書も重要な情報である。

4.3.2 個々の研究に対する評価 STEP 1

アウトカムごとにまとめられた文献集合の個々での論文について、研究デザインごとに**テンプレート【SR-5】、【SR-6】**を用いて、バイアスリスク、非直接性を評価し、対象人数を抽出する。その際、効果指標の提示方法が異なる場合は、リスク比、リスク差などに統一して記載する。RevMan (Review Manager)などを用いてメタアナリシスと併せて計算すると簡便である。RCT でもバイアスの影響を受けている可能性があるので (Chess LE 2013, Hamm MP 2010), 個々の研究についていくつかのドメイン・項目別にバイアスリスクの評価を行う。

(1) バイアスリスク：個々の研究についての評価 (Higgins JPT 2011b: 8.8-9 を引用改変)

1) 原則

以下の Cochrane バイアスリスクツール (Cochrane Risk of Bias Tool) による原則 (Higgins JP 2011a) を参考とする。

1. 質評価用のチェックリストあるいは尺度は用いない。
2. 内的妥当性は、非直接性とは分けて評価する。
3. 方法の記載や報告の不備ではなく、研究結果から推測して評価する (Mhaskar R 2012)。
4. 評価者の判断が必要であり、機械的には決められない。
5. バイアスリスクのうち、重要視すべきドメインあるいは項目は一様ではない。
6. 研究対象は患者全体を代表しているものとして評価し、症例を除外することによってバイアスが小さくなる操作が行われていないか注意する。
7. アウトカムごとに分けて評価を行う。

出典：Higgins JP, et al. (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343:d5928.
By courtesy of Dr. Higgins より作成

また、評価者間の一致率は項目によって異なり、全体として中等度であることが知られている (Hartling L 2011, Hartling L 2013, Vale CL 2013), 2名の評価者で異なる結果の場合は、コンセンサスを得て決定する。

2) バイアスリスクのドメイン・項目と定義

バイアスリスクはドメインによって分類され、さらに各ドメインに評価項目が含まれる。現時点では、各バイアスリスクのドメインには1つまたは2つの評価項目が含まれている (表4-7 参照)。

表 4-7 RCT の場合のバイアスリスク

ドメイン	評価項目
選択バイアス (selection bias)	ランダム化 (ランダム配列の生成, ベースライン不均衡)
	割り付けの隠蔽 (コンシールメント)
実行バイアス (performance bias)	参加者と医療提供者の盲検化
検出バイアス (detection bias)	アウトカム測定者の盲検化
症例減少バイアス (attrition bias)	不完全アウトカムデータ
	ITT (intention-to-treat) 解析非実施
その他のバイアス	選択的アウトカム報告
	早期試験中止バイアス
	その他のバイアスの可能性

出典：Higgins JP, et al. (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343:d5928. By courtesy of Dr. Higgins より作成

各ドメインの概念は RCT だけでなく観察研究にも適用される。

臨床研究におけるバイアスは数多く報告されており (Berkman ND 2014, Dechartres A 2011, Viswanathan M 2012), 研究ごとに異なるバイアスが異なる影響を与えている可能性がある。可能であれば、バイアスがアウトカムに影響する大きさ (magnitude), 方向 (direction) を評価すべきであるが、臨床的な知識・経験がないと判定が難しいことが多いので、大きなバイアスのリスクの程度、すなわちバイアスリスクとして判定することが一般的である。バイアスリスクが高い研究は過大評価になっているという報告が多い (van Tulder MW 2009), 疾患, 介入, アウトカムによって異なる可能性があり (Fenwick J 2008), 個別の判断が重要である。

① 選択バイアス

研究対象の選択の偏りにより生じるバイアス。特に、比較される群の研究対象が介入や危険因子への曝露以外の点で異なることによってアウトカムが影響を受けるバイアス (Armijo-Olivo S 2015, Britton A 1999, Savović J 2012)。

例) 比較試験でランダム割り付けが行われていないため、比較される群の年齢構成が異なり、平均年齢が高い方のアウトカムが悪くなる。

- ・アウトカムの差が介入あるいはリスクファクターへの曝露によるものかどうかを考える。
- ・非ランダム化比較試験 (non-randomized controlled trial : non-RCT) や歴史的対照群を用いる場合など、比較される群のさまざまな特性がもともと異なる場合には、選択バイアスが生じる。

評価項目：ランダム化

- ・ランダム配列の生成

ランダム配列生成：患者の割り付けがランダム化されているか、さらにランダム化の方法が乱数表やコンピューターランダム化など適正なものかについて、詳細に記載されているかを検討する（Kjaergard LL 2001）。

Chalmers TC ら（1983）の145件の急性心筋炎のRCTを検討した報告では、割り付けを盲検化したRCTでは死亡率が8.8%であったのに対し、患者を盲検化しなかったRCTでは24.4%，ランダム化が行われなかった試験では58.1%であったとされている。

- ・ベースライン不均衡

症例登録時に、アウトカムに影響を与える患者属性（年齢、重症度、病期など）に比較する群間で差がないかを検討する。これらの情報は、患者背景（patients' profile）の表にまとめられていることが多い。

ランダム配列の生成とベースライン不均衡はランダム化に関わるバイアスリスクとして、区別して評価してもよい。

評価項目：割り付けの隠蔽（コンシールメント）

患者を組み入れる担当者に組み入れる患者の隠蔽化がなされているかを検討する。介入現場での割り付けではなく登録センターや中央化などの方法が有用である（Kunz R 2007, Herbison P 2011, Nüesch E 2009a, Wood L 2008, Pildal J 2007, Schulz KF 1995）。

Herbison F ら（2011）は、389件のRCTを解析した結果、「double blind」の記載があるても、割り付けの隠蔽化が不明瞭な場合はバイアスリスクが高くなり、RORs（pooled ratios of odds ratios）は0.86（95%信頼区間0.78～0.96）であったと報告している。

- ・観察研究の場合は、比較される群のアウトカムに影響を与える得る背景因子がそろっているかを評価する。また、傾向係数を用いた解析（プロペンシティー解析）が行われているかを評価する。
- ・診断法の研究の場合は、実臨床でその診断法が実施される群が対象になっているか（single-gate study）、任意に集められた群が比較されているか（two-gate study）を評価する。

② 実行バイアス

評価項目：参加者と医療提供者の盲検化

比較される群で、介入・ケアの実行に系統的な差がある場合に生じるバイアス (Armijo-Olivo S 2017, Krogh J 2015, Gurusamy KS 2009)。

例) RCT で割り付けが分かってしまい、医療提供者がケアを変えてしまう、あるいは患者が他の治療を受けてしまうなどによってアウトカムに差が出る。

- ・介入群か非介入群か、あるいはどの介入が行われているのかを、患者から分からなくする（単盲検）ことに加えて、医療提供者にも分からなくしているか（二重盲検）を評価する。患者のプラセボ効果や医療提供者のバイアスを排除することを目的とする。
- ・盲検化されていない場合は、それが結果に及ぼす影響を評価する。

③ 検出バイアス

評価項目：アウトカム測定者の盲検化

比較される群で、アウトカム測定に系統的な差がある場合に生じるバイアス (Hróbjartsson A 2012, Liu CJ 2011)。

例) RCT で測定者が割り付けを知ってしまい、新しい治療法に有利な測定結果を出してしまう。

- ・アウトカム測定者が盲検化されているかどうかを評価する。
- ・盲検化されていない場合は、それが結果に及ぼす影響を評価する。
- ・観察研究の場合は、アウトカム測定が正確で、適切なタイミングで行われているか、測定記録が正確かなどを評価する。

④ 症例減少バイアス

比較される群で、解析対象となる症例の減少に系統的な差がある場合に生じるバイアス。

- ・それぞれの主アウトカムに対するデータが完全に報告されているか（解析における採用および除外データを含めて）、アウトカムのデータが不完全なため、症例を除外していないかを評価する (Nüesch E 2009b)。
- ・症例の除外が結果に影響するほど大きいかを評価する。

評価項目：不完全アウトカムデータ

症例の減少はアウトカムが不明であることに根拠があると考えられるので、不完全アウトカムデータとして捉えられる（Riley RD 2004）。

例) 患者が通院を止めてしまい脱落症例となる、副作用により治療を中断し脱落症例となるような場合アウトカムのデータが不完全となるが、その程度が比較する群で異なり、それが群間のアウトカム評価の差に影響する。

評価項目：ITT 解析非実施

ITT 解析とは治療企図分析のこと、RCT の統計解析において、脱落例やプロトコル非合致例を無効例あるいは何らかの方法で補完されたアウトカムの例として割り付けどおりに解析することである。評価統合する報告が RCT であり ITT 解析が行われていない場合には、バイアスが生じている可能性を疑う。

⑤ その他のバイアス

評価項目：選択的アウトカム報告

測定された複数のアウトカムのうち一部しか報告されていない場合、効果の大きい都合のいい結果だけが報告されるというバイアス（報告バイアス）を生じる可能性がある（Bialy L 2014, Norris SL 2012, Ioannidis JP 2011, Dwan K 2010, Kirkham JJ 2010, Kyzas PA 2005, Chan AW 2004a, Chan AW 2004b）。

- 登録された研究プロトコル（研究計画書）に記載されたアウトカムが全て報告されているかどうかを評価する。

評価項目：早期試験中止バイアス

中間解析が計画されたデザインでない、あるいは適切に計画された研究（adaptive study design：適応的研究デザイン）でないにもかかわらず、当初計画されたサンプルサイズを満たす前に効果が証明されたとして中止された臨床試験の場合、効果が過大評価されるバイアスが生じる可能性がある（Bassler D 2013, Schou IM 2013, Hughes MD 1992）。

- あらかじめ多段階の試験が計画されたかどうかを評価する。
- O'Brien-Flemming 法、ペイジアン解析などの方法が採用されているかどうか評価する。

評価項目：その他のバイアスの可能性

上記のバイアス以外のバイアスの可能性。

- ・利益相反 (conflict of interest : COI) とは、教育・研究に携わる専門家としての社会的責任と、産学連携活動に伴い生じる利益などが衝突・相反する状態である。COI の開示・管理などの記載から、結果が COI から影響を受けていないか評価する。
- ・多変量解析により交絡因子 (confounding factor) の調整が行われているかを評価する。
- ・その他のバイアスがあり得るか評価する。例えば、単施設での臨床試験は多施設の臨床試験と異なり (Inaba Y 2009)，前者でより過大評価になるバイアスも報告されている (Unverzagt S 2013, Bafeta A 2012)。また、営利組織から資金提供を受けた臨床試験はポジティブな結果が多いことが報告されている (Als-Nielsen B 2003)。

3) Cochrane バイアスリスクツール v.2.0

RCT のバイアスリスク評価ツールとして発表された Cochrane バイアスリスクツール^{*14}は、アップデートにより従来のバイアスドメイン名から、より直接的表現を用いた長い名称に変更されたが、概念的には従来のバイアスドメインと同じである。ドメインの数も 5 個に限定された。**表 4-8** に、従来のドメイン名と AHRQ の分類を比較して示す。

もうひとつの変更点は、シグナリングクエスチョンに Yes / Probably Yes / Probably No / No / No information または Not applicable で答えることによって、ドメインごとの判定が決定できるようになったことである。Higgins JPT ら (2019) は、その答えの組み合わせで判定が可能な Excel のマクロが付随したファイルをウェブサイトで提供しているが、使用者の判定アルゴリズムを優先することを認めているため、シグナリングクエスチョンに対する回答の組み合わせの結果によるドメインの評価と評価者の評価が異なる場合は、評価者の評価を記録できるようになっている。

また、判定の分類が High, Low, Unclear から High, Some concerns, Low とされた。Unclear がなくなり、3 段階の判定であることがより明確になっている。

* 14 2018 年に Version 2.0 が発表され、2019 年 8 月に若干のアップデートが発表された (Higgins JPT 2019)。

表 4-8 バイアスの名称

従来 ^{*15} のドメイン名／現在の Minds のドメイン（評価項目）	Cochraneバイアスリスクツール v.2.0 の分類	AHRQ の分類
Selection bias 選択バイアス (ランダム化, コンシールメント)	Bias arising from the randomization process ランダム化の過程から生じるバイアス	Bias arising in the randomization process or due to confounding ランダム化の過程あるいは交絡で生じるバイアス
Performance bias 実行バイアス (盲検化)	Bias due to deviations from intended interventions 治療企図からの乖離のため生じるバイアス	Bias due to departures from intended interventions 治療企図からの乖離によるバイアス
Attrition bias 症例減少バイアス (ITT 解析, 不完全アウトカムデータ)	Bias in measurement of the outcome アウトカムデータの欠損のため生じるバイアス	Bias from missing data データ欠損によるバイアス
Detection bias 検出バイアス (盲検化)	Bias in measurement of the outcome アウトカム測定におけるバイアス	Bias in measurement of outcomes アウトカム測定におけるバイアス
Reporting bias 報告バイアスまたは出版バイアス (出版バイアス)	Bias in selection of the reported result 報告結果の選択におけるバイアス	Bias in reporting results selectively 結果を選択的に報告することによるバイアス
Other bias その他のバイアス (選択的アウトカム報告, 早期試験中止, COI によるバイアス, その他)	なし	観察研究の場合以下の 2 ドメインが追加： Bias in selecting participants into the study 研究への参加者の選択におけるバイアス Bias in classifying interventions 介入の分類におけるバイアス

Cochrane バイアスリスクツールのような方法では、評価対象の多面的な属性をドメインおよびドメインの下位の概念である項目として設定し、それぞれの属性を評価し、全体をまとめて判定し、対象の評価を決める。判定の方法には、ドメインベース (domain-based) と目標評価基準の 2 つの手法があり、Cochrane バイアスリスクツールのような方法は、ドメインベースとされている (IOM 2011: 136)。例えば、選択バイアスのランダム配列の生成の判定では、コンピューター生成ランダム配列、封筒シャッフル法は低バイアスリスク、それ以外の方法、例えば、ID 番号や誕生日の末尾の奇数・偶数での割り付けは高バイアスリスクと判定することになっている。これは、項目が該当するかどうかで判定するものに当たるので、ドメインベースと考えられる。一方で、ベースライン不均衡という項目については、介入群と対照群の年齢や重症度にバランスがとれていない場合、それが結果に大きな影響を及ぼすかどうかを判断し、大きな影響を及ぼすと判断したら、高バイアスリスク、その可能性が低いと判断したら低バイアスリスクと判定することになる。この場合は、そのバイアス項目のアウトカムへの影響を判断するものに当たるため、目標評価基準と考えられる。すなわち、評価項目が本来の目標を達成しているかどうかを判断している。

また、Unclear の判定項目が 3 つ以上のときにその研究のバイアスリスクのまとめを高リスクバイアスと判定するというルールを設定すれば、目標評価基準になる。このように、

* 15 Cochrane バイアスリスクツール v.1.0

Cochrane バイアスリスクツールを用いる場合、評価者が意識するしないにかかわらず、ドメインベースと目標評価基準が混在して判定が行われているのが現状である。

Minds の方法は原則的に目標評価基準であり、各バイアスの項目のアウトカムへの影響の大きさの判断に基づいて、高、中、低の3段階の判定を行い、研究ごとのバイアスリスクのまとめも同じく3段階で行う。

Cochrane バイアスリスクツール v.2.0 については、評価が定まるまで時間を要するが、診療ガイドライン作成グループによっては採用を検討してもよい。

4) 観察研究のバイアスリスク評価 (ROBINS-I 含む)

ROBINS-I (Risk of Bias in Non-randomised Studies for Interventions) は観察研究のバイアスリスク評価のツールである (Sterne JA 2016)。そのガイダンスによると観察研究の中でも non-RCT、コホート研究で介入の効果を比較する研究を主な対象としている。

ROBINS-I では target trial として理想の RCT を想定し、それと評価対象の観察研究を比較しながら、バイアスリスクの程度を評価する。ターゲットトライアル (target trial) は実行可能性がないものも想定可能である。したがって、エビデンス総体の評価の際には、観察研究であってもエビデンスの確実性は A (高) からスタートしてグレードダウンする方向で考えることになる。

評価の分類は Low, Moderate, High, Critical の4段階となっている。もし、複数の観察研究のバイアスリスク評価の結果、エビデンス総体としてバイアスリスクのドメインが High になった場合は、エビデンス総体の他の評価ドメイン、非直接性、不精確性、非一貫性、出版（報告）バイアスと併せて評価しても A (高) から C (低) へとグレードダウンすることが考えられる。

GRADE ワーキンググループは従来の GRADE アプローチにおける観察研究のバイアスリスク評価およびエビデンス総体の評価方法と ROBINS-I を用いた場合を比較し、いずれの方法でも結論は同じになる可能性が高いと述べているが、交絡で生じるバイアスの取り扱いについて注意が必要としている (Schünemann HJ 2019a)。GRADE アプローチではエビデンス総体の評価は C (低) からスタートし、大きな効果、量反応関係、効果減弱交絡因子がある場合、グレードアップするルールになっている。C (低) からスタートする理由として、未知の交絡因子の影響を受けている可能性を排除することができないことが挙げられている。

バイアス効果の大きさと方向を評価せずにバイアスのリスクしか評価しない場合、次のような問題が生じる。高リスクと判定した場合、効果減弱に働く交絡因子と、効果増強に働く交絡因子のいずれの場合もエビデンス総体の確実性をグレードダウンの方向で評価してしまうことになる。しかし、効果減弱に働く交絡因子の場合は、確実性をグレードアップする方向で評価しなければならないので、バイアス効果の大きさと方向を評価せずにバイアスのリスクのみを評価すると、エビデンスの確実性の判定を間違えることになる。ROBINS-I を用いる場合は、この点に注意が必要である。それを防ぐためにはバイアス効果の大きさと方向を評価する必要があり、ROBINS-I ではそれがオプションとなっているので、GRADE アプ

ローチの枠組みで用いる場合は必須となるであろう。

また ROBINS-I では、どのような交絡因子があり得るかを最初に疾患専門家と議論して可不足なくリストアップする作業を勧めている。この段階の交絡因子は候補であって、研究結果に影響する交絡因子と決定されたわけではない。交絡因子は、どの介入を受けるかとアウトカムの両方に関連がある因子として同定される。したがって、介入群と対照群のベースラインで頻度あるいは値に差があることが前提となる。時間経過の中で変動する因子を取り上げるかどうかは評価者の方針によるが、観察研究で十分なデータが得られる可能性は低いと考えられる。

Minds では観察研究のバイアスの評価ドメインは RCT と同じ、選択バイアス、実行バイアス、検出バイアス、症例減少バイアス、その他のバイアスを取り上げており、それらの項目として、比較される群間の背景因子の差、ケアの差、不適切なアウトカム測定、不完全なフォローアップ、不十分な交絡の調整、その他のバイアスを取り上げている。他のバイアスに選択的アウトカム報告を含めることによって、ROBINS-I あるいは AHRQ の方法と評価項目が同じになる。比較される群間の背景因子の差は上記のごとくベースラインで差があれば交絡因子の可能性が高いと考えられるので、交絡で生じるバイアスの評価と同じと考えられる。Minds では観察研究のエビデンス総体の評価の際、ROBINS-I のように、A（高）からスタートする方法と、GRADE のように C（低）からスタートする方法のいずれも対応できる。観察研究でもデザインにより、例えば傾向スコア分析法や変数操作法あるいはメンデリアンランダム化法のような RCT に近いものから、症例対照研究、症例集積（case series）、症例報告（case reports）までさまざまな確実性のものがある。未知の交絡因子の影響が大きいと考えられる場合は C（低）からスタートし、その可能性が低いと考えられる場合は A（高）からスタートするという方法も取り得る。

また、AHRQ は Cochrane バイアスリスクツール v.2.0 および ROBINS-I に対応し、さらに RCT と観察研究へ共通して適用できるバイアスリスク評価法を発表している（Viswanathan M 2017）。研究者が治療の割り付けを行う non-RCT と、前向きおよび後ろ向きコホート研究で比較されるアームのある観察研究、そして症例対照研究を対象としている。

ここまで述べた観察研究のバイアスリスク評価方法について、**表 4-9** にまとめた。なお、観察研究の評価については、Cochrane から 2011 年に発表された、RoBANS^{*16} もある。

* 16 RoBANS (Risk of Bias Assessment tool for Non-randomized Studies)

<https://abstracts.cochrane.org/2011-madrid/risk-bias-assessment-tool-non-randomized-studies-robans-development-and-validation-new>

表 4-9 観察研究のバイアスリスク評価方法

Minds 観察研究 (2017 年)	AHRQ ^{a)} (2017 年)	ROBINS-I ^{b)} (2016 年)
non-RCT, コホート研究, 症例対照研究, その他の観察研究	RCT / 非ランダム化研究 ^{c)} / 症例対照研究	non-RCT, コホート研究, その他の観察研究 ^{d)}
選択バイアス	・ランダム化の過程あるいは交絡で生じるバイアス ^{e)} ・研究への参加者の選択におけるバイアス	・交絡で生じるバイアス ・研究への参加者の選択におけるバイアス
比較される群間の背景因子の差 背景因子の割合や値に差がある場合、アウトカムへの影響が想定されれば交絡因子の可能性がある。	アウトカムに影響する予後因子が介入の選択にも影響する場合。 介入 / 要因曝露後に観察された属性により選択された場合（観察研究のみ）。	ベースラインでアウトカムの予知因子で受ける介入の予知ができる場合。 適格例の一部の除外、一部の参加者の最初の経過観察期間、いくつかのアウトカムの生起が介入とアウトカムの両方に関係するとき、効果が同じ介入でも介入とアウトカムに関連が生じる。（例：すでにその治療を受けていた者が参加する。）
実行バイアス	治療企図からの乖離によるバイアス	治療企図からの乖離のため生じるバイアス
ケアの差 比較する群はそれぞれの比較する介入が一定でなく、アドヒアランスが異なったり、他の治療を受けたり、ケアに差がないか。	企図された介入と実際に受けた介入が異なる場合。	介入群と対照群で受けるケアに系統的な差があり、企図された介入からの乖離を表す。このドメインのバイアス評価は、効果のタイプ（介入への割り付けの効果か、介入の開始とアドヒアランスの効果か）による。
検出バイアス	アウトカム測定におけるバイアス	アウトカム測定におけるバイアス
不適切なアウトカム測定 アウトカム測定者が介入に気付いたり、アウトカム測定や評価の信頼性が低かったりしないか。	全体としての、あるいは系統的なアウトカム評価における群間の差。	アウトカムデータ測定の差別的あるいは非差別の過誤によってもたらされる。そのようなバイアスは、もし、異なる介入群で異なる方法が用いられたり、測定過誤が介入の状態あるいは効果に関連していたりする場合、アウトカム評価者が介入の状態に気付くことで起き得る。
症例減少バイアス	データ欠損によるバイアス	データ欠損のため生じるバイアス
不完全なフォローアップ フォローアップが不完全で、アウトカムデータの記録が不正確で、不完全なアウトカムデータか。	全体としての、あるいは系統的な解析に含まれない参加者の減少による群間の差。	最初に対象に含めフォローされた個人の、その後のフォローアップができなかったとき（予後因子に影響されるフォローアップの差別的逸失のような場合）にバイアスが生じる。すなわち、介入の状態あるいは交絡因子のようなその他の変数に関する情報の逸失のあった個人の除外によるバイアス。
	結果を選択的に報告することによるバイアス	報告結果の選択におけるバイアス
	結果に基づいてアウトカムを選択的に報告する。	知見に基づくような、あるいは、メタアナリシス（あるいはその他の統合）による推定値が得られなくなるような結果の選択的報告。
その他のバイアス	介入の分類におけるバイアス	介入の分類におけるバイアス
・不十分な交絡の調整 ・その他のバイアス 多変量解析、傾向スコア法、操作変数法、メンデリアンランダム化法などにより交絡の対処が十分か。 選択的アウトカム報告によるバイアスはその他のバイアスに含める。	介入開始時に、介入の状態が有効で信頼できる方法で記録されず、参加者の介入の状態が間違って分類される。	差別的な、あるいは非差別的な介入状態の誤った分類により引き起こされるバイアス。非差別的な誤った分類はアウトカムとは関連がなく、通常介入の効果を無効果の方向へのバイアスとなる。差別的な誤った分類は介入の誤った分類がアウトカムあるいはアウトカムのリスクと関係があるときに起き、バイアスを引き起こす可能性がある。

太字：ドメイン名（概念的な単位）、下線：項目名

- a) バイアスリスクの評価の分類は、Low, Moderate, High, Unclear である。AHRQ ではバイアス効果の向き（過大評価・過小評価）と大きさの 2 つを評価することを求めている。
- b) バイアスリスクの評価の分類は、Low, Moderate, Serious, Critical, No information となっている。
- c) 研究者が治療の割り付けを行う non-RCT と前向きおよび後ろ向きコホート研究で比較されるアームのある観察研究。
- d) ROBINS-I は non-RCT, 準ランダム化比較試験、コホート研究、症例対照研究に適用される。なお、ガイダンス^{*17} では研究への参加者が比較される異なる介入を受け、経過観察されるコホート型研究を対象としている。
- e) アウトカムに影響を及ぼす予知変数が受けた介入とも関連している場合をいう。AHRQ の分類では、非ランダム化研究=研究者が割り付けを行う non-RCT と比較対照がある前向きおよび後ろ向きコホート研究に適用されるが、RCT には適用されない。AHRQ の分類では、これらのバイアスの適用は、RCT、非ランダム化研究、症例対照研究の 3 つに分類されている。概念的には選択バイアスに属する。AHRQ はランダム化の過程あるいは交絡で生じるバイアスと 1 つにまとめて取り扱っている。

* 17 ガイダンス（ROBINS-I の解説）[https://sites.google.com/site/riskofbiastool/welcome/home/current-version-of-robbins-i/robbins-i-detailed-guidance-2016](https://sites.google.com/site/riskofbiastool/welcome/home/current-version-of-robins-i/robbins-i-detailed-guidance-2016)

5) 定量的バイアス評価

Cochrane バイアスリスクツール (Higgins JP 2011a) では研究結果あるいは結論に顕著な影響を与えるバイアスを実質的なバイアスと定義し、実質的バイアスの可能性が高い場合にバイアスリスクは「高」、可能性がないか低い場合に「低」、可能性が不明か可能性はあってもその効果あるいは結果への影響が不明の場合は「不明」の判定をする。ただし、どの程度の大きさのバイアスを実質的バイアスとするかはレビューが決める必要がある。また、バイアスの方向と大きさの判定が困難な場合は、不明とする。研究ごとのバイアスリスクのまとめは推奨レベルになっており、今までの Cochrane review では必須事項になつてない。現在 Cochrane バイアスリスクツールは v2.0 が発表されており、それぞれのバイアスのドメインについてシグナリングクエスチョンに No, Probably No, No information, Probably Yes, Yes のいずれかで答えながら、最終的に高 (High), いくらかの懸念 (Some concerns), 低 (Low) の 3 段階で評価する方法へ改訂された (Higgins JPT 2019)。

Cochrane バイアスリスクツールでは 1 つの研究でバイアスリスクの 1 つのドメインあるいは項目が「高」の判定をする場合、その研究のバイアスリスクのまとめは「高」と判定することを前提とする。複数の研究のまとめは各研究のサンプルサイズとイベント率で重み付けをして、エビデンス総体として評価する。

Minds の判定方法は以下のとおりである。

1. それぞれの論文についての評価

バイアスは、上記の各要素について、評価者の判断によって、高リスク (-2)、中／疑い (-1)、低リスク (0) の 3 段階で評価する。その判断は評価者の知識、経験、専門領域などの影響を受けるため、評価者によって異なる判定がなされることがあり得る。できるだけ、2 名の評価者により、判定が異なる場合には、意見を調整し統一する。

バイアスのリスクがあるかないか、リスクが高いか低いか、リスクが高い場合どの程度高いかという評価が行われる。

2. バイアスリスク「まとめ」の判定 (STEP 1 表 4-7)

ほとんどが -2 …… 「まとめ」 ⇒ とても深刻なリスク (-2)

3 種が混じる …… 「まとめ」 ⇒ 深刻なリスク (-1)

ほとんどが 0 …… 「まとめ」 ⇒ リスクなし (0)

(2) 非直接性：個々の研究についての評価

非直接性にはさまざまな名称が用いられてきたが^{*} (Turner RM 2009, Atkins D 2011)^{*18}、ある研究から得られた結果が、現在考えている CQ や臨床状況・集団・条件へ適用し得る程度を示す。CQ の PICO の各項目と研究の PICO とのずれが結果にどのように影響するかを推定し、CQ で問われている効果推定値を明らかにする。

* 18 非直接性のうち、対象者の相違について適用可能性として分離する考え方もある。AHRQ はそのような考え方を採用しており、非直接性の概念はアウトカムが代理アウトカムであるような場合と、介入の比較が直接的な比較ではなく間接的に行われている場合に限定して用い、適用可能性については臨床において医療者が判断することを前提としている。

非直接性の評価は疾患専門家でないと困難である。日本人を対象とした診療ガイドラインで欧米人を対象にした研究を利用する場合、遺伝的背景、医療環境、生活習慣などの違いがアウトカムへ与える影響を推定する必要がある。個々の研究についての評価を行った上で、エビデンス総体の評価を行う。

以下の項目について検討する。

・研究対象集団の違い（適用可能性：applicability）

例）年齢が異なる。

CQ では 85 歳未満を対象にしているが、研究論文には年齢制限を設けていない。

・介入の違い（適用可能性）

例）薬剤の投与量、投与方法が異なる。

欧米の報告では 100mg の NSAIDs 坐薬を推奨しているが、日本には 50mg までの剤形しかない。

・比較の違い

例）設定した CQ と検索収集した研究報告で、対照か別の介入があるかどうかが異なる。

介入の直接比較でなく、共通の対照を介した間接比較である。

・アウトカム測定の違い（surrogate outcomes）

例）設定した CQ と検索収集した研究報告で、アウトカム測定が異なる。

〈判定方法〉

非直接性が高（-2）

非直接性が中／疑い（-1）

非直接性が低（0）

なお、もし個々の研究については非直接性の評価を実施せず、エビデンス総体の評価の際に実施するのであれば、**テンプレート【SR-3】**を作成する際に、PICO の各項目に非直接性の判断に必要な情報を記載しておく必要がある。その情報は研究対象者の適格基準・除外基準ではなく、実際に解析対象となった症例の属性を含む必要がある。介入についても、日本で実行可能な介入との差を判断できる情報が必要である。

（3）値の抽出

アウトカムが名義変数で二値変数の場合、例えば治癒と非治癒、発症と非発症、改善と非改善などのような場合は、それぞれの群の症例数（分母＝リスク人数）とアウトカム生起例数（分子＝イベント数）を抽出する。これらの値からは、リスク比、リスク差、オッズ比などの効果指標の統合値がメタアナリシスによって算出できる。

アウトカムが、連続変数、例えば入院日数、臨床検査の測定値などのような場合は、それぞれの群の平均値と標準偏差の値を抽出する。これらの値からは、平均値差または標準化平均値差が統合値として算出される。

生存分析の結果の場合は、各群の症例数とハザード比とその標準誤差を抽出する。ハザード比とその標準誤差からメタアナリシスが行われるが、これらの値が直接記載されていない場合には計算が必要となる（Tierney JF 2007）。ハザード比と95%信頼区間、ログランク検定の結果、ハザード比と総イベント数などさまざまな指標から計算することが可能である。

● 手順

- (1) 論文の選出・選択
- (2) 個々の研究に対する評価
 - (2)-1 バイアスリスクの評価
 - (2)-2 非直接性の評価
 - (2)-3 値の抽出

● テンプレート

- 【SR-5 評価シート 介入研究】(RCT評価用)
- 【SR-6 評価シート 観察研究】(観察研究評価用)

● テンプレート 記入方法

- (なし)

● テンプレート 記入例

- 【SR-5 評価シート 介入研究 記入例】(RCT評価用)
- 【SR-6 評価シート 観察研究 記入例】(観察研究評価用)

4.4 ベースライン調査 STEP 4：エビデンス総体の評価 STEP 2

エビデンス総体とは、ある臨床上の問題（CQ）に対して収集し得た全ての研究報告を、アウトカムごと、研究デザインごとに評価し、その結果をまとめたものである。このアウトカムは、スコープ作成時に1つのCQに対して「患者にとって重要、重大な各アウトカム」として列挙されたものである（Guyatt GH 2011b）。

エビデンス総体をアウトカム横断的に統合した全体を「エビデンス総体の総括」と呼ぶ。診療ガイドライン作成過程においてCQに対する推奨を提示する場合（第6章参照）に、エビデンス総体の総括としての強さを1つだけ決定する（Guyatt G 2011a）。

4.4.1 エビデンス総体評価

1つのアウトカムで選択抽出された複数の論文をまとめて、エビデンス総体を評価する。RCT、観察研究など研究デザインごとにそれぞれの文献集合をまとめ直し、あらためてバイアスリスク、非直接性を評価し、これに加え、非一貫性、不精確性、出版（報告）バイアスなどを評価する。

システムティックレビューを疾患専門家以外が行った場合、非直接性の評価には限界がある。また、推奨の方向や強さが想定できない状態で、エビデンスの確実性（強さ）を決定するのも困難であろう。そのような場合は、システムティックレビューでのエビデンス総体の評価と推奨作成時の最終的なエビデンス総体の評価は異なる可能性があることを前提に、前者をエビデンスの確実性（質）、後者をエビデンスの確実性（強さ）として記録することが望ましい。

- 現在目的としている1つのアウトカムが、評価項目（エンドポイント）として記載されている論文を全て抽出する。さらに、これらの論文を研究デザインごとにまとめると、以降、まとめられた論文を対象として作業を行う。
- 4.3節 STEP 1 で行った「バイアスリスク評価」結果を、あるアウトカムに関連して抽出し、まとめられた「論文の全て」を対象として統合し、「バイアスリスク評価」のまとめを記載する。非直接性については、個別の論文に対して行われ、かつ、1つのアウトカムに関してまとめられた複数の論文間の評価にも用いられる。
- 同様の作業を、非一貫性、不精確性、出版（報告）バイアスに対しても行い、バイアスリスク、非直接性と合わせて計5項目を評価する。
- この5項目を用いて、1つのアウトカムに対する「エビデンスの総体」を評価し、記載する。

- アウトカムごとにまとめられたエビデンス総体の確実性（質）はシステムティックレビュ一段階で、**表4-5**で示す分類で評価される。推奨作成の段階では、エビデンスの確実性（強さ）として評価する。本マニュアルが推奨するエビデンス総体のエビデンスの確実性（強さ）は**表4-10**のとおりであり、GRADEアプローチと同じ分類である。

表4-10 エビデンス総体のエビデンスの確実性（強さ）

A（強）：	効果の推定値が推奨を支持する適切さに強く確信がある
B（中）：	効果の推定値が推奨を支持する適切さに中程度の確信がある
C（弱）：	効果の推定値が推奨を支持する適切さに対する確信は限定的である
D（とても弱い）：	効果の推定値が推奨を支持する適切さにほとんど確信できない

- なお、この後、推奨を決定する作業として、各アウトカムに対するエビデンス総体の評価結果を見渡して、アウトカム全般に関する全体的なエビデンスの強さ（エビデンス総体の総括）を決定することになる（**第6章**参照）。この過程はいわば **STEP 3** である。この作業はシステムティックレビューには含まれない。
- 決定されたエビデンスの強さは、CQとその推奨に対するエビデンスの強さを意味し、「検索し統合されたエビデンス総体が、推奨を支えるものとしてどれほど適切か」を示している。決定にあたっては、益と害も考慮し、推奨文作成に最も重要な要素となったエビデンス総体のレベルを選択する。

4.4.2 エビデンスの確実性の評価を下げる5項目

エビデンス総体について次の5つの項目で評価を行う。

(1) バイアスリスク

4.3節 STEP 1 の結果をまとめて、バイアスリスク評価として記載する。

(2) 非直接性

4.3節 STEP 1 で評価された個々の研究に対する非直接性をまとめた上で、プールされた論文全体が現在検討中のCQや臨床状況・集団・条件へ適用し得る程度を再度検討する。

- 研究対象集団の違い

例) 年齢が異なる。

CQでは85歳未満を対象にしているが、研究論文には年齢制限を設けていない。

- 介入の違い

- 比較の違い

例) 設定したCQと検索収集した研究報告で、対照か別の介入があるかどうかが、異なる。

- ・アウトカム測定の違い

例) 設定したCQと検索収集した研究報告で、アウトカム測定が異なる。

〈判定方法〉

とても深刻な非直接性あり（-2）

深刻な非直接性あり（-1）

非直接性なし（0）

(3) 非一貫性

アウトカムに関連して抽出された全て（複数）の研究を見ると、報告によって治療効果の推定値が大きく異なる（すなわち、結果に異質性または、ばらつきが存在する）ことを指し、根本的な治療効果に真の差異が存在することを意味する。

〈判定方法〉

とても深刻な非一貫性あり（-2）

深刻な非一貫性あり（-1）

非一貫性なし（0）

研究間の結果のばらつきは、サンプリングエラーに伴う偶然の偏り、バイアスによる偏り、非直接性による偏り、説明できない偏りによって生じる。これらの偏りの程度が研究ごとに異なるため、結果にばらつきが生じる。ここでいう非一貫性は説明できない結果のばらつきのことである（Guyatt GH 2011f）。

非一貫性の判定は、メタアナリシスの結果として得られるフォレストプロットを見て、リスク比、ハザード比などの相対効果指標の点推定値の分布のばらつきと、95%信頼区間の重なりの程度を評価して行う（ばらつきが大きく出やすいので、オッズ比、リスク差は用いるべきでない）。

統計学的な異質性 (I^2 値, τ^2 値, Q 統計値の P 値, H^2 値を指標とする) は臨床的な異質性と研究方法の異質性が原因で生じるが、かならずそれらを検出できるとは限らない。あくまで非一貫性の根拠の1つとして使用すべきである。Cochrane ハンドブック^{*19}あるいはGRADE ハンドブック (Schünemann H 2013) では、 I^2 値について以下のような基準が提示されているが、重なりが多いので I^2 値だけで非一貫性の判定をするのは困難であり、個別の事例に応じて考える。

< 40% おそらく小さい

30 ~ 60% おそらく中等度

50 ~ 90% おそらくかなり大きい

75 ~ 100% おそらく大きい

* 19 Cochrane Handbook for Systematic Reviews of Interventions (<https://training.cochrane.org/handbook/current>)

臨床的な異質性と研究方法の異質性の判断を行った上で、統計学的な異質性の指標を参考にし、フォレストプロットで95%信頼区間が全ての研究で重なっているかを見て、全体として判断することが望ましい。

(4) 不精確性

サンプルサイズやイベント数が少なく、そのために効果推定値の信頼区間が幅広いことを指す。プロトコルに示された予定症例数が達成されていることが必要である。

〈判定方法〉

とても深刻な不精確性あり（-2）

深刻な不精確性あり（-1）

不精確性なし（0）

不精確性の判断方法については議論の多いところである（Schünemann HJ 2016）。GRADEアプローチでは、システムティックレビューの段階と推奨作成の段階で異なる不精確性の定義を用いている（Schünemann H 2013, Guyatt GH 2011e）。

前者の場合、メタアナリシスの結果を1つの臨床試験の結果と見なした場合に必要なサンプルサイズである最適情報量（optimal information size : OIS）を満たすかどうか、次に無効果の効果推定値（例えばリスク比なら1.0、リスク差なら0）が95%信頼区間に含まれるかを判定する。OISが満たされ、95%信頼区間に含まれない場合はレートダウンしない。もしOISの基準が満たされ、95%信頼区間が無効果を含み（例えば、リスク比の信頼区間が1.0を含む）、重要な益あるいは重要な害を除外することができない場合には、不精確性によりレートダウンする。この場合でも、効果が大きく、重要な益あるいは害があると考えられる場合にはレートダウンしない。

推奨作成の段階では、信頼区間の上・下限値が決断閾値の片側にあるかどうかを考慮する。治療推奨と非推奨の臨床決断閾値を信頼区間がまたぐかどうかを考慮し、もしそうなら（すなわち、信頼区間が閾値と交差する）、点推定値と信頼区間がどこにあっても不精確性によりレートダウンする。もし、信頼区間が交差しなければ、OISを満たすかどうかを見る。満たされない場合は、不精確性によりレートダウンする。

ただし、例外としてイベント率が低く、サンプルサイズが大きい場合は、OISの基準は適用せず信頼区間の基準だけで判定する。

OISの計算にはRのTrialSizeパッケージ（Zhang E 2017）、ウェブツールとして無料で提供されているCCRB（Centre for Clinical Research and Biostatistics）（CCRB 2017）、CRAB（Cancer Research And Biostatistics）：Statistical Tools（CRAB web）を利用することができる。

最近、Hultcrantz Mら（2017）は益のアウトカムに対するエビデンス総体の確実性の決定には、害のアウトカムに対するエビデンス総体の確実性が影響を与えるので、複数のアウトカム全体の文脈を織り込んだ確実性の決定が、推奨作成のためには必要だと述べている。そのためには、1つのアウトカムに対するエビデンス総体の確実性の評価のために臨床閾値を設定する際

に、他のアウトカムに対するその介入の効果の大きさと不確実性を考慮しながら決めた上で、不精確性の評価を決め、そのエビデンス総体の確実性を決める必要がある。臨床閾値の設定は価値観に影響されるので、エビデンスの確実性の概念に含めるべきかという議論も今後必要である。

(5) 出版（報告）バイアス：複数の研究の総合評価

研究が選択的に出版されることによって、根底にある益と害の効果が系統的に過小評価または過大評価されることをいう（Guyatt GH 2011d, Hedin RJ 2016, Driessen E 2015, Roest AM 2015, Niemeyer H 2012, Poorolajal J 2010, McGuire W 2010, Lee A 2006, McLeod BD 2004, Song F 2000, Song F 2010）。臨床試験計画時のサンプルサイズ設定で過小に設定するため（Clark T 2013），研究結果が有意にならず出版されにくくなる可能性も報告されている（Aberegg SK 2010）。サブグループ解析として出版された研究は過大評価が多いことも報告されている（Hahn S 2000）。

なお、個別の論文評価としての、選択的アウトカム報告バイアスは、前述のバイアスリスクとして評価する。

〈判定方法〉

とても深刻な出版（報告）バイアスあり（-2）

深刻な出版（報告）バイアスあり（-1）

出版（報告）バイアスなし（0）

出版（報告）バイアスの評価は、メタアナリシスの際に作成するファンネルプロット（funnel plot）の目視、および Egger 検定、Begg 検定の結果で $P < 0.1$ を基準として総合的に判断する。ファンネルプロットではサンプルサイズの小さい、精度の低い研究の効果指標が左右均等に分布していない場合、出版（報告）バイアスを疑う。

出版（報告）バイアスへの対処法として trim and fill method（van Aert RC 2016, Watzlawick R 2014, Schwarzer G 2010, Sterne JA 2001）などが用いられているが、感度分析の一環として施行することが勧められている（Peters JL 2007）。

4.4.3 観察研究に対するエビデンスの確実性の評価

観察研究に対するエビデンス評価の考え方も、RCT と同様である。ただし、エビデンスとして対象となる論文は、前向きコホート研究や各国の国民規模の症例対照研究など、背景が十分調整された研究や大規模な研究である。

観察研究の場合は、エビデンス総体のエビデンスの強さは「弱」から評価を開始する。しかし、まれながら研究結果によっては、観察研究のエビデンスの強さの評価を「中」あるいは「強」に上げることもある。いろいろな理由により RCT を行うことが難しいが、その介入の効果が極めて大きく、明らかに有効で、日常診療で通常行われている場合などが該当する。基本原則は、エビデンスがある

推奨を支持するのにどれだけ確かに、確信できるかによって判定されるということであり、これはRCTと同様である。

観察研究のエビデンスの強さを評価する場合には、**4.4.2 項**で挙げた項目で強さを下げるることを考慮するほかに、以下の項目で強さを「中」に上げることも検討する（Guyatt GH 2011h）。

(1) 介入による大きな効果

飛行機から降下する場合、パラシュートを使うとほとんど救命され、使わないと死亡することが多いが、このように介入の効果が極めて大きい場合にはエビデンスの強さを上げることを考慮してもよい。介入（治療）を行うとほとんど救命され、行わないとほとんど死亡するような場合が相当する。

〈判定方法〉

効果が非常に大きい（very large） RR > 5 または < 0.2 (+2)

効果が大きい（large） RR > 2 または < 0.5 (+1)

上記以外 (0)

(2) 用量 - 反応勾配

報告では介入の量（回数、投与法）が少ないが、より多くの量の介入を行えば有意差が出ていたと考えられる場合には、エビデンスの強さを上げることを考慮してもよい。

〈判定方法〉

非常に大きな用量 - 反応勾配が考えられる (+2)

大きな用量 - 反応勾配が考えられる (+1)

上記以外 (0)

(3) 可能性のある交絡因子による効果の減弱

報告の中に調整できていない交絡因子があり、その交絡因子が報告されている効果を減弱させていると想定される場合には、エビデンスの強さを上げることを考慮してもよい。例えば、介入を行った群には高齢者、糖尿病の患者が多くいたため、効果としての死亡率がわずかしか改善しなかったが、もし、背景が均一化されていれば大きな有意差が出ていただろうと想定される場合などが相当する。

〈判定方法〉

可能性のある交絡因子が、効果を大きく減弱させていると考えられる (+2)

可能性のある交絡因子が、効果を減弱させていると考えられる (+1)

上記以外 (0)

なお、観察研究の評価を省略する場合はその理由を記載し、エビデンスの強さは「弱」のままとする。

4.4.4 エビデンスの統合

(1) 定性的システムティックレビュー

エビデンス総体を質的に統合することを定性的システムティックレビューという。定性的システムティックレビューは各研究のバイアスリスクの評価と非直接性の評価、それらを反映したエビデンス総体のバイアスリスクと非直接性の評価、エビデンス総体を構成する研究間の非一貫性、不精確性、報告（出版）バイアスなどの評価と臨床的文脈の評価も含まれる。**4.4.1～4.4.3項**までをまとめた作業のことである。

なお、定量的システムティックレビューを行った場合でも、効果指標の統合値と信頼区間だけでエビデンスの強さを決定せず、定性的システムティックレビューの結果もエビデンスの強さの評価に反映させるため、診療ガイドライン作成において必須の作業となる。

(2) 定量的システムティックレビュー（メタアナリシス）

研究結果をまとめる際に統計学的に効果指標の値を統合し、その信頼区間とともに提示するのが定量的システムティックレビュー、すなわちメタアナリシスである。

診療ガイドライン作成のためのシステムティックレビューで、研究デザインが同じで、PICOの各項目の類似性が高い場合には、効果指標を量的に統合するメタアナリシスが可能となる。メタアナリシスの結果、効果指標の統合値と信頼区間が得られるとともに、フォレストプロット、ファンネルプロットが得られる。これらのプロットは非一貫性、不精確性、出版（報告）バイアスの判定にも有用となる。診療ガイドライン作成におけるシステムティックレビューは定量的システムティックレビュー、すなわちメタアナリシスだけではなく、定性的システムティックレビューを同時に行う必要がある。

メタアナリシスが実行できる場合には、その結果はエビデンス総体の強さを検討する1つの項目となる。例えば、ある介入が統計学的有意に、そして顕著によりアウトカムを導くという結果が出ている場合には、エビデンスの強さを上げることを考慮してもよい。また、小さな差しかないという結果が出ている場合には、エビデンスの強さを上げる結果ではないと判断してもよい。そして、有意な差がない場合には、エビデンスの強さを下げることを考慮してもよい。

1) メタアナリシスが省略できる場合

メタアナリシスは必須ではない。次の場合にはメタアナリシスは省略できる可能性がある。

① 定性的にエビデンスの強さが保証できる場合

すなわち、エビデンスの強さが定性的評価から論理的に説明ができ、効果の確実性が保証されていると評価可能である場合は、メタアナリシスを省略してもよい。例えば、対象とした論文が全て同じ結論である場合などがこれに相当する。

② 同じ研究デザインの研究報告が1つしかない場合

あるアウトカム・介入群に関する報告のうち、同じ研究デザインについての報告が1つしかない場合には、その報告自体の値しか統合に利用することができないため、メタアナリシスの必要がない。

③ 診療ガイドライン作成グループによって定められたCQおよび方法と同様のメタアナリシスがある場合

例えばCochrane reviewや先行する診療ガイドラインのシステムティックレビューが診療ガイドライン作成グループの定めた方法とほぼ同様であり、その結果を利用できると判断する場合には、そのメタアナリシスの結果をエビデンスの評価に用いることが可能である。ただし、これらのレビュー報告後に新たな研究報告がなされている可能性があり、さらなる注意深い検索が必要である。

2) メタアナリシスのためのソフトウェア

メタアナリシスのためのソフトウェアは無料のものも含め多数存在するので、それぞれが使いやすいと思われるものを必要な機能に応じて使用すればよい。

Cochraneはシステムティックレビューを行うためのRevMan（現バージョンは5.4）と呼ばれるソフトウェアを無料で提供しており、その中にメタアナリシスのプログラムが含まれている。ウェブサイトからダウンロードして自由に使用することができる。RevManでは各研究の名称や介入などを順次入力し、テーブルを作成してからデータを入力し、メタアナリシスを実行する。同じデータからリスク比、オッズ比、率差などを指標としたメタアナリシスを行うことができ、フォレストプロットをWordなどに貼り付け可能な形で出力できる（[4.6節](#)参照）。

また、統計解析のオープンプラットホームであるRでもメタアナリシス用のさまざまなパッケージが公開されており、これらを利用することも可能である。リスク比、リスク差、オッズ比、ハザード比、平均値差などの効果指標を用いるのであれば、metaforパッケージを用いることができる。Excelでデータを用意し、簡単な操作でメタアナリシスを実行し、フォレストプロット、ファンネルプロットなどの結果を得ることができる（[4.7節](#)参照）。

3) 効果指標

メタアナリシスでは複数の研究結果を統合するが、統合されるのは効果指標の値である。効果指標にはさまざまなものがあるが、リスク比[risk ratio: RR、または相対危険度(relative risk: RR)]、オッズ比(odds ratio: OR)、率差[rate difference: RD、またはリスク差(risk difference: RD)]、平均値差(mean difference: MD)、標準化平均値差(standardized mean difference: SMD)、ハザード比(hazard ratio: HR)、その他が用いられている。

効果指標の値は、効果の強さあるいは大きさを定量的に表すものであることから、効果サ

イズ (effect size) とも呼ばれる。また、効果指標は一定の分布に従うが正規分布に従う場合には、ばらつきの指標として標準偏差 (standard deviation) を用いることができる。効果指標の分布の標準偏差は、標準偏差とも標準誤差 (standard error) とも呼ばれるが、同じ意味で用いられている。

以下、表 4-11 の四分表を基として、これら効果指標の算出について解説する。

表 4-11 効果指標算出の基になる四分表

	アウトカム (+)	アウトカム (-)	合計
治療:介入 (+)	r_i^T*	$n_i^T - r_i^T$	n_i^T
対照:介入 (-)	r_i^C	$n_i^C - r_i^C$	n_i^C

*その研究の治療群におけるイベント数。 r : アウトカム陽性（イベント生起）例の人数、 n : 各群の総症例数、 i : 研究番号。
2群の率（割合）から RR, OR などが算出される。

① リスク比

リスク比は $(r_i^T/n_i^T)/(r_i^C/n_i^C)$ で求められる。リスク比の自然対数は正規分布に従い、その分布の標準偏差（標準誤差）は $\sqrt{1/r_i^T + 1/r_i^C - 1/n_i^T - 1/n_i^C}$ で計算される。リスク比はそれぞれの群の率の比に相当するが、対数変換するとそれぞれの群の率の対数の差になるため、モデル化が容易である。対数化したリスク比、標準誤差を用いて、統合値および 95% 信頼限界を算出し、指数変換 (exponential transformation) して元のスケールに戻すことが行われる。その際には、標準誤差の平方 = 分散の逆数で重み付けした平均値を求める（固定効果モデル）。分散に研究間の分散を加算して重み付けするとランダム効果モデルとなる。なお、割り算の分母が 0 になる場合には、 r および $n - r$ に 0.5 を加算する。

② オッズ比

オッズ比は $[r_i^T(n_i^C - r_i^C)]/[r_i^C(n_i^T - r_i^T)]$ で求められる。オッズ比の自然対数は正規分布に従い、その分布の標準偏差（標準誤差）は $\sqrt{1/r_i^C + 1/(n_i^C - r_i^C) + 1/r_i^T + 1/(n_i^T - r_i^T)}$ で計算される。なお、割り算の分母が 0 になる場合には、 r および $n - r$ に 0.5 を加算する。

③ 率差（リスク差）

率差は $r_i^T/n_i^T - r_i^C/n_i^C$ で求められる。2群の率の差は正規分布に従い、その分布の標準偏差（標準誤差）は $\sqrt{[r_i^T(n_i^T - r_i^T)/(n_i^T)^2] + [r_i^C(n_i^C - r_i^C)/(n_i^C)^2]}$ で計算される。

④ 標準化平均値差

標準化平均値差として、Hedge's unbiased estimator が推奨されているが、Cohen's d, Hedge's g, Glass's Δ なども用いられている。連続変数の場合の効果指標の例を以下に示す。

・ Cohen's $d = (M_1 - M_2)/SD_{\text{pooled}}$

$$SD_{\text{pooled}} = \sqrt{\frac{\sum (X_A - \bar{X}_A)^2 + \sum (X_B - \bar{X}_B)^2}{n_A + n_B - 2}}$$

・ Glass's $\Delta = (M_1 - M_2)/SD_{\text{control}}$

・ Hedge's $g = (M_1 - M_2)/SD_{\text{pooled}}^*$

$$SD_{\text{pooled}}^* = \sqrt{\frac{(n_A - 1)SD_A^2 + (n_B - 1)SD_B^2}{n_A + n_B - 2}}$$

連続変数に対する d ファミリーの効果量。pooled：プール，control：対照。

⑤ ハザード比

ハザード比はリスク比と類似した概念であるが、時間イベントアウトカムの場合、すなわち生存分析の場合に適用される。Cox の比例ハザード解析、カプラン・マイヤー生存解析、ログランク検定などの結果からハザード比と信頼区間を算出することが可能である。

⑥ その他

リスク比の信頼区間が 1.0 を含んでいる場合には、益に対する治療必要人数 (number needed to treat for benefit : NNTB) と害に対する治療必要人数 (number needed to treat for harm : NNTH) の値が∞を挟んだ形になる。生存分析に基づく必要人数 (number needed to treat : NNT) の計算も可能である。

アウトカムが害の場合には害のための必要人数 (number needed to harm : NNH) が、介入がスクリーニング検査の場合にはスクリーニングのための必要人数 (number needed to screen : NNS) が用いられる。

リスク比、オッズ比はログ変換（自然対数）することによって、正規分布に従うので、ログ変換後に正規分布を前提とした統合を行うことが多い。

4) 統合のモデル：固定効果モデル／ランダム効果モデル

固定効果モデルでは統合の対象となった研究以外の研究は想定せずに、それらの研究の効果指標の平均値を求める。ランダム効果モデルでは、実際に統合の対象となった研究以外の研究が母集団として想定され、母集団からランダムに抜き出されたものが統合の対象となつた研究であると考える。

したがって、研究間の異質性があることを理由にランダム効果モデルを用いるという考えは誤りであり、一方で、固定効果モデルによる効果指標の統合値は対象となった研究だけをまとめたものとしては全く正しいといえる。しかしながら、臨床研究は多くの異質性を生む要素があるので、ランダム効果モデルによる統合値を算出することが推奨されている (Morton SC 2018)。

固定効果モデルの計算法の一例を示す。分散に基づく方法である。分散の逆数 (inverse variance) を重み付けに用いて効果指標の平均値を算出する。固定効果モデルでは、各研究の分散だけが重み付けに用いられる。

$$M = \frac{\sum W_i M_i}{\sum W_i}, \quad W_i = \frac{1}{V_i} = \frac{N_i}{s_i^2}$$

$$V_M = \frac{1}{\sum W_i}, \quad M \pm 1.96 \times \sqrt{V_M}, \quad Z = \frac{|\ln(M)|}{\sqrt{V_M}}$$

Inverse variance-based method とも呼ばれる。M: 統合値, W: 重みであり分散の逆数, V: 分散, Z: 有意差検定に用いられる値 (1.96 以上で有意), i: 研究番号, s: 標準偏差 (標準誤差)。

研究間の異質性の検出には Q 統計値や I^2 統計値が用いられる。検出力が低いため, $P < 0.1$ で有意とする場合も多い。これら 2 つの統計値の算出法に関しても一般化 Q 統計値を標準的に用いるべきであるという考え方もある。

ランダム効果モデルと Q 統計値, I^2 統計値について、以下に示す。

$$Q = \sum W_i M_i^2 - \frac{(\sum W_i M_i)^2}{\sum W_i}$$

$$Q > k - 1 = df \rightarrow \text{heterogeneity}(+)$$

$$I^2 = \frac{Q - df}{Q} \times 100\%, \quad T^2 = \frac{Q - df}{\sum W_i - \frac{\sum W_i^2}{\sum W_i}}$$

$$W_i^* = \frac{1}{V_i^*} = \frac{1}{V_i + T^2}$$

$T^2: \tau^2$ の推定値 (データに基づき算出された効果指標の研究間の分散), i: 研究番号, M: 効果指標の値, V: 分散, W: 重み。統合値は上記の式で W の代わりに W^* を用いて計算する。DerSimonian-Laird の方法に基づく。

また、統計学的異質性に基づいて固定効果モデル ($P > 0.10$), ランダム効果モデル ($P < 0.10$) とすることは推奨されない。メタアナリシスの対象とした研究以外の研究は想定せずに、対象研究の効果指標の平均値を知りたいのであれば、固定効果モデルの統合値がそれに相当する。また、ランダム効果モデルで研究間の分散である τ^2 が 0 になる場合は、固定効果モデルと同じ結果となる。

また、ランダム効果モデルの結果は効果指標の平均値 μ とその信頼区間および効果指標の分布の標準偏差すなわち上記の τ の 2 つのパラメータの解釈が必要になる。多くの場合、 μ

の値とその95%信頼区間が評価の対象とされ、95%信頼区間が1を含んでいなければ有意であり、効果があるというように解釈されている。しかしながら、このような結果は全ての研究の効果指標の平均値のとり得る値についての解釈であり、個々の研究の効果指標のばらつきは別に考える必要がある。したがって、将来同じような研究が行われた場合には、得られる効果指標の値がどのように分布するかについては、 μ と τ の両者の値に基づいたベイジアン推測が必要になる。そのためには、ランダム効果モデルベイジアンメタアナリシスが必要となる。

5) 感度分析

メタアナリシスを行う際に、ある仮定を設定する必要があったり、任意の選定あるいは決断が必要であったりする場合、その仮定や選定を変えた場合に結果がどのように変動するかを見るのが感度分析である。

6) システマティックレビューチームとメタアナリシス

システムティックレビューの実行には文献検索のスキルおよびメタアナリシスのための統計学的知識とスキルが必要とされる。また、文献の選定、感度分析、メタリグレッションの妥当性を確保するためには実診療に精通している必要がある。したがって、多くの分野の専門家から構成されるチームでシステムティックレビューを実行することが望まれる。

しかしながら、わが国の現状では、システムティックレビューチームに医学統計学の専門家をそれぞれ配置することは多くの場合困難であり、メタアナリシスの実行は、その知識・スキルを持つメタアナリシス担当者と各領域の専門家の協力の下に実行することが望まれる。

4.4.5 エビデンスの強さに関する判定および表記方法

(1) 総合的評価 (cf. 算術和)

エビデンスの強さに関する各要素・項目は、「-2」、「-1」、「0」、「+1」、「+2」といった数値を用いて評価を行っているが、「-2」は「-1」の2倍低いという意味ではなく、「-2（とても深刻な問題）」は「-1（深刻な問題）」という程度を示す指標として用いる。つまり、「-1 - 2 = -5」になるという計算式ではない。それぞのまとめの評価は、それぞれの評価項目の算術和ではなく、評価者の総合的判断によって決定する (Jüni P 1999, Linde K 1999)。例えば、観察研究のエビデンスの評価を上げる3項目とも「+1」であれば、エビデンスの強さが「弱」から「中（+1）」に上昇判定される。

(2) エビデンス総体としてのエビデンスの強さの評価

エビデンス総体としてのエビデンスの強さの評価は、研究デザインをベースにして評価を開始するが、評価を下げる項目、評価を上げる項目を考慮し、最終的な強さを決定する。

具体的には、RCTのみでまとめられたエビデンス総体の評価は「強」、観察研究（コホート、

症例対照研究）のみでまとめられたエビデンス総体の評価は「弱」、症例報告、症例集積研究のみでまとめられたエビデンス総体の評価は「とても弱」として初期評価を与える。次に、エビデンスの強さの評価を下げる5項目、上げる3項目の検討を行う。そして、メタアナリシスが実行可能である場合にはメタアナリシスを実行する。初期評価と項目評価とメタアナリシスの結果を総合して、エビデンスの強さをA「強」、B「中」、C「弱」、D「非常に弱」に分類する。

(3) エビデンスの強さの判定作業に関する注意事項

次のような場合には特別な対応を検討してもよい。

① RCT が 1 つしかない場合

原則としてエビデンス総体は「弱」と判定するが、その内容（バイアスリスク、非直接性やエビデンスの強さの評価を上げる項目）を吟味して総合判定する。

② Cochrane review で内容が評価されている論文を用いる場合

Cochrane review のエビデンス評価〔バイアスリスク、非一貫性、不精確性、出版（報告）バイアスなど〕を利用することができる（= Cochrane review の多くは、GRADE アプローチで評価されている）。

③ 海外の診療ガイドラインを利用する場合

海外の診療ガイドラインが、診療ガイドライン作成グループによって決められた方法と同様の方法によって作成されている（例えばGRADE アプローチで作成されているなどの）場合は、それぞれのエビデンス評価〔バイアスリスク、非一貫性、不精確性、出版（報告）バイアスなど〕を利用してもよい。ただし、診療ガイドラインの推奨内容に関しては、コンセンサスや海外の医療事情が盛り込まれているので、そのままエビデンスとして利用することは難しい。

④ 1 つの CQ に対して RCT と観察研究がある場合

RCT と観察研究は別シートで評価し、最初から一括して評価しないことが重要である。作業としては、RCT のエビデンス総体と観察研究のエビデンス総体をそれぞれ評価し、必要がある場合にはそれぞれの評価結果を総合して評価する。

例えば、RCT では追跡期間が短く有意差なしとされた内容が、コホート研究では長い追跡期間の結果、大きな有意差が示されることがある。このような場合は、コホート研究のバイアスが影響しているためなのか、追跡期間が十分長いために有意差が出たのかについては診療ガイドライン作成グループで検討を行う必要がある。その上で、エビデンスの強さを決定する。

⑤ 専門家の意見／コンセンサス

専門家の意見やコンセンサスはエビデンスとしては用いないが、重要なエビデンスの引用や、方向性、追試の必要な臨床研究の考え方など重要な情報源になる可能性がある。

⑥ 先行する診療ガイドライン

過去に出版された診療ガイドラインを取り上げる場合、診療ガイドライン全体（特に推奨文）としては、コンセンサスが含まれるのでエビデンスとしては用いない。一方、その診療ガイドライン作成で行われた、それぞれのCQに対する文献収集や選択、統合などによって作成されたシステムティックレビューの結果やエビデンス評価〔バイアスリスク、非一貫性、不正確性、出版（報告）バイアスなど〕については、必要に応じて利用してもよい。

⑦ 害の評価と症例集積研究／症例報告

症例集積研究や症例報告は原則としてエビデンスの強さは「非常に低い」と判定される。しかし、前述のごとく「害」の評価では重要な情報となる場合があり、注意が必要である。

4.4.6 決断分析における効果推定値の取り扱い

臨床研究では介入あるいは要因曝露の効果を明らかにするため、対照群との相対的な効果を表す効果指標が用いられる。リスク差は絶対リスク減少（absolute risk reduction）とも呼ばれ、絶対効果指標と呼ばれるが、介入群と対照群のイベント率の差なので、相対的な効果を表す指標である。なお、単一群のイベント率や平均値、例えば介入群、対照群それぞれのイベント率や平均値は絶対リスク（absolute risk）と呼ばれる。

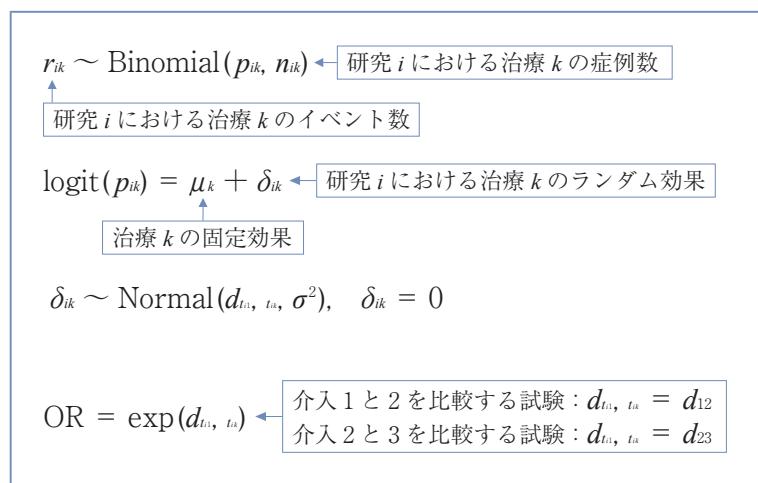
一方、多基準決断分析（multi-criteria decision analysis : MCDA）のうち、NCI/Gail法ではリスク差を用いて正味の益が計算されるが、swing-weighting その他の多くの多基準決断分析の方法で総スコア（aggregate score）を算出する場合は1つの群の絶対リスク、すなわち介入群と対照群、またネットワークメタアナリシスのような複数の介入を比較するような場合は、それぞれの介入群におけるイベント率や平均値の値が用いられる。

メタアナリシスで得られる効果推定値は通常、相対的な効果指標であり、各群のイベント率の統合値は得られない。単一群のイベント率の統合値を分散逆数法などで算出する必要がある。また、ネットワークメタアナリシスでも「アームベースモデル」(Zhang J 2014, Hong H 2015, Hong H 2016)のような手法を用いる必要がある (Hawkins N 2016)。アームベースモデルでは単一群の研究の結果を取り込む方法も開発されている (Zhang J 2019)。Rのpcnetmetaパッケージ (Lin L 2017) はベイジアンの手法でアームベースネットワークメタアナリシスを実行するプログラムである。

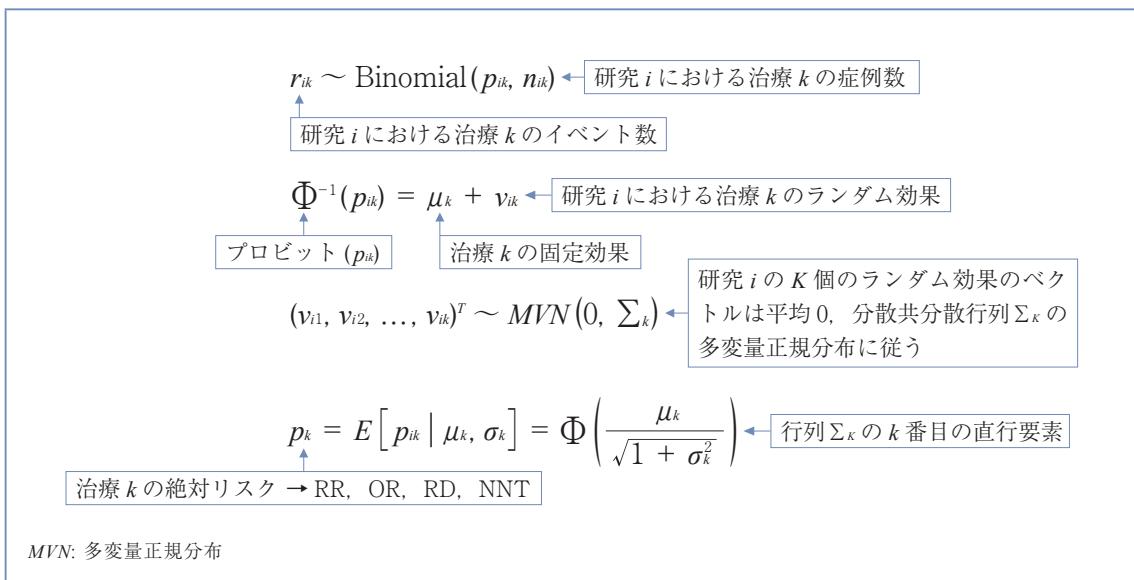
従来のメタアナリシスあるいはネットワークメタアナリシスでは「コントラストベースモデル」(contrast-based model) が用いられ、これは相対的効果指標を統合するモデルであった。一方、アームベースモデルは、各治療群のイベント率や平均値が統合され、それらから相対的効果指標を算出するモデルである。コントラストベースモデルでは、ベースラインリスクが変動しても相対的効果指標はほとんど変動しないと見なされ、ランダム割り付けが崩されることはない。一方で、ベースラインリスクの推定値はコホート研究、疾患レジストリなどから得るか、個別に設定することが望ましいとされている。通常RCTは参加者の選択の時点で、リアルワールド（real world）の患者を必ずしも代表しなくなる場合が多いこと、治療も厳密に管理されており、リアルワールドの効果を必ずしも正確に表しているとはいえないと考えられ、RCTの対照群のデータからベースラインリスクの推定値

を求めるることは必ずしも適切でないとされてきた。一方、アームベースモデルはランダム割り付けが保証されなくなり、ばらつきが大きくなると批判されてきた。しかし、最近必ずしもそうではなく、いずれのモデルもそれぞれのバイアスの影響を受ける可能性が報告されている（White IR 2019）。いずれもモデルでも、それぞれの特性を理解して用いる必要がある。

以下に、メタアナリシスにおけるコントラストベースモデルを示す。



続いて、メタアナリシスにおけるアームベースモデルを示す。



また、マルコフモデル、discrete event simulation、マイクロシミュレーションなどで1つの健康状態から次の健康状態へ推移する確率の値を設定する場合、それらの値をメタアナリシスで直接算出することは困難である。ハザード率やイベント率から指数関数モデルなどを前提に確率を算出する必要がある。

推薦作成の際に、益と害の大きさとバランス、正味の益を定量的に評価する際にどのような効果推

定値を用いるかを十分検討しておく必要がある。確率的シミュレーションを実施する場合には、平均値だけでなく標準誤差などの不確実性の指標の値も必要となる。

● 手順

- (1) エビデンスの確実性の評価を下げる5項目を評価する。
(1)-1 観察研究の場合は、エビデンスの確実性の評価を上げる3項目も評価する。
- (2) 定性的システムティックレビューをまとめる。
- (3) 定量的システムティックレビュー（メタアナリシス）を行える場合はメタアナリシスを行い、結果をまとめる。
- (4) エビデンスの強さを評価する。

● テンプレート

[SR-7 評価シート エビデンス総体]

[SR-8 評価シート エビデンス総体（絶対効果指標の結果を記入する場合）]

[SR-9 定性的システムティックレビュー]

[SR-10 メタアナリシス]

● テンプレート 記入方法

[SR-9 定性的システムティックレビュー 記入方法]

● テンプレート 記入例

[SR-7 評価シート エビデンス総体 記入例（介入研究）]

[SR-7 評価シート エビデンス総体 記入例（観察研究）]

[SR-10 メタアナリシス 記入例]

4.5 ベースライン調査の実施

4.5.1 システマティックレビューレポートのまとめ方

システムティックレビューチームは、定性的または定量的システムティックレビューの結果をエビデンス総体の強さとしてシステムティックレビューレポート（サマリーレポートともいう）にまとめ、推奨作成を行う診療ガイドライン作成グループに提出する。

CQ のうち、適切な論文が検索されなかった場合、または検索された全ての論文の質が高くなかった場合など、将来的な研究が必要と判断される場合は、Future Research Question として記載する。

診療ガイドライン作成グループは、提出されたシステムティックレビューレポートを基に推奨作成のステップに進む。したがって、添付の資料も含めて、推奨決定の作業において必要・十分な情報を提示することに留意する。

システムティックレビューレポートには GRADE アプローチで用いられている結果のまとめ表、SoF 表（Guyatt G 2011a, Guyatt GH 2012, Guyatt GH 2013b, Carrasco-Labra A 2015, Carrasco-Labra A 2016）を作成してもよい。SoF は以下の 7 つの要素を含む。

- ① 全ての重要・重大な望ましい・望ましくないアウトカム
- ② これらのアウトカムのベースラインリスク
- ③ 介入群のリスク（イベント率）あるいは絶対リスク減少（絶対効果指標）
- ④ リスク比、ハザード比などの相対効果指標
- ⑤ 参加者の人数（総数）と研究数
- ⑥ アウトカムごとのエビデンス総体の確実性
- ⑦ コメントあるいは説明

4.5.2 SoF 表の作成

テンプレート【SR-12】を参考に、上記の 7 項目を含めるように作成する。

4.5.3 協働意思決定を考慮した結果の提示法

医療を受ける者のヘルスリテラシー、数量リテラシーを考慮し、理解しやすい表現を用いることやグラフを作成することを考慮する。AHRQ の SHARE アプローチ^{*20}などを参考に資料を作成する。

* 20 <https://www.ahrq.gov/health-literacy/professional-training/shared-decision/index.html>

● テンプレート

- 【SR-11 システマティックレビューレポートのまとめ】
- 【SR-12 結果のまとめ（SoF表）（ペア比較のメタアナリシス）】
- 【SR-13 結果のまとめ（SoF表）（ネットワークメタアナリシス：様式1）】
- 【SR-14 結果のまとめ（SoF表）（ネットワークメタアナリシス：様式2）】
- 【SR-15 Future Research Question】

● テンプレート 記入方法

- 【SR-11 システマティックレビューレポートのまとめ 記入方法】
- 【SR-15 Future Research Question 記入方法】

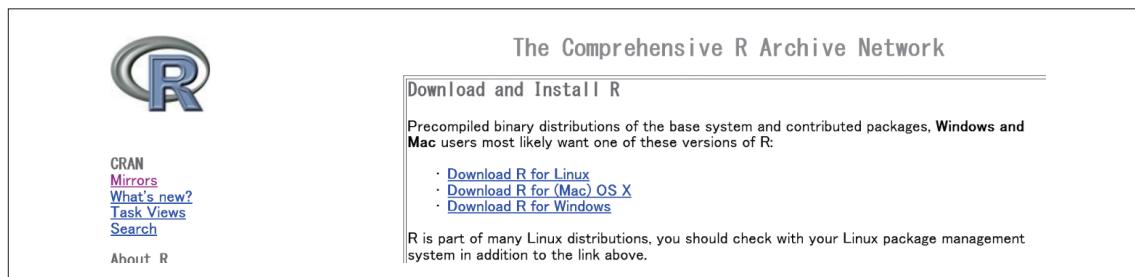
● テンプレート 記入例

- 【SR-12 結果のまとめ（SoF表）（ペア比較のメタアナリシス） 記入例】

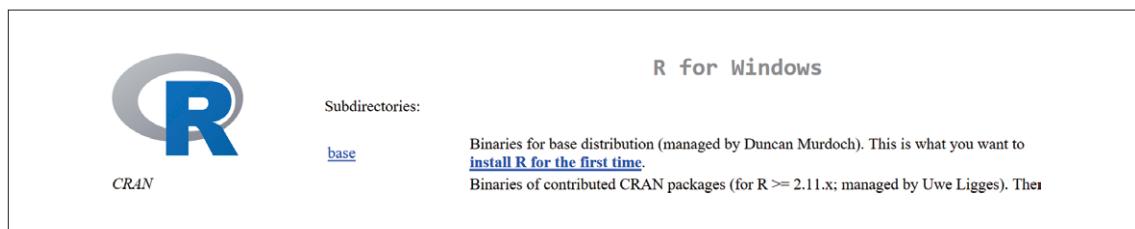
4.6 補足資料 1：R と metafor パッケージを用いたメタアナリシス

(1) R のインストール

R のプログラムは The Comprehensive R Archive Network (CRAN) のウェブサイト^{*21}からダウンロード可能である。OS のプラットホームとして Windows, Linux, Mac に対応したバージョンが用意されている。



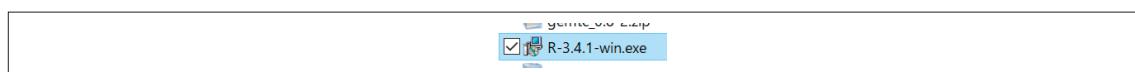
Windows の場合であれば、「Download R for Windows」をクリックし、次の画面で「install R for the first time」をクリックする。



次の画面で「Download R “バージョン No.” for Windows」をクリックして R-3.4.1-win.exe^{*22}ファイルをダウンロードし、任意のフォルダに保存する。



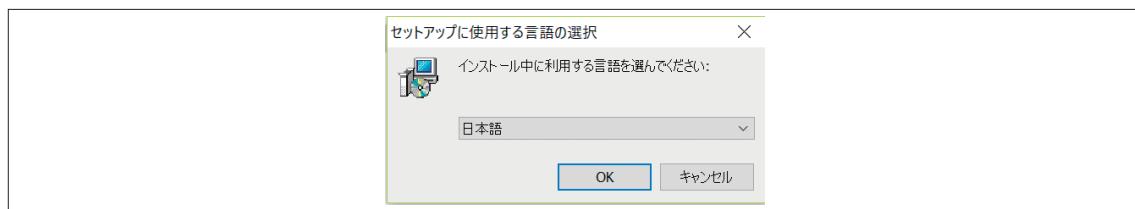
ダウンロードしたファイルをダブルクリックして、インストールを開始する。



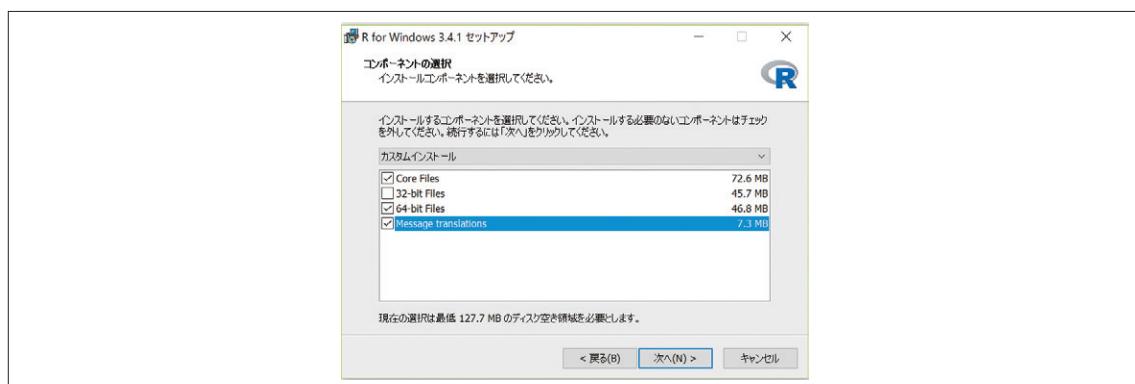
コンピューターへの変更を確認する画面が出るので「[はい]」をクリックする。言語を選択する画面となるので日本語を選択する。

* 21 CRAN (The Comprehensive R Archive Network) <https://cran.r-project.org/>

* 22 ファイル名の数字は参照時点のもの。



セットアップウィザードの指示に従ってインストールを進める。途中、「コンポーネントの選択」画面で「Message translations」にチェックを入れておくと、インストールされたRのメニューが日本語表示となる。

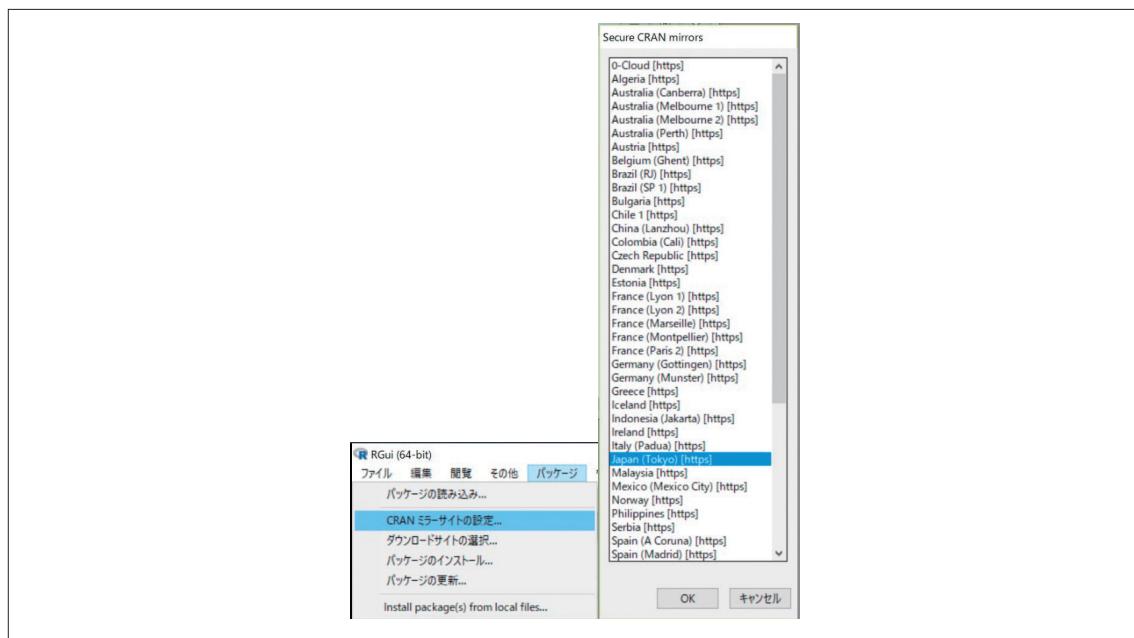


32ビット版と64ビット版が選択できるので、各自のパソコンに合わせていずれかをインストールすればよい。両方をインストールすることもできる。

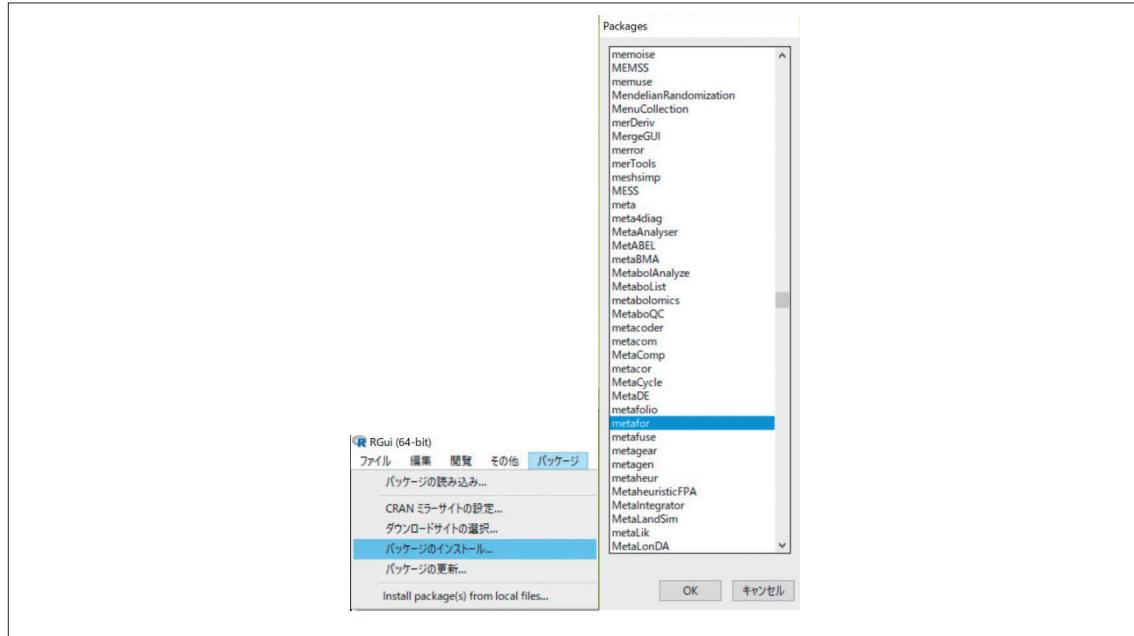
(2) パッケージのインストール

Rでは特定の解析を目的とした「パッケージ」と呼ばれるプログラムが世界中の開発者により提供されている。ここでは、メタアナリシスのために開発された metafor パッケージ (Viechtbauer W 2020) と、フォレストプロット解析のために開発された forestplot パッケージ (Gordon M 2020) を用いる。

Rを起動後、パッケージメニューから「CRAN ミラーサイトの設定」を選択し、一覧から「Japan (Tokyo) [https]」を選択して、[OK] をクリックする。



次に、「パッケージ」メニューから「パッケージのインストール」を選択し、表示されたパッケージの一覧から目的のパッケージを選択して、[OK] をクリックする。forestplot パッケージについても、同様の操作を行う。



通常、管理者権限で Windows を使用している場合は「Program Files」フォルダ>「R」フォルダ>「library」フォルダにパッケージがインストールされる。R の「パッケージ」メニューから「パッケージの読み込み」を選択するとインストール済みのパッケージを確認することができる。

(3) メタアナリシスのためのデータの準備

Excelで用意したデータが使用できるので、テンプレートなどの評価シートの一部を用いることができる。データのラベル名はそのまま、ラベルのカラムのテキストは内容に応じて変更する。効果指標の種類は略称によって判断してメタアナリシスを実行する。

アウトカムが二値変数で効果指標がリスク比（RR）の場合の例を以下に示す。同じデータで「RR」を「RD」に、「Risk Ratio」を「Risk Difference」に変更すれば、リスク差（RD）を効果指標としたメタアナリシスを行うことができる。同様に、それぞれ「OR」、「Odds Ratio」に変更すると、オッズ比（OR）を効果指標としたメタアナリシスを行うことができる。

アウトカム 治癒					
	対照群		介入群		
研究コード	総例数	アウトカム +例数	総例数	アウトカム +例数	ラベル
author	nc	cpos	nt	tpos	label
Salminen P 2015	273	272	257	186	Author Year
Vons C 2011	119	104	120	72	Surgery
Hansson J 2009	167	142	202	83	Antibiotics
Styrud J 2006	124	124	128	97	Healing
Eriksson S 1995	20	20	20	13	Risk Ratio
				RR	この行のラベル名は変更不可
					研究ID
					対照
					介入
					アウトカム
					効果指標
					略称(RR, OR, RD, HR,MD)
					表示用の効果指標名
					効果指標の種類を指定。略称のいづれか。

効果指標がハザード比（HR）の場合の例を以下に示す。

アウトカム 死亡(ハザード比)					
	介入群		対照群		
研究コード	総例数	総例数	効果指標 値	標準誤差	ラベル
author	nt	nc	yi	sei	label
Doffoel M 2008	62	61	-0.0726	0.1972	Author Year
Lo CM 2002	40	40	-0.6931	0.245	BSC
Llovet JM 2002	40	35	-0.755	0.4661	TACE
Pelletier G 1998	37	36	-0.0834	0.2659	Death
Groupe d'Etude 1995	50	46	-0.3425	0.2303	Hazard Ratio
Pelletier G 1990	21	21	0.3946	0.343	HR
					研究
					対照
					介入
					アウトカム
					効果指標
					略称(RR, OR, RD, HR,MD)

アウトカムが連続変数で効果指標が平均値差（MD）の場合の例を以下に示す。

アウトカム 入院日数					
	対照群		介入群		
研究コード	総例数	平均値	標準偏差	総例数	平均値
author	n2i	m2i	sd2i	n1i	m1i
Eriksson 1995	20	3.4	19	20	3.1
Styrud 2006	124	2.6	1.2	128	3
Turhan 2009	183	2.4	0.14	107	3.14
Vons 2011	119	3.04	1.5	120	3.96
					3 Author Year
					Appendectomy
					Antibiotics
					Days in hospital
					Mean Difference
					MD
					研究
					対照
					介入
					アウトカム
					効果指標
					略称(RR, OR, RD, HR,MD)

(4) メタアナリシスの実行

インターネットに接続された状態でRを起動し、「ファイル」メニューから「新しいスクリプト」でエディター画面を開き、以下のスクリプトを書き込む、あるいはExcelのデータ画面からコピーして貼り付ける（Excelでデータを用意した場合）*²³。

* 23 あとで利用することを考えて、「ファイル」メニューから「保存」を選択して「metaclip.R」というファイル名で保存しておく。以降使用する際は「ファイル」メニューから「スクリプトを開く」で開いて用いる

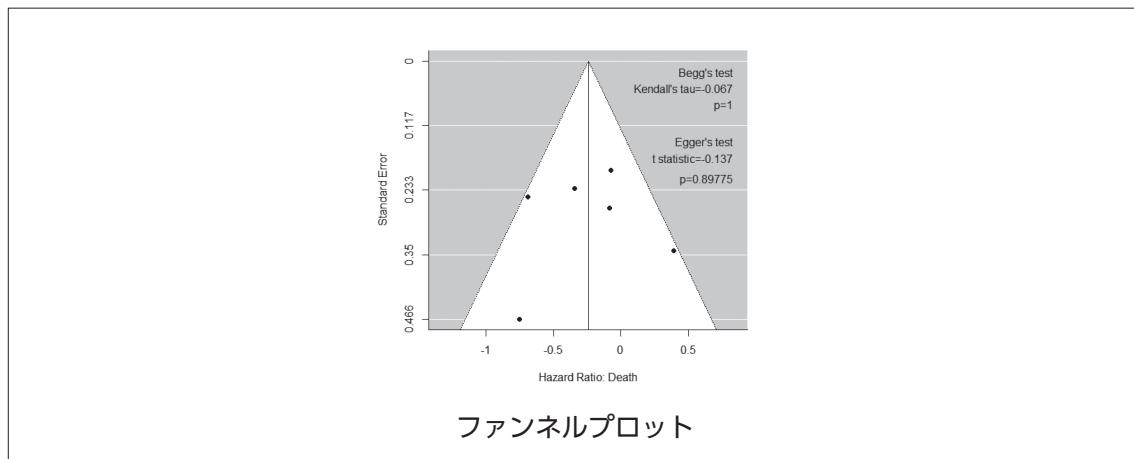
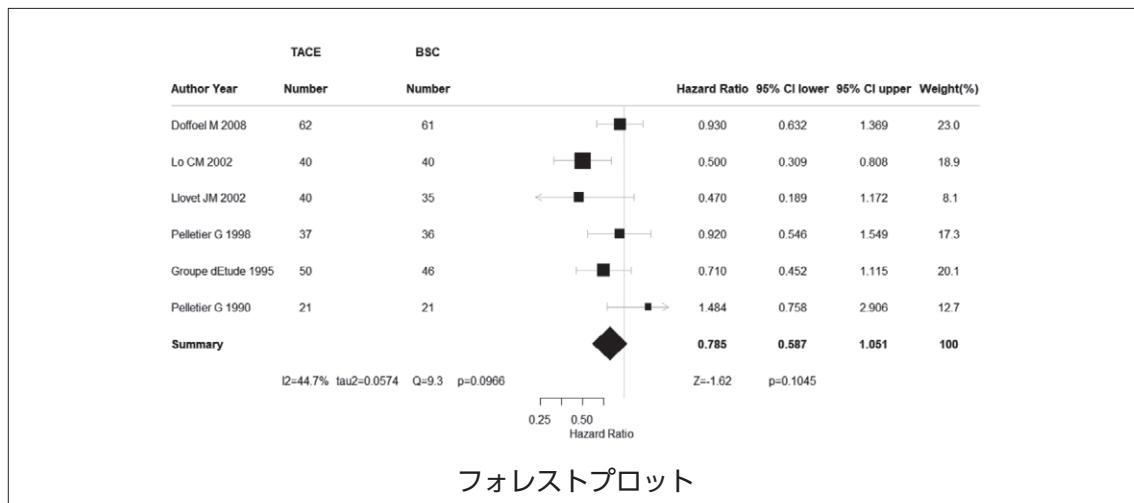
metaclip.R

```
#1. クリップボード経由でデータ読み込み：ヘッダーあり
exdat=read.delim("clipboard",sep="\t",header=TRUE) # Windows
exdat=read.delim(pipe("pbpaste"),sep="\t",header=TRUE) # Mac
#2. スクリプトを読み込んでメタアナリシスを実行。
source("http://zanel.biz/med/useRj/scripts/forest_metafor_b.R")
```

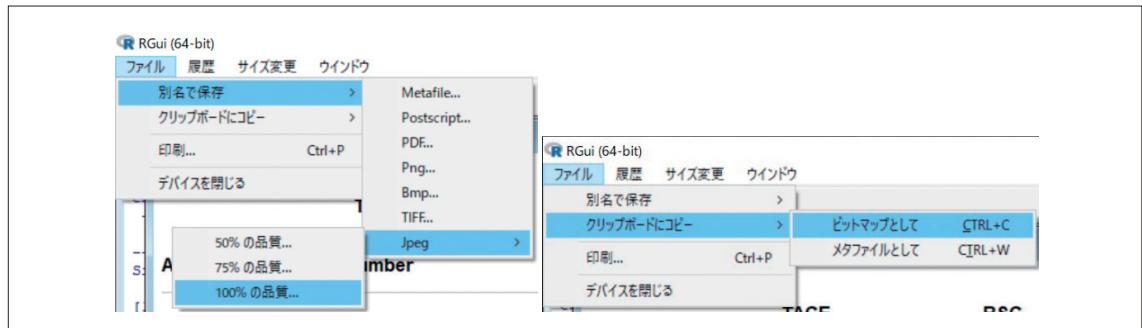
Excel を起動して、(3) メタアナリシスのためのデータの準備で用意したデータファイルを開き、青枠の範囲を選択してコピーする。

R に戻り、上記スクリプトの #1 を実行する。Windows の場合は、#Windows と書いてある行にカーソルを置き、実行ボタンをクリックする（または [Ctrl] + [R] キー）。Mac の場合は、#Mac と書いてある行にカーソルを置き、[Command] + [Enter] キーを押す。この操作により、データがクリップボードを経由して R に手渡される。

続いて、上記スクリプトの #2 を実行する。これは上記スクリプトの source(……) の行にカーソルを置くことで同様に実行でき、フォレストプロットとファンネルプロットが描画される。フォレストプロットはウィンドウの大きさを変更して、全体が見えるように調整する。



プロットを選択した状態で、「ファイル」メニューからファイルとして保存したり、「クリップボードにコピー」より Excel, Word, PowerPoint に貼り付けたりすることができる。



4.7 準備資料2：診断精度研究のメタアナリシス（RとmadaおよびOpenBUGS）

詳細は以下を参照されたい。

森實敏夫 他. (2015) 特別寄稿 5 診断に関する診療ガイドライン (CPG) の作成.
https://minds.jcqhc.or.jp/s/guidance_special_articles5_1 (参照 2021/01/28)

ここでは、参考基準が不完全で、インデックス診断法と相関があり、異なる参考基準が用いられている場合に対応するDendukuri NらによるBUGSコードを提示する。

```
#Imperfect reference standard, covariance-adjusted (T1 dependent on T2).
#l: the number of studies, k: the number of references.
model {

  for(i in 1:l) {
    theta[i] ~ dnorm(THETA,prec[1])
    alpha[i] ~ dnorm(LAMBDA,prec[2])
    p[1,i] <- phi(-(theta[i] - 0.5*alpha[i])/exp(beta/2))
    p[2,i] <- phi(-(theta[i] + 0.5*alpha[i])*exp(beta/2))
    prob[i,1] <- pi[i]*( p[1,i] * s2[ref[i]] +covs12[i] ) + (1-pi[i])*( p[2,i] * (1-c2[ref[i]]) +covc12[i] )
    prob[i,2] <- pi[i]*( p[1,i] * (1-s2[ref[i]]) -covs12[i] ) + (1-pi[i])*( p[2,i] * c2[ref[i]] -covc12[i] )
    prob[i,3] <- pi[i]*( (1-p[1,i]) * s2[ref[i]] -covs12[i] ) + (1-pi[i])*( (1-p[2,i]) * (1-c2[ref[i]]) -covc12[i] )
    prob[i,4] <- pi[i]*( (1-p[1,i]) * (1-s2[ref[i]]) +covs12[i] ) + (1-pi[i])*( (1-p[2,i]) * c2[ref[i]] +covc12[i] )
    results[i,1:4] ~ dmulti(prob[i,1:4],n[i])
    n[i] <- sum(results[i,1:4])
    pi[i] ~ dbeta(1,1)
    se[i] <- p[1,i]
    sp[i] <- 1-p[2,i]
    alpha.covs12[i] <-1
    beta.covs12[i] <-1
    alpha.covc12[i] <-1
    beta.covc12[i] <-1
    us[i] <- min(se[i],s2[ref[i]])-( se[i]*s2[ref[i]] );
    uc[i] <- min(sp[i],c2[ref[i]])-( sp[i]*c2[ref[i]] );
    u.covs12[i] ~ dbeta(alpha.covs12[i],beta.covs12[i]);
    u.covc12[i] ~ dbeta(alpha.covc12[i],beta.covc12[i]);
    covs12[i] <- u.covs12[i]*us[i];
    covc12[i] <- u.covc12[i]*uc[i];
  }

  for(j in 1:2) {
    prec[j] <- pow(sigma[j],-2)
    sigma[j] ~ dunif(0,2)
  }

  THETA ~ dunif(-1.5,1.5)
  LAMBDA ~ dunif(-3,3)
  beta ~ dunif(-0.75,0.75)
  S_overall<-phi(-(THETA-LAMBDA)/exp(beta/2))
  C_overall<-phi( (THETA+LAMBDA)/exp(beta/2))
  theta_new ~ dnorm(THETA,prec[1])
}
```

By courtesy of Dr. Dendukuri

4.8 補足資料 3：ネットワークメタアナリシス

(1) OpenBUGS を用いるベイジアンアプローチによるコントラストベースモデルによるネットワークメタアナリシス

通常のメタアナリシスで行われる治療1と治療2の直接比較だけでなく、さらに治療3と間接比較する場合、治療1と治療3の間接比較と、治療2と治療3の間接比較ではそれぞれの治療効果 δ に対して次の分布を想定できる。 i は研究番号を表す。

$$\begin{aligned}\delta_{i,13} &\sim N(d_{13}, \sigma_{13}^2) \\ \delta_{i,23} &\sim N(d_{23}, \sigma_{23}^2)\end{aligned}$$

これらから、 $\delta_{i,23} = \delta_{i,13} - \delta_{i,12}$ が成立する場合、1つの研究内での共通の基準との効果の比較は互いに加算・減算できるということになり、移行性が成立するという。ネットワークメタアナリシスではこの移行性が成立することを前提としている。すなわち、治療1と治療3の差から治療1と治療2の差を除くと、治療2と治療3の差が得られることが正しいことを前提としている。パラメータである d についても同様の式が成立し、 $d_{23} = d_{13} - d_{12}$ となる。

分散については、1つの研究で得られた効果指標の間には相関があることを想定し、相関係数 ρ で調整が行われる。すなわち $\sigma_{23}^2 = \sigma_{12}^2 + \sigma_{13}^2 - \rho_{23}^{(1)}\sigma_{12}^2\sigma_{13}^2$ となる。ここでは治療1を基準として設定している。

分散の評価における $\rho_{23}^{(1)}$ は、1つの研究における治療1と比較した治療3の効果と治療1と比較した治療2の効果の相関である。すなわち、1つの研究で3つの治療がRCTで比較され、そのうち1つ（治療1）がプラセボのような場合、対象者が全体として軽症者であれば、治療2、治療3のいずれの治療も効果が高めに出る可能性がある。逆に重症者が多ければ、効果が低めに出る可能性がある。したがって、治療2と治療3に相関があることを前提とする必要がある。

以下に紹介するBUGSのコードで用いるDiasら（2011）のモデルでは、分散に関しては同じ値、すなわち $\sigma_{12}^2 = \sigma_{13}^2 = \sigma_{23}^2 = \sigma^2$ と見なし、全ての比較されるペアの間で $\sigma^2 = 0.5$ としている。

このモデルを適用することによって、治療1と治療2の比較あるいは治療1と治療3の比較が異なる研究で行われていた場合、間接的に治療2と治療3を比較することが可能になる。その場合にそれぞれの研究の交換可能性が前提となる。

また、間接的な比較で得られる効果については次の式が成立する必要があり、これらが成立する状態を一致性（consistency）またはコヒレンス（coherence）と呼ぶ。 s は治療番号を示す。

$$\begin{aligned}d_{23} &= d_{13} - d_{12} \\ d_{24} &= d_{14} - d_{12} \\ &\vdots \\ d_{(s-1),s} &= d_{1s} - d_{1,(s-1)}\end{aligned}$$

言い換えると、直接比較の結果と間接比較の結果の間に乖離がないということが必要な条件となる。実際には乖離することもあり得るので、何らかの方法、例えば後述するネットヒートプロット（net heat plot）やノード分割モデルプロットなどで点検が行われる。

k 種類の治療、すなわち治療 $1 \sim k$ を比較する場合についても適用可能な表記を用いてランダム効果モデルを記述すると、以下のようなになる。 t_{ik} は研究 i の治療 k という意味である。 σ^2 は研究間のばらつき、すなわち研究間の異質性を表す項である。

$$\delta_{i,1k} \sim N(d_{t_{i,ta_i}}, \sigma^2)$$

さらに、4種類以上の治療を比較した臨床試験の結果を取り込んでネットワークメタアナリシスを行う際に Dias S らの述べているモデルは以下のとおりである。

$$\delta_i = \begin{bmatrix} \delta_{i,12} \\ \delta_{i,13} \\ \vdots \\ \delta_{i,1a_i} \end{bmatrix} \sim N_{a-1} \begin{bmatrix} d_{i1,t_2} \\ d_{i1,t_3} \\ \vdots \\ d_{i1,t_{a_i}} \end{bmatrix}, \begin{bmatrix} \sigma^2 & \frac{\sigma^2}{2} & \dots & \frac{\sigma^2}{2} \\ \frac{\sigma^2}{2} & \sigma^2 & \dots & \frac{\sigma^2}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma^2}{2} & \frac{\sigma^2}{2} & \dots & \sigma^2 \end{bmatrix}$$

研究 i における効果指標 δ_i は治療の種類分の値からなるベクトルで表され [a_i は治療の個数 ($2, 3, \dots, k$)]、それぞれ正規分布に従う。研究間の分散は同じ値であるとしている。

また、治療 2 から治療 $k - 1$ の効果指標が与えられた際の $\delta_{i,1k}$ の分布は次の式で示される。

$$\delta_{i,1k} \begin{bmatrix} \delta_{i,12} \\ \delta_{i,13} \\ \vdots \\ \delta_{i,1(k-1)} \end{bmatrix} \sim N \left[(d_{1,t_k} - d_{1,t_1}) + \frac{1}{k-1} \sum_{j=1}^{k-1} [\delta_{i,1j} - (d_{1,t_j} - d_{1,t_1})], \frac{1}{2(k-1)} \sigma^2 \right]$$

これら 2 つの式のいずれかを用いることによって、治療間の相関を取り込んだランダム効果モデルによるネットワークメタアナリシスが可能になる。

実際の BUGS 用のコードは、以下のとおりである (Dias S 2011: Appendix)。

```
### Network meta-analysis with logit link regression, random-effects model###

model
{
for(i in 1:ns)           # ns は研究数
{
w[i,1]<-0               # 対照のアームの調整は 0 に設定。
delta[i,1]<-0            # 対照のアームの治療効果は 0 に設定。
mu[i]~dnorm(0,0.0001)    # ベースライン（対照の治療効果）のあいまい事前分布の設定。
for(k in 1:na[i])        # 研究 i の治療アーム数だけ以下を繰り返す。
{
r[i,k]~dbin(p[i,k],n[i,k]) # イベント数（アウトカム+例数）は二項分布に従う。二項分布尤度。
logit(p[i,k])<-mu[i]+delta[i,k] # 直線回帰のロジットモデル。k=1 つまり対照の場合 delta[i,k]=0 となる。
rhat[i,k]<-p[i,k]*n[i,k]      # イベント数の期待値 (Deviance 算出に必要)。
dev[i,k]<-2*(r[i,k]*(log(r[i,k])-log(rhat[i,k]))+(n[i,k]-r[i,k])*(log(n[i,k]-r[i,k])-log(n[i,k]-rhat[i,k]))) # Deviance 要素算出
}
resdev[i]<-sum(dev[i,1:na[i]]) # 研究 i の Deviance 合計。
for(k in 2:na[i])           # 治療 1 を対照に、治療アーム分繰り返す。
{
delta[i,k]~dnorm(md[i,k],taud[i,k]) # 研究 i の δ は ln(OR) の平均値 md、分散 taud の正規分布よりサンプリング。
md[i,k]<-d[t[i,k]]-d[t[i,1]]+sw[i,k] # アーム数で調整した ln(OR) の平均値
taud[i,k]<-tau^2*(k-1)/k          # 分散 σ^2 * 1/[2(k-1)] の逆数 precision を計算。
w[i,k]<-(delta[i,k]-d[t[i,k]]+d[t[i,1]]) # アーム数の調整分の計算
sw[i,k]<-sum(w[i,1:k-1])/(k-1)       # 調整分の合計の計算
}
}
totresdev<-sum(resdev[])    # Total residual deviance
d[1]<-0                      # 参照アームの効果を 0 に設定。
for(k in 2:nt)
{
d[k]~dnorm(0,0.0001)         # 治療効果のあいまい事前分布の設定。
}
sd~dunif(0,5)                 # 研究間分布の標準偏差にあいまい事前分布の設定。
tau<-pow(sd,-2)              # 標準偏差から分散の逆数を求める tau に代入する。
### sampling effect measures OR, RR, RD, NNT
for(c in 1:(nt-1)){
for(k in (c+1):nt){
OR[c,k]<- exp(d[k] - d[c])           # 統合値のオッズ比をサンプリング。
lor[c,k]<- d[k]-d[c]                  # オッズ比の自然対数をサンプリング。
RR[c,k]<-exp(d[k)*(1+exp(d[c]))/exp(d[c])/(1+exp(d[k])) # リスク比のサンプリング。
RD[c,k]<-exp(d[k])/(1+exp(d[k]))-exp(d[c])/(1+exp(d[c])) # リスク差のサンプリング。
NNT[c,k]<-1/abs(RD[c,k])             # NNT のサンプリング。
}
}
###
# ranking on relative scale
for(k in 1:nt)
{
rk[k]<- nt+1-rank(d[],k)            # アウトカムがよい事象の場合。
#rk[k]<- rank(d[],k)                # アウトカムが悪い事象の場合。
best[k]<- equals(rk[k],1)           # 治療 k が最善である確率。
for(h in 1:nt){
prob[h,k]<- equals(rk[k],h)         # 治療 k が h 番目による治療である確率。
}
}
```

出典 : Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; last updated September 2016; available from <http://www.nicedsu.org.uk>. p.59 より作成

なお、将来同じような研究が行われた場合のオッズ比の予測分布を得たい場合には、

```
lor[c,k] <- d[k]-d[c]      # オッズ比の自然対数をサンプリング
```

の行に続けて以下の行を追加し、node として lornew と ORnew を設定する。

`ornew[c,k]~dnorm(lor[c,k],taud)`
`ORnew[c,k]~exp(lornew[c,k])`

(2) BUGS 用のデータの準備について

BUGS でコードを走らせる際には、「Model」メニューから「Specification…」を選択し、「Specification Tool」で「check model」, 「load data」, 「compile」, 「load inits」, 「gen inits」の順に操作する。データは Excel で用意して必要な範囲を選択してコピーし、BUGS の「Edit」メニューから「Paste Special」を選び、貼り付ける形式で「Plain Text」を選択して [OK] ボタンをクリックする。このようにすることで、Excel から直接コピー＆ペーストでデータを取り込める。

例えば、Excel で次のようなデータを用意したとする。

	A	B	C	D
1	rc[]	nc[]	rt[]	nt[]
2	4	11	5	13
3	3	14	1	13
4	14	19	2	17
5	1	12	0	13
6	9	13	7	17
7	13	47	18	48
8	END			

セル A1 から D8 の範囲を選択してコピーし、BUGS のコード画面に貼り付けることができる。

```

DATA #k is the number of studies.
list(k=6, rc=c(4,3,14,1,9,13),nc=c(11,14,19,12,13,47),rt=c(5,1,2,0,7,18),nt=c(13,13,17,13,17,48))

list(k=6)
rc[] nc[] rt[] nt[]
4 11 5 13
3 14 1 13
14 19 2 17
1 12 0 13
9 13 7 17
13 47 18 48
END

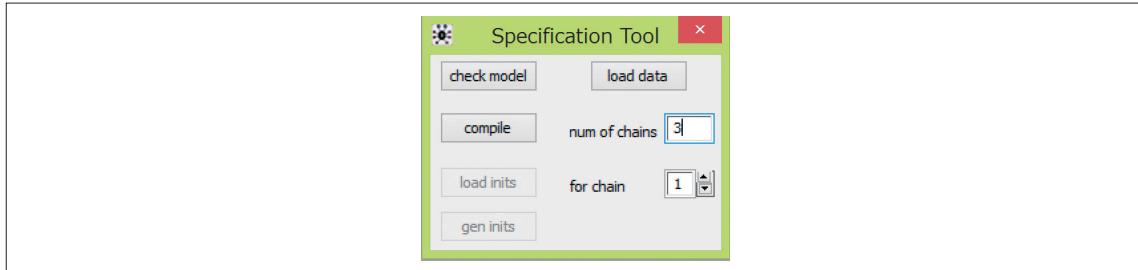
```

さらに、`list(k=6)` を書き加えたのが上記の例である。ここには 2 種類のデータの記述法を示している。上部の記述法の場合は、`list` を反転させて [load data] をクリックする。下部の記述法の場合は、`list` を反転させて [load data] をクリックし、次に `rc[]` から `nt[]` の行を選択して再度 [load data] をクリックする。データとして読み込まれるものはいずれの操作の場合も同じである。例えば、`rc[2]` で指定される値は 3 となりいずれの場合も同じになる。

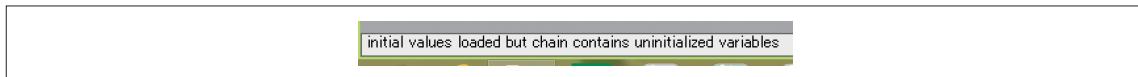
このように、配列のデータを作成する場合に、列名を変数名として用いることによって、Excel で用意したデータを利用利用することができる。

なお、上記のネットワークメタアナリシスの解析データは後述する `netmeta` パッケージのデータセット「smokingcessation」と同じものである。R でいったん読み込んで、`write.csv()` 関数でファイルとして書き出し、それを Excel で開いて、形式を整えてから、同様な手順で用意した。

さて、それではこのデータに対してネットワークメタアナリシスを実行してみよう。この例では、初期値を3セット用意してあり、3チェーン実行させる。「check model」、「load data」の実行後、「num of chains」を3に変更し、[compile] をクリックする。



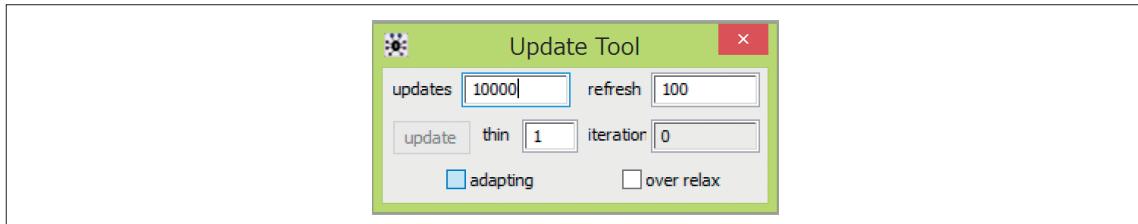
次に、#INITS の下の初期値を順に設定する。最初の list を選択して、[load inits] ボタンをクリックする。すると、「num of chains」が1から2に変わる。次の list を選択して同じことを繰り返す。3種類の初期値の「load inits」が終了すると「num of chains」が1に戻る。その際に、一部の変数の初期値が設定されていないというメッセージがウィンドウの下枠に表示されるので、[gen inits] ボタンをクリックする。



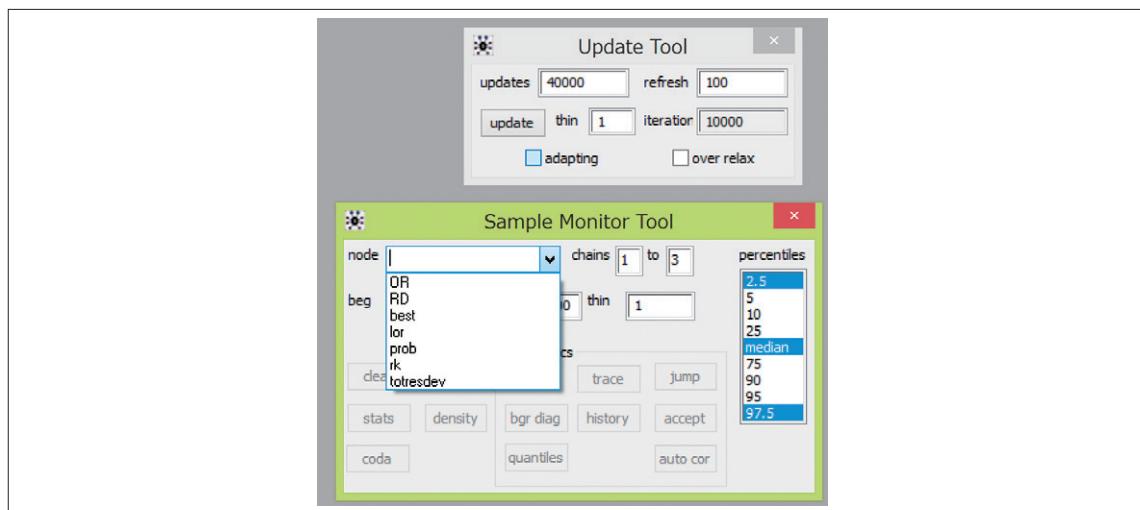
すると、メッセージが以下のように変わる。



これで MCMC (Markov chain Monte Carlo：マルコフ連鎖モンテカルロ) を走らせる準備ができた。「Model」メニューから「Update…」を選択し「Update Tool」を表示させ、「updates」を「10000」に設定して、burn-in を1万回実行する。



次に、「Inferences」メニューから「Samples…」を選択し「Sample Monitor Tool」を表示させ、「node」にモニター（記録）したい変数名を入力して [set] ボタンをクリックする。必要なものについてこれを繰り返す。例えば、「OR」、「Ior」、「RD」、「totresdev」、「rk」、「best」、「prob」などを必要に応じて設定する。「node」の設定が終わったら、「Update Tool」で「updates」を「40000」に設定して [update] ボタンをクリックする。この時点では「Sample Monitor Tool」で「chains」が「1 to 3」に設定されているのが分かる。



「Update Tool」の「iteration」の値が「50000」に達すると MCMC が停止する。

「Sample Monitor Tool」で「node」に「*」を入力して [stats] ボタンをクリックすると次の結果が表示される。

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
OR[1,2]	1.772	0.7805	0.003913	0.7476	1.623	3.66	10001	120000
OR[1,3]	2.391	0.6026	0.003552	1.474	2.304	3.812	10001	120000
OR[1,4]	3.311	1.642	0.009828	1.298	2.963	7.36	10001	120000
OR[2,3]	1.55	0.6875	0.002731	0.6319	1.421	3.248	10001	120000
OR[2,4]	2.073	1.129	0.005164	0.7133	1.827	4.893	10001	120000
OR[3,4]	1.409	0.6359	0.003022	0.5726	1.285	2.968	10001	120000
RD[1,2]	0.1157	0.09116	4.523E-4	-0.07221	0.1188	0.2854	10001	120000
RD[1,3]	0.1966	0.04965	2.803E-4	0.09583	0.1974	0.2922	10001	120000
RD[1,4]	0.2413	0.08086	4.774E-4	0.06482	0.2477	0.3804	10001	120000
RD[2,3]	0.08088	0.09221	3.794E-4	-0.09219	0.0779	0.2724	10001	120000
RD[2,4]	0.1256	0.1002	4.387E-4	-0.07149	0.1253	0.3247	10001	120000
RD[3,4]	0.04471	0.07826	3.579E-4	-0.1231	0.04947	0.1861	10001	120000
best[1]	3.333E-5	0.005773	1.662E-5	0.0	0.0	0.0	10001	120000
best[2]	0.05717	0.2322	7.834E-4	0.0	0.0	1.0	10001	120000
best[3]	0.2381	0.4259	0.00165	0.0	0.0	1.0	10001	120000
best[4]	0.7047	0.4562	0.0018	0.0	1.0	1.0	10001	120000
lor[1,2]	0.4896	0.4016	0.002007	-0.2909	0.4843	1.297	10001	120000
lor[1,3]	0.8424	0.2397	0.001375	0.3881	0.8348	1.338	10001	120000
lor[1,4]	1.098	0.4393	0.002655	0.2608	1.086	1.996	10001	120000
lor[2,3]	0.3528	0.4117	0.001678	-0.459	0.3512	1.178	10001	120000
lor[2,4]	0.6082	0.4866	0.002235	-0.3378	0.6027	1.588	10001	120000
lor[3,4]	0.2554	0.4158	0.001989	-0.5576	0.251	1.088	10001	120000
prob[1,1]	3.333E-5	0.005773	1.662E-5	0.0	0.0	0.0	10001	120000
prob[1,2]	0.05717	0.2322	7.834E-4	0.0	0.0	1.0	10001	120000
prob[1,3]	0.2381	0.4259	0.00165	0.0	0.0	1.0	10001	120000
prob[1,4]	0.7047	0.4562	0.0018	0.0	1.0	1.0	10001	120000
prob[2,1]	0.002383	0.04876	1.533E-4	0.0	0.0	0.0	10001	120000
prob[2,2]	0.1723	0.3777	0.001169	0.0	0.0	1.0	10001	120000
prob[2,3]	0.6001	0.4899	0.001685	0.0	1.0	1.0	10001	120000
prob[2,4]	0.2251	0.4177	0.001472	0.0	0.0	1.0	10001	120000
prob[3,1]	0.1055	0.3072	0.001133	0.0	0.0	1.0	10001	120000
prob[3,2]	0.6671	0.4712	0.001501	0.0	1.0	1.0	10001	120000
prob[3,3]	0.1616	0.3681	0.001288	0.0	0.0	1.0	10001	120000
prob[3,4]	0.06575	0.2478	8.077E-4	0.0	0.0	1.0	10001	120000
prob[4,1]	0.8921	0.3103	0.001159	0.0	1.0	1.0	10001	120000
prob[4,2]	0.1033	0.3044	0.001126	0.0	0.0	1.0	10001	120000
prob[4,3]	1.667E-4	0.01291	3.667E-5	0.0	0.0	0.0	10001	120000
prob[4,4]	0.004408	0.06625	2.006E-4	0.0	0.0	0.0	10001	120000
rk[1]	3.89	0.3212	0.001207	3.0	4.0	4.0	10001	120000
rk[2]	2.817	0.6861	0.002631	1.0	3.0	4.0	10001	120000
rk[3]	1.924	0.6281	0.002438	1.0	2.0	3.0	10001	120000
rk[4]	1.37	0.6253	0.002455	1.0	1.0	3.0	10001	120000
totresdev	54.06	10.15	0.05225	35.95	53.47	75.61	10001	120000

「Node statistics」を見ると sample が 120000 になっており、4万回の MCMC が初期値を変えて3チェーン分実行されたことが分かる。

ここでは実際の結果は示さないが、「Sample Monitor Tool」で「density」, 「history」および「auto cor」も見ることができる。

「best」の結果を見ると、best[4] の mean が 0.7047 となっている。これは、治療 4 が最も効果が高い可能性が高く、その確率が 0.7047 であることを示している。また、rk の結果を見ると median が rk[4] が 1, rk[3] が 2, rk[2] が 3, rk[1] が 4 で、治療効果がこの順で高いことが分かる。また prob の項はそれぞれの治療が特定の順位になる確率を示しており、例えば prob[1,4] は治療 4 が 1 番である確率であり、mean を見ると 0.7047 で best[4] と同じ値になっていることが分かる。

(3) gemtc パッケージによるネットワークメタアナリシス

gemtc パッケージを使用する際は、R とは別に JAGS をインストールしておく必要がある。さらに、R の rjags パッケージと gemtc パッケージをインストールする。gemtc パッケージはコントラストベースモデルを用いたベイジアンの手法によるネットワークメタアナリシスを実行するパッケージである。

次に、Excel で次のような形式でデータを用意する (smokingcessation のデータ)。以下の例では、4種類の治療のうち 1 つの RCT で直接比較されたのは 3 種類であり、「1」が対照として用いられている。

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]
9	140	23	140	10	138	1	3	4	3
11	78	12	85	29	170	2	3	4	3

(以下省略)

R のスクリプトは次のとおりである。データを Excel で用意し、データの範囲をラベルの行も含めてコピーし、R でスクリプトの全てを選択して実行する。

```
####Network meta-analysis with gemtc of Bayesian model#####
#Using JAGS, rjags, Rcpp#####
library("rjags")
library("gemtc")

#Read and convert BUGS type (eg. smoking cessation) data.
if(colnames(exdat)[2]=="y..1."||colnames(exdat)[1]=="y..1."){dtype="continuous"}
if(colnames(exdat)[2]=="r..1."||colnames(exdat)[1]=="r..1."){dtype="binary"}
dtype

#Binary data:
if(dtype=="binary"){
{
gdat=mtc.data.studyrow(exdat,armVars=c('treatment'='t', 'sampleSize'='n', 'responders'='r'))
lh="binom";lk="logit"           #responders, sampleSize --> Odds Ratio (default)
}
#Continuous data:
if(dtype=="continuous"){
{
gdat=mtc.data.studyrow(exdat,armVars=c('sampleSize'='n','treatment'='t',mean='y','std.dev'='sd'))
lh="normal";lk="identity"      #mean, std.dev, sampleSize OR mean, std.err --> Mean difference
}
#----->>>>> To gemtc.
burnin=2000
cycl=20000
thn=1

netw=mtc.network(data.ab=gdat, treatments=NULL,description="Network",data.re=NULL,studies=NULL,data=NULL)
summary(netw)
plot(netw)

#lh="normal";lk="identity"    #mean, std.dev, sampleSize OR mean, std.err --> Mean difference
#lh="binom";lk="log"          #responders, sampleSize --> Risk Ratio
#lh="binom";lk="logit"        #responders, sampleSize --> Odds Ratio (default)
#lh="binom";lk="cloglog"      #responders (for the same peirod), sampleSize --> Hazard Ratio
#lh="poisson";lk="log"        #responders, exposure --> Hazard Ratio

model=mtc.model(netw,type="consistency",likelihood=lh, link=lk, linearModel="random")#Random effects model

#Run MCMC for the model with JAGS via rjags. n.adapt=burn-in
res=mtc.run(model, sampler=NA, n.adapt=burnin,n.iter=cycl,thin=thn)
summary(res)

#plot(res)
dev.new();forest(res)

rank.p=rank.probability(res,preferredDirection=1,covariate=NA)
#rank.p=rank.probability(res,preferredDirection=-1,covariate=NA)
print(rank.p)
dev.new();plot(rank.p)

relative.e=relative.effect.table(res,covariate=NA)
print(relative.e)

gelman.diag(res)
mtc.deviance(res)
```

```
res$dic #DIC

#####
#Node-splitting analysis:
res.ns=mtc.nodesplit(netw,likelihood=lh, link=lk, linearModel="random",n.adapt=burnin,n.iter=cycl,thin=thn)
#dev.new();plot(res.ns)
summary.ns=summary(res.ns)
summary.ns
dev.new();plot(summary.ns)

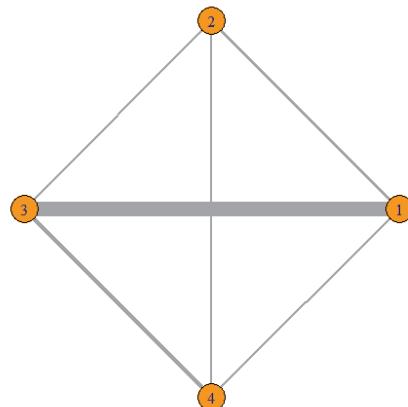
###SUCRA plot function modified from the one of Salanati's script.
sucra = function(effectiveness)
{
  # effectiveness: rank.p from GeMTC output.
  #Creates cumulative ranking curves extrapolated at the middle of each bar

  names <- names(effectiveness)
  nr.of.treat <- dim(effectiveness)[2]
  pn=round(sqrt(nr.of.treat))
  cumeffectiveness <- apply(effectiveness, 2, cumsum)

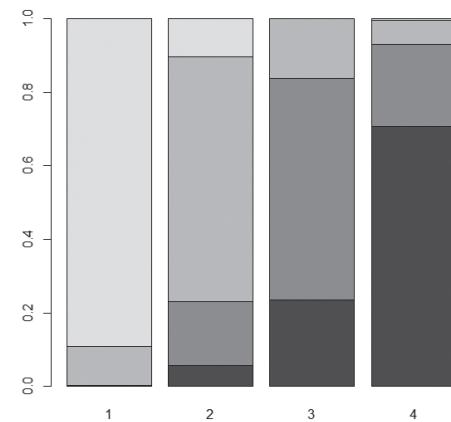
  dev.new()
  par(mfrow = c(pn,pn))
  for(i in 1:nr.of.treat) {
    plot(1:nr.of.treat, type = "n", ylim = c(0, 1), xlab =
      paste("Rank of", as.character(names[i])), ylab =
      "Cumulative Probability",xaxp=c(1,nr.of.treat,nr.of.treat-1))
    y=cumeffectiveness[c(1, 1:c(nr.of.treat - 1), c(nr.of.treat - 1)), i]
    x=c(1, c(1:c(nr.of.treat - 1)) + 0.5, nr.of.treat)
    lines(lwd = 2, x, y,type="l")

  }
  dev.new()
  par(mfrow = c(pn,pn))
  for(i in 1:nr.of.treat) {
    plot(1:nr.of.treat, type = "n", ylim = c(0, 1), xlab =
      paste("Rank of", as.character(names[i])), ylab =
      "Probability",xaxp=c(1,nr.of.treat,nr.of.treat-1))
    x=c(1, c(1:c(nr.of.treat - 1)) + 0.5, nr.of.treat)
    y=effectiveness[c(1, 1:c(nr.of.treat - 1), c(nr.of.treat - 1)), i]
    lines(lwd = 2, x, y,type="l")
  }
}
names(rank.p)=row.names(rank.p)
sucra(rank.p)
#####End to the script#####
```

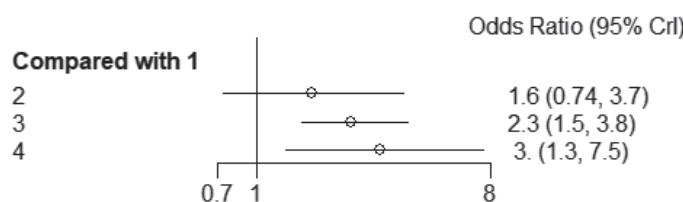
結果として、ネットワークジオメトリプロット、フォレストプロット、順位グラフ、ノード分割モデルプロット、SUCRA プロットなどが得られる。



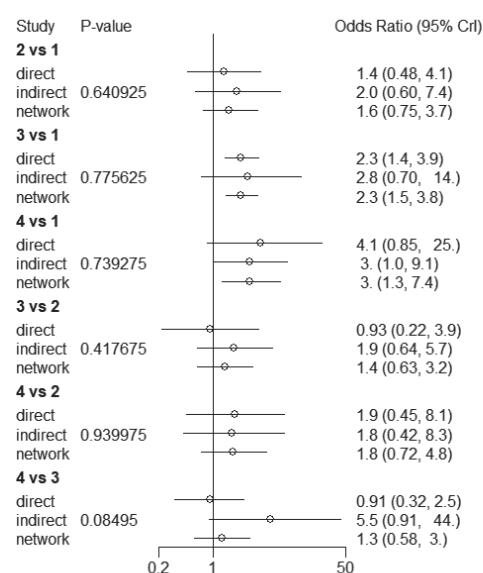
ネットワークジオメトリプロット



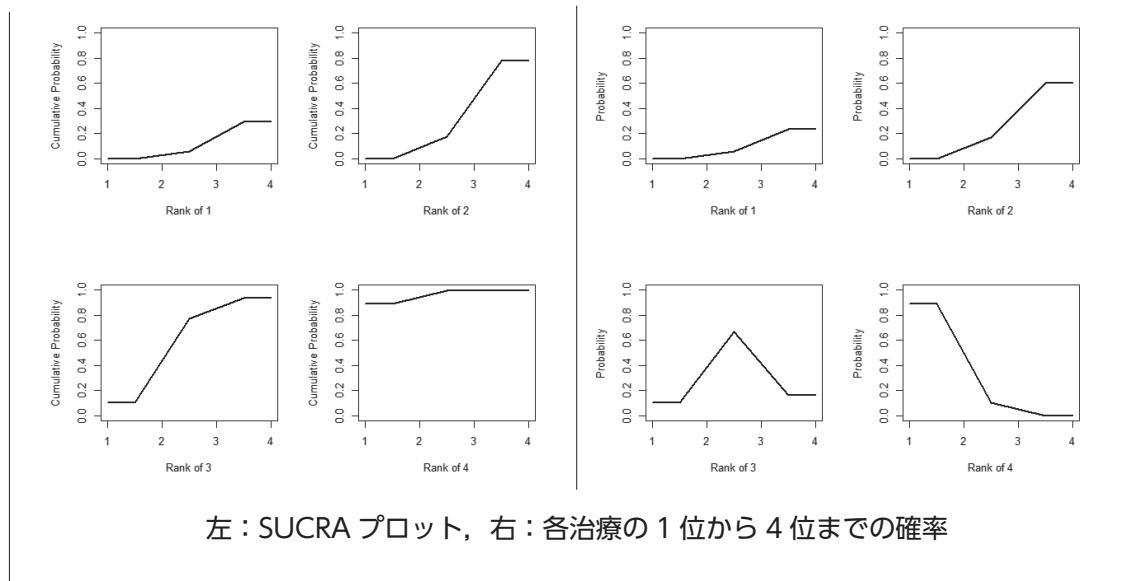
順位グラフ



フォレストプロット



ノード分割モデルプロット



この結果からは、治療 4 と 3 のペア比較でインコヒレンスが疑われるが、間接比較の信頼区間が大きく、ネットワーク推定値への貢献度は低いと考えられる。順位からは治療 4 が最善である確率が 0.7 程度で最も高いことが分かる。

上記の例は効果指標としてオッズ比を用いているが、gemtc パッケージは mtc.model 関数の引数 `likelihood` の設定を変えることで、連続変数、リスク比、ハザード比などにも対応できるようになっている。スクリプト中の変数 `lh` の設定を変更する。

連続変数の場合のデータの形式は以下のようになる。この例では、治療が 3 種類あり、1～3 の番号で示されている。`n` は症例数、`y` は平均値、`sd` は標準偏差である。10 番目の研究では 3 種類の介入が直接比較されているが、それ以外では治療 3 の群がないため、`n[3]`, `y[3]`, `sd[3]` が NA となっている。このような形式のデータを Excel などで用意すれば、その後の操作は前述と同様である。

<code>n[1]</code>	<code>y[1]</code>	<code>sd[1]</code>	<code>t[1]</code>	<code>n[2]</code>	<code>y[2]</code>	<code>sd[2]</code>	<code>t[2]</code>	<code>n[3]</code>	<code>y[3]</code>	<code>sd[3]</code>	<code>t[3]</code>	<code>na[]</code>
70	-10.26	13.43	2	66	-4.88	11.64	1	1	NA	0	NA	2
54	-14.78	12.49	2	56	-8.13	12.72	1	1	NA	0	NA	2
125	-13.36	8.78	2	123	-10.39	10.45	3	1	NA	0	NA	2
220	-13.11	8.53	2	114	-9.1	9.36	1	1	NA	0	NA	2
55	-16.6	12.49	2	60	-14.9	11.39	3	1	NA	0	NA	2
58	-15.49	8.148877	2	59	-15.25	8.372449	1	1	NA	0	NA	2
187	-11.5	10.9	3	177	-9	10.9	1	1	NA	0	NA	2
144	-10.1	10.9	3	78	-8.7	10.9	1	1	NA	0	NA	2
100	-11.71	7	2	101	-8.97	6.934414	1	1	NA	0	NA	2
201	-9.4	8.506468	2	186	-8.2	8.455673	3	99	-7.4	7.9599	1	3

(4) OpenBUGS を用いるベイジアンアプローチによるアームベースモデルによるネットワークメタアナリシス

ここでは、二値変数アウトカムの場合の BUGS コードとデータの例を示す。データのフォーマットは 1 行が 1 治療アームとなる。サンプルデータは (3) 項の smokingcessation と同じものでアウトカムは 1 年後の禁煙であり、治療数は 4 種類ある。治療 1 が参照治療である。変数 probt が各治療群の絶対リスク、各ペア比較のオッズ比は OR、リスク比は RR、リスク差は RD で得られる。best は各治療が最善である確率、rk はそれぞれの順位になる治療番号である。

```
#Zhang J, et al: Network meta-analysis of randomized clinical trials: reporting the proper summaries. Clin Trials 2014;11:246-62.
PMID: 24096635
#Arm-based model.
model {
for(i in 1:tS) {
p[i]<-phi(mu[t[i]]+ vi[s[i], t[i]]) # model
r[i]~dbin(p[i], totaln[i]) # binomial likelihood
}
for(j in 1:sN){
vi[j, 1:tN]~dmnorm(mn[1:tN], T[1:tN,1:tN]) # multivariate normal distribution
}
invT[1:tN, 1:tN]<-inverse(T[,])
for (j in 1:tN){
mu[j]~dnorm(0, 0.001)
sigma[j]<-sqrt(invT[j, j])
probt[j]<-phi(mu[j]/sqrt(1+invT[j, j]))
# population-averaged treatment specific event rate
}
T[1:tN,1:tN]~dwish(R[1:tN, 1:tN], tN) # Wishart prior
for (k in 1:tN) {
rk[k]<- tN + 1 - rank(probt[,k]) # ranking
best[k]<-equals(rk[k],1) # prob {treatment k is best}
}
for (j in 1:tN){ # calculation of RR, RD and OR
for (k in (j+1):tN){
RR[j, k] <- probt[k]/probt[j]
RD[j, k] <- probt[k]-probt[j]
OR[j, k] <- probt[k]/(1-probt[k])/probt[j]*(1-probt[j])
}
}
}

#DATA
list(tN=4,sN=24,tS=50,mn=c(0, 0, 0, 0),R=structure(.Data=c(1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1),.Dim=c(4,4)))
s[]      t[]      r[]      totaln[]
1       1        9       140
1       3       23       140
1       4       10       138
2       2       11        78
2       3       12        85
2       4       29       170
3       1       75       731
3       3      363       714
4       1        2       106
4       3        9       205
```

5	1	58	549
5	3	237	1561
6	1	0	33
6	3	9	48
7	1	3	100
7	3	31	98
8	1	1	31
8	3	26	95
9	1	6	39
9	3	17	77
10	1	79	702
10	2	77	694
11	1	18	671
11	2	21	535
12	1	64	642
12	3	107	761
13	1	5	62
13	3	8	90
14	1	20	234
14	3	34	237
15	1	0	20
15	4	9	20
16	1	8	116
16	2	19	146
17	1	95	1107
17	3	134	1031
18	1	15	187
18	3	35	504
19	1	78	584
19	3	73	675
20	1	69	1177
20	3	54	888
21	2	20	49
21	3	16	43
22	2	7	66
22	4	32	127
23	3	12	76
23	4	20	74
24	3	9	55
24	4	3	26

END

OpenBUGS で「burn in」を 10000 回、「MCMC」を 50000 回実行後「stats」で得られた結果は以下のとおりである。

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
OR[1,2]	2.517	1.284	0.07639	0.8763	2.229	5.924	10001	50000
OR[1,3]	2.629	0.7801	0.04756	1.404	2.517	4.508	10001	50000
OR[1,4]	3.782	1.753	0.1029	1.591	3.388	8.563	10001	50000
OR[2,3]	1.244	0.6155	0.03546	0.4579	1.129	2.845	10001	50000
OR[2,4]	1.819	1.132	0.06503	0.4546	1.529	4.995	10001	50000
OR[3,4]	1.483	0.6152	0.03562	0.6914	1.342	3.027	10001	50000
RD[1,2]	0.09202	0.06456	0.003849	-0.01078	0.08138	0.2531	10001	50000
RD[1,3]	0.1012	0.03234	0.001993	0.03745	0.1008	0.169	10001	50000
RD[1,4]	0.1576	0.06774	0.003864	0.04821	0.1484	0.317	10001	50000
RD[2,3]	0.00922	0.06626	0.003951	-0.1459	0.01715	0.1222	10001	50000
RD[2,4]	0.06561	0.09525	0.005655	-0.1434	0.06561	0.2656	10001	50000
RD[3,4]	0.05639	0.06979	0.003991	-0.05385	0.04746	0.2138	10001	50000
RR[1,2]	2.18	0.886	0.05251	0.8875	2.025	4.418	10001	50000
RR[1,3]	2.315	0.6025	0.03674	1.345	2.235	3.744	10001	50000
RR[1,4]	3.016	1.048	0.0613	1.5	2.83	5.734	10001	50000
RR[2,3]	1.194	0.491	0.02823	0.5561	1.106	2.46	10001	50000
RR[2,4]	1.572	0.7569	0.04348	0.5501	1.404	3.559	10001	50000
RR[3,4]	1.33	0.4038	0.0233	0.7376	1.262	2.258	10001	50000
best[1]	0	0	4.47E-13	0	0	0	10001	50000
best[2]	0.1794	0.3837	0.01995	0	0	1	10001	50000
best[3]	0.1539	0.3608	0.0161	0	0	1	10001	50000
best[4]	0.6667	0.4714	0.0228	0	1	1	10001	50000
probt[1]	0.08299	0.01644	0.000991	0.05367	0.08177	0.1177	10001	50000
probt[2]	0.175	0.06423	0.003843	0.07964	0.1626	0.3368	10001	50000
probt[3]	0.1842	0.02659	0.001635	0.1384	0.1818	0.2409	10001	50000
probt[4]	0.2406	0.06579	0.003732	0.1378	0.2308	0.3934	10001	50000
rk[1]	3.954	0.2098	0.009321	3	4	4	10001	50000
rk[2]	2.436	0.8324	0.04505	1	3	4	10001	50000
rk[3]	2.167	0.6688	0.0326	1	2	3	10001	50000
rk[4]	1.442	0.6841	0.03449	1	1	3	10001	50000

これより、gemtc パッケージのコントラストベースモデルの結果とほぼ同じ結果が得られた。 median の列を見ると、rk[4] すなわち治療 4 の順位が 1 で 95% 確信区間 (val2.5pc と val97.5pc) が 1～3 となっており、また、治療 4 が最善である確率は best[4] の mean を見ると 0.6667 であることが分かる。参照治療 1 に対して治療 4 のオッズ比 3.388 (95% 確信区間 1.591～8.563) となっており、gemtc パッケージの結果よりも全体として少し大きめの値になっていることが分かる。

OpenBUGS を、R から BRugs パッケージを介して動かすことも可能である。以下にその一例を示す。

```
####Arm-based Network Meta-analysis for binary outcomes#####
#####based on the BUGS code by Zhang: J Clin Trials 2014;11:246-62. PMID: 24096635#####
#You need OpenBUGS installed.(You can get it at http://openbugs.net/w/FrontPage)
#-----
packneed=c("BRugs", "forestplot", "tcltk2");current=installed.packages();addpack=setdiff(packneed, rownames(current));url
="https://cran.rsm.ac.jp/";if(length(addpack)>0){install.packages(addpack, repos=url)};if(length(addpack)==0){print("Already
installed.")}
#-----
###To use BRugs, tcltk, and forestplot.
library(BRugs)
library(tcltk2)
library(forestplot)
#-----
###Set the working directory where you have data, inits text files, and Results folder.
path0=tk_choose.dir(default="",caption="Select a folder to save data files")
path0=paste(path0, "/", sep="")
setwd(path0)
#Create a folder for results in the working directory if there is not.
paths=paste(path0, "Results/", sep="")
if(!dir.exists(paths)){dir.create(paths)}

#Set a name to data and result files.
titl="NWM_smoke" #Set title as you want..
datfilename=paste(titl, "_dat.txt", sep="") #File name for data.
numfilename=paste(titl, "_num.txt", sep="") #File name for the number of studies and others.

#Open and read a model text and save as a text file in the working directory you set.
filnam=tclvalue(tkgetOpenFile(initialfile="", filetypes="{{Text Files} {.txt}} {{All files} {*}}"));if(filnam!=""){nma_
model=readLines(filnam)}
#nma_model==readLines("https://stat.zanet.biz/bugs/nma/ab_model.bi.txt")
sink(modelfilename) #The model file name to be used.
writeLines(nma_model)
sink()

#-----
#Make data to be analyzed and save as text files.
dat=read.delim("clipboard",sep="\t",header=FALSE) #Select only number cells in Excel and copy and get data via
clipboard into dat.

colnames(dat)=c("s[]", "t[]", "r[]", "totaln[]") #Add column names as the names of variables.
tS=length(dat[,1]) #The number of all comparisons.
sN=length(unique(dat[,1])) #The number of studies.
#Add "END" to the data. dat holds the case numbers data to be analyzed.
tN=length(unique(dat[,2])) #The number of treatment options.
sN=length(unique(dat[,1])) #The number of studies.
dat=rbind(dat,c("END", "", "", "", ""))

write.table(dat,datfilename,quote=FALSE,sep="\t",row.names=FALSE,col.names=TRUE) #Save a text file of the data.

numex=paste("list(tN=",tN,",sN=",sN,",tS=",tS,sep="")
#tN the number of
treatment options; sN the number of studies; tS the number of all comparisons.
zeros=rep(0,tN)
```

```
numex=paste(numex,",mn=c(",toString(zeros),")"," ,R=structure(.Data=c("sep=""  
rmat=rep(0,tN*tN)  
for(i in 1:tN){  
rmat[tN*(i-1)+i]=1  
}  
numex=paste(numex,toString(rmat)," ,Dim=c(" ,tN," ,tN,"))","sep=""  
  
write.table(numex,numfilename,quote=FALSE,row.names=FALSE,col.names=FALSE)      #Save a text file of the data in the  
working directory.  
  
###Check the model, load the data from the data files, compile,  
#read the inits file and set inits values and generate inits for not-given inits.  
modelCheck(modelfilename)  
modelData(datfilename)  
modelData(numfilename)  
  
modelCompile(numChains=1)  
#modelInits("inits.txt")  
modelGenInits()  
  
#Burn-in MCMC, set variables for sampling, and do MCM  
modelUpdate(10000)      #<--You can change iteration number.  
  
#For imperfect reference standards with covariance-adjustments (model #3).  
samplesSet(c("OR","RR","RD","probt","rk","best"))  
  
dicSet()                #To get deviance data.  
modelUpdate(50000)      #<--MCMC iteration number you can change.  
  
#Show stats data in the console window.  
res=samplesStats("*")  
print(res)  
#print(modelname)  
print(res)  
  
#-----  
#-----  
###Save files of the results.  
setwd(paths)  
  
#Make file names for stats, density, DIC, History.  
resstats=paste(modelname,titl,"Stats.csv",sep="_")  
resstattxt=paste(modelname,titl,"Stats.txt",sep="_")  
resdensity=paste(modelname,titl,"Density.pdf",sep="_")    #This file would be large in size, several hundred KB.  
resdic=paste(modelname,titl,"DIC.txt",sep="_")  
reshistory=paste(modelname,titl,"History.pdf",sep="_")  
resmodelname=paste(modelname,titl,"model.txt",sep="_")  
  
#Save model name, path, and model script as a text file in the "Results" directory.  
sink(resmodelname)  
cat("# ")  
writeLines(modelname)  
cat("# ")  
writeLines(modelfilename)  
setwd(path0)  
modlines=readLines(modelfilename)  
setwd(paths)
```

```

writeLines(nma_model)
sink()

#Save stats as a csv file.
write.csv(modelname,resstats)
write.csv(samplesStats("*"),resstats,append=TRUE)

#Save density plots as a PDF file.
pdf(resdensity)
samplesDensity("*")
dev.off()

#Save deviance data as a text file.
sink(resdic)
print(dicStats())
sink()

#Save history data file.
pdf(reshistory)
samplesHistory("*")
dev.off()

#Set working directory back to path0.
setwd(path0)
#-----
#-----

```

#Open and read a model text の部分で、BUGS コードの部分のみを記述したテキストファイルを開くと、その結果は設定したフォルダ内の「Results」フォルダに出力される。

(5) pcnetmeta パッケージによるアームベースモデルによるネットワークメタアナリシス

pcnetmeta パッケージは、R から rjags を介して JAGS を動かしてベイジアンの手法でネットワークメタアナリシスを実行するパッケージである。

```

##########
###pcnetmeta for network meta-analysis###
library(pcnetmeta)

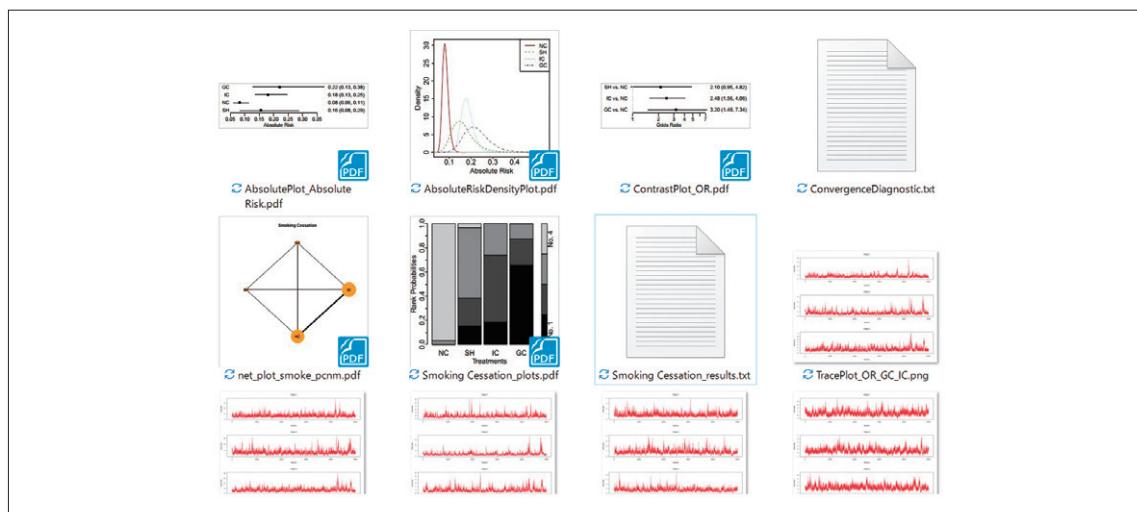
###Setting the working directory.
library("tcltk2")
#Set a folder to save files.
paths=tk_choose.dir(default="",caption="#1 Select a folder to save data files")
paths=paste(paths,"/",sep="")
setwd(paths)

###Use smoke cessation data.
data("smoke")
exdat=smoke
#or ####Read data via clipboard. Remove the head # to use the line.
#exdat=read.delim("clipboard",sep="\t",header=TRUE)

```

```
#Set the names of treatments and the reference treatment name.  
tx.name=c("NC", "SH", "IC", "GC")  
cont="NC"  
  
#Set title and study name, and make a network geometry plot.  
title="Smoking Cessation"  
stnam="smoke_pcnm"  
#nma.networkplot(s.id=s.id, t.id=t.id, data=exdat,trtnname = tx.name, title = title)  
#Save the geometry plot in a pdf file.  
filpdf=paste("net_plot_",stnam,".pdf",sep="")  
pdf(filpdf) #File name to save plots in a PDF file.  
nma.networkplot(s.id=s.id, t.id=t.id, data=exdat,trtnname =tx.name, title = title)  
dev.off()  
  
#---Run MCMC. This process takes time.  
modelnam="het_cor" #The default model for the variance-covariance matrix.  
#modelnam="hom_eqcor"  
  
#Run MCMC: burn-in 10000, MCMC 50000.  
set.seed(12345)  
exdat.out<- nma.ab.bin(s.id, t.id, r, n, data = exdat, trtnname = tx.name,  
param = c("AR", "OR", "RR", "LOR", "LRR", "RD", "rank.prob"),  
model = modelnam, higher.better = TRUE, digits = 3,  
n.adapt = 10000, n.iter = 50000, n.thin = 1, conv.diag = TRUE,  
dic = TRUE, trace = "OR", postdens = TRUE)  
  
#---  
#Save the results as a text file.  
resfilnam=paste(title,"_results.txt",sep="") #Set the file name (xxx.txt)  
sink(resfilnam)  
print(exdat.out$model)  
print(exdat.out$AbsoluteRisk)  
print("Odds Ratio")  
print(exdat.out$OddsRatio)  
print("Risk Ratio")  
print(exdat.out$RelativeRisk)  
print("Risk Difference")  
print(exdat.out$RiskDifference)  
print(exdat.out$TrtRankProb)  
print(exdat.out$DIC)  
sink()  
  
#Save plots as a PDF file.  
filpdf=paste(title,"_plots.pdf",sep="") #File name to save plots in a PDF file.  
pdf(filpdf)  
absolute.plot(exdat.out, width = 5, height = 1.5)  
contrast.plot(exdat.out, reference = cont, width = 5, height = 1.5)  
rank.prob(exdat.out, cex.axis = 2, cex.lab = 2)  
dev.off()  
  
#####
```

結果として以下のファイルが出力される。



(6) netmeta パッケージによる頻度論派の手法によるネットワークメタアナリシス

netmeta パッケージは、アウトカムが二値変数の場合はオッズ比、リスク比、率差を、連続変数の場合は平均値差、ハザード比などを取り扱えるとともに、それらのサンプルデータが提供されている。今回は率差を効果指標とした解析例を示す。

はじめに、R を起動して以下のスクリプトを実行する。

```
library(netmeta)      #netmeta パッケージを読み込む。
data(smokingcessation) # データセット smokingcessation を読み込む。
dat=smokingcessation # データセットの内容を変数 dat に格納する。
```

必要なデータ形式は次のようになる。

	A	B	C	D	E	F	G	H	I	J
1	study	event1	n1	event2	n2	event3	n3	treat1	treat2	treat3
2	1	9	140	23	140	10	138	A	C	D
3	2	11	78	12	85	29	170	B	C	D
4	3	75	731	363	714	NA	NA	A	C	
5	4	2	106	9	205	NA	NA	A	C	
6	5	58	549	237	1561	NA	NA	A	C	
7	6	0	33	9	48	NA	NA	A	C	
8	7	3	100	31	98	NA	NA	A	C	
9	8	1	31	26	95	NA	NA	A	C	
10	9	6	39	17	77	NA	NA	A	C	
11	10	79	702	77	694	NA	NA	A	B	
12	11	18	671	21	535	NA	NA	A	B	

これは netmeta パッケージのデータセットに含まれる smokingcessation のデータの一部であるが、4つの治療法の名称は A, B, C, D である。これらのうち3つを直接比較した研究は1, 2 の2つであり、それ以外の研究では2つの治療法が比較されている。したがって、1つの研究で比較された群は最大で3つあり、それぞれの群の症例数は n1, n2, n3 で表されている。それぞれの群のアウトカム生起症例数は event1, event2, event3 で表されている。2群の比較しか行われていない研究の場合、データなしとしてセルには NA が入力されている。どの治療法が用いられたかは、treat1, treat2, treat3 のカラムに入力されている。こちらは、該当する治療なしの場合は空欄となっている。treat1 の内容は研究によって同じではなく、treat1, treat2, treat3 はそれぞれの研究の群1, 群2, 群3 の受けた治療に対応している。

Excel で用意したデータをクリップボード経由で読み込んで解析する場合には、セル B1 から右下のデータの部分をコピーする（カラム A は不要）。そして、R の側で以下のスクリプトを実行する。R のエディターにこれを書き込み、Excel から必要なデータ範囲を選択してコピーし、R に戻ってこのスクリプトの行にカーソルを置き [実行ボタン] をクリックする。Mac の場合は [Command] + [Enter] キーを押す。Mac の場合は Numbers からクリップボード経由でデータを読み込むこともできる。

```
dat=read.delim("clipboard",sep="\t",header=TRUE)      #Windows の場合
exdat=read.delim(pipe("pbpaste"),sep="\t",header=F);reiv()    #Mac の場合
```

このデータはアウトカムが二値変数で表され、効果指標としてオッズ比、リスク比、リスク差が用いられる場合の例となる。pairwise() 関数の sm 引数に「OR」、「RR」、「RD」 と設定することで、それぞれ対応した解析が行われる。

まず、pairwise() 関数を実行し、結果を変数 p3rd に格納する（変数名は任意）。以下の例では、リスク差を指標としている。次に、p3rd の最初の 6 行を head() 関数で表示させる。

```
p3rd=pairwise(list(treat1,treat2,treat3),list(event1,event2,event3),list(n1,n2,n3),data=dat,sm="RD")
head(p3rd)

> p3rd=pairwise(list(treat1,treat2,treat3),list(event1,event2,event3),list(n$)
> head(p3rd)
   TE      seTE studlab treat1  treat2  event1  n1  event2  n2
1 -0.1000000000 0.03755462     1      A      C      9 140     23 140
2 -0.0081780538 0.03027728     1      A      D      9 140     10 138
3  0.0918219462 0.03831102     1      C      D     23 140     10 138
4 -0.0001508296 0.05458446     2      B      C     11  78     12  85
5 -0.0295625943 0.04883979     2      B      D     11  78     29 170
6 -0.0294117647 0.04752580     2      C      D     12  85     29 170
```

引き続いて、pairwise() 関数の実行結果に対して、netmeta() 関数を実行する。comb.random 引数を TRUE に設定するとランダム効果モデル、FALSE に設定すると固定効果モデルが用いられる。続いて net3r を実行して結果を表示させる。

```
net3r=netmeta(TE,seTE,treat1,treat2,studlab,comb.random=TRUE,data=p3rd)
net3r
```

```
Random effects model

Treatment estimate (sm='RD'):
      A      B      C      D
A   . -0.0412 -0.0944 -0.1291
B   0.0412   . -0.0531 -0.0879
C   0.0944  0.0531   . -0.0347
D   0.1291  0.0879  0.0347   .

Lower 95%-confidence limit:
      A      B      C      D
A   . -0.1294 -0.1438 -0.2237
B  -0.0469   . -0.1456 -0.1956
C   0.0449 -0.0394   . -0.1268
D   0.0345 -0.0199 -0.0573   .

Upper 95%-confidence limit:
      A      B      C      D
A   .  0.0469 -0.0449 -0.0345
B   0.1294   .  0.0394  0.0199
C   0.1438  0.1456   .  0.0573
D   0.2237  0.1956  0.1268   .

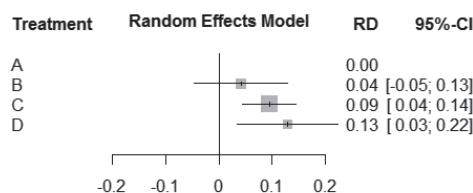
Quantifying heterogeneity/inconsistency:
tau^2 = 0.0092; I^2 = 94%

Test of heterogeneity/inconsistency:
    Q d.f. p.value
 380.61   23 < 0.0001
```

`comb.random=TRUE` の設定では、固定効果モデルの計算も行われるので、両方のモデルの結果がコンソールに出力される。

続いて、`forest()` 関数でフォレストプロットを作成する。この際には、`ref`引数に基準とする治療法の名称を設定する。

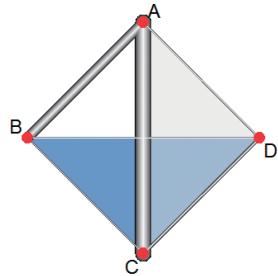
```
dev.new();forest(net3r,ref="A")
```



A を基準に設定したので、A の RD は 0, B, C, D に対しては A を基準とした RD の統合値と 95 %信頼区間が表示される。この結果から治療 A と C および D では差があるが、A と B では差がなく、B, C, D の間でも差がないと判定される。また、A と C の比較が最も信頼区間が狭く、症例数が多く、かつ研究数も多いことが反映されていると考えられる。次に示すネットワークグラフにおいて A と C を結ぶ線が一番太いことと対応している。

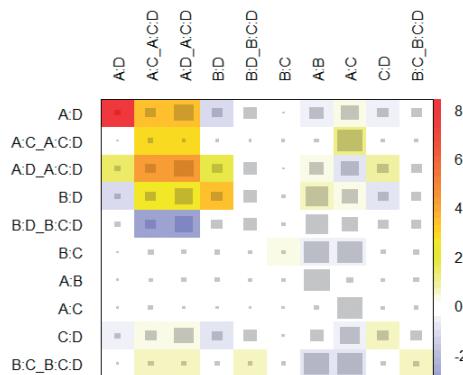
続いて、`netgraph()` 関数でネットワークグラフを作成する。点の大きさは `cex` で指定でき、点を結合する線の太さは標準誤差の逆数に比例させるように設定してある。

```
dev.new();netgraph(net3r,points=TRUE,cex.points=3,cex=1.5,thickness="se.random")
```



次に、ネットヒートプロットを作成する。

```
dev.new();netheat(net3r)
```



ネットヒートプロットは Krahn U ら (2013) が提唱した方法で、直接比較の結果と間接比較の結果の乖離の程度すなわち不一致の程度を表したものである。左側の列に研究で実際に直接比較された治療のペアが示されている。「_」の右側は研究で直接比較された治療を示している。上の行には間接比較も含めたネットワークメタアナリシスで比較された治療のペアを示している。グレーの四角はネットワークメタアナリシスの結果への直接比較の影響 (contribution) を表している。直接比較の結果を除いた場合に不一致の程度が減少するか、増加するかを色で示しており、暖色は減少、寒色は増加することを示し、その程度は色の濃さで表されている。色が濃いほど不一致の程度の変動が大きいことを示す。この例では、治療 A と D の比較は、直接比較を除いてしまうと、不一致の程度が大きく減少することを示しているので、間接比較を含めたネットワークメタアナリシスの結果は異質性が高いことを示している。

最後に、`dnetmeasures()` 関数で、各指標の値をコンソールに出力する。

```
dnetmeasures(net3r)
```

4.9 補足資料 4：システマティックレビューツール

• GRADEpro GDT

診療ガイドライン作成作業全般を支援するウェブツールとして、McMaster University and Evidence Prime Inc. が提供する GRADEpro GDT がある (GRADEpro GDT 2021)。これは GRADE ワーキンググループの Schünemann H らが開発に参加しており、GRADE アプローチに沿った作業をクラウドの機能を利用して行えるように、スコープ作成から推奨作成まで、チーム作業を支援する形でさまざまな機能が提供されている。診療ガイドライン作成グループが利用する場合は有料となっている。個人の使用は無料、2名のチームは \$29 (12CQ まで)、15名までのチームは \$2,000 (25CQ まで)、それ以上の場合は個別に利用料が設定される (2021年2月24日現在)。

• Rayyan

文献検索の結果から選定する作業は時間と労力を要するが、無料のクラウドサービスとして Qatar Computing Research Institute (Data Analytics) が提供している Rayyan (レイヤン) がある (Ouzzani M 2016)。文献リストの作成を共同作業で進めることができる。また、JavaScript で作成された PubMed 検索用ウェブツール pmSearch も利用可能である (pmSearch 2019)。

• RevMan web

現在の RevMan (使用方法は『Minds 診療ガイドライン作成マニュアル 2017』第4章4.6節を参照) はダウンロードしてスタンドアロンのアプリケーションとして動かすことができるが、Cochrane は今後クラウドを利用したウェブツールに変更する計画である。そのため、今後自由に利用できなくなる可能性がある。

• 機械学習を利用したシステマティックレビュー支援ツール

機械学習を利用し、文献選定作業を省力化するツールとして SWIFT-Review (Howard BE 2016, O'Mara-Eves A 2015) が発表されているが、感度を 100% にすることは困難であることが示されている。また、バイアスリスクの評価を自動化することも試みられており、バイアスドメインの評価結果を High, Low の二値変数として扱ったときに、Cochrane の熟練したレビューの評価結果との一致率が 70% 程度となることが示されている [Robot Reviewer (Marshall IJ 2016), robot analysis 4.8.7 (Millard LA 2016)]。

また、PubMed では Best match という機能が提供されており、アブストラクト付き、フルテキストなどに限定するメニューの中から選択できる。Best match では各論文のタイトルとアブストラクトの単語の出現頻度のデータを参照し、検索式の語句の出現頻度が高い論文から順番に表示する。その機能は感度 100% には到達していないため、PubMed 自体はシステマティックレビューの文献検索には用いないことを推奨している。

• PICORON

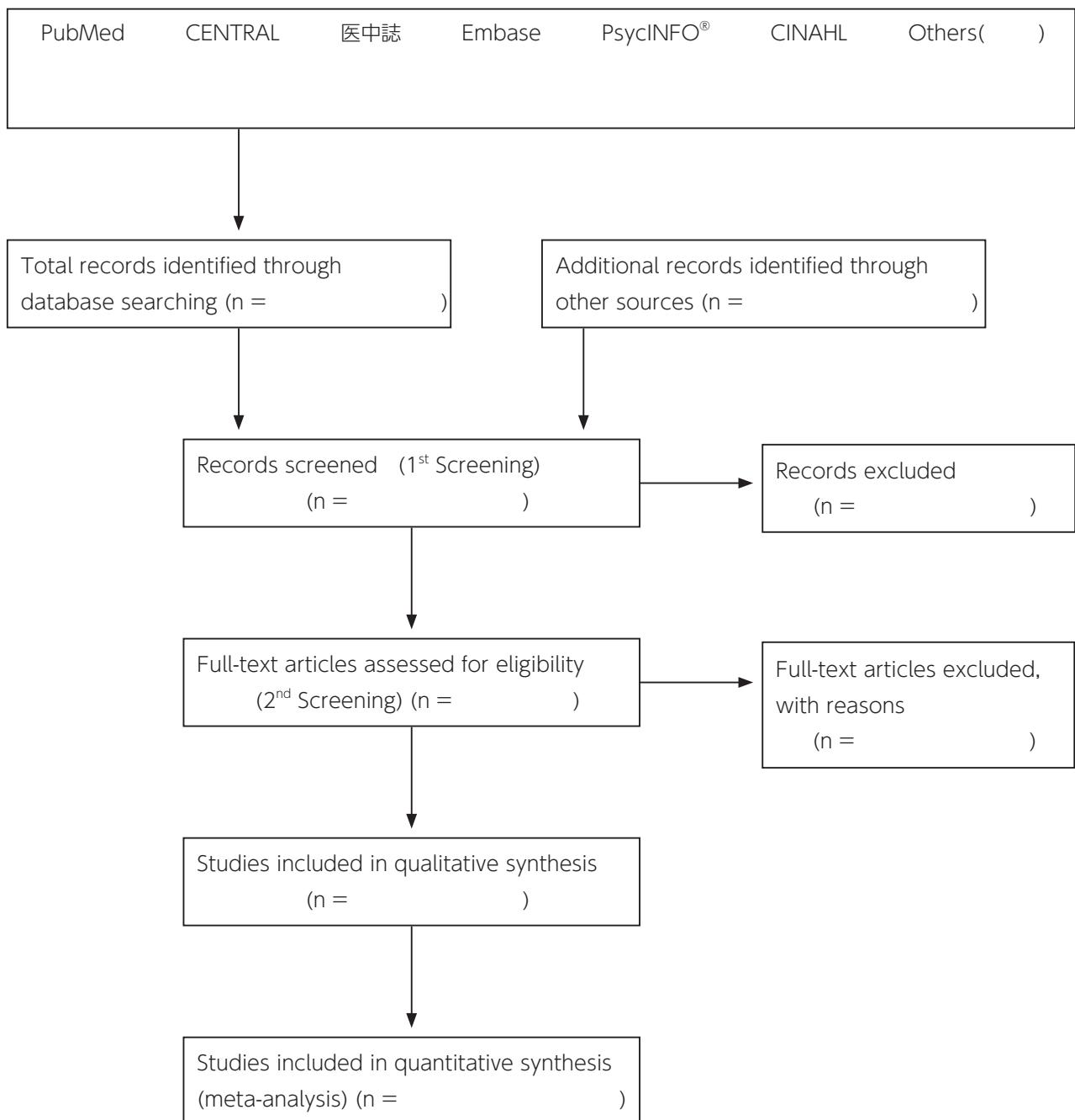
PDF ファイルの論文中で「randomization」、「outcome」などのキーワードをハイライトする機能を持つ。

4章 テンプレート -

[SR-1 データベース検索結果]

タイトル :
CQ :
データベース :
日付 :
検索者 :

【SR-2 文献検索フローチャート】



【SR-3 二次スクリーニング後の一覧表】

【SR-4 引用文献リスト】

	文献 ID	書誌情報
採用論文		
不採用論文		
その他の引用論文		

【SR-5 評価シート 介入研究】

診療ガイド ライン		
対象		
介入		
対照		

*各項目の評価は「高（-2）」、「中（-1）」、「低（0）」の3段階。
まとめは「高（-2）」、「中（-1）」、「低（0）」の3段階でエビデンス総体に反映させる。

アウトカムごとに別紙にまとめる。

コメント（該当するセルに記入）

【SR-6 評価シート 観察研究】

*バイアスリスク、非直接性
各ドメインの評価は「高 (-2)」、「中／疑い (-1)」、「低 (0)」の3段階。
まとめは「高 (-2)」、「中 (-1)」、「低 (0)」の3段階でエビデンス総体に反映させる。

* * 上昇要因 各項目の評価は「高(++)」、「中(+1)」、「低(0)」の3段階。また項目は「高(++)」、「中(+1)」、「低(0)」の3段階でエビデンス総体に反映される。

診療ガイド ライン			
対象			
介入			
対照			

アウトカム	バイアスリスト*
-------	----------

コメント（該当するセルに記入）

【SR-7 評価シート エビデンス総体】

* バイアスリスク、非直接性

各ドメインの評価は「高（-2）」、「中（-1）」、「低（0）」の3段階。まとめは「高（-2）」、「中（-1）」、「低（0）」の3段階でエビデンス総体に反映させる。

* * 上昇要因

診療ガイド ライン			
対象			
介入			
対照			

エビデンス総体

コメント（該当するセルに記入）

[SR-8 評価シート エビデンス総体（絶対効果指標の結果を記入する場合）]

診療ガイド ライン			
対象			
介入			
対照			

リスク人数（アウトカム率）

エビデンスの強さについて、RCTは「強(A)」からスタート、観察研究は「弱(C)」からスタート。

*各ドメインは「高 (-2)」, 「中／疑い (-1)」, 「低 (0)」の3段階。

***工ビデシスの強さは「強(A)」、「中(B)」、「弱(C)」、「非常に弱(D)」の4段階。

* * * 重要性はアドガムの重要性 (1~9)。

【SR-9 定性的システムティックレビュー】

CQ		
P		
I		
C		
臨床的文脈		

01	
非直接性のまとめ	
バイアスリスクのまとめ	
非一貫性その他のまとめ	
コメント	

02	

03	
:	

【SR-10 メタアナリシス】

CQ						
P			I			
C			O			
研究デザイン			文献数			
コード						
モデル		方法				
効果指標		統合値	(-) P=			
フォレスト プロット						
コメント :						
ファンネル プロット						
コメント :						
その他の解析				コメント :		
メタリグ レッション						
感度分析						

【SR-11 システマティックレビューレポートのまとめ】

【SR-12 結果のまとめ (SoF 表) (ペア比較のメタアナリシス)】

重要臨床課題 :						
疾患／対象者 :						
セッティング :						
介入 :						
対照 :						
アウトカム 対象者数 (研究数)	相対効果 (95%信頼区間)	期待される絶対効果* (95%信頼区間)			エビデンス 確実性	何が 起きるか?
		対照	介入	差		
アウトカム 1 : 介入 経過観察期間 :						
対象者 1 対象者数 (研究数)		%	% (~)	% 少ない (~ 少ない)	⊕⊕⊕⊕ 高 Due to	
対象者 1 対象者数 (研究数)		%	% (~)	% 少ない (~ 少ない)	⊕⊕⊕○ 中 Due to	
アウトカム 2 : 経過観察期間 : 対象者数 (研究数)		%	% (~)	% 少ない (~ 少ない)	⊕⊕○○ 低 Due to	
アウトカム 3 : 経過観察期間 : 対象者数 (研究数)		%	% (~)	% 少ない (~ 少ない)	⊕○○○ 非常に低 Due to	
アウトカム 4 : 経過観察期間 : 対象者数 (研究数)		%	% (~)	% 少ない (~ 少ない)	⊕○○○ 非常に低 Due to	
* **						
解説						
1 2 3 4 5 6 7 8 9 10 11 12						

出典 : Guyatt GH, et al. (2012) GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. J Clin Epidemiol 66:158-172.
By courtesy of Dr. Guyatt より作成

[SR-13 結果のまとめ (SoF表) (ネットワークメタアナリシス : 様式 1)]

一番左の列には Geometry plot への参照を含める（ノードの色など）。

【SR-14 結果のまとめ (SoF 表) (ネットワークメタアナリシス : 様式 2)】

益/害													
患者/参加者 (P) :													
介入 (I) :													
対照 (参照) (C) :	ネットワークメタアナリシス Geometry plot*												
アウトカム (O) :													
セッティング (S) :													
アウトカム	効果推定値と確信区間												
	治療 1	治療 2	治療 3	コメント									
アウトカム 1													
参考治療 人／人 (%)***	ORなど相対 効果 (95%CrI) ネットワーク 推定値	絶対効果 (95%CrI)	ORなど相対 効果 (95%CrI) ネットワーク 推定値	絶対効果 (95%CrI)	ORなど相対 効果 (95%CrI) ネットワーク 推定値	絶対効果 (95%CrI)							
	効果推定値に対する確信度	高 +++++		中 ++++−									
順位 (95%CrI)	順位(95%CrI)	3(2~9)		3(1~10)									
	参加者総数(研究数)	823(3RCT)		427(1RCT)									
アウトカム 1													
参考治療 人／人 (%)	ORなど相対 効果 (95%CrI) ネットワーク 推定値	絶対効果 (95%CrI)	ORなど相対 効果 (95%CrI) ネットワーク 推定値	絶対効果 (95%CrI)	ORなど相対 効果 (95%CrI) ネットワーク 推定値	絶対効果 (95%CrI)							
	効果推定値に対する確信度	効果推定値に対する確信度		効果推定値に対する確信度									
順位 (95%CrI)	順位(95%CrI)	順位(95%CrI)		順位(95%CrI)									
	参加者総数(研究数)	参加者総数(研究数)		参加者総数(研究数)									
NMA-SoF 表の定義													
* 実線は直接比較													
** 推定値はオッズ比。CrI は確信区間 (ベイジアンアプローチによる)。													
*** 想定されるリスク (例えば、研究横断的な対照群の中央値) は脚注に記載。													
CrI : Credible Interval, OR : Odds Ratio													
GRADE によるエビデンスのグレーディング													
高い質：将来の研究が効果推定値に対する我々の確信度を変える可能性は非常に低い。													
中等度の質：将来の研究が効果推定値に対する我々の確信度へ重要な影響を持つ可能性があり、推定値を変えるかもしれない。													
低い質：将来の研究が我々の効果推定値に対する確信度へ重要な影響を持つ可能性がかなりあり、推定値を変える可能性がある。													
非常に低い質：推定値に対して我々は非常に不確かである。													
解説脚注：													
1 ベースラインリスク (想定される対照群のリスク) は National Cancer Institute の集約プログラムから得た。													
2 95% 確信区間が同等の値と交差するので非常に重篤な不精確性があり、広い確信区間から害の可能性が高いことを示す。													
3 概念的にはあり得る修飾因子分布は異なる化学療法薬を用いる研究間で差がないので、明らかな非移行性はない。													
4 順位は中央値 (順位 1 ~ 10) と 95% 確信区間を示す。													

一番左の列には Geometry plot への参照を含める (ノードの色など)。

[SR-15 Future Research Question]

4章 テンプレート 記入方法

[SR-1 データベース検索結果 記入方法]

タイトル：混同のないよう全てのテンプレートにタイトルを記載する。

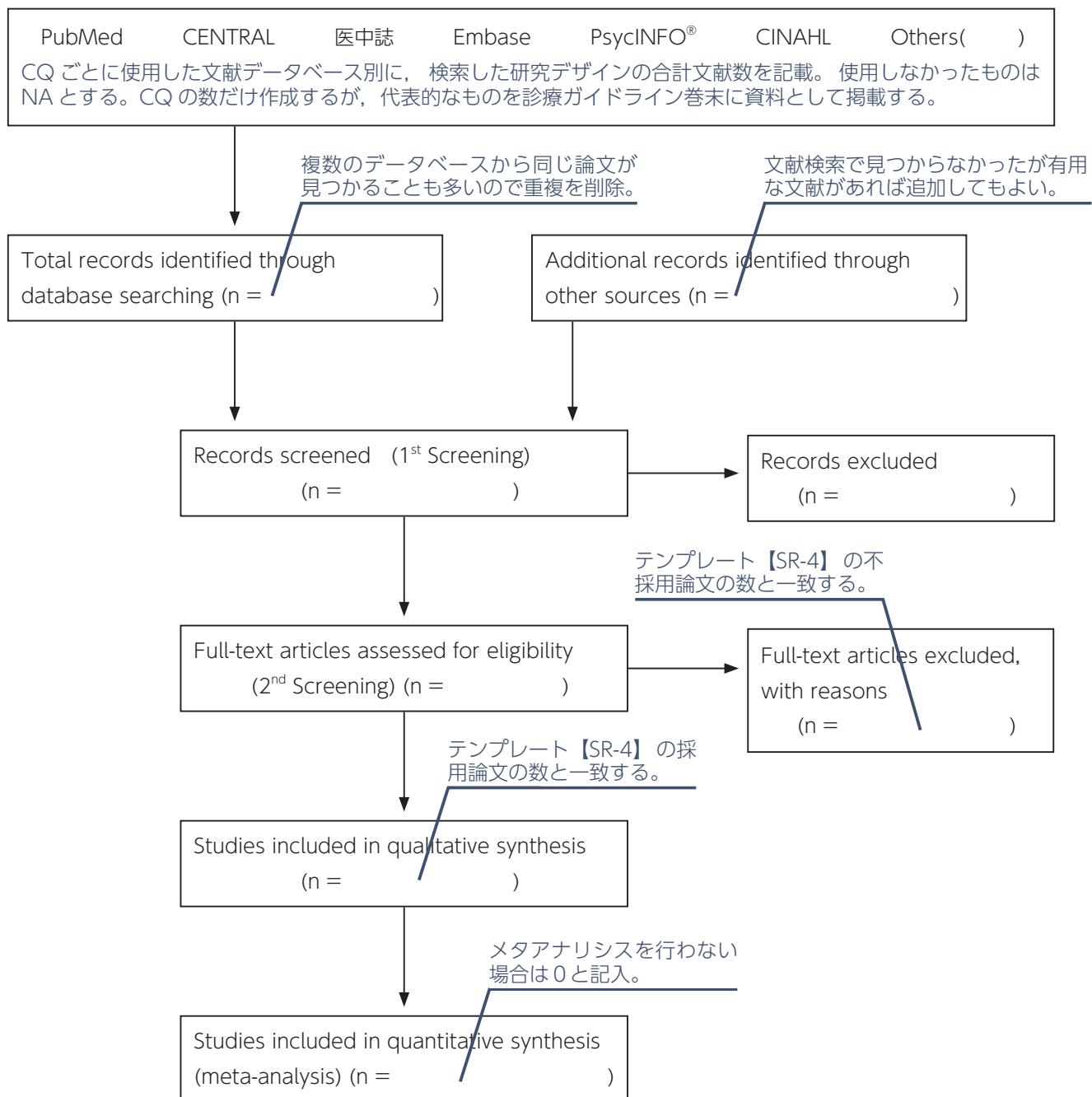
CQ：混同のないよう全てのテンプレートにCQ番号とCQ文を記載する。

データベース：データベース名とその検索期間を記載する。

日付：検索日を記載する。

検索者：2名の検索者を記載する。公開時には削除または、イニシャルに変更すること。

[SR-2 文献検索フローチャート 記入方法]



【SR-3 二次スクリーニング後の一覧表 記入方法】

【SR-4 引用文献リスト 記入方法】

	文献 ID	書誌情報
採用論文	ハーバード方式による文献整理 ID で、本文と対応させる。	テンプレート【SR-3】に残った、システムティックレビューに使用した論文の書誌情報を記載する。
不採用論文	ハーバード方式による文献整理 ID で、本文と対応させる。	テンプレート【SR-3】から、その後のステップで削除された論文の書誌情報を記載する。
その他の引用論文	ハーバード方式による文献整理 ID で、本文と対応させる。	総論や疫学など、システムティックレビューを行わない章において引用した論文の書誌情報を記載する。

【SR-9 定性的システムティックレビュー 記入方法】

CQ	管理番号	CQ の文章を記入する。
P		CQ の文章を補足する十分に詳細な記述をする。タイミングやセッティングも必要に応じて記述する。
I		介入の詳細を記述する。
C		1：1 の比較が望ましいが、必要に応じて複数の比較を記述する。
臨床的文脈		診療のプロセスのどこに位置付けられるか、診断、治療、予防、予後予測、その他のいずれに分類されるかなどを記述する。

01	害のように CQ の文章に表れない場合もあるが、エビデンス総体の対象が分かるように記述する。
非直接性のまとめ	CQ と得られたエビデンス総体の間の乖離について記述する。介入の比較が非直接的である場合や日本人患者への適用に問題がある場合には必ず記述する。
バイアスリスクのまとめ	特に問題となるバイアスリスク、推奨の決定に影響を及ぼすバイアスリスクを記述する。
非一貫性その他のまとめ	複数の研究間の相違の大きさ、信頼区間やバイアスリスクから判定した不確実性、効果の大きさなどについてまとめを記述する。研究デザインが異なるものをまとめた場合はそれについて記述する。
コメント	このアウトカムに関するエビデンス総体の評価において特に注意すべき点があればそれを記述する。

02	
----	--

03	
----	--

⋮

【SR-11 システマティックレビューレポートのまとめ 記入方法】

テンプレート【SR-9】、【SR-10】の結果をまとめる。

[SR-15 Future Research Question 記入方法]

現在の研究では解明されていない課題から、将来の研究で回答が得られるであろう CQ を作成する。以下の項目を記述する。

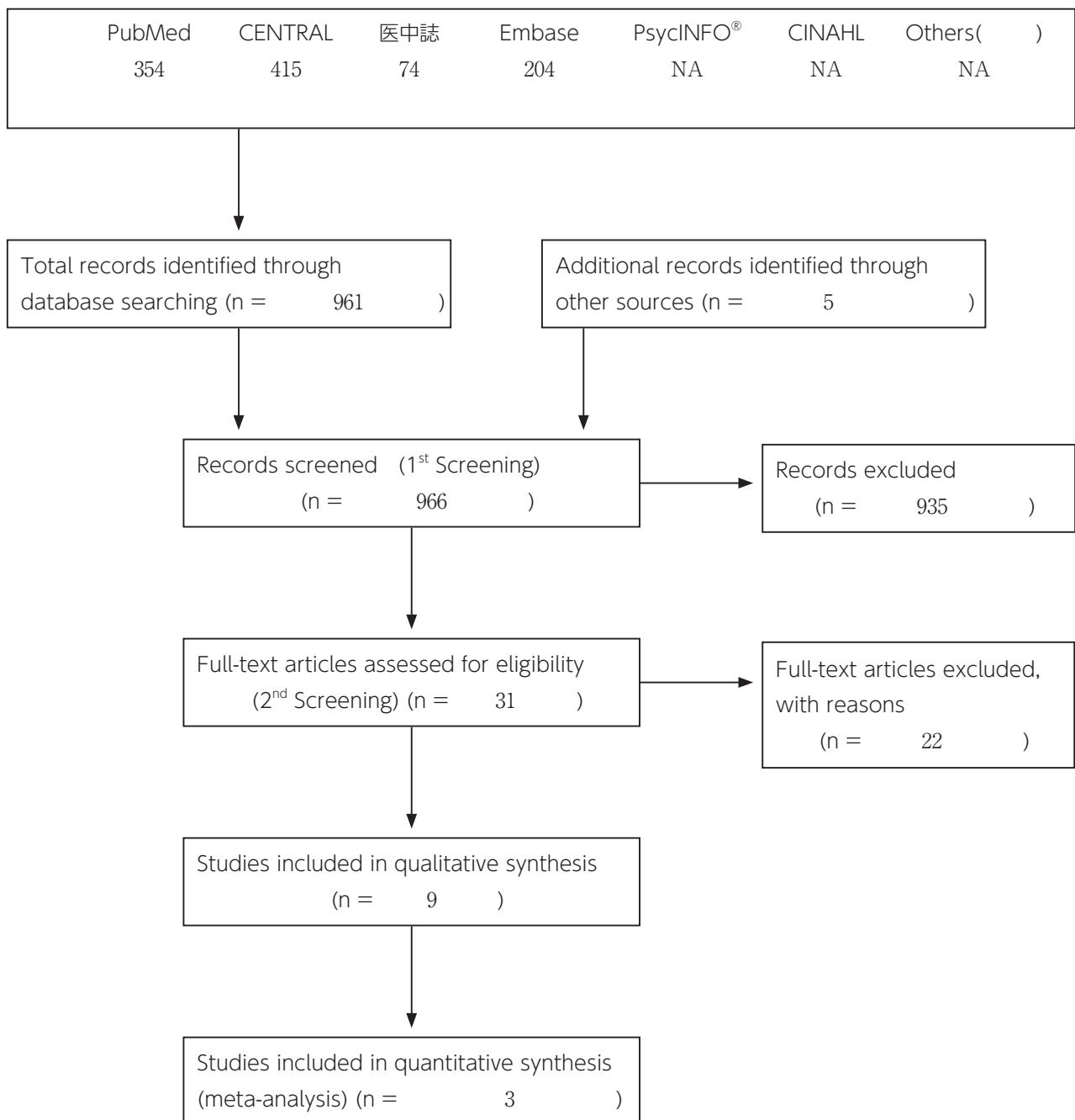
- ・リサーチクエスチョン
- ・背景
- ・可能な研究計画の概略

4章 テンプレート 記入例

[SR-1 データベース検索結果 記入例]

タイトル : Stroke and t-PA
CQ : 発症 6 時間以内の 80 歳未満の急性脳塞栓患者に対して発症 3 時間以降の t-PA 静脈投与は推奨されるか。
データベース : The Cochrane Library
日付 : 2019/9/21
検索者 : NK / FK

【SR-2 文献検索フローチャート 記入例】



【SR-3 二次スクリーニング後の一覧表 記入例】

【SR-4 引用文献リスト 記入例】

	文献 ID	書誌情報
採用論文	The IST-3 collaborative group, 2012	The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [IST-3]): a randomised controlled trial. Lancet. 2012. 379(9834):2352-2363.
不採用論文		
その他の引用論文		

【SR-5 評価シート 介入研究 記入例】

診療ガイド ライン	虚血性脳卒中にに対するアルテープ ラー
対象	虚血性脳卒中（79歳までの成人）
介入	6時間以内のアルテープラーゼ 0.6mg/kg 静注
対照	プラセボ

*各項目の評価は「高（-2）」、「中／疑い（-1）」、「低（0）」の3段階。
まとめは「高（-2）」、「中（-1）」、「低（0）」の3段階でエビデンス総体に反映させる。

メシト（該当するヤツルに記入）

【SR-6 評価シート 観察研究 記入例】

診療ガイドライン	自己免疫性肝炎
対象	自己免疫性肝炎
介入	PSL 単独またはAZP併用療法
対照	無治療

アウトカム		生化学的・組織学的窓解			
		バイアスリスク			
個別研究選択バイアス	実行バイアス	検出バイアス	症例減少バイアス	その他	
背景因子の差	ケアの差	ア不適切なアウトカム測定	バイアスの調整	その他のバイアス	

研究コード	研究デザイン	上昇要因**				非直接性*				リスク人數（アウトカム率）					
		まとめ	量反応関係	効果減弱交絡	効果の大ささ	対象	介入	対照	アウトカム	まとめ	対照群分子	介入群分子	介入群分子 (%)	効果指標（種類）	効果指標（値）
Czaja 2008 コホート研究	-1	-1	0	-1	-2	0	0	0	-2	-2	0	-2	87.5	RR	0.52
Czaja 2008 コホート研究	-2	-1	0	-1	-2	0	0	0	-2	-2	0	-2	NA	NA	NA
Al-Chalabi 2008 コホート研究	-2	-1	0	-1	-2	0	0	0	-1	-2	0	-2	NA	NA	NA
Feld JJ 2005	-2	-1	0	-1	-2	0	0	0	-1	-2	0	-2	NA	NA	NA

コメント（該当するセルに記入）

後ろ向き 群集研究	群集例でマッチさせている 同一群の大きいとみなす 群集研究の群 一群	組織学的 窓われる 窓われる 同一群の 大きいとみな す	病理評価 あるが アリで測定 て一方と えられる	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	アザチオ ブリジング は使用され ない	アザチオ ブリジング が国では使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない
後ろ向き 群集研究	同一群の 大きいと みなす 群集研究の 群 一群	窓われる 同一群の 大きいと みなす 群集研究の 群 一群	病理評 価あるが アリで測定 て一方と えられる	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない
後ろ向き コホート 研究の群 一群	同一群の 大きいと みなす 群集研究の 群 一群	窓われる 同一群の 大きいと みなす 群集研究の 群 一群	病理評 価あるが アリで測定 て一方と えられる	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない
後ろ向き コホート 研究の群 一群	同一群の 大きいと みなす 群集研究の 群 一群	窓われる 同一群の 大きいと みなす 群集研究の 群 一群	病理評 価あるが アリで測定 て一方と えられる	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	フォローアップ間 隔が不定 期	多量の腫 瘍所では未施 行	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない	アザチオ ブリジング は日本で使 えない

【SR-7 評価シート エビデンス総体 記入例（介入研究）】

診療ガイドライン	虚血性脳卒中に対するアルテプラーゼ
対象	虚血性脳卒中（79歳までの成人）
介入	6時間以内のアルテプラーゼ0.6mg/kg 静注
対照	プラセボ

- *バイアスリスク、非直接受
各ドメインの評価は「高（-2）」、「中／疑い（-1）」、「低（0）」の3段階。
まとめは「高（-2）」、「中（-1）」、「低（0）」の3段階。
- **上昇要因
各項目の評価は「高（+2）」、「中（+1）」、「低（0）」の3段階。
まとめは「高（+2）」、「中（+1）」、「低（0）」の3段階。
- アウトカムごとに別紙にまとめる。
アウトカムごとに別紙にまとめる。

エビデンス総体

アウトカム	研究デザイン／ 数	バイアスリスク	非一貫性	不正確性	非直接性	その他のバイアスなど	リスク人数（アウトカム率）					
							対照群分子	介入群分子	介入群分子 (%)	介入群分子 (%)	効果指標（種類）	効果指標統合値
6ヶ月後の総死亡	RCT ／3	0	0	0	-1	0	1960	456	0.223	1.976	464	0.23
6ヶ月後の依存性 (mRS0-2)	RCT ／3	0	0	0	-1	0	1960	1226	0.626	1976	1176	0.6
1週間以内の脳 出血	RCT ／3	0	0	0	-1	0	1961	29.5	0.015	1977	144.5	0.07

総死亡



依存性



頭蓋内出血



【SR-7 評価シート エビデンス総体 記入例（観察研究）】

診療ガイド ライン	自己免疫性肝炎
対象	自己免疫性肝炎
介入	PSL+AZA または PSL 単独
対照	無治療

* バイアスリスク、非直接性

各メインの評価は「高 (-2)」、「中 (-1)」、「低 (0)」の3段階。また各メインの評価は「高 (-2)」、「中 (-1)」、「低 (0)」の3段階。

* * 上引要因

各項目の評価は「高(+2)」「中(+1)」「低(0)」の3段階

まとめは「高(+2)」、「中(+)」
「低(-2)」に別紙にまとめる。

エビデンス総体

コメント（該当するセルに記入）

Czaja 2008 の研究 の介入 は1つ にまと める								軽症例 のみの 値	

【SR-10 メタアナリシス 記入例】

CQ		成人の急性虫垂炎において抗菌薬投与による保存的治療は外科的虫垂切除術と比べ1年以内の再発を伴わない治癒が改善するか？																																																																																																																		
P	成人の急性虫垂炎				I	抗菌薬投与による保存的治療																																																																																																														
C	外科手術				O	1カ月以降1年以内の再発																																																																																																														
研究デザイン	RCT				文献数	4																																																																																																														
コード	Salminen P 2015 Vons C 2011 Styrud J 2006 Eriksson S 1995																																																																																																																			
モデル	ランダム効果モデル	方法	分散逆数法																																																																																																																	
効果指標	リスク比 (RR)	統合値	41.399 (10.260-167.05) P=<0.0001																																																																																																																	
効果指標	リスク差 (RD)	統合値	0.221 (0.14-0.303) P=<0.0001																																																																																																																	
フォレストプロット	<p>Antibiotics vs Surgery</p> <table border="1"> <thead> <tr> <th>Author Year</th> <th>Number</th> <th>Recurrence</th> <th>Number</th> <th>Recurrence</th> <th>Risk Ratio</th> <th>95% CI lower</th> <th>95% CI upper</th> <th>Weight(%)</th> </tr> </thead> <tbody> <tr> <td>Salminen P 2015</td> <td>227</td> <td>55</td> <td>215</td> <td>0</td> <td>105.158</td> <td>6.535</td> <td>1691.808</td> <td>25.2</td> </tr> <tr> <td>Vons C 2011</td> <td>111</td> <td>30</td> <td>104</td> <td>0</td> <td>57.187</td> <td>3.542</td> <td>923.443</td> <td>25.2</td> </tr> <tr> <td>Styrud J 2006</td> <td>128</td> <td>16</td> <td>124</td> <td>0</td> <td>31.977</td> <td>1.939</td> <td>527.304</td> <td>24.8</td> </tr> <tr> <td>Eriksson S 1995</td> <td>20</td> <td>7</td> <td>20</td> <td>0</td> <td>15.000</td> <td>0.914</td> <td>245.210</td> <td>24.9</td> </tr> <tr> <td>Summary</td> <td></td> <td></td> <td></td> <td></td> <td>41.399</td> <td>10.260</td> <td>167.050</td> <td>100</td> </tr> </tbody> </table> <p>Z=0.0% I²=0.0000 Q=1.0 p=0.7958</p> <p>Z=5.23 p<0.0001</p> <p>1.0 4.0 Risk Ratio</p> <table border="1"> <thead> <tr> <th>Author Year</th> <th>Number</th> <th>Recurrence</th> <th>Number</th> <th>Recurrence</th> <th>Risk Diff.</th> <th>95% CI lower</th> <th>95% CI upper</th> <th>Weight(%)</th> </tr> </thead> <tbody> <tr> <td>Salminen P 2015</td> <td>227</td> <td>55</td> <td>215</td> <td>0</td> <td>0.241</td> <td>0.195</td> <td>0.297</td> <td>31.6</td> </tr> <tr> <td>Vons C 2011</td> <td>111</td> <td>30</td> <td>104</td> <td>0</td> <td>0.268</td> <td>0.184</td> <td>0.351</td> <td>26.8</td> </tr> <tr> <td>Styrud J 2006</td> <td>128</td> <td>16</td> <td>124</td> <td>0</td> <td>0.124</td> <td>0.005</td> <td>0.183</td> <td>31.2</td> </tr> <tr> <td>Eriksson S 1995</td> <td>20</td> <td>7</td> <td>20</td> <td>0</td> <td>0.333</td> <td>0.118</td> <td>0.548</td> <td>10.4</td> </tr> <tr> <td>Summary</td> <td></td> <td></td> <td></td> <td></td> <td>0.221</td> <td>0.140</td> <td>0.303</td> <td>100</td> </tr> </tbody> </table> <p>Z=75.2% I²=0.0047 Q=12.5 p=0.0000</p> <p>0.075 0.2 0.3 0.4 0.5 Risk Diff.</p>	Author Year	Number	Recurrence	Number	Recurrence	Risk Ratio	95% CI lower	95% CI upper	Weight(%)	Salminen P 2015	227	55	215	0	105.158	6.535	1691.808	25.2	Vons C 2011	111	30	104	0	57.187	3.542	923.443	25.2	Styrud J 2006	128	16	124	0	31.977	1.939	527.304	24.8	Eriksson S 1995	20	7	20	0	15.000	0.914	245.210	24.9	Summary					41.399	10.260	167.050	100	Author Year	Number	Recurrence	Number	Recurrence	Risk Diff.	95% CI lower	95% CI upper	Weight(%)	Salminen P 2015	227	55	215	0	0.241	0.195	0.297	31.6	Vons C 2011	111	30	104	0	0.268	0.184	0.351	26.8	Styrud J 2006	128	16	124	0	0.124	0.005	0.183	31.2	Eriksson S 1995	20	7	20	0	0.333	0.118	0.548	10.4	Summary					0.221	0.140	0.303	100							
Author Year	Number	Recurrence	Number	Recurrence	Risk Ratio	95% CI lower	95% CI upper	Weight(%)																																																																																																												
Salminen P 2015	227	55	215	0	105.158	6.535	1691.808	25.2																																																																																																												
Vons C 2011	111	30	104	0	57.187	3.542	923.443	25.2																																																																																																												
Styrud J 2006	128	16	124	0	31.977	1.939	527.304	24.8																																																																																																												
Eriksson S 1995	20	7	20	0	15.000	0.914	245.210	24.9																																																																																																												
Summary					41.399	10.260	167.050	100																																																																																																												
Author Year	Number	Recurrence	Number	Recurrence	Risk Diff.	95% CI lower	95% CI upper	Weight(%)																																																																																																												
Salminen P 2015	227	55	215	0	0.241	0.195	0.297	31.6																																																																																																												
Vons C 2011	111	30	104	0	0.268	0.184	0.351	26.8																																																																																																												
Styrud J 2006	128	16	124	0	0.124	0.005	0.183	31.2																																																																																																												
Eriksson S 1995	20	7	20	0	0.333	0.118	0.548	10.4																																																																																																												
Summary					0.221	0.140	0.303	100																																																																																																												
コメント：																																																																																																																				
ファンネルプロット	<p>Begg's test: Kendall's tau=-0.067 p=0.33333</p> <p>Egger's test: t=2.190 p=0.15902</p> <p>Begg's test: Kendall's tau=-0.333 p=0.75</p> <p>Egger's test: t=0.737 p=0.53708</p>																																																																																																																			
コメント：ファンネルプロットからは出版バイアスは見られない。																																																																																																																				
その他の解析							コメント：																																																																																																													
メタリグレッション																																																																																																																				
感度分析																																																																																																																				

【SR-12 結果のまとめ (SoF 表) (ペア比較のメタアナリシス) 記入例】

重要臨床課題：成人の急性虫垂炎において抗菌薬投与による保存的治療は外科的虫垂切除術と比べ推奨できるか？						
疾患／対象者：18歳以上の成人で、急性虫垂炎で穿孔、膿瘍形成などの合併症をともなわないもの						
セッティング：医療体制の確立した地域						
介入：アモキシシリン+クラブラン酸の抗菌薬投与による保存的治療						
アウトカム 対象者数 (研究数)	相対効果 (95%信頼区間)	期待される絶対効果* (95%信頼区間)			エビデンス 確実性	何が 起きるか？
		対照	介入	差		
虫垂炎の再発 1年 949名(4件)	リスク比 41.39 (10.26 ~ 167.05)	100.0%	77.80%	22.1% 多い (14.0 ~ 30.3% 多い)	⊕⊕⊕⊕ 高 バイアスリスク、非直接性が認められる1, 2が絶対リスク減少は0.14 ~ 0.303で、臨床的閾値を0.05に設定した場合確実性は高い	抗菌薬投与は治癒率が低く、1ヵ月以内に手術を受けないで済んでも、1年以内に再発して手術を受けることになる例が14 ~ 30%程度ある。
虫垂切除術 (1ヵ月以内) 1年 1061名(4件)	リスク比 0.09 (0.061 ~ 0.143)	99.8%	8.60%	90.4% 多い (86.3% ~ 94.4% 少ない)	⊕⊕⊕⊖ 中 バイアスリスク、非直接性があるが1, 2閾値を85%少ないと設定すると確実性は中等度と考える。	抗菌薬投与で最初の1ヵ月間手術を受けないで済む者が90%程度いる。
主要な合併症 1年 969名(4件)	リスク比 0.58 (0.355 ~ 0.955)	8.0%	5.00%	2.5% 少ない (6.3% 少ない ~ 1.2% 多い)	⊕⊕⊖⊖ 低 バイアスリスク、非直接性があり1, 2閾値2%少ない~2%多いに設定すると、範囲から外れる確率がかなり高く、確実性は低いと考える。	各割り付け群全体でみると、抗菌薬投与群のほうが手術を受けないで済む者が多いので、合併症は少なくなる。
その他の合併症 1年 949名(4件)	リスク比 0.22 (0.088 ~ 0.540)	13.0%	2.00%	9% 少ない (17.2% 少ない ~ 2% 少ない)	⊕⊕⊕⊖ 中 バイアスリスク、非直接性があるが1, 2閾値を2%少ないに設定すると確実性は中等度と考える。	各割り付け群全体でみると、抗菌薬投与群のほうが手術を受けないで済む者が多いので、合併症は少なくなる。
入院期間 0 ~ 14日 1050名(4件)	—	平均値 2.83日	平均値 3.23日	0.39日長い (0.25日延長 ~ 0.59日延長)	⊕⊕⊕⊖ 中 バイアスリスク、非直接性1, 2が認められるが、1日以上の延長の可能性は低い。	抗菌薬投与の方が入院期間がわずかに長くなる。
病休期間 0 ~ 14日 1017名(3件)	—	平均値 13.8日	平均値 9.043日	3.6日短い (7.93日短縮 ~ 0.74日延長)	⊕⊕⊖⊖ 低 バイアスリスク、非直接性、非一貫性1, 2, 3が認められる。	抗菌薬投与の方が病休時間が短くなる可能性がある
* 略語：CI (信頼区間)						
解説						
<p>1 ランダム割り付け、コンシールメントに問題がある研究が多く、また、手術の施行に伴う実行・検出の盲検化ができないため、バイアスリスクに問題がある。</p> <p>2 合併症を伴う症例が多く含まれているために、古い研究では診断法に問題があり、日本の現状と比べ、より進行した例が多く対象に含まれている可能性が高いため、非直接性に問題がある研究が多い。抗菌薬の種類も研究により異なる。</p> <p>3 Forest plotで信頼区間の重ならない研究があり、I^2統計量が大きく非一貫性に問題がある。</p>						

4章 文献

- Aberegg SK, et al. (2010) Delta inflation: a bias in the design of randomized controlled trials in critical care medicine. Crit Care 14:R77. doi: [10.1186/cc8990](https://doi.org/10.1186/cc8990). PMID: [20429873](#)
- Agency for Healthcare Research and Quality (AHRQ). (2012) Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(12)-EHC063-EF. <https://effectivehealthcare.ahrq.gov/> (参照 2021/02/24)
- Als-Nielsen B, et al. (2003) Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 290:921-928. doi: [10.1001/jama.290.7.921](https://doi.org/10.1001/jama.290.7.921). PMID: [12928469](#)
- Arber M, et al. (2017) Which information sources should be used to identify studies for systematic reviews of economic evaluations in healthcare? Value in Health 20:PA738.
- Armijo-Olivo S, et al. (2015) What is the influence of randomisation sequence generation and allocation concealment on treatment effects of physical therapy trials? A meta-epidemiological study. BMJ Open 5:e008562. doi: [10.1136/bmjopen-2015-008562](https://doi.org/10.1136/bmjopen-2015-008562). PMID: [26338841](#)
- Armijo-Olivo S, et al. (2017) Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. Am J Phys Med Rehabil 96:34-44. doi: [10.1097/PHM.0000000000000521](https://doi.org/10.1097/PHM.0000000000000521). PMID: [27149591](#)
- Atkins D, et al. (2011) Assessing the Applicability of Studies When Comparing Medical Interventions. Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews. AHRQ Publication No. 11-EHC019-EF. <https://effectivehealthcare.ahrq.gov/> (参照 2021/02/24)
- Bafeta A, et al. (2012) Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. BMJ 344:e813. doi: [10.1136/bmj.e813](https://doi.org/10.1136/bmj.e813). PMID: [22334559](#)
- Balshem H, et al. (2011) GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 64:401-406. doi: [10.1016/j.jclinepi.2010.07.015](https://doi.org/10.1016/j.jclinepi.2010.07.015). PMID: [21208779](#)
- Bassler D, et al. (2013) Reflections on meta-analyses involving trials stopped early for benefit: is there a problem and if so, what is it? Stat Methods Med Res 22:159-168. doi: [10.1177/0962280211432211](https://doi.org/10.1177/0962280211432211). Epub 2011 Dec 13. PMID: [22170891](#)
- Berkman ND, et al. (2014) The Empirical Evidence of Bias in Trials Measuring Treatment Differences. Agency for Healthcare Research and Quality. PMID: [25392898](#)
- Bialy L, et al. (2014) A meta-epidemiological study to examine the association between bias and treatment effects in neonatal trials. Evid Based Child Health 9:1052-1059. doi: [10.1002/ebch.1985](https://doi.org/10.1002/ebch.1985). PMID: [25504975](#)
- Brignardello-Petersen R, et al. (2019) GRADE approach to rate the certainty from a network meta-analysis: addressing incoherence. J Clin Epidemiol 108:77-85. doi: [10.1016/j.jclinepi.2018.11.025](https://doi.org/10.1016/j.jclinepi.2018.11.025). Epub 2018 Dec 5. PMID: [30529648](#)
- Britton A, et al. (1999) Threats to applicability of randomised trials: exclusions and selective participation.

- J Health Serv Res Policy 4:112-121. doi: [10.1177/135581969900400210](https://doi.org/10.1177/135581969900400210). PMID: [10387403](https://pubmed.ncbi.nlm.nih.gov/10387403/)
- Cancer Research and Biostatistics (CRAB). Statistical Tools. <https://stattools.crab.org/> (参照 2021/02/24)
- Carrasco-Labra A, et al. (2015) Comparison between the standard and a new alternative format of the Summary-of-Findings tables in Cochrane review users: study protocol for a randomized controlled trial. Trials 16:164. doi: [10.1186/s13063-015-0649-6](https://doi.org/10.1186/s13063-015-0649-6). PMID: [25873338](https://pubmed.ncbi.nlm.nih.gov/25873338/)
- Carrasco-Labra A, et al. (2016) Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. J Clin Epidemiol 74:7-18. doi: [10.1016/j.jclinepi.2015.12.007](https://doi.org/10.1016/j.jclinepi.2015.12.007). Epub 2016 Jan 11. PMID: [26791430](https://pubmed.ncbi.nlm.nih.gov/26791430/)
- Centre for Clinical Research and Biostatistics (CCRB). (2017) Sample Size Calculation. <http://www2.ccrb.cuhk.edu.hk/web/> (参照 2021/02/24)
- Chalmers TC, et al. (1983) Bias in treatment assignment in controlled clinical trials. N Engl J Med 309:1358-1361. doi: [10.1056/NEJM198312013092204](https://doi.org/10.1056/NEJM198312013092204). PMID: [6633598](https://pubmed.ncbi.nlm.nih.gov/6633598/)
- Chan AW, et al. (2004a) Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 291:2457-2465. doi: [10.1001/jama.291.20.2457](https://doi.org/10.1001/jama.291.20.2457). PMID: [15161896](https://pubmed.ncbi.nlm.nih.gov/15161896/)
- Chan AW, et al. (2004b) Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 171:735-740. doi: [10.1503/cmaj.1041086](https://doi.org/10.1503/cmaj.1041086). PMID: [15451835](https://pubmed.ncbi.nlm.nih.gov/15451835/)
- Chang SM, et al. (2012) Methods Guide for Medical Test Reviews [Internet]. Agency for Healthcare Research and Quality (US) PMID: 22834019. <https://www.ncbi.nlm.nih.gov/books/NBK98241/> (参照 2021/02/24)
- Chess LE, et al. (2013) Risk of bias of randomized controlled trials published in orthopaedic journals. BMC Med Res Methodol 13:76. doi: [10.1186/1471-2288-13-76](https://doi.org/10.1186/1471-2288-13-76). PMID: [23758875](https://pubmed.ncbi.nlm.nih.gov/23758875/)
- Clark T, et al. (2013) Sample size determinations in original research protocols for randomised clinical trials submitted to UK research ethics committees: review. BMJ 346:f1135. doi: [10.1136/bmj.f1135](https://doi.org/10.1136/bmj.f1135). PMID: [23518273](https://pubmed.ncbi.nlm.nih.gov/23518273/)
- Dechartres A, et al. (2011) Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. J Clin Epidemiol 64:136-144. doi: [10.1016/j.jclinepi.2010.04.015](https://doi.org/10.1016/j.jclinepi.2010.04.015). PMID: [20705426](https://pubmed.ncbi.nlm.nih.gov/20705426/)
- Deeks JJ, et al. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration. <http://methods.cochrane.org/sdt/handbook-dta-reviews> (参照 2021/02/24)
- Dias S, et al. TSDs completed or currently in progress. <http://nicedsu.org.uk/technical-support-documents/technical-support-documents/> (参照 2021/02/24)
- Dias S, et al. (2010) Checking consistency in mixed treatment comparison meta-analysis. Stat Med 29:932-944. doi: [10.1002/sim.3767](https://doi.org/10.1002/sim.3767). PMID: [20213715](https://pubmed.ncbi.nlm.nih.gov/20213715/)
- Dias S, et al. (2011) NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. <http://www.nicedsu.org.uk> (参照 2021/02/24)

- Dias S, et al. (2018) Network Meta-Analysis for Decision Making. John Wiley & Sons Ltd.
- Driessen E, et al. (2015) Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of US national institutes of health-funded trials. PLoS One 10:e0137864. doi: [10.1371/journal.pone.0137864](https://doi.org/10.1371/journal.pone.0137864). PMID: [26422604](#)
- Dwan K, et al. (2010) Assessing the potential for outcome reporting bias in a review: a tutorial. Trials 11:52. doi: [10.1186/1745-6215-11-52](https://doi.org/10.1186/1745-6215-11-52). PMID: [20462436](#)
- Fenwick J, et al. (2008) The effect of bias on the magnitude of clinical outcomes in periodontology: a pilot study. J Clin Periodontol 35:775-782. doi: [10.1111/j.1600-051X.2008.01291.x](https://doi.org/10.1111/j.1600-051X.2008.01291.x). PMID: [18840153](#)
- Gagnier JJ, et al. (2012) Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. BMC Med Res Methodol 12:111. doi: [10.1186/1471-2288-12-111](https://doi.org/10.1186/1471-2288-12-111). PMID: [22846171](#)
- Gordon M, et al. (2020) forestplot: Advanced Forest Plot Using ‘grid’ Graphics. <https://cran.r-project.org/web/packages/forestplot/index.html> (参照 2021/02/24)
- GRADEpro GDT. (2021) <https://gradepro.org/> (参照 2021/02/24)
- Gurusamy KS, et al. (2009) Assessment of risk of bias in randomized clinical trials in surgery. Br J Surg 96:342-349. doi: [10.1002/bjs.6558](https://doi.org/10.1002/bjs.6558). PMID: [19283747](#)
- Guyatt G, et al. (2011a) GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 64:383-394. doi: [10.1016/j.jclinepi.2010.04.026](https://doi.org/10.1016/j.jclinepi.2010.04.026). PMID: [21195583](#)
- Guyatt GH, et al. (2011b) GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 64:395-400. doi: [10.1016/j.jclinepi.2010.09.012](https://doi.org/10.1016/j.jclinepi.2010.09.012). PMID: [21194891](#)
- Guyatt GH, et al. (2011c) GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 64:407-415. doi: [10.1016/j.jclinepi.2010.07.017](https://doi.org/10.1016/j.jclinepi.2010.07.017). PMID: [21247734](#)
- Guyatt GH, et al. (2011d) GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol 64:1277-1282. doi: [10.1016/j.jclinepi.2011.01.011](https://doi.org/10.1016/j.jclinepi.2011.01.011). PMID: [21802904](#)
- Guyatt GH, et al. (2011e) GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 64:1283-1293. doi: [10.1016/j.jclinepi.2011.01.012](https://doi.org/10.1016/j.jclinepi.2011.01.012). PMID: [21839614](#)
- Guyatt GH, et al. (2011f) GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol 64:1294-1302. doi: [10.1016/j.jclinepi.2011.03.017](https://doi.org/10.1016/j.jclinepi.2011.03.017). PMID: [21803546](#)
- Guyatt GH, et al. (2011g) GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol 64:1303-1310. doi: [10.1016/j.jclinepi.2011.04.014](https://doi.org/10.1016/j.jclinepi.2011.04.014). PMID: [21802903](#)
- Guyatt GH, et al. (2011h) GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 64:1311-1316. doi: [10.1016/j.jclinepi.2011.06.004](https://doi.org/10.1016/j.jclinepi.2011.06.004). PMID: [21802902](#)
- Guyatt GH, et al. (2012) GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. J Clin Epidemiol 66:158-172. doi: [10.1016/j.jclinepi.2012.01.012](https://doi.org/10.1016/j.jclinepi.2012.01.012). PMID: [22609141](#)
- Guyatt G, et al. (2013a) GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol 66:151-157. doi: [10.1016/j.jclinepi.2012.01.006](https://doi.org/10.1016/j.jclinepi.2012.01.006). PMID: [22542023](#)
- Guyatt GH, et al. (2013b) GRADE guidelines: 13. Preparing summary of findings tables and evidence

- profiles-continuous outcomes. J Clin Epidemiol 66:173-183. doi: [10.1016/j.jclinepi.2012.08.001](https://doi.org/10.1016/j.jclinepi.2012.08.001). PMID: [23116689](https://pubmed.ncbi.nlm.nih.gov/23116689/)
- Hahn S, et al. (2000) Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. Stat Med 19:3325-3336. doi: [10.1002/1097-0258\(20001230\)19:24<3325::aid-sim827>3.0.co;2-d](https://doi.org/10.1002/1097-0258(20001230)19:24<3325::aid-sim827>3.0.co;2-d). PMID: [11122498](https://pubmed.ncbi.nlm.nih.gov/11122498/)
- Hamm MP, et al. (2010) A descriptive analysis of a representative sample of pediatric randomized controlled trials published in 2007. BMC Pediatr 10:96. doi: [10.1186/1471-2431-10-96](https://doi.org/10.1186/1471-2431-10-96). PMID: [21176224](https://pubmed.ncbi.nlm.nih.gov/21176224/)
- Hartling L, et al. (2011) Applying the risk of bias tool in a systematic review of combination long-acting beta-agonists and inhaled corticosteroids for persistent asthma. PLoS One 6:e17242. doi: [10.1371/journal.pone.0017242](https://doi.org/10.1371/journal.pone.0017242). PMID: [21390219](https://pubmed.ncbi.nlm.nih.gov/21390219/)
- Hartling L, et al. (2013) Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. J Clin Epidemiol 66:973-981. doi: [10.1016/j.jclinepi.2012.07.005](https://doi.org/10.1016/j.jclinepi.2012.07.005). PMID: [22981249](https://pubmed.ncbi.nlm.nih.gov/22981249/)
- Hawkins N, et al. (2016) 'Arm-based' parameterization for network meta-analysis. Res Synth Methods 7:306-313. doi: [10.1002/jrsm.1187](https://doi.org/10.1002/jrsm.1187). PMID: [26610409](https://pubmed.ncbi.nlm.nih.gov/26610409/)
- Haynes RB, et al. (2004) Optimal search strategies for retrieving scientifically strong studies of diagnosis from MEDLINE: analytical survey. BMJ 328:1040. doi: [10.1136/bmj.38068.557998.EE](https://doi.org/10.1136/bmj.38068.557998.EE). PMID: [15073027](https://pubmed.ncbi.nlm.nih.gov/15073027/).
- Hedin RJ, et al. (2016) Publication bias and nonreporting found in majority of systematic reviews and meta-analyses in anesthesiology journals. Anesth Analg 123:1018-1025. doi: [10.1213/ANE.0000000000001452](https://doi.org/10.1213/ANE.0000000000001452). PMID: [27537925](https://pubmed.ncbi.nlm.nih.gov/27537925/)
- Herbison P, et al. (2011) Different methods of allocation to groups in randomized trials are associated with different levels of bias. A meta-epidemiological study. J Clin Epidemiol 64:1070-1075. doi: [10.1016/j.jclinepi.2010.12.018](https://doi.org/10.1016/j.jclinepi.2010.12.018). PMID: [21474279](https://pubmed.ncbi.nlm.nih.gov/21474279/)
- Higgins JP, et al. (2011a) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343:d5928. doi: [10.1136/bmj.d5928](https://doi.org/10.1136/bmj.d5928). PMID: [22008217](https://pubmed.ncbi.nlm.nih.gov/22008217/)
- Higgins JPT, et al. (2011b) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. <http://handbook-5-1.cochrane.org/> (参照 2021/02/24)
- Higgins JPT, et al. (2019) RoB 2 tool A revised Cochrane risk of bias tool for randomized trials. <https://www.riskofbias.info/welcome/rob-2-0-tool> (参照 2021/02/24)
- Hong H, et al. (2015) Incorporation of individual-patient data in network meta-analysis for multiple continuous endpoints, with application to diabetes treatment. Stat Med 34:2794-2819. doi: [10.1002/sim.6519](https://doi.org/10.1002/sim.6519). PMID: [25924975](https://pubmed.ncbi.nlm.nih.gov/25924975/)
- Hong H, et al. (2016) A Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons. Res Synth Methods 7:6-22. doi: [10.1002/jrsm.1153](https://doi.org/10.1002/jrsm.1153). PMID: [26536149](https://pubmed.ncbi.nlm.nih.gov/26536149/)
- Howard BE, et al. (2016) SWIFT-review: a text-mining workbench for systematic review. Syst Rev 5:87. doi: [10.1186/s13643-016-0263-z](https://doi.org/10.1186/s13643-016-0263-z). PMID: [27216467](https://pubmed.ncbi.nlm.nih.gov/27216467/)
- Hróbjartsson A, et al. (2012) Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. BMJ 344:e1119. doi: [10.1136/bmj.e1119](https://doi.org/10.1136/bmj.e1119)

e1119. PMID: [22371859](#)

Hughes MD, et al. (1992) The impact of stopping rules on heterogeneity of results in overviews of clinical trials. *Biometrics* 48:41-53. PMID: [1581492](#)

Hultcrantz M, et al. (2017) The GRADE working group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 87:4-13. doi: [10.1016/j.jclinepi.2017.05.006](#). PMID: [28529184](#)

Hutton B, et al. (2015) The PRISMA extension statement. *Ann Intern Med* 163:566-567. doi: [10.7326/L15-5144-2](#). PMID: [26436629](#)

Inaba Y, et al. (2009) Impact of single or multicentre study design on the results of trials examining the efficacy of adjunctive devices to prevent distal embolisation during acute myocardial infarction. *EuroIntervention* 5:375-383. doi: 10.4244/v5i3a59. PMID: [19736164](#)

Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press

Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for systematic reviews. National Academies Press

Ioannidis JP. (2011) Excess significance bias in the literature on brain volume abnormalities. *Arch Gen Psychiatry* 68:773-780. doi: [10.1001/archgenpsychiatry.2011.28](#). PMID: [21464342](#)

Jüni P, et al. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 282:1054-1060. doi: [10.1001/jama.282.11.1054](#). PMID: [10493204](#)

Kirkham JJ, et al. (2010) The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 340:c365. doi: [10.1136/bmj.c365](#). PMID: [20156912](#)

Kjaergard LL, et al. (2001) Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 135:982-989. doi: [10.7326/0003-4819-135-11-200112040-00010](#). PMID: [11730399](#)

Knudsen AB, et al. (2016) Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the US preventive services task force. *JAMA* 315:2595-2609. doi: [10.1001/jama.2016.6828](#). PMID: [27305518](#)

Krahn U, et al. (2013) A graphical tool for locating inconsistency in network meta-analyses. *BMC Med Res Methodol* 13:35. doi: [10.1186/1471-2288-13-35](#). PMID: [23496991](#)

Krogh J, et al. (2015) DEPERROR: Risks of systematic errors in drug and non-drug randomized clinical trials assessing intervention effects in patients with unipolar depression. *J Affect Disord* 179:121-127. doi: [10.1016/j.jad.2015.03.042](#). PMID: [25863907](#)

Kunz R, et al. (2007) Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev* MR000012. doi: [10.1002/14651858.MR000012.pub2](#). PMID: [17443633](#)

Kyzas PA, et al. (2005) Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst* 97:1043-1055. doi: [10.1093/jnci/dji184](#). PMID: [16030302](#)

Lee A, et al. (2006) Publication bias affected the estimate of postoperative nausea in an acupoint stimulation systematic review. *J Clin Epidemiol* 59:980-983. doi: [10.1016/j.jclinepi.2006.02.003](#). PMID: [16895822](#)

- Lefebvre C, et al. (2020) Chapter 4: Searching for and selecting studies. In: Higgins JPT, et al. editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.1. <https://training.cochrane.org/handbook/current> (参照 2021/02/24)
- Lin L, et al. (2017) Performing arm-based network meta-analysis in R with the pcnetmeta package. *J Stat Softw* 80. doi: [10.18637/jss.v080.i05](https://doi.org/10.18637/jss.v080.i05). PMID: [28883783](https://pubmed.ncbi.nlm.nih.gov/28883783/) [ノックページ：<https://cran.r-project.org/web/packages/pcnetmeta/>] (参照 2021/02/24)
- Linde K, et al. (1999) Impact of study quality on outcome in placebo-controlled trials of homeopathy. *J Clin Epidemiol* 52:631-636. doi: [10.1016/s0895-4356\(99\)00048-7](https://doi.org/10.1016/s0895-4356(99)00048-7). PMID: [10391656](https://pubmed.ncbi.nlm.nih.gov/10391656/)
- Liu CJ, et al. (2011) Do unblinded assessors bias muscle strength outcomes in randomized controlled trials of progressive resistance strength training in older adults? *Am J Phys Med Rehabil* 90:190-196. doi: [10.1097/PHM.0b013e31820174b3](https://doi.org/10.1097/phm.0b013e31820174b3). PMID: [21173683](https://pubmed.ncbi.nlm.nih.gov/21173683/)
- Marshall IJ, et al. (2016) RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials. *J Am Med Inform Assoc* 23:193-201. doi: [10.1093/jamia/ocv044](https://doi.org/10.1093/jamia/ocv044). PMID: [26104742](https://pubmed.ncbi.nlm.nih.gov/26104742/)
- McDonagh M, et al. (2013) Avoiding Bias in Selecting Studies. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Oregon Health & Science University, McMaster University, and Southern California Evidencebased Practice Centers under Contract No. 290-2007-10057-I) AHRQ Publication No. 13- EHC045-EF. Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/methods-guidance-bias_methods.pdf (参照 2021/02/24)
- McGowan J, et al. (2016) PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E). Ottawa: CADTH. https://www.cadth.ca/sites/default/files/pdf/CP0015_PRESS_Update_Report_2016.pdf (参照 2021/03/08)
- McGuire W, et al. (2010) The effect of fluconazole prophylaxis in very low birthweight infants is overestimated by before-after studies. *Arch Dis Child Fetal Neonatal Ed* 95:F182-187. doi: [10.1136/adc.2009.160796](https://doi.org/10.1136/adc.2009.160796). PMID: [19822526](https://pubmed.ncbi.nlm.nih.gov/19822526/)
- McInnes MDF, et al. (2018) Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 319:388-396. doi: [10.1001/jama.2017.19163](https://doi.org/10.1001/jama.2017.19163). PMID: [29362800](https://pubmed.ncbi.nlm.nih.gov/29362800/)
- McLeod BD, et al. (2004) Using dissertations to examine potential bias in child and adolescent clinical trials. *J Consult Clin Psychol* 72:235-251. doi: [10.1037/0022-006X.72.2.235](https://doi.org/10.1037/0022-006X.72.2.235). PMID: [15065958](https://pubmed.ncbi.nlm.nih.gov/15065958/)
- Mhaskar R, et al. (2012) Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. *J Clin Epidemiol* 65:602-609. doi: [10.1016/j.jclinepi.2011.10.016](https://doi.org/10.1016/j.jclinepi.2011.10.016). PMID: [22424985](https://pubmed.ncbi.nlm.nih.gov/22424985/)
- Millard LA, et al. (2016) Machine learning to assist risk-of-bias assessments in systematic reviews. *Int J Epidemiol* 45:266-277. doi: [10.1093/ije/dyv306](https://doi.org/10.1093/ije/dyv306). PMID: [26659355](https://pubmed.ncbi.nlm.nih.gov/26659355/)
- Mills EJ, et al. (2013) Demystifying trial networks and network meta-analysis. *BMJ* 346:f2914. doi: [10.1136/bmj.f2914](https://doi.org/10.1136/bmj.f2914). PMID: [23674332](https://pubmed.ncbi.nlm.nih.gov/23674332/)
- Morton SC, et al. (2018) Quantitative Synthesis-An Update. Methods Guide for Comparative Effectiveness

- Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 18-EHC007- EF. Agency for Healthcare Research and Quality. Posted final reports are located on the Effective Health Care Program search page. https://doi.org/10.23970/AHRQE_PCMETHGUIDE3 (参照 2021/02/24)
- Niemeyer H, et al. (2012) Publication bias in meta-analyses of the efficacy of psychotherapeutic interventions for schizophrenia. *Schizophr Res* 138:103-112. doi: [10.1016/j.schres.2012.03.023](https://doi.org/10.1016/j.schres.2012.03.023). PMID: [22484024](https://pubmed.ncbi.nlm.nih.gov/22484024/)
- Norris SL, et al. (2012) Selective outcome reporting as a source of bias in reviews of comparative effectiveness. Agency for healthcare research and quality. PMID: [22993870](https://pubmed.ncbi.nlm.nih.gov/22993870/)
- Nüesch E, et al. (2009a) The importance of allocation concealment and patient blinding in osteoarthritis trials: a meta-epidemiologic study. *Arthritis Rheum* 61:1633-1641. doi: [10.1002/art.24894](https://doi.org/10.1002/art.24894). PMID: [19950329](https://pubmed.ncbi.nlm.nih.gov/19950329/)
- Nüesch E, et al. (2009b) The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 339:b3244. doi: [10.1136/bmj.b3244](https://doi.org/10.1136/bmj.b3244). PMID: [19736281](https://pubmed.ncbi.nlm.nih.gov/19736281/)
- O'Mara-Eves A, et al. (2015) Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Syst Rev* 4:5. doi: [10.1186/2046-4053-4-5](https://doi.org/10.1186/2046-4053-4-5). PMID: [25588314](https://pubmed.ncbi.nlm.nih.gov/25588314/)
- Ouzzani M, et al. (2016) Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 5:210. doi: [10.1186/s13643-016-0384-4](https://doi.org/10.1186/s13643-016-0384-4). PMID: [27919275](https://pubmed.ncbi.nlm.nih.gov/27919275/). <https://rayyan.qcri.org/welcome> (解説動画：<https://www.youtube.com/watch?v=irAOQgzFMs4>) (参照 2021/02/24)
- Peters JL, et al. (2007) Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Stat Med* 26:4544-4562. doi: [10.1002/sim.2889](https://doi.org/10.1002/sim.2889). PMID: [17476644](https://pubmed.ncbi.nlm.nih.gov/17476644/)
- Pildal J, et al. (2007) Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 36:847-857. doi: [10.1093/ije/dym087](https://doi.org/10.1093/ije/dym087). PMID: [17517809](https://pubmed.ncbi.nlm.nih.gov/17517809/)
- Plummer M. (2017) JAGS. <http://mcmc-jags.sourceforge.net/> (参照 2021/02/24)
- Plummer M, et al. (2019) jags: Bayesian Graphical Models using MCMC. <https://cran.r-project.org/web/packages/rjags/> (参照 2021/02/24)
- pmSearch (2019) <https://stat.zanet.biz/sr/pmsearch.htm> (参照 2021/02/24)
- Poorolajal J, et al. (2010) Capture-recapture method for assessing publication bias. *J Res Med Sci* 15:107-115. PMID: [21526067](https://pubmed.ncbi.nlm.nih.gov/21526067/)
- Puhan MA, et al. (2014) A GRADE working group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 349:g5630. doi: [10.1136/bmj.g5630](https://doi.org/10.1136/bmj.g5630). PMID: [25252733](https://pubmed.ncbi.nlm.nih.gov/25252733/)
- Riley RD, et al. (2004) Sensitivity analyses allowed more appropriate and reliable meta-analysis conclusions for multiple outcomes when missing data was present. *J Clin Epidemiol* 57:911-924. doi: [10.1016/j.jclinepi.2004.01.018](https://doi.org/10.1016/j.jclinepi.2004.01.018). PMID: [15504634](https://pubmed.ncbi.nlm.nih.gov/15504634/)
- Roest AM, et al. (2015) Reporting bias in clinical trials investigating the efficacy of second-generation antidepressants in the treatment of anxiety disorders: a report of 2 meta-analyses. *JAMA Psychiatry* 72:500-510. doi: [10.1001/jamapsychiatry.2015.15](https://doi.org/10.1001/jamapsychiatry.2015.15). PMID: [25806940](https://pubmed.ncbi.nlm.nih.gov/25806940/)
- Rücker G, et al. (2019) Network meta-analysis of multicomponent interventions. *Biometrical Journal*

- 62:808-821. doi: [10.1002/bimj.201800167](https://doi.org/10.1002/bimj.201800167). PMID: 31021449 [R のパッケージ : <https://cran.r-project.org/web/packages/netmeta/> (参照 2021/03/08)]
- Salanti G, et al. Multiple-Treatments Meta-Analysis. <http://www.mtm.uoi.gr/> (参照 2021/02/24)
- Salanti G, et al. (2011) Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 64:163-171. doi: [10.1016/j.jclinepi.2010.03.016](https://doi.org/10.1016/j.jclinepi.2010.03.016). PMID: 20688472
- Savovi J, et al. (2012) Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess* 16:1-82. doi: [10.3310/hta16350](https://doi.org/10.3310/hta16350). PMID: 22989478
- Schou IM, et al. (2013) Meta-analysis of clinical trials with early stopping: an investigation of potential bias. *Stat Med* 32:4859-4874. doi: [10.1002/sim.5893](https://doi.org/10.1002/sim.5893). PMID: 23824994
- Schulz KF, et al. (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 273:408-412. doi: [10.1001/jama.273.5.408](https://doi.org/10.1001/jama.273.5.408). PMID: 7823387
- Schünemann H, et al. (2013) GRADE Handbook. <http://gdt.guidelinedevelopment.org/app/handbook/handbook.html> (参照 2021/02/24)
- Schünemann HJ (2016) Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision? *J Clin Epidemiol* 75:6-15. doi: [10.1016/j.jclinepi.2016.03.018](https://doi.org/10.1016/j.jclinepi.2016.03.018). PMID: 27063205
- Schünemann HJ, et al. (2019a) GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol* 111:105-114. doi: [10.1016/j.jclinepi.2018.01.012](https://doi.org/10.1016/j.jclinepi.2018.01.012). PMID: 29432858
- Schünemann HJ, et al. (2019b) GRADE guidelines: 22. The GRADE approach for tests and strategies—from test accuracy to patient-important outcomes and recommendations. *J Clin Epidemiol* 111:69-82. doi: [10.1016/j.jclinepi.2019.02.003](https://doi.org/10.1016/j.jclinepi.2019.02.003). PMID: 30738926
- Schwarzer G, et al. (2010) Empirical evaluation suggests Copas selection model preferable to trim-and-fill method for selection bias in meta-analysis. *J Clin Epidemiol* 63:282-288. doi: [10.1016/j.jclinepi.2009.05.008](https://doi.org/10.1016/j.jclinepi.2009.05.008). PMID: 19836925
- Selva A, et al. (2017) Development and use of a content search strategy for retrieving studies on patients' views and preferences. *Health Qual Life Outcomes* 15:126. doi: [10.1186/s12955-017-0698-5](https://doi.org/10.1186/s12955-017-0698-5). PMID: 28851437
- Shea BJ, et al. (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 7:10. doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10). PMID: 17302989
- Shea BJ, et al. (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomized or non-randomised studies of healthcare interventions, or both. *BMJ* 358:j4008. doi: [10.1136/bmj.j4008](https://doi.org/10.1136/bmj.j4008). PMID: 28935701
- Shemilt I, et al. (2019) Chapter 20: Economic evidence. In: Higgins JPT, et al. editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0*. <https://training.cochrane.org/handbook/archive/>

v6 (参照 2021/02/24)

Song F, et al. (2000) Publication and related biases. *Health Technol Assess* 4:1-115. PMID: [10932019](#)

Song F, et al. (2010) Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess* 14:iii, ix-xi, 1-193. doi: [10.3310/hta14080](#). PMID: [20181324](#)

Sox HC, et al. (2013) Medical decision making, 2nd ed. Wiley-Blackwell

Sterne JA, et al. (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 355:i4919. doi: [10.1136/bmj.i4919](#). PMID: [27733354](#). <https://www.riskofbias.info/welcome/home/current-version-of-robins-i> (参照 2021/02/24)

Sterne JA, et al. (2001) Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 323:101-105. doi: [10.1136/bmj.323.7304.101](#). PMID: [11451790](#)

Sterne JAC, et al. (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 366: i4898. doi: [10.1136/bmj.i4898](#). PMID: [31462531](#)

Tierney JF, et al. (2007) Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 8:16. doi: [10.1186/1745-6215-8-16](#). PMID: [17555582](#)

Turner RM, et al. (2009) Bias modelling in evidence synthesis. *J R Stat Soc Ser A Stat Soc* 172:21-47. doi: [10.1111/j.1467-985X.2008.00547.x](#). PMID: [19381328](#)

Unverzagt S, et al. (2013) Single-center trials tend to provide larger treatment effects than multicenter trials: a systematic review. *J Clin Epidemiol* 66:1271-1280. doi: [10.1016/j.jclinepi.2013.05.016](#). PMID: [23972520](#)

Vale CL, et al. (2013) Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. *BMJ* 346:f1798. doi: [10.1136/bmj.f1798](#). PMID: [23610376](#)

van Aert RC, et al. (2016) Conducting meta-analyses based on p values: reservations and recommendations for applying p-uniform and p-curve. *Perspect Psychol Sci* 11:713-729. doi: [10.1177/1745691616650874](#). PMID: [27694466](#)

van Hoorn R, et al. (2016) The development of PubMed search strategies for patient preferences for treatment outcomes. *BMC Med Res Methodol* 16:88. doi: [10.1186/s12874-016-0192-5](#). PMID: [27473226](#)

van Tulder MW, et al. (2009) Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine (Phila Pa 1976)* 34:1685-1692. doi: [10.1097/BRS.0b013e3181ab6a78](#). PMID: [19770609](#)

van Valkenhoef G, et al. gemtc. <https://cran.r-project.org/web/packages/gemtc/> (参照 2021/02/24)

van Valkenhoef G, et al. (2016) Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis. *Res Synth Methods* 7:80-93. doi: [10.1002/jrsm.1167](#). PMID: [26461181](#)

Viechtbauer W. (2020) metafor: Meta-Analysis Package for R. <https://cran.r-project.org/web/packages/metafor/index.html> (参照 2021/02/24)

Viswanathan M, et al. (2012) Development of the RTI item bank on risk of bias and precision of observational studies. *J Clin Epidemiol* 65:163-178. doi: [10.1016/j.jclinepi.2011.05.008](#). PMID: [21959223](#)

Viswanathan M, et al. (2017) Assessing the Risk of Bias in Systematic Reviews of Health Care Interventions. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 17(18)-EHC036-EF. Rockville, MD: Agency for Healthcare Research and Quality. Posted final reports are located on the Effective Health Care Program search page. <https://doi.org/10.23970/AHQEPCMETHGUIDE2> (参照 2021/02/24)

Watzlawick R, et al. (2014) Effect and reporting bias of RhoA/ROCK-blockade intervention on locomotor recovery after spinal cord injury: a systematic review and meta-analysis. JAMA Neurol 71:91-99. doi: [10.1001/jamaneurol.2013.4684](https://doi.org/10.1001/jamaneurol.2013.4684). PMID: [24297045](#)

Wessels M, et al. (2016) How to identify existing literature on patients' knowledge, views, and values: the development of a validated search filter. J Med Libr Assoc 104:320-324. doi: [10.3163/1536-5050.104.4.014](https://doi.org/10.3163/1536-5050.104.4.014). PMID: [27822157](#)

White CM, et al. (2009) Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. In: Agency for Healthcare Research and Quality (AHRQ). Methods Guide for Comparative Effectiveness Reviews. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/methods-guidance-de-novo-processes_methods.pdf (参照 2021/02/24)

White IR, et al. (2019) A comparison of arm-based and contrast-based models for network meta-analysis. Stat Med 38:5197-5213. doi: [10.1002/sim.8360](https://doi.org/10.1002/sim.8360). PMID: [31583750](#)

Whiting PF, et al. (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529-536. doi: [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009). PMID: [22007046](#)

World Health Organization (WHO). (2019) WHO Handbook for Guideline Development. 2nd edition. 17: Developing guideline recommendations for tests or diagnostic tools. <https://apps.who.int/iris/bitstream/handle/10665/145714/9789241548960-chap17-eng.pdf?sequence=20&isAllowed=y> (参照 2021/02/24)

Wood L, et al. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 336:601-605. doi: [10.1136/bmj.39465.451748.AD](https://doi.org/10.1136/bmj.39465.451748.AD). PMID: [18316340](#)

Yepes-Nuñez JJ, et al. (2019) Development of the summary of findings table for network meta-analysis. J Clin Epidemiol 115:1-13. doi: [10.1016/j.jclinepi.2019.04.018](https://doi.org/10.1016/j.jclinepi.2019.04.018). PMID: [31055177](#)

Zhang J, et al. (2014) Network meta-analysis of randomized clinical trials: reporting the proper summaries. Clin Trials 11:246-262. doi: [10.1177/1740774513498322](https://doi.org/10.1177/1740774513498322). PMID: [24096635](#)

Zhang J, et al. (2019) Bayesian hierarchical methods for meta-analysis combining randomized-controlled and single-arm studies. Stat Methods Med Res 28:1293-1310. doi: [10.1177/0962280218754928](https://doi.org/10.1177/0962280218754928). PMID: [29433407](#)

Zhang Y, et al. (2017) Using patient values and preferences to inform the importance of health outcomes in practice guideline development following the GRADE approach. Health Qual Life Outcomes 15:52. doi: [10.1186/s12955-017-0621-0](https://doi.org/10.1186/s12955-017-0621-0). PMID: [28460638](#)

Zhang E, et al. (2017) TrialSize. <https://cran.r-project.org/web/packages/TrialSize/index.html> (参照

2021/02/24)

上岡洋晴 他. (2018) 「AMSTAR 2：ヘルスケア介入のランダム化または非ランダム化研究あるいは両方を含むシステムティック・レビューのための厳密な評価ツール」の紹介と解. 薬理と治療 46:111-113.

河合富士美 他. (2017) システマティックレビュー検索式査読のためのガイドライン：PRESS Peer Review of Electronic Search Strategies. 医学図書館 64:75-78.

森實敏夫 他. (2015) 特別寄稿 5 診断に関する診療ガイドライン (CPG) の作成. https://minds.jcqhc.or.jp/s/guidance_special_articles5_1 (参照 2021/1/28)

第5章

医療経済評価

Minds

5章 Key Words

CHEERS 声明 (Consolidated Health Economic Evaluation Reporting Standards)
◎ Embase
EQ-5D (EuroQol 5 dimensions)
◎ EtD フレームワーク (evidence to decisions framework)
◎ GRADE (The Grading of Recommendations Assessment, Development and Evaluation)
HUI (Health Utilities Index)
◎ individual perspective (個人の立場)
◎ MEDLINE
◎ MeSH (Medical Subject Headings)
◎ National Institute for Health and Care Excellence (NICE)
◎ population perspective (集団の立場)
◎ PubMed
SF-6D (short form 6 dimensions)
◎ SoF 表 (summary of findings table)
◎ 医中誌 Web
◎ 益と害 (benefit and harm)
◎ エビデンス総体 (body of evidence)
◎ 感度分析 (sensitivity analysis)
基準的賭け (standard gamble : SG) 法
◎ クリニカルクエーション (clinical question : CQ)
国際医薬経済・アウトカム研究学会 (International Society for Pharmacoeconomics and Outcomes Research : ISPOR)
時間得失 (time trade-off : TTO) 法
◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
◎ システマティックレビューチーム (systematic review team : SR チーム)
質調整生存年 (quality-adjusted life year : QALY)
◎ 診療ガイドライン (clinical practice guideline : CPG)
◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
◎ 診療ガイドライン統括委員会
◎ 推奨
◎ スコープ (scope)
◎ 増分費用効果比 (incremental cost-effectiveness ratio : ICER)

- | |
|---|
| ◎ バイアスリスク (risk of bias) |
| ◎ 非一貫性 (inconsistency) |
| ◎ 非直接性 (indirectness) |
| 評価シート (evidence profile) |
| ◎ 不精確性 (imprecision) |
| ◎ メタアナリシス (meta-analysis, メタ解析と同義) |
| ◎ ランダム化比較試験 (randomized controlled trial : RCT) |
| 離散選択実験 (discrete choice experiment : DCE) |
| レファレンスケース (reference case) |

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】

SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】

RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】

PO : post 【第8章 診療ガイドライン公開後の取り組み】

5.0 はじめに

少子高齢化社会と社会保障費の増大を受け、近年、より効果的・効率的な医療の在り方が問われるようになってきている。2019年4月からは中央社会保険医療協議会（中医協）において、一部の医薬品、医療機器に対して費用対効果評価が行われるようになり、わが国の医療保険制度に系統的な医療経済評価が導入された。また、医療経済評価を診療ガイドライン（clinical practice guideline：CPG）に組み入れる国もあることから、わが国の診療ガイドライン作成者からも、診療ガイドラインにおける費用効果分析を含む医療経済評価についての方針を示すことが求められている。

Mindsでは、これまでに特別寄稿1「推奨作成における医療制度・経済的視点の考慮（2015年12月1日）」を企画・掲載し、また厚生労働行政推進調査事業費補助金（地域医療基盤開発推進研究事業）「診療ガイドラインの担う新たな役割とその展望に関する研究」の成果を踏まえ、診療ガイドラインと費用対効果の在り方検討会で議論を行った。その中で、費用が重要な情報となることが考えられる場合などには医療経済評価を組み入れることが望ましいこと、費用効用分析は必須としないこと、本マニュアルなどに具体的な方針を示すことなどが提案された。また、患者中心の医療が揺るがないような情報提供の方法、個人または集団の視点のどちらに立っているかを明示することなどへの留意を求めた。

本マニュアルでは、上記提案におおむね沿う形での方針の記載に努めた。特に、individual perspective（個人の視点）、population perspective（集団の視点）のどちらでも資源利用および費用対効果を考慮して診療ガイドライン作成が行えるように手順を整理した。また、『Minds 診療ガイドライン作成マニュアル2017』において、すでにEtDフレームワーク（evidence to decisions framework）などに必要資源量と費用対効果の評価が含まれていたが、具体的な手順や説明が限定的であった。そのため、本章ではそれらに対応できる形で具体的な手順の例を示した。

経済学的エビデンスには大きく分けて、①資源利用に関するもの、②費用対効果に関するものの2つがある。本マニュアルでは、①、②のいずれに関しても推奨作成時に考慮することを必須としていない。

本マニュアルでは、**第3章**に示したとおり、診療ガイドライン作成の計画段階において、資源利用・費用対効果を考慮するかどうかは診療ガイドライン統括委員会または診療ガイドライン作成グループ（guideline development group：GDG、診療ガイドライン作成パネルともいう）で決定することを推奨しており、決定以降の手順の詳細について本章で解説する。

経済学的エビデンスの解釈を行ったのち、推奨作成に至るまでの手順については**第6章**に記載する。

5.1 資源利用

資源利用とは、保健医療サービスを生産するために費やされた物的・人的資源投入量のことである。金銭の単位だけでなく、時間や個数などの物理的単位で表されることもある。本章では、資源利用が金銭の単位で表されるときには「費用」と呼ぶ。

診療ガイドラインまたは各クリニカルクエスチョン（clinical question : CQ）について資源利用を考慮するかどうかは、スコープ（scope）確定前までに診療ガイドライン統括委員会または診療ガイドライン作成グループで決定する。

資源利用の考慮は必須ではないが、特に資源利用が重要と思われるCQに関しては、有効性・安全性などとまとめて、資源利用をアウトカムの1つとしてシステムティックレビュー（systematic review : SR、系統的レビューと同義）のスコープに含める。

ただし、新規に資源利用の評価をする場合については**5.1.9項**で説明する。資源利用のシステムティックレビューの具体的な手順は**5.1.2項～5.1.8項**に示す（図5-1参照）。

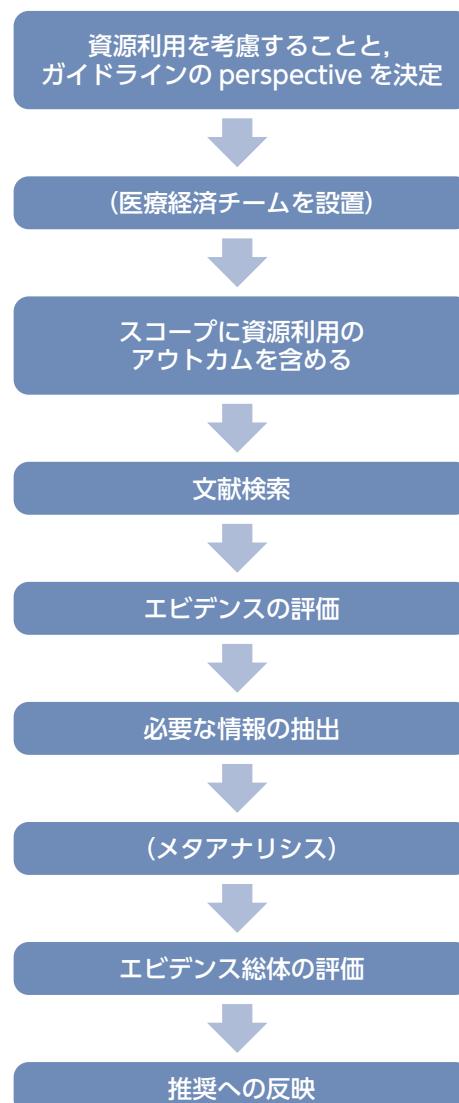


図5-1 資源利用のシステムティックレビューの手順

5.1.1 Perspective (分析の視点) を決める

推奨決定において、各基準に照らしてエビデンスを検討する際に、それが誰にとってのものであるかを決定するのが perspective (分析の視点) である。また、GRADE Handbook^{*1}では、あるアウトカムの重要性は、(文化または) どの perspective から評価するのかによって異なる可能性が高いとしている。Perspective に関する詳細は**第3章**を参照されたい。また、各推奨決定の基準が perspective によりどのように異なるかに関しては、EtD フレームワークが参考にできる (Alonso-Coello P 2016)。

資源利用の評価を行うことを診療ガイドライン統括委員会または診療ガイドライン作成グループが決定した場合、perspective によってスコープに含めるべき重要な資源利用が変わるために、事前に診療ガイドライン（または CQ）の perspective を診療ガイドライン統括委員会または診療ガイドライン作成グループが決定している必要がある。Individual perspective と population perspective の違いについては、**第3章**および**5.3 節**も参照されたい。

5.1.2 スコープの決定

資源利用について検討することを決めた場合は、システムティックレビュー全体のスコープを決定する際に、資源利用に関するアウトカムを含めておく必要がある（**第3章**も参照）。全ての資源利用を含める必要はなく、重要な資源利用だけを選ぶ。重要な資源利用は CQ ごとに異なる。資源利用のアウトカムの例を**表5-1**に示す。

GRADE (Brunetti M 2013) では、重要な資源利用を決める際の考え方として以下を提案している。

- ① 資源利用が、推奨決定に関して重要（または決定的）かどうかを検討する。
- ② 特定の資源利用の項目とそれが各治療法などに与える影響を検討する。

また、患者にとって重要なアウトカムの一部は資源利用の間接指標としても位置付けることができる。
(例：入院日数、合併症発生率)

(1) 分析期間 (time horizon) の決定

選択肢の間に存在すると予想される資源利用の差を検出するのに十分な長さの期間を定める。

(2) その他の包含基準

資源利用のアウトカムを測定しているランダム化比較試験 (randomized controlled trial : RCT) の数は限られる。十分に多くのエビデンスがあると予想される場合は RCT に限定することが可能であるが、それ以外のデザインの研究を組み入れることも可能である。

* 1 GRADE Handbook (<https://gdt.gradepro.org/app/handbook/handbook.html>)

表 5-1 資源利用の違いを同定するには（資源利用のアウトカムの例）

1. 医療資源の利用の違い (Luce BR 1996 より改変)
 - ・介入 (例：薬剤, 手術, カウンセリング, 理学療法)
 - 土地, 建物, 設備
 - 人材, 時間
 - 消耗品
 - ・検体検査
 - ・診察
 - ・緊急搬送
 - ・救急外来の受診
 - ・入院
 - ・専門医の受診
 - ・一般医の受診
 - ・往診, 訪問看護
2. 医療資源以外の資源利用の違い
 - ・家の改修
 - ・食事療法
 - ・医療施設への移動
 - ・福祉サービス (例：住居, 職業訓練)
 - ・犯罪 (例：窃盗, 詐欺, 暴力, 警察の捜査, 裁判)
3. 患者自身や家族などの介護者の資源利用の違い
 - ・訪問
 - ・入院
 - ・セルフケアにかかる時間
 - ・家族などによる介護にかかる時間
4. 生産性の違い
 - ・疾患, 治療などのために仕事を休んだ時間

出典：Brunetti M, et al. (2013) GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol 66:140-150. より作成

5.1.3 文献検索

MEDLINE, CINAHL (Cumulative Index to Nursing and Allied Health Literature), Embaseなどの医学全般のデータベースを検索する際、研究デザインの包含基準が資源利用以外のシステムティックレビューと同じとする場合は、基本的に資源利用のために別に検索を行う必要はなく、全体の検索結果の中からスコープで選んだ資源利用のアウトカムが含まれている研究を同定する。

必須ではないが、一般的なデータベースに加えて、経済学データベースを検索することができる。この場合は基本的に経済学フィルターは不要である。経済学データベースの種類については、**5.6節**を参考にされたい。上記により得られるエビデンスでは不十分と判断される（あるいは、そう予想される）場合は、包含基準と検索データベースの範囲を再検討することができる。

5.1.4 エビデンスの評価

健康アウトカムと同様の方法で各研究のバイアスリスク (risk of bias) の評価を行う（**第4章 4.3節**参照）。特に、非直接性 (indirectness) は資源利用のエビデンスの評価においては鍵となり、場合によってはメタアナリシス (meta-analysis, メタ解析と同義) の結果よりも、最も直接性の高い研究のデータを強く考慮することもあり得る。不精確性 (imprecision) については、RCT が資源利用に関する検出力不足の場合があることや、P 値が報告されないことがあることに注意を要する。非直接性、不精確性、非一貫性 (inconsistency) の詳しい説明は**第4章 (4.3.2項 (2) および 4.4.2項)**を参照されたい。

5.1.5 必要な情報の抽出

事前にスコープで定めた重要な資源利用のアウトカムだけを抽出する。**表 5-2** の例のように、介入群と対照群における資源利用の差の推定値が示されている必要がある。1人当たりの費用よりも時間数や個数などの自然単位を優先して記載する。その場合、可能な限り単価も記載する。

5.1.6 メタアナリシス

資源利用の項目に関してもメタアナリシスを行うことができる。行うかどうかは、他のアウトカムと同じ考え方方が適用できる（**第4章 4.4.4項** 参照）。判断基準については、GRADE guidelines: 10 (Brunetti M 2013: Box 4) も参考にできるが、特に費用のメタアナリシスについては実験的、あるいは適切でないという考え方もあり、費用、自然単位で表された資源利用のいずれについても、メタアナリシスは必須ではない。

5.1.7 エビデンス総体の評価

(1) 評価シート (evidence profile)

他のアウトカムと同じ基準に従ってエビデンス総体 (body of evidence) の質を評価する。評価シートでも、自然単位を優先して記載し、可能なら単価をつける（**表 5-2** 参照）。

(2) SoF表 (summary of findings table)

GRADE guidelines: 10 (Brunetti M 2013: Table 2) を参考に、資源利用を含めたSoF表を作成することができる。もし日本国内かつ最新の単価のデータが入手できれば、その情報を書き入れることで診療ガイドライン統括委員会および診療ガイドライン作成グループが資源利用の金銭的価値をより適切に評価できるようになる。

5.1.8 推奨への反映

詳細は第6章を参照のこと。

5.1.9 新規の評価を行う場合

資源利用、特に費用について新規の評価を行う場合は、5.2節以降に記載されている医療経済評価における手法や考え方などを適用する。ただし、通常の医療経済評価と異なり、資源利用を考慮する際に割引を行うべきかについては議論がある。新規の医療技術の導入により、医療関係者の負担が増大する、病床が逼迫する、既存の設備（医療機器等含む）では需要の増加に対応できないなど、医療資源を圧迫する可能性がある場合（あるいは医療資源の消費が大きく減少する場合）は、医療機関の立場から資源消費の分析を行うこともできる。

以上により得られた結果についても、第7章7.1節のエビデンス総体に含めて扱う。

表5-2 資源利用の評価シートの記入例

Question: Should buprenorphine maintenance flexible doses vs. methadone maintenance flexible doses be used for opioid maintenance treatment?

対象：オピオイド依存症患者

介入：buprenorphine 維持療法、可変用量

対照：methadone 維持療法、可変用量

Perspective：societal

アウトカム 研究（追跡期間）	Quality assessment						資源、費用の要約		Overall quality		
	研究 デザイン	限界	非一貫性	非直接性	不精確性	その他	患者数	資源または費用 (患者1人当たり(1999 AU \$))			
								Methadone	Buprenorphine		
薬剤（6ヶ月） 1つの研究 (Doran, 2003)	RCT	No	低	中	低	低	405	資源（1日当たり）平均		中 ⊕⊕⊕○	
								57mg	11mg		
								費用（6ヶ月）			
								37 (33 SD)	459 (461 SD)		
他の医療費 (6ヶ月) 1つの研究 (Doran, 2003)	RCT	No	低	中	低	低	405	資源		中 ⊕⊕⊕○	
								NA	NA		
								費用（6ヶ月）			
								1,378 (NA)	1,270 (NA)		
犯罪に関連する費用 情報なし											

出典：Brunetti M, et al. (2013) GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol 66:140-150. より作成

5.2 概論：診療ガイドラインにおける費用対効果の考慮について

費用対効果とは、医療技術にかかる費用（将来に発生する費用も含む）とそこから得られるアウトカムのバランスを考えるものであり、その分析手法は医療経済評価と呼ばれる。一般的に医療経済評価は、費用最小化分析、費用効果分析、費用効用分析、費用便益分析の4種類に分類される。このうち、費用効用分析は費用効果分析の一種と見なされることから、本章では費用効用分析を独立して取り扱うことはしない。また、費用便益分析はアウトカムの価値を金銭で評価するものであり、経済学が伝統的に用意してきた政策の評価手法として環境や交通分野では一般的である。一方、健康状態を金銭化することには、倫理的な課題も伴うことから保健医療上の意思決定にはあまり用いられていない。

5.2.1 診療ガイドラインと費用対効果

医療経済評価においては、個々の患者の立場（診療ガイドラインにおける individual perspective に対応するもの）ではなく、公的医療や社会など何らかの集団の立場（population perspective）^{*2} から分析を実施することが一般的である。そこから得られる結果は、患者個人ではなく、公的医療や社会などにおける費用対効果であり、それらを用いて何らかの社会的な視点からの意思決定を行うことを目的とする。

例えば、日本において患者個人の立場から医療経済評価を実施する場合は、費用として患者の自己負担分のみを用いて分析を行う必要がある。しかし年齢や所得の多寡、公費診療の有無などにより、患者の自己負担が細かく変わることから、平均的な患者の立場による分析を実施することは困難を伴う。また、患者の自己負担分を除いた費用については保険料や税（公費）から支出されていることを考慮すれば、患者の自己負担分のみを取り扱う評価に基づいて社会的な意思決定をすることは十分とはいえない。例えば、公費負担により自己負担がない（患者の立場から見れば費用は0円である）状況を理由に、医療技術を無制約に使用することが、社会的な視点からの意思決定として推奨されるわけではないだろう。

一方で、診療ガイドラインは、individual perspective により個々の患者の立場で作成されることが少なくない。これら2つの perspective の違いは、以下のような問題をはらむ可能性がある。すなわち、費用対効果とは、医療技術の費用（将来に発生する費用も含む）と、そこから得られる効果のバランスを考えることを目的とする。費用対効果の悪い医療技術とは、一定の追加的有用性は認められるものの、それを得るための費用が高額すぎる（あるいは医療サービスを行うために費やされた人的・物的資源が多すぎる）と考えられるものであり、臨床的に意味のない「無駄」な状況とは異なる。

例えば、日本では一部の諸外国とは異なり、医療提供者は地域や医療機関の予算管理の責任を負っていない。そのような医療資源の制約が明確に設定されていない状況においては、個々の患者にとって、費用対効果にかかわらず最も有効性の高い医療技術を使用することが、治療効果を最大化する。つまり、用対効果が悪いことのみを理由として、診療ガイドラインにおいてその医療技術の使用を制約す

* 2 「立場（perspective）」という用語は、医療経済評価における分析の「立場」と、診療ガイドライン作成における「視点」を表す場合がある。本章においては、診療ガイドラインでの視点の場合は、英語の「perspective」をそのまま用いた（individual perspectiveなど）。一方で医療経済評価における立場を表す場合は、漢字で「立場」と表記し（公的医療の立場など）、必要に応じて「分析の立場」を用いる、あるいは同一文中に「分析」の語を補うなどとした。

することは、診療ガイドラインが individual perspective の場合の考え方とは整合しない可能性がある。

5.2.2 Individual perspective で作成される診療ガイドラインにおける費用対効果の取り扱い

5.2.1 項での議論から、individual perspective で作成される診療ガイドラインにおける費用対効果の取り扱いについては、個々の患者の立場に焦点をあてる individual perspective と矛盾しない、例えば以下のような場面において活用することを推奨する。

① 患者にとっての益と害を勘案した「正味のアウトカム」が同程度と見なせる医療技術が複数ある場合

これは、統計学的に厳密な同等性が証明されているものに限らず、臨床現場で使用される際に「益と害 (benefit and harm) の両者を勘案した観点からは無差別 (どちらも等価に使用され得る)」と判断される可能性があるものを指す。また、一部の患者集団に限定すれば前記の条件が満たされるものも含み得る。

「同程度」であるかどうかの判断は、個々の患者の状況によって異なる可能性もあり、また最終的には診療ガイドラインの使用者にゆだねられている。よって、状況や考え方によっては無差別と判断される「可能性がある」ものを含めて、情報提供という観点から幅広く取り扱うことを推奨する。

例えば、診療ガイドライン上で複数の医療技術が併記して推奨されているもの、複数の選択肢があるが診療現場ではほぼ等価であると見なされて使用されているもの、あるいは無効例・忍容性のない患者などにのみ高価な医療技術を使用するなどしても臨床的に問題がないもの（使用順序の設定）、一部の集団に限定すれば正味のアウトカム^{*3}が同等と考えられるものなどが、評価対象となり得る。これらについては効果が同等との前提の下で費用の比較を実施し、最も安価なもののが推奨される。

② Individual perspective の立場からも増分費用効果比を用いた意思決定を実施することが適切と判断される医療技術の場合

例えば、効果が十分に高い既存技術が存在しており、患者にとって新規医療技術の追加的アウトカムが小さい、あるいは大きな不確実性を有すると判断されるなど、患者にとっての追加的なアウトカムが限制的なものなどが挙げられる。これらについては費用効果分析^{*4}を実施し、増分費用効果比 (incremental cost-effectiveness ratio : ICER) を算出する。

なお、診療ガイドラインにおいて費用対効果を検討する場合に、本来的には上記①と②を包括的に評価すべきだろうが、実施可能性を考慮して上記の一部、例えば評価の簡易な①から対応するということもあり得る。

* 3 本章では、患者にとっての益と害の両者を勘案したものを「正味のアウトカム」と呼び、医療経済評価や費用対効果におけるアウトカム（指標）は「効果（指標）」とする。

* 4 以降、本章で用いる「費用効果分析」には、質調整生存年 (quality-adjusted life year : QALY) を用いる費用効用分析も含んでいる。

5.2.3 Population perspective で作成される診療ガイドラインにおける費用対効果の取り扱い

原則的に全ての介入、診断技術などに関して評価の対象となり得る。中医協では、「高額な医療技術」「医療財源への影響が大きい医療技術」などを評価の対象としているが、診療ガイドラインにおいては、それらに加えて**5.5 節表 5-6**などを参考に選ぶ、または、費用対効果により推奨が変わることの可能性があるCQ、あるいは臨床的に重要性が高いCQを選ぶことが考えられる。

5.3 レファレンスケース

レファレンスケース (reference case) とは、医療経済評価を行う際に、基準となる手法を規定したものである。学術的観点からは、医療経済評価についてもさまざまな手法や考え方があり得るので、レファレンスケースと異なる手法を用いることを否定するものではない。ただし、分析の標準化や結果の比較可能性の観点から、レファレンスケース分析を含めて実施することが望ましい。

レファレンスの要約を以下表5-3、表5-4に示す。

表5-3 レファレンスケースの要約：正味のアウトカムが同程度と見なせる一群の治療

対象となり得る医療技術	患者にとっての正味の益と害（正味のアウトカム）が同程度と見なせると判断される一群の治療 ・診療ガイドライン上で複数の治療が併記して推奨されているもの ・複数の選択肢があるが診療現場では無差別に使用されているもの ・無効例・忍容性のない患者などにのみ、より価格の高い治療を使用しても臨床的に問題がないもの ・一部の集団に限定すれば、正味のアウトカムが同等と考えられるものなど。
評価対象集団	原則として保険適用となるもの。結果に異質性があり得る場合は、各集団で評価を実施する。
分析の立場（5.3.1項）	公的医療の立場
分析手法（5.3.2項）	費用最小化分析
資源消費量と単価（5.3.2項）	・「日本における平均的な使用量や標準的な診療過程等が反映」されたものの。単価は最新時点の診療報酬点数表や薬価基準などを使用する。 ・後発医薬品の費用を考慮する。
比較対照技術（5.3.3項）	相互に費用を比較する。
分析期間（5.3.4項）	「評価対象技術の費用や効果に及ぼす影響を評価するのに十分な長さの分析期間」だが、多くの場合は、効果が同様であるので、短期間の分析で十分と考えられる。
効果指標（5.3.5項）	(効果同等と判断されていることが前提なので、該当しない。)
割引率（5.3.6項）	短期の場合は実施せず。
結果の解釈	原則として最も安価なものが最も費用対効果に優れると解釈する。

表 5-4 レファレンスケースの要約：ICER を用いた意思決定を実施することが適切と判断される場合

対象となり得る医療技術	Individual perspective あるいは population perspective を問わず ICER を用いた意思決定を実施することが適切と判断される場合 ・治療効果が十分に高い既存技術がすでに存在しており、新規治療の追加的効果が小さいあるいは不確実と判断される場合 ・ワクチンや検診などの公衆衛生的介入 ・Population perspective で作成される診療ガイドラインなど。
評価対象集団	原則として保険適用となる（あるいはワクチンや検診などそれに準ずる）もの。結果に異質性があり得る場合は、各集団で評価を実施する。
分析の立場（5.3.1 項）	公的医療の立場
分析手法（5.3.2 項）	費用効果分析（費用効用分析：QALY を用いた費用効果分析）
資源消費量と単価（5.3.2 項）	・「日本における平均的な使用量や標準的な診療過程等が反映」されたもの。単価は最新時点の診療報酬点数表や薬価基準などを使用する。 ・後発医薬品の費用を考慮する。
比較対照技術（5.3.3 項）	現時点で当該治療が導入されることにより最も代替されると想定されるもの（標準治療など）。
分析期間（5.3.4 項）	評価対象技術の費用や効果に及ぼす影響を評価するのに十分な長さの分析期間。
効果指標（5.3.5 項）	QALY を基本とする。
割引率（5.3.6 項）	2%（年率）
結果の解釈	中医協における基準値 500 万円 /QALY は参考し得るが、意思決定現場の状況に依存するものであり、総合的に判断する。

5.3.1 分析の立場

分析の立場は、「公的医療の立場」を含めることとする。「公的医療の立場」とは、費用について公的医療費の範囲のみを考慮するものである。自己負担分のみならず、保険者負担分や公費分など医療費の総額となる。公的介護費用（介護保険による支出）や、生産性損失（病気などにより働くことができないため生じる社会的な損失）は含めない。

一方で、それらの費用を含めた立場からの分析が有用な場合も想定される。参考として、公的介護費用や社会の立場から生産性損失を加えた分析を追加的に実施することは可能だが、その際には過大推計や公平性などの問題に配慮する必要がある。

生産性損失は一般的には病気などにより働けない日数に賃金などをかけることにより算出する（これを人的資本法と呼ぶ）。しかし、賃金などから生産性損失を推計する場合、賃金の多寡に影響を与える要因、例えば性差や年齢などで生産性損失の大きさが異なってしまう。給与の低い集団は、生産性損失の値が小さくなるので、同じ医療技術であっても費用対効果が悪くなる。実際には賃金が発生しない家事労働などの損失をどのように推計するかについても課題があることに留意が必要である。また、公的介護費用を費用に含める場合には、インフォーマルケアの費用などについても検討が必要になるかもしれない。なお、賃金のデータソースとして日本では「賃金構造基本統計調査（賃金センサス）」が用いられることが多い。賃金センサスでは性年齢階級別に職種ごとの税引き前賃金が計算されている。

患者の自己負担の観点から分析を実施することは、患者により自己負担率が変わること、患者の自己負担分を除いた費用については保険料や税金から支出されていることを考慮すれば、推奨されない。公費負担などにより自己負担額が免除されれば、どのような医療資源の制約も考慮しなくてよいという帰結を導きかねない。

5.3.2 分析手法

(1) 「正味のアウトカムが同程度と見なせる」と臨床的に判断される医療技術

「正味のアウトカムが同程度と見なせる」と臨床的に判断される医療技術を評価する場合は、

1) 医療技術の公定価格（診療報酬点数、薬価、保険医療材料価格など）

あるいは

2) 一定期間における関連する総医療費

を比較することにより、費用最小化分析（費用の比較）を実施する。後発医薬品が存在する場合は、先発品・後発品それぞれを用いた分析を実施することが原則である。同一成分・用量の後発医薬品として複数の価格帯が存在する場合は、価格帯の中央値（例えば3価格帯の場合は2番目に高い価格、2価格帯の場合は両価格の平均値）を用いることを推奨する。

1) 医療技術の公定価格（診療報酬点数、薬価、保険医療材料価格など）

費用最小化分析とは、効果が同等であるという前提の下で、治療に要する費用の比較をする分析法である。費用の比較においては、必ずしも全ての治療費用を考慮する必要はなく、例えばある薬物治療（治療Aと治療B）の費用対効果を考える場合、両治療に同様に含まれる費用は群間差を検討する際に相殺されるので、考慮しなくても結論は変わらない。効果は両者変わらない前提であるので、例えば継続的に使用し、診療頻度などその他の要因に伴う費用が変わらないと想定されれば、最も単純には1日薬剤費のみの比較で十分である。

ただし、複数の用量が設定されている場合は、どの用量同士を比較するかによって結果が異なる可能性がある。その場合、臨床的な有効性が同等と考えられる用量同士、あるいは実際に使用されている平均用量や用量の分布を用いて検討することなどが考えられる。単純に最大（あるいは最小）用量を用いて計算することは、効果が同等でない可能性もあるので推奨されないが、薬価算定における汎用規格^{*5}を用いることは可能かもしれない。

2) 一定期間における関連する総医療費

評価対象技術と比較対照技術間で、その他の診療行為の頻度など（例：入院期間や通院間隔）が異なる場合は、その他の技術料や治療費用などを考慮すればよい。ただし、分析期間は両群でそろえる必要がある。例えばA群の費用が1ヵ月分、B群の費用は来院頻度に合わせて3週間分となると費用の比較ができなくなるので、同一期間における費用を比較する（例えば、1サイクルの日数比で補正するなど）。

費用の算出方法については、診療報酬を出来高で積み上げていく手法と、レセプトなど実際の診療情報を用いる方法がある。それほど複雑でない治療プロセスなどであれば前者の手法で十分である。レセプトなどの情報を用いて費用推計する手法は、実際に医療現場で使用された医療資源消費量が反映されているという点で有益であるが、入手などを含め解析するまでに手間やコストがかかること、また当該疾患と関連しない費用（例：高血圧治療中におけるインフルエンザ治療の費用）が含まれる可能性があること、限定された施設で収集されたレセプトの場合、結果の一般化可能性が問題となることなどの限界がある。診療報酬を出来高で積み上げていく場合は、診療報酬項目ごとに表5-5のように「単価」と「消費数量」を分けて報告することが望ましい。

なお、いずれの場合においても、費用は評価対象の治療にかかるもののみならず、副作用や将来に起これ得る合併症など、評価対象技術に関連するものは全て含めることが原則である。ただし、将来発生する費用についても、群間差がないと考えられるものについては、相殺される（「正味のアウトカムが同程度と見なせる」医療技術についてはそのようなケースが多いと想定される）。また、結果に与える影響が軽微なものについては、考慮しなくともよい。

* 5 汎用規格とは、組成および剤形が同一の類似薬の年間販売量を規格別に見て、最もその合計量が多い規格をいう。

表 5-5 診療報酬を出来高で積み上げていく場合の例

診療報酬項目	単価（点数×10）	消費量	小計
再診料	720	1	720
鼻腔・咽頭拭い液採取	50	1	50
インフルエンザウイルス抗原定性	1430	1	1,430
免疫学的検査判断料	1440	1	1,440
処方箋料（その他）	680	1	680
調剤基本料1	410	1	410
後発医薬品調剤体制加算1	180	1	180
内服薬調剤料	50	5	250
頓服薬調剤料	210	1	210
薬剤服用歴管理指導料	410	1	410
オセルタミビル	129.7	10	1,297
アセトアミノフェン	6.9	5	34
		計	7,111

(2) ICER を用いた意思決定を実施することが適切と判断される医療技術

Individual perspective あるいは population perspective によらず「ICER を用いた意思決定を実施することが適切と判断される医療技術」の場合、分析結果は、増分費用を増分効果で除した ICER を用いて表す。

ICER は下式で定義される。これは分母の単位が QALY であれば「1QALY（完全に健康な生存期間 1 年分）を獲得するのに追加的にかかる費用」と解釈できる。

$$\text{ICER} = \frac{\text{費用（評価対象）} - \text{費用（比較対照）}}{\text{効果（評価対象）} - \text{効果（比較対照）}} = \frac{\text{増分費用}}{\text{増分効果}}$$

ただし、費用が安くて効果が大きい場合、計算上は ICER が負の値をとるが、その場合は ICER を算出せずに優位（dominant）であるとする。一方で、費用が高くて効果が小さい場合は同様に ICER を算出せずに劣位（dominated）と表記する。

ICER の算出にあたっては、医療経済学的なモデル構築あるいはパラメータ設定に関する専門的知識が必要になることが多い。それらの分析経験を有する者、あるいは医療経済評価の専門家が実施することを検討する必要がある。

効果を考慮しない場合、あるいは定量的に考慮することが困難な場合など（例えば獲得できる効果が QALY 単位で極めて小さい場合）については、医療費以外の支出、インフォーマルな介護や看護による機会損失、労働生産性損失などを金銭的側面から考慮した費用の比較（あるいは費用便益分析の一種）を併せて実施することも可能である。

日本で分析する際に、海外の臨床試験や疫学データを用いて分析を行うことは、データの外挿可能性を考慮する必要があるが、許容されることも多い。一方で費用については、(そのことにより結論が変わり得る可能性がある場合は最新の)日本における診療報酬点数や薬価・保険医療材料価格などを反映させる必要がある。

得られた結果の不確実性が大きい場合は、感度分析（sensitivity analysis、シナリオ分析を含む）などの結果を併せて提示することも有効である。感度分析とは、あるパラメータを動かした場合に、それがどの程度結果に影響を及ぼすかを検討するものである。単一のパラメータが与える影響について検討するものを一元感度分析、分布を当てはめて複数のパラメータを同時に検討するものを確率感度分析と呼ぶ。感度分析などにより結果の不確実性が大きいと判断される場合は、結果として複数の ICER を表記する、ICER を範囲で示すなどの対応についても検討する必要がある。

5.3.3 比較対照技術

比較対照技術については、「正味のアウトカムが同程度と見なせる」と臨床的に判断される医療技術を評価する場合は、「同程度と見なせる」医療技術間で相互に費用を比較する。その場合、比較対照としては単一の技術のみではなく、一連の治療方法や医療技術の使用の優先順序（一次治療、二次治療）なども検討できる。また、全体での使用を最適化するために、適応集団の一部についてのみ、異なる比較対照技術を用いることもあるかもしれない。

それ以外の ICER を算出する場合には、当該医療技術が導入される（された）ことにより、臨床現場で幅広く使用され最も代替される（された）と想定されるものを含むことを推奨する。しかし上記の定義は概念的なものであり、定義に従って厳密に比較対照技術を探索することは困難であることが多い^{*6}。よって、多くの場合は標準治療や最も効果の大きな治療法などが比較対照技術の候補になり得ると想定される。それらが一義的に決まらない場合は、費用対効果の大きさや比較試験における対照技術なども考慮して決定する、複数の比較対照技術に対して評価を行うなどの対応をする。比較対照技術の費用対効果がよくない場合や必ずしも適正とはいえない治療が広く行われている場合などは、それらを比較対照技術から除外することも検討できる。

5.3.4 分析期間

分析期間とは、その間に発生する効果や治療費用などを計算上考慮する期間であり、時間地平（time horizon）とも呼ばれる。評価対象技術が費用や効果に及ぼす影響を評価するのに十分な長さの分析期間とすることが原則である。「正味のアウトカムが同程度と見なせる治療」を評価する場合は、効果が同様であることが前提であるので、多くの場合において短期間の分析で十分と考えられる。

医療経済評価においては、特に ICER を算出する場合は、臨床研究などで得られているよりもより長期の分析期間が設定されることがあり、そのような場合は医療経済学的なモデル（マルコフモデルなど）を用いて、長期的な予後や費用の予測を行うことがある。

長期の分析期間（例えば生涯）を設定することが評価において必要な場合、そのような医療経済学

* 6 すでに評価対象となった医療技術が導入された世界において、導入されなかったときに使用されていた医療技術は観測できない。

的モデルを用いた分析を行うことは可能である。ただし、モデルを用いた分析を行う場合は、効果や費用のパラメータを含めて、それに伴うさまざまな不確実性が存在することに留意する。必要に応じて、感度分析（シナリオ分析を含む）を実施し、得られた結果の頑健さなどについて検討する必要がある。

5.3.5 効果指標

「正味のアウトカムが同程度と見なせる治療」を評価する場合は、効果に差がないことが前提であるので、特に効果指標を設定しない。

一方で、「ICER を用いた意思決定を実施することが適切と判断される治療」を評価する場合は、効果指標として QALY を用いることを推奨する。QALY は生存期間に QOL 値（「効用値」とも呼ぶ）をかけ合わせたもので、例えば 0.8 の QOL 値で 3 年間生存できれば $0.8 \times 3 = 2.4\text{QALY}$ と計算できる。QOL 値は 0（死亡）と 1（完全な健康状態）として定義されるので、 $1\text{QALY} = 1 \times 1\text{年}$ 、すなわち完全な健康状態で 1 年間生存した状態である。

QALY 以外の効果指標を用いることも可能であるが、算出した ICER の結果をどのように解釈するか、事前に検討が必要である。

QOL 尺度には患者の健康状態を測定するための「プロファイル型尺度」と医療経済評価に用いる「選好に基づく尺度」が存在するが、QALY を算出する際に活用できるのは後者の「選好に基づく尺度」による測定値であり、前者のスコアを用いることはできない。

具体的な「選好に基づく尺度」としては、EQ-5D (EuroQol 5 dimensions), HUI (Health Utilities Index), SF-6D (short form 6 dimensions) などが存在する。これらの尺度は患者から得られた回答を、一般の人々の選好が反映された換算表を用いて、スコアに変換することにより QOL 値を得るものである。この換算表は尺度ごとに異なっている。また国によって各項目や水準の重みなどが異なることから、換算表が国ごとに作成されている尺度が多い。

日本国内で新たに QOL 値を収集する際には、国内データに基づき換算表が開発された選好に基づく尺度を使用することが推奨される。なお、適切な研究データが存在する場合、国内で測定された QOL 値を優先的に使用することを推奨する。海外で測定されたデータの使用を否定するものではないが、結果の解釈に注意が必要である。その他、基準的賭け (standard gamble : SG) 法、時間得失 (time trade-off : TTO) 法、離散選択実験 (discrete choice experiment : DCE) など、QOL 尺度を用いる以外の測定手法もある。

5.3.6 割引率

医療経済評価においては、人々の時間選好（将来得られる利得よりも現在の利得を選好すること）を反映するために、将来において得られる効果や発生する費用を割り引くことが一般的である。割引率としては、当面は年率 2% を用いて、以下のような指數型割引を行うことを推奨する。

$$C_p = \frac{C_i}{(1+0.02)^{i-1}}$$

(C_i : i 年目に発生する費用, C_p : i 年目に発生する費用の現在の価値)

例えば、10年後にかかる100万円の医療費は、 $100\text{万円}/(1+0.02)^9 = \text{約} 82\text{万円}$ となる。この計算方法は将来時点で得られる効果についても同様である。

特に分析期間が長期に及ぶ場合には、割引率が結果に大きな影響を与えることも多いため、0%から4%の範囲で感度分析を行うこともある。ただし、「正味のアウトカムが同程度と見なせる治療」を費用最小化分析する場合に想定されるように、分析期間が短期間である場合は、割引を行わないことも可能とする。

5.4 費用対効果の評価手順

費用対効果のエビデンスを診療ガイドラインに採り入れる手順については、以下が考えられる（図5-2および図5-3）。また、本章や他章の参照すべき部分を示す。

① Perspective および費用対効果を当該診療ガイドラインで考慮するかを診療ガイドライン統括委員会または診療ガイドライン作成グループで決定

診療ガイドライン統括委員会、あるいは診療ガイドライン作成グループにおいて、perspectiveを決定し、さらに資源利用、費用対効果、それぞれを考慮するかどうかを決定する。考慮しない場合は、その旨と理由を記載する。

② 医療経済チームを設置

費用対効果を考慮する場合は、システムティックレビューチームとは別に医療経済チームを設置することを推奨する。特に新規の医療経済評価により後述のICERを算出する場合などには、医療経済評価の専門家を含めることが望ましい。

③ 費用対効果プラン作成（5.5節参照）

診療ガイドライン統括委員会または診療ガイドライン作成グループと協議の上作成する。

④（行う場合は）医療経済評価のシステムティックレビューの実施（5.6節参照）

⑤ 新規の医療経済評価の実施（5.7節参照）

⑥ 評価結果のまとめ、診療ガイドライン統括委員会または診療ガイドライン作成グループへの報告（5.7.3項、5.8節、および第6章参照）

実施した場合は、テンプレート【EC-1】にシステムティックレビューの結果をまとめる。新規解析の結果はテンプレート【EC-2】のフォーマットなどを用いて、結果をとりまとめ、診療ガイドライン統括委員会または診療ガイドライン作成グループへ報告する。

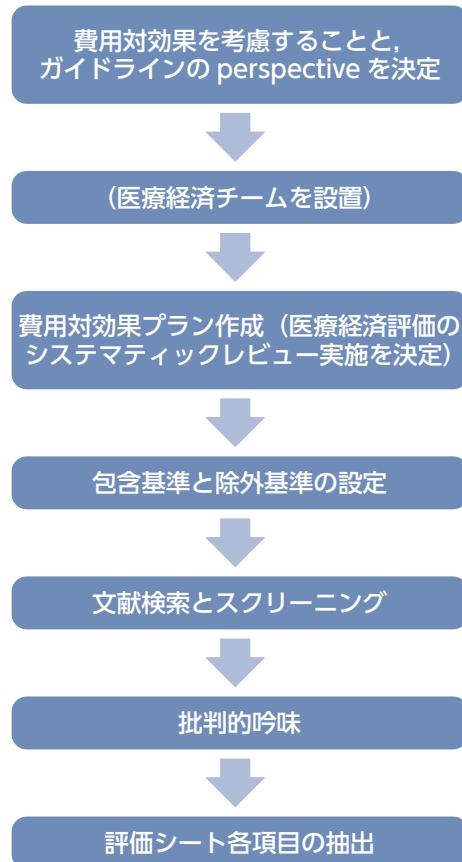


図 5-2 医療経済評価のシステムティックレビューの手順

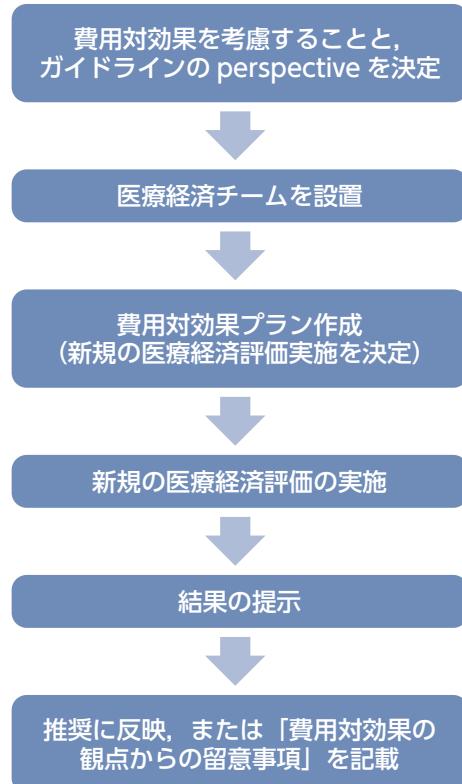


図 5-3 新規の医療経済評価の手順

● テンプレート

【EC-1 医療経済評価からのデータ抽出のための表（経済学的評価シート）】

【EC-2 分析結果のまとめ用フォーマット】

● テンプレート 記入方法

(なし)

● テンプレート 記入例

【EC-2 分析結果のまとめ用フォーマット 記入例】

5.5 費用対効果プラン

費用対効果プランには、医療経済評価のシステムティックレビューと、新規の医療経済評価の計画が含まれる。新規に医療経済モデルを作成し分析する場合など、全てのCQについて医療経済評価を行うことはリソースの面から見て、難しいことがある。このため、必要があれば表5-6を参考に、CQの重要性も加味して、最も優先順位の高い問い合わせ可能な範囲で評価を行う計画を立てるとよい。

表5-6 医療経済評価を実施する際の優先順位付けアルゴリズム

期待増分効果	期待増分費用	経済的エビデンスが現在の意思決定を変え得る可能性	経済的エビデンスを検討する優先度
小	小	可能性は低い	優先度は低い
小	小	可能性は高い	優先度は中程度
大	小	可能性は低い	優先度は高い
大	小	可能性は高い	優先度は低い
小	大	可能性は低い	優先度は中程度
小	大	可能性は高い	優先度は高い
大	大	可能性は低い	優先度は低い
大	大	可能性は高い	優先度は中程度

出典：Frick K, Neissen L, Bridges J, Walker D, Wilson RF, Bass EB. Usefulness of Economic Evaluation Data in Systematic Reviews of Evidence. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Oct. Report No: 12(13)-EHC114-EF. p.19 より作成

【費用対効果プラン（例）】

第〇回 診療ガイドライン作成グループ会議

診療ガイドライン名：周産期メンタルヘルス（update）

1. 費用対効果研究のシステムティックレビューの計画

対象 CQ：全ての CQ

組み入れ基準：費用便益分析、費用効用分析、費用効果分析、費用最小化分析

2. 新規解析対象のクエスチョン

A

対象	産後の女性
解析に含む予定の介入	1. EPDS のみ 2. Whooley questions の後に EPDS 3. Whooley questions の後に PHQ-9 4. 標準的な同定方法
解析の種類	費用効用分析（CUA）

B

対象	閾値下から軽症～中等症のうつ状態の産後女性
解析に含む予定の介入	1. サポート付きのセルフヘルプ 2. 傾聴的な訪問（listening visits） 3. 標準的ケア
解析の種類	費用効用分析（CUA）

出典：National Institute for Health and Care Excellence (NICE). Antenatal and postnatal mental health (update): health economic plan, Antenatal and postnatal mental health: clinical management and service guidance, 2014 (<https://www.nice.org.uk/guidance/cg192/documents/antenatal-and-postnatal-mental-health-update-health-economic-plan2>) ©National Collaborating Centre for Mental Health (NCCMH), ©Copyright National Institute for Health and Care Excellence. より作成

5.6 医療経済評価のシステムティックレビュー

経済的モデルを用いて評価を行う場合、必要に応じて医療経済評価のシステムティックレビューを行うことを推奨する。その際には、システムティックレビューから得られた結果を単純にわが国の環境に外挿するのではなく、新規に医療経済学的モデルを構築する際などにそれらの結果を活用することなどが想定される。

また、新規の医療経済評価の結果に、医療経済評価のシステムティックレビューの結果を併せて提示することができる。

医療経済研究のシステムティックレビューは、①「医療経済チーム」がまとめて担当する方法と、②トピックごとに各レビューチームが行う方法のどちらでも可能である。

包含基準は、費用効果分析、費用効用分析、費用便益分析、費用最小化分析とする方法が考えられる。除外基準には、費用分析を含めることを提案する（NICE 2020）。

5.6.1 医療経済評価の検索方法

近年の研究からは MEDLINE/PubMed、Embase、HTA Database の 3 つのデータベースを利用するabilia ことが有用とされ（Arber M 2018）、そのうち MEDLINE/PubMed、international HTA database^{*7} は無料で利用できる。また、2013 年以前の MEDLINE/PubMed および Embase の論文は NHS EED（Economic Evaluation Database）に収載されているため、NHS EED を検索対象に加えた場合は、2013 年以前の MEDLINE および Embase 収載論文の検索を省略できる。

以上に加えて、日本または日本語の医療経済評価を同定するため、①中医協による医療経済評価の報告書について、厚生労働省および保健医療科学院ウェブサイト内の検索を行うこと、および、②医療経済研究機構による「医療経済研究論文検索」で検索することが望ましい。より完全な、医療経済評価を含むデータベースの一覧は SuRe Info^{*8} から入手可能である。

(1) 包括的検索

診療ガイドラインのテーマを包括するキーワード群（通常は網羅的な疾患名）を用いる。経済学的研究に特化していないデータベースを検索する場合はそれらキーワード群と経済学的研究フィルターを AND で連結して検索を行う。

(2) 経済学的研究フィルターの例

通常は MEDLINE については、NHS EED フィルター（感度 0.999、適合度 0.040）または Wilczynski best optimization of sensitivity and specificity フィルター（感度 0.923、適合度 0.093）、Embase については NHS EED フィルター（感度 0.997、適合度 0.029）または McKinlay best optimization of sensitivity and specificity フィルター（**5.9 節**参照、感度 0.986、適合度 0.064）（Glanville J 2009）が、感度と適合度のバランスが良好なため勧められるが、目的に応じて他のフィルターを使用することも可能である。

* 7 international HTA database (<https://database.inahta.org>)

* 8 SuRe Info (<http://vortal.htai.org/?q=sure-info>)

上記以外では、医中誌 Web、および厚生労働省研究成果データベースを検索する方法も考えられるが、フィルターの性能を評価した研究が同定されず、効率的な検索方法の開発は発展途上である。HTA Database や、経済学的研究に特化したデータベースを検索する際には、経済学的研究フィルターも使用する必要はなく、診療ガイドラインのテーマを包括的に検索可能なキーワード群または特定の CQ に関連するキーワード群で検索を行うことができる。費用対効果研究のシステムティックレビューにおいて、優先して検索すべきデータベースを表5-7に示す。

検索に用いたデータベースとキーワードおよびフィルターは診療ガイドラインの作成方法の中、または別添で報告することが望ましい。

表 5-7 優先して検索すべきデータベース（医療経済評価）

- | |
|------------------|
| 1 MEDLINE/PubMed |
| 2 Embase（有料） |
| 3 HTA Database |
| 4 NHS EED |

5.6.2 医療経済評価の批判的吟味

レファレンスケースを参照しながら、テンプレート【EC-3】に従って批判的吟味を行う。GRADE アプローチでは費用効果分析のエビデンスの質を評価する枠組みがまだ発表されておらず、基本的には NICE の批判的吟味チェックリストの内容を採用した。チェックリストの前半（適用可能性）部分で、日本の状況に適用不可能と判断された場合は、その研究を除外する。もし少なくとも「ある程度適用可能」と評価された場合に、後半の「研究の限界・研究方法の質」の評価を行い、さらに、テンプレート【EC-1】の各項目について情報を抽出する。

● テンプレート

【EC-1 医療経済評価からのデータ抽出のための表（経済学的評価シート）】

【EC-3 医療経済評価論文のチェックリスト】

● テンプレート 記入方法

(なし)

● テンプレート 記入例

(なし)

5.7 新規の医療経済評価

新規の分析では、定式化されたリサーチクエスチョン（RQ）に基づき、日本における最新の状況を反映した費用の比較や費用効果分析を新たに実施することを推奨する。その場合、医療経済評価の専門家が評価を実施するか、専門家と協働して行う必要がある。

新規の解析では、レファレンスケースに定められた評価手法を含めることが推奨される。本マニュアルに記載のない詳細な分析手法などについては「中央社会保険医療協議会における費用対効果評価の分析ガイドライン 第2版」*9 が参照可能である。

5.7.1 費用対効果プランと新規解析の詳細計画

新規の解析を行うのに先立ち、**5.2節**の方針を参考して評価対象となる医療技術、対象となる（患者）集団、比較対照技術などについて、診療ガイドライン統括委員会または診療ガイドライン作成グループと協議し費用対効果プランを立てる。**テンプレート【EC-4】**を用いて、レファレンスケースを参照しながら、実施する医療経済評価の分析枠組みについて設定する。

臨床的な有効性・安全性や費用対効果から、分析結果が異なることが予測される集団があり、またそれらが生物学的あるいは臨床的に妥当なメカニズムや他の要因によって支持される場合には、それらの対象集団ごとに分析を行うことが望ましい。

検討過程での熟慮により、異なる対象集団を設定することなど（費用対効果プランからの変更）も許容される。複数の対象集団に分ける場合は、患者のベースライン情報に基づく場合と、治療への反応性に基づく場合とがあり得る（無効例については新しい医療技術に切り替える、有害事象が既存治療と同程度と考えられる集団を別の分析対象とする、など）。

5.7.2 評価対象となる医療技術の選定

医療経済評価の対象とする医療技術については、事前に明確に基準などを設定して選定する必要がある。関心のある技術のみを恣意的に対象とするのは避けるべきである。評価対象となる医療技術は、個々の医療技術である場合と、複数の医療技術を含む一連の治療（「Aが無効である場合はB」など）の場合があり得る。

Individual perspectiveで作成される診療ガイドラインにおいて、「正味のアウトカムが同程度と見なせる医療技術」を評価する場合には、どの医療技術が相互に「同程度」と見なせる可能性があるのか、あるいはどのような集団（一部の限定した集団など）に対してであれば「同程度」であると考えられるのか、費用対効果を検討する前のリサーチクエスチョンを作成する段階で臨床的に判断する必要がある。

「同程度」であるかどうかは、個々の患者の状況によって異なる可能性もあり、また最終的には診療ガイドラインの使用者の判断に委ねられている。よって、状況や考え方によっては「同程度」であると判断される「可能性がある」ものを含めて、情報提供という観点から幅広く取り扱うことを推奨する。

* 9 https://c2h.niph.go.jp/tools/guideline/guideline_ja.pdf

同様に、費用効果分析の実施までターゲットに含める場合は、「医療技術効果が十分に高い既存技術が存在しており、患者にとって新規医療技術の追加的アウトカムが小さい、あるいは大きな不確実性を有する」ものについても対象医療技術を同定する必要がある。

Population perspectiveに基づいた診療ガイドラインの場合は、より広範な医療技術を評価対象とすることが可能である。ただし、診療ガイドライン上のCQ全てに対応して、医療経済評価を実施することは現実的でない場合も多い。Individual perspectiveでも評価対象となるような「正味のアウトカムが同程度と見なせる」あるいは「新規医療技術の追加的アウトカムが小さい、あるいは大きな不確実性を有すると判断される医療技術」に加えて、「高額な医療技術」「予算への影響が大きい医療技術」「公衆衛生的な影響の大きい医療技術」「医療現場への負荷が大きい医療技術」など、集団の観点から必要性が高いものに限定して評価を行うことも検討できる。評価対象となる医療技術の選定については、恣意的にならないよう一定の基準を作成して、適用することが望ましい。

5.7.3 新規の医療経済評価の実施と結果の提示

費用最小化分析（効果が同等との前提の下での費用の比較、詳細は**5.3.2 項**参照）を含めて、新規の医療経済評価を実施する場合は、**5.3 節**（レファレンスケース）に定められた解析方法を含めることが推奨される。

診療ガイドラインにおいて、医療経済評価の情報を提示する際には、定式化されたリサーチクエスチョンに基づき、日本における最新の状況を反映した費用の比較や費用効果分析を新たに実施する（あるいは該当する文献がある場合はそれを参照する）ことが望ましい。

分析結果の概要については、**テンプレート【EC-2】**に記載する。結果の解釈に用いるための費用やICERの情報などを含める。結果の不確実性が大きいと考えられる場合は、ICERを1点ではなく、一定の範囲で示すこと、あるいは複数の値を提示することも検討する^{*10}。

また、中医協などの公的な機関により、評価対象となる医療技術に関連した評価が実施されている場合、その（追加的有効性や費用対効果に関する）評価内容や結論についても**テンプレート【EC-2】**にまとめる。

分析結果の詳細（用いた手法なども含め）について特に定まった書式などはないが、医療経済評価における報告様式については、CHEERS声明（Consolidated Health Economic Evaluation Reporting Standards）（Husereau D 2013）^{*11}と呼ばれるものが国際医薬経済・アウトカム研究学会（International Society for Pharmacoeconomics and Outcomes Research : ISPOR）により作成されており、参考となる。

*10 ICERは多くの場合に各種データを統合した結果として得られるものであり、通常の統計的推測におけるように信頼区間は算出しない（できない）ことが一般的である。

*11 CHEERS声明 日本語版 (<https://www.niph.go.jp/journal/data/62-6/201362060009.pdf>)

● テンプレート

【EC-2 分析結果のまとめ用フォーマット】

【EC-4 新規評価実施時の分析枠組みの設定】

● テンプレート 記入方法

(なし)

● テンプレート 記入例

【EC-2 分析結果のまとめ用フォーマット 記入例】

5.8 費用対効果のエビデンスの解釈

得られた結果について、「正味のアウトカムが同程度と見なせる」とされるものは、効果が同等であることが前提である。実際の患者にとって治療上同等であると考えられる場合には、一群の医療技術のうち最も安価なものの使用が費用対効果に優れると解釈される。また、全体での使用を最適化するために、適応集団の一部についてのみ、分析結果や推奨内容が異なっていてもよい。

一方で、individual perspectiveあるいはpopulation perspectiveにかかわらず、「ICER を用いた意思決定を実施することが適切と判断される治療」については、算出された ICER の値に基づき結果を解釈する。なお、評価対象技術のアウトカムが比較対照技術と比べて大きく、費用も小さい場合は優位とされ、ICER は算出しない。この場合、評価対象技術は費用対効果に優れていると判断できる。診療ガイドライン上で ICER（円/QALY）の閾値（基準値）、すなわちこの値以下なら費用対効果がよいと判断される基準値をいくらと設定するべきかについては、さまざまな議論があるため、今後の検討課題と考えられる。

ただし、十分な既存治療が存在するような状況では、中央社会保険医療協議会（中医協）で用いられている基準値の 500 万円/QALY は参考し得るものであるかもしれない。この値と比べて、ICER が極端に大きいもの、あるいは十分に小さいものについては、費用対効果のよしあしを判断し得るだろう。また、費用が小さく効果が大きい、いわゆる優位の状態、あるいはその逆に費用が大きく効果が小さい、いわゆる劣位の状態の場合は、意思決定は容易である。

ICER（円/QALY）の基準値は一律の値が機械的に適用されるものではなく、疾患領域（重症度や患者数など）や予算の制約、その領域すでに使用されている医療技術などの費用など諸状況により異なるべきという議論もあることに留意が必要である。

結果の解釈の次のステップとして、推奨作成時に費用対効果を考慮することになるが、それについては**第6章**に記載する。

5.9 補足資料：フィルターの例

• NHS EED MEDLINE フィルター (OvidSP)

```
1 Economics/
2 exp "costs and cost analysis"/
3 Economics, Dental/
4 exp economics, hospital/
5 Economics,cdxz Medical/
6 Economics, Nursing/
7 Economics, Pharmaceutical/
8 (economic$ or cost or costs or costly or costing or price or prices or pricing or
pharmacoconomic$).ti,ab.
9 (expenditure$ not energy).ti,ab.
10 value for money.ti,ab.
11 budget$.ti,ab.
12 or/1-11
13 ((energy or oxygen) adj cost).ti,ab.
14 (metabolic adj cost).ti,ab.
15 ((energy or oxygen) adj expenditure).ti,ab.
16 or/13-15
17 12 not 16
18 letter.pt.
19 editorial.pt.
20 historical article.pt.
21 or/18-20
22 17 not 21
23 exp animals/ not humans/
24 22 not 23
25 bmj.jn.
26 "cochrane database of systematic reviews".jn.
27 health technology assessment winchester england.jn.
28 or/25-27
29 24 not 28
30 limit 29 to yr="2010 -Current"
```

出典 : Centre for Reviews and Dissemination. NHS EED MEDLINE using OvidSP (<https://www.crd.york.ac.uk/crdweb/searchstrategies.asp>)
Copyright © 2020 University of York

• Wilczynski best optimization of sensitivity and specificity フィルター (Ovid MEDLINE)

```
cost-benefit analysis.sh. OR costs.tw. OR cost effective.tw.
```

出典 : Wilczynski NL, et al. (2004) Optimal search strategies for detecting health services research studies in MEDLINE. CMAJ 171:1179-1185.

• NHS EED Embase フィルター (OvidSP)

1. Health Economics/
2. exp Economic Evaluation/
3. exp Health Care Cost/
4. pharmacoeconomics/
5. 1 or 2 or 3 or 4
6. (econom\$ or cost or costs or costly or costing or price or prices or pricing or pharmacoconomic\$).ti,ab.
7. (expenditure\$ not energy).ti,ab.
8. (value adj2 money).ti,ab.
9. budget\$.ti,ab.
10. 6 or 7 or 8 or 9
11. 5 or 10
12. letter.pt.
13. editorial.pt.
14. note.pt.
15. 12 or 13 or 14
16. 11 not 15
17. (metabolic adj cost).ti,ab.
18. ((energy or oxygen) adj cost).ti,ab.
19. ((energy or oxygen) adj expenditure).ti,ab.
20. 17 or 18 or 19
21. 16 not 20
22. animal/
23. exp animal experiment/
24. nonhuman/
25. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).ti,ab,sh.
26. 22 or 23 or 24 or 25
27. exp human/
28. human experiment/
29. 27 or 28
30. 26 not (26 and 29)
31. 21 not 30
32. 0959-8146.is.
33. (1469-493X or 1366-5278).is.
34. 1756-1833.en.
35. 32 or 33 or 34
36. 31 not 35
37. conference abstract.pt.
38. 36 not 37
39. limit 38 to yr="2010 -Current"

出典 : Centre for Reviews and Dissemination. NHS EED EMBASE using OvidSP (<https://www.crd.york.ac.uk/crdweb/searchstrategies.asp>)
Copyright © 2020 University of York

• McKinlay best optimization of sensitivity and specificity フィルター (Embase)

cost.tw. OR costs.tw.

出典 : McKinlay RJ, et al. (2006) Optimal search strategies for detecting cost and economic studies in EMBASE. BMC Health Serv Res 6:67.

• CADTH's Database search filters: economic evaluations/cost/economic models—PubMed

出典 : CADTH's Database search filters (economic evaluations/cost/economic models- Pubmed: <https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#eco>). Reproduced with permission by CADTH.

Economics[Mesh:NoExp] OR "Costs and Cost Analysis"[mh] OR Economics, Nursing[mh] OR Economics, Medical[mh] OR Economics, Pharmaceutical[mh] OR Economics, Hospital[mh] OR Economics, Dental[mh] OR "Fees and Charges"[mh] OR Budgets[mh] OR budget*[tiab] OR economic*[tiab] OR cost[tiab] OR costs[tiab] OR costly[tiab] OR costing[tiab] OR price[tiab] OR prices[tiab] OR pricing[tiab] OR pharmacoeconomic*[tiab] OR pharmaco-economic*[tiab] OR expenditure[tiab] OR expenditures[tiab] OR expense[tiab] OR expenses[tiab] OR financial[tiab] OR finance[tiab] OR finances[tiab] OR financed[tiab] OR value for money[tiab] OR monetary value*[tiab] OR models, economic[mh] OR economic model*[tiab] OR markov chains[mh] OR markov[tiab] OR monte carlo method[mh] OR monte carlo[tiab] OR Decision Theory[mh] OR decision tree*[tiab] OR decision analy*[tiab] OR decision model*[tiab]

• MacMaster University Health Information Research Unit: Evidence-Based Health Informaticsによる Economics フィルター (MEDLINE) の PubMed 版

(感度最大化)

randomized controlled trial[Publication Type] OR cost effectiveness[Title/Abstract] OR cost effective[Title/Abstract]

(特異度最大化)

cost effective[Title/Abstract] OR sensitivity analys*[Title/Abstract]

(感度・特異度ベストバランス)

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH Term] OR health care costs[MeSH:noexp]

出典 : McMaster University Health Information Research Unit Search Filters for MEDLINE in Ovid Syntax and the PubMed translation.
"Economics - MEDLINE". https://hiru.mcmaster.ca/hiru/HIRU_Hedges_MEDLINE_Strategies.aspx (2020年10月15日閲覧) Copyright © 2004 - 2021 · Health Information Research Unit より作成

5章 テンプレート

【EC-1 医療経済評価からのデータ抽出のための表(経済学的評価シート)】

(システムティックレビューと新規解析の結果まとめ)

論文 ID /タイトル	限界 (軽度の限界がある/潜在的に大きな問題を含む/重大な限界がある)	適用可能性 (適用可能/ある程度適用可能/適用不可能)	その他 コメント	増分費用	増分効果	増分費用 効果	不確実性

【EC-2 分析結果のまとめ用フォーマット】

対応する CQ	
設定した費用対効果のリサーチクエスチョン (RQ)	
評価対象技術	[RQ の I を記載する]
評価対象集団	[RQ の P を記載する]
分析の立場	<input type="checkbox"/> 公的医療の立場 <input type="checkbox"/> その他（括弧内に具体的に記載する）
分析手法	<input type="checkbox"/> 費用最小化分析 [効果を同等とおく場合] <input type="checkbox"/> 費用効果分析 [ICER を算出する場合]
費用の算出方法	<input type="checkbox"/> 医療技術の公定価格（単価）のみ <input type="checkbox"/> 一定期間における関連する医療費（積み上げ） <input type="checkbox"/> 一定期間における関連する医療費（レセプト） <input type="checkbox"/> モデリングを用いた推計
単価（価格）の設定時点	年　　月　　時点の価格
比較対照技術	[RQ の C を記載する]
分析期間	<input type="checkbox"/> 単価の比較、あるいは短期間における医療費の比較 <input type="checkbox"/> 生涯 <input type="checkbox"/> その他（括弧内に具体的に記載する）
効果指標	<input type="checkbox"/> 特に設定せず [効果同等の場合] <input type="checkbox"/> QALY <input type="checkbox"/> その他（括弧内に具体的に記載する）
割引率	<input type="checkbox"/> 割引せず [短期間での分析] <input type="checkbox"/> 年率 2% <input type="checkbox"/> その他（括弧内に具体的に記載する）
分析結果の概要	
関連する公的機関等による評価結果	<input type="checkbox"/> なし <input type="checkbox"/> あり（以下に具体的な評価結果を記載する）
分析結果の解釈	

【EC-3 医療経済評価論文のチェックリスト】

著者名、タイトル、出典、出版年		
CQ :	CQ 番号	
チェックリストの記入者：		
セクション1：適用可能性 それぞれのCQ、およびレファレンスケースに 対して適切か。 (このセクションをまず用いて適さない研究を除外する。)	はい／一部該当／いいえ／ 不明／当てはまらない	コメント
1.1 研究対象がCQに対して適切か。ベース ラインリスクが日本で推奨の対象となる 集団の値と異なっていないか。		
1.2 CQに対して適切な全ての介入が含まれ ているか。 ・全ての適切な（より費用対効果に優れている 可能性のある）介入が含まれているか。 ・不適切な介入との比較により ICER が計算さ れていないか。		
1.3 研究が行われた地域の医療システムが十 分に現在の日本の状況に類似しているか。		
1.4 費用の立場（perspective）がCQに対 して適切か。		
1.5 アウトカムの視点（perspective）が CQに対して適切か。また、その視点に おいて含まれるべきアウトカムが過不足 なく含まれているか。		
1.6 全ての未来の費用とアウトカムが適切に 割引されているか。		
1.7 QALYの値はレファレンスケースで示さ れた方法で測定されているか。そうでな い場合は、分析の立場に則して、使われ たアウトカムと使われた理由を説明する。		
総合判断：適用可能／ある程度適用可能／適用不可能 (適用不可能と判断された場合はセクション2による評価は不要。)		
その他コメント：		

セクション2：研究の限界・研究方法の質 (このセクションは十分に適用可能と判断された研究にのみ用いる。)	はい／一部該当／いいえ／ 不明／当てはまらない	コメント
2.1 モデルの構造がCQの特徴を適切に反映しているか。 ・仮定／前提 (assumptions) は適切か。		
2.2 分析期間は重要な全ての費用とアウトカムを反映するのに十分な期間か。		
2.3 全ての適かつ重要なアウトカムが分析に含まれているか (副作用によるものを含む)。		
2.4 ベースラインのアウトカムの推定値は入手可能で最も適切な情報源から得ているか。		
2.5 介入の比較効果を入手可能で最も適切な情報源から得ているか。		
2.6 全ての重要かつ適した費用が解析に含まれているか (副作用によるものを含む)。		
2.7 資源の利用を入手可能で最も適切な情報源から得ているか。		
2.8 資源の単価を入手可能で最も適切な情報源から得ているか。		
2.9 適切な増分分析の結果が示されているか。あるいは、データから計算が可能か。		
2.10 不確実性を持つ重要なパラメータ全てが適切な感度分析の対象になっているか。各パラメータが持つ不確実性が示され、非精確性 (imprecision) がエビデンスに基づいているか。		
2.11 未公表の潜在的な経済的利益相反がないか。		
総合評価：軽度の限界がある／潜在的に大きな問題を含む／重大な限界がある		
その他コメント：		

出典：National Institute for Health and Care Excellence (NICE). Process and methods guides Developing NICE guidelines: the manual, 2014 (<https://www.nice.org.uk/media/default/about/what-we-do/our-programmes/developing-nice-guidelines-the-manual.pdf>) ©Copyright National Institute for Health and Care Excellence, 2014. より作成

【EC-4 新規評価実施時の分析枠組みの設定】

評価対象技術	
評価対象集団	
分析の立場	
比較対照技術	
分析期間	
効果指標	
割引率	

5章 テンプレート 記入例

【EC-2 分析結果のまとめ用フォーマット 記入例】

対応する CQ	その他の併存疾患がない慢性疾患 X の患者に対して Y 薬効群を用いた薬物療法は有効か。
設定した費用対効果のリサーチクエスチョン (RQ)	その他の併存疾患がない慢性疾患 X の患者に対して、Y 薬効群の医薬品 (A, B, C) のうち、どれを用いるのが最も安価であるか。
評価対象技術	Y 薬効群の医薬品である A, B, C
評価対象集団	併存疾患がない慢性疾患 X の患者
分析の立場	<input checked="" type="checkbox"/> 公的医療の立場 <input type="checkbox"/> その他（括弧内に具体的に記載する）
分析手法	<input checked="" type="checkbox"/> 費用最小化分析 [効果を同等とおく場合] <input type="checkbox"/> 費用効果分析 [ICER を算出する場合]
費用の算出方法	<input checked="" type="checkbox"/> 医療技術の公定価格（単価）のみ <input type="checkbox"/> 一定期間における関連する医療費（積み上げ） <input type="checkbox"/> 一定期間における関連する医療費（レセプト） <input type="checkbox"/> モデリングを用いた推計
単価（価格）の設定時点	2020年 5月 時点の価格
比較対照技術	A, B, C を相互に比較する
分析期間	<input checked="" type="checkbox"/> 単価の比較、あるいは短期間ににおける医療費の比較 <input type="checkbox"/> 生涯 <input type="checkbox"/> その他（括弧内に具体的に記載する）
効果指標	<input checked="" type="checkbox"/> 特に設定せず [効果同等の場合] <input type="checkbox"/> QALY <input type="checkbox"/> その他（括弧内に具体的に記載する）
割引率	<input checked="" type="checkbox"/> 割引せず [短期間での分析] <input type="checkbox"/> 年率 2% <input type="checkbox"/> その他（括弧内に具体的に記載する）
分析結果の概要	医薬品 A, B, C の 1 日薬価を比較したところ、それぞれ 350.8 円、257.6 円、432.9 円であった。
関連する公的機関等による評価結果	<input type="checkbox"/> なし <input checked="" type="checkbox"/> あり (3つの医薬品のうち、医薬品 C は費用対効果評価の対象となっていた。医薬品 C は医薬品 B と比較してイベント回避を効果指標とした追加的有効性が示されておらず、費用は高かった。)
分析結果の解釈	患者にとって治療上同等であると考えられる場合には、まずは最も安価な医薬品 B の使用が最も費用対効果に優れる。

5章 文献

- Alonso-Coello P, et al. (2016) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 353:i2016. doi: [10.1136/bmj.i2016](https://doi.org/10.1136/bmj.i2016). PMID: [27353417](#)
- Ara R, et al. (2017) The identification, review and synthesis of health state utility values from the literature. Pharamcoeconomics 35:43-55. doi: [10.1007/s40273-017-0547-8](https://doi.org/10.1007/s40273-017-0547-8). PMID: [29052156](#)
- Arber M, et al. (2016) Performance of search filters to identify health state utility studies. Value in Health 19:A390-A391.
- Arber M, et al. (2017) Performance of ovid MEDLINE search filters to identify health state utility studies. Int J Technol Assess Health Care 33:472-480. doi: [10.1017/S0266462317000897](https://doi.org/10.1017/S0266462317000897). PMID: [29065942](#)
- Arber, M. et al. (2018) Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care 34:547-554. doi: [10.1017/S0266462318000636](https://doi.org/10.1017/S0266462318000636). PMID: [30442221](#)
- Arber M, et al. (2018) Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care 34:547-554. doi: [10.1017/S0266462318000636](https://doi.org/10.1017/S0266462318000636). PMID: [30442221](#)
- Brunetti M, et al. (2013) GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol 66:140-150. doi: [10.1016/j.jclinepi.2012.04.012](https://doi.org/10.1016/j.jclinepi.2012.04.012). PMID: [22863410](#)
- Drummond MF, et al. (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ 313:275-283. doi: [10.1136/bmj.313.7052.275](https://doi.org/10.1136/bmj.313.7052.275). PMID: [8704542](#)
- Drummond MF, et al. (2015) Methods for Economic Evaluation of Health Care Programmes. Fourth ed. Oxford University Press
- European network for Health Technology Assessment. (2020) METHODOLOGICAL GUIDELINES: Practical considerations when critically assessing economic evaluations. https://eunethta.eu/wp-content/uploads/2020/03/EUNetHTA-JA3WP6B2-5-Guidance-Critical-Assessment-EE_v1-0.pdf (参照 2020/10/25)
- Frick K, et al. (2012) Usefulness of Economic Evaluation Data in Systematic Reviews of Evidence [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US) <https://www.ncbi.nlm.nih.gov/books/NBK114533/> (参照 2020/10/15)
- Glanville J, et al. (2009) How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. Int J Technol Assess Health Care 25:522-529. doi: [10.1017/S0266462309990523](https://doi.org/10.1017/S0266462309990523). PMID: [19845982](#)
- Husereau D, et al. (2013) Consolidated Health Economic Evaluation Reporting Standards (CHEERS)-- explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 16:231-250. doi: [10.1016/j.jval.2013.02.002](https://doi.org/10.1016/j.jval.2013.02.002). PMID: [23538175](#)

ISSG Search Filters Resource. <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home> (参照 2020/10/25)

Luce BR, et al. (1996) Estimating costs in cost-effectiveness analysis. In: Gold MR, et al. editors. Cost-effectiveness in health and medicine. Oxford University Press

McKinlay RJ, et al. (2006) Optimal search strategies for detecting cost and economic studies in EMBASE. BMC Health Serv Res 6:67. doi: [10.1186/1472-6963-6-67](https://doi.org/10.1186/1472-6963-6-67). PMID: [16756662](https://pubmed.ncbi.nlm.nih.gov/16756662/)

National Collaborating Centre for Mental Health (NCCMH). (2014) Antenatal and postnatal mental health (update): clinical management and service guidance. <https://www.nice.org.uk/guidance/cg192/evidence> (参照 2020/10/25)

National Institute for Health and Care Excellence (NICE). (2020) Developing NICE guidelines: the manual. [PMG20] <https://www.nice.org.uk/process/pmg20/> (参照 2020/10/15)

Paisley S. (2016) Identification of evidence for key parameters in decision-analytic models of cost effectiveness: a description of sources and a recommended minimum search requirement. Pharmacoeconomics 34:597-608. doi: [10.1007/s40273-015-0372-x](https://doi.org/10.1007/s40273-015-0372-x). PMID: [26861793](https://pubmed.ncbi.nlm.nih.gov/26861793/)

Papaioannou D, et al. (2010) NICE DSU Technical support document 9: the identification, review and synthesis of health state utility values from the literature. <https://pubmed.ncbi.nlm.nih.gov/28481496/> (参照 2020/10/15)

Philips Z, et al. (2006) Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics 24:355-371. doi: [10.2165/00019053-200624040-00006](https://doi.org/10.2165/00019053-200624040-00006). PMID: [16605282](https://pubmed.ncbi.nlm.nih.gov/16605282/)

Shemilt I, et al. (2019) Chapter 20: Economic evidence. In: Higgins JPT, et al. editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). <https://www.training.cochrane.org/handbook> (参照 2020/10/15)

ScHARRHUD (Health Utilities Database) <http://www.scharrhud.org/> (参照 2021/03/06)

Thielen FW, et al. (2016) How to prepare a systematic review of economic evaluations for clinical practice guidelines: database selection and search strategy development (part 2/3). Expert Rev Pharmacoecon Outcomes Res 16:705-721. doi: [10.1080/14737167.2016.1246962](https://doi.org/10.1080/14737167.2016.1246962). PMID: [27805466](https://pubmed.ncbi.nlm.nih.gov/27805466/)

van Mastrigt GA, et al. (2016) How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3). Expert Rev Pharmacoecon Outcomes Res 16:689-704. doi: [10.1080/14737167.2016.1246960](https://doi.org/10.1080/14737167.2016.1246960). PMID: [27805469](https://pubmed.ncbi.nlm.nih.gov/27805469/)

Wilczynski NL, et al. (2004) Optimal search strategies for detecting health services research studies in MEDLINE. CMAJ 171:1179-1185. doi: [10.1503/cmaj.1040512](https://doi.org/10.1503/cmaj.1040512). PMID: [15534310](https://pubmed.ncbi.nlm.nih.gov/15534310/)

Wijnen B, et al. (2016) How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: data extraction, risk of bias, and transferability (part 3/3). Expert Rev Pharmacoecon Outcomes Res 16:723-732. doi: [10.1080/14737167.2016.1246961](https://doi.org/10.1080/14737167.2016.1246961). PMID: [27762640](https://pubmed.ncbi.nlm.nih.gov/27762640/)

第6章 推奨

Minds

6章 Key Words

- ◎ Consensus Development Conference
- ◎ Delphi (デルファイ) 法
- ◎ EtD フレームワーク (evidence to decisions framework)
- ◎ GRADE (The Grading of Recommendations Assessment, Development and Evaluation)
- ◎ GRADE グリッド (GRADE grid)
- ◎ individual perspective (個人の立場)
- ◎ NGT 法 (Nominal Group Technique)
- ◎ PICO
- ◎ population perspective (集団の立場)
- ◎ SoF 表 (summary of findings table)
- ◎ WHO-INTEGRATE フレームワーク (WHO-INTEGRATE Evidence framework)
- ◎ アウトカムの重要性
- ◎ 益と害 (benefit and harm)
- ◎ 益と害のバランス
- ◎ エビデンスの確実性
- ◎ エビデンスの確実性 (強さ)
- ◎ エビデンス総体 (body of evidence)
- ◎ 外部評価
- ◎ 感度分析 (sensitivity analysis)
- ◎ 観察研究 (observational study)
- ◎ クリニカルクエション (clinical question : CQ)
 - 決定木 (decision tree)
- ◎ 今後の研究 (future research)
 - 今後の研究が推奨される臨床疑問 (future research question)
- ◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
- ◎ システマティックレビューチーム (systematic review team : SR チーム)
 - システムティックレビューレポート (サマリーレポートともいう)
- ◎ 重要臨床課題 (key clinical issues)
- ◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
- ◎ 推奨
- ◎ スコープ (scope)

◎ 多基準決断分析 (multi-criteria decision analysis : MCDA)

フォーマルな合意形成 (formal consensus formation)

◎ メタアナリシス (meta-analysis, メタ解析と同義)

◎ ランダム化比較試験 (randomized controlled trial : RCT)

◎ リスク差 (risk difference : RD)

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】

SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】

RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】

PO : post 【第8章 診療ガイドライン公開後の取り組み】

6.0 概要

診療ガイドライン作成グループ（guideline development group : GDG, 診療ガイドライン作成パネルともいう）は、システムティックレビューチーム（systematic review team : SR チーム）がエビデンス評価を行った後、その結果やその他の資料を用いて推奨、およびその強さを決定する。その際、エビデンスの確実性（強さ）がそのまま推奨の強さになるのではない。システムティックレビュー（systematic review : SR, 系統的レビューと同義）の結果やその他の資料を用いて、原則として合意形成のための診療ガイドライン作成パネル会議を行う。推奨作成に重要なことは、診療ガイドライン作成グループ参加者の判断の偏りを防ぐ合意形成方法を用いることである。

この場合、診療ガイドライン作成委員は、合意形成過程で行われた投票等の結果の提示や、推奨を作成する方法と最終的な決定に至るまでの方法、経過を診療ガイドライン（clinical practice guideline : CPG）に記載することが望ましい。

○ ステップ――

- (1) 推奨作成の具体的方法の決定
- (2) 推奨文草案の作成
- (3) 推奨の作成、推奨の強さの判定
- (4) 解説の執筆
- (5) 一般向けサマリーの執筆
- (6) GRADE アプローチを用いる場合の資料
- (7) その他の意思決定に関するフレームワーク

6.1 ステップ1：推奨作成の具体的方法の決定

「推奨作成の基本方針」についてはスコープ（scope）で決められているが、推奨作成の作業に入る前に、ここでさらに具体的、詳細な方法について決定しておく。

● 手順

(1) 推奨、推奨の強さの表現方法について確認・決定する

診療ガイドライン作成グループは、委員全員の共通認識として、推奨文、および、推奨の強さの表現について、あらかじめ確認しておく。

推奨について、特定の介入の実施／非実施が問題となっている場合は、「実施することを推奨する」もしくは「実施しないことを推奨する」、また、3つ以上の選択肢（A, B, C, ……）が問題となる場合は、「Aを実施することを推奨する」「Bを実施することを推奨する」もしくは、「Cを実施することを推奨する」という提示方法とする。

推奨の強さ、および、推奨文の記載方法の例を以下に挙げておく。（ ）内は、各診療ガイドライン作成グループの判断で記載可能とする。

① 推奨の強さの記載方法

推奨の強さ：強く推奨する
推奨の強さ：弱く推奨する（提案する、条件付きで推奨する）
(推奨の強さ「なし」：明確な推奨ができない)

② それぞれの診療ガイドライン利用者に対する強い推奨と弱い推奨の意味するところ

(Andrews JC 2013b)

1. 患者に対して

- (1) **強い推奨**：その状況において、ほとんどの患者はこの推奨された行動方針を望むだろう。希望しない患者は少数にすぎない。
(2) **弱い推奨**：その状況において、多数の患者はこの提案された行動方針を望むだろう。しかし、そうでない患者も少なくない。

2. 臨床家に対して（臨床に用いるには）

- (1) **強い推奨**：臨床家は、原則としてこの推奨された行動方針を受け入れるべきである。

診療ガイドラインに基づく推奨の活用度合いは、（医療の）質指標として、あるいは診療ガイドラインの有効性の指標として用いられることがある。個々の臨床家が彼らの価値観や希望に合わせるのを援助するための公式の意思決定の方法（ガイド）は、不要である。

(2) **弱い推奨**：異なる患者に対しては、異なる選択肢が必要であることを認識する。そして、臨床家は、それぞれの患者が自分の価値観や希望に合った診療を決定できるように援助しなければならない。個々の臨床家が彼らの価値観や希望に合わせた意思決定をするのを援助するための公式の意思決定の方針（ガイド）は、有用かもしれない。治療方針を決定するときには、臨床家は患者とともに多くの時間を費やすことを考慮すべきである。

3. 政策決定者に対して（政策に用いるには）

(1) **強い推奨**：ほとんどの状況において政策に適用可能である。これには、診療ガイドラインの有効性の評価指標として用いることも含まれる。

(2) **弱い推奨**：政策立案に用いる場合には、多くの利害関係者の参加と相当な議論を要する。

また、政策には、地域による差が存在すると思われる。診療ガイドラインの有効性の指標として用いるには、診療の選択肢のひとつとして十分な議論がなされた結果かということに注意する必要がある。

③ 推奨文の記載方法

推奨文は、上記推奨の強さ①に、エビデンスの確実性（強さ）を併記する。

推奨の強さ	エビデンスの確実性
強い 弱い	強い 中程度 弱い 非常に弱い

以上の内容が読者に理解されるように推奨文を記載する。

例 1) 患者 P に対して治療 I を行うことを推奨する

= (強い推奨、エビデンスの確実性が強い)

例 2) 患者 P に対して治療 I を行うことを条件付きで推奨する

= (弱い推奨、エビデンスの確実性が弱い)

例 3) 患者 P に対して治療 I を行わないことを提案する

= (弱い推奨、エビデンスの確実性が非常に弱い)

例 4) 患者 P に対して治療 I を行わないことを強く推奨する

= (強い推奨、エビデンスの確実性が中程度)

推奨文の内容によっては、このような記載方法が難しい場合がある。そのようなときは、臨床場面で誤解なく理解されるように、文脈に沿った自然な表現になるよう工夫する必要がある。

なお、推奨の提示においては、「患者 P に対して治療 I を行うことを推奨しない」という表現は用いない。実施すべきでないという推奨を提示するのであれば、「患者 P に対して治療 I を行わないことを推奨する」と記載する。また、推奨文は診療行為をするかどうかの指針を示すものであるので、「患者 P に対して治療 I は有効である」という文言も不適切である。

診療ガイドラインの perspective を個人の立場 (individual perspective) とする場合, [集団の立場 (population perspective) で評価した] 費用対効果の観点からの留意事項は推奨の強さには直接反映しないことを原則とする (第5章参照)。

(2) 推奨を決定する方法（方式）の確認

決定方法は、フォーマルな合意形成方法 (formal consensus formation) (Delphi 法, NGT 法, GRADE グリッド, 他) やその他の合意形成方法のいずれを用いてもよい。この場合、合意しなかった部分やその解決法も明記しなければならない。

推奨決定会議進行の具体的な方法については、6.3節を参照のこと。

以下では、推奨を決定する方法の例を挙げる。

① Delphi (デルファイ) 法 (Dalkey N 1969)

まずは検討すべき事項について適切な情報を与えられた専門家が個別に評価し、その結果を匿名で共有する。その後、再度個別に評価を行い、その結果からコンセンサス、および、合意／不一致などを決定する。途中でパネルの討議を行う機会は設けない (第3章 3.3節参照 : Delphi 法)。

② 修正 Delphi 法 : RAND/UCLA Appropriateness Method (RAM) (Fitch K 2001)

まずは検討すべき事項について適切な情報を与えられた専門家が個別に評価を行い (第1ラウンド), その評価結果を資料とした会議での討議後に、再度個別に評価を行う (第2ラウンド)。第2ラウンドの結果として得られた中央値を基に、推奨についてのコンセンサスを決定する。その際、二極化した場合などには、合意ではなく不一致と判断する。

③ NGT 法 (Nominal Group Technique) (Delbecq AL 1971)

検討すべき事項について、パネル全体が参加する会議でそれぞれの考えを発表し、そこで取り上げられたことの重要性をパネルが個別に評価する。その結果を集計して、コンセンサスを形成する。

④ GRADE グリッド (GRADE grid) (Jaeschke R 2008)

検討すべき事項について、診療ガイドライン作成グループの委員全体会議でそれぞれの考えを発表し、その結果を基に投票を行う (6.6節参照)。投票結果の判定方法は事前に検討し決定しておく (下記 (3) 項参照)。投票項目は強く推奨するか弱く推奨するかの2段階で、さらに「実施する」ことを推奨するか、実施しないことを推奨するかの分類となる。複数回の投票を行っても意見の集約が得られない場合は、まれに推奨の強さを提示できないことも想定されるが、可能な限り明確に提示することが望ましい。

⑤ Consensus Development Conference

検討すべき事項についてパネル全体が参加する会議で、互いに許容可能なコンセンサスを作る義務が負わされて、文献レビュー、プレゼンテーション、聴衆による討議を経て、報告書に対する参加者の合意をもって終了とする。NIH Consensus Development Conference が代表的なものであるが、その方法は各国で独自に進化している。

⑥ その他の合意形成会議

全員参加の検討会議を経て、合議制で総合的な結果を導く。発展的な討議により、予想を超えた推奨文草案が提示される可能性もある。

(3) 投票などによる基準を確認・決定する

投票などによって推奨の強さなどを決定する場合には、診療ガイドライン作成グループの全会一致とはならないことが多い。したがって、診療ガイドライン作成グループ全体の意見として決定するための基準をあらかじめ確認・決定しておくことが望ましい。

推奨の強さを決定する際には、全体としての同意が得られなかった場合にやむを得ず「推奨なし」と決定せざるを得ないこともある。しかし、医療現場では何らかの決断を行う必要があり、可能な限り推奨を提示する努力をすることが望ましい。

例) 草案を決定する基準は、委員会メンバーの 75 %以上が投票し、その 80 %以上が賛成するものとした。コンセンサスが得られなかつた草案は、草案委員会メンバーがフィードバックを行い、3 回まで検討を繰り返す (Andrew R 2017)。

例) 多数の投票が特定の方向に賛成であり、20%以下が反対の方向に投票された（中立的な投票も含む）場合、多数の投票が得られた案（特定の診療を実施すること、もしくは実施しないこと）を推奨する。特定の推奨を強いとするには、70%以上が「強い」と投票する必要がある。もし、「強い」を支持する投票が 70%未満であった場合は、推奨は強さの分類において「弱い」とする (Jaeschke R 2008, Dellinger RP 2008)。

(参考)

EtD フレームワーク (evidence to decisions framework) を用いて、以下の方法で推奨のタイプを決定する (2019 年 11 月第 1 回 Minds Tokyo GRADE Center Workshop の方法を一部改変)。

- ① 80%以上の票が「強い」推奨に集中したら、「強い」推奨とする。
- ② ①の条件は満たさないが、80%以上の票が特定の方向に集中した場合、一方の「条件付き」推奨とする。
- ③ ①②の条件は満たさないが、「当該介入または比較対照のいずれかについての条件付き推奨」に 80% 以上の票が集中した場合、「当該介入または比較対照のいずれかについての条件付き推奨」とする。
- ④ ①～③ともに条件を満たさない場合は、結果を公表した上で討論し、再投票する。
- ⑤ 本行程を 3 回繰り返しても決定できない場合は、「推奨なし」とする。

6.2 ベース 2：推奨文草案の作成

診療ガイドライン作成グループで推奨、推奨の強さを決定するための資料として、**テンプレート【RC-1】**（または**【RC-2】**）を作成する。

テンプレート【RC-1】（または**【RC-2】**）は推奨作成、推奨の強さの決定を行う際の資料となるものである。本資料の作成を誰が担当するべきかについては作成組織の構成などによって異なるが、システムティックレビューの結果を受けてクリニカルクエスチョン（clinical question : CQ）の担当者が作成する、システムティックレビューを担当した人が委員長や診療ガイドライン作成グループの中の担当者などのチェックを受けつつ作成する、などの方法が考えられる。

推奨文を執筆する際には、診療ガイドラインが活用される場面、利用者が望む情報を想定した記載が望まれる。とりわけ、明確さ（SIGN 2011: 40）は重要であり、標準的な形式で記述し、さらに強い推奨を提示する場合には、その実施割合の程度が評価できるように、明確にされるべきである（IOM 2011: 131）。

● 手順

（1）推奨に関する価値観や意向を提示する

推奨文草案【テンプレート【RC-1】（または【RC-2】）】を作成するにあたっては、当該CQに対して提示された全アウトカムの中で、エビデンスの評価結果を受けて、あらためてアウトカムの重要性について相対的に評価し、重大（main outcomes）と考えられる価値を有するアウトカムを中心として作成する。場合によっては、全アウトカムの総まとめを推奨文として表記してもよい。

例) 癌疾患において、致死率減少というアウトカムを重要視した。

（2）CQに対するエビデンス総体の総括を記載する

システムティックレビューチームが作成したエビデンス総体（body of evidence）の作業シートを用い、アウトカムごとに評価されたエビデンス総体の「エビデンスの確実性（強さ）」を統合して、CQに対するエビデンス総体の総括（overall evidence）を提示する（表6-1 参照）。さらに、その判定の具体的な資料となった事項や効果推定値などについても記載する。

エビデンス総体の総括は、CQとその推奨に対するエビデンスの強さを意味し、「検索し統合されたエビデンス総体が、推奨を支えるものとしてどれほど適切か」を示している。

エビデンス総括評価は、上記の推奨文草案作成時に用いた「重大と考えられる価値を有するアウトカム」を用いて、エビデンス総体のエビデンスの確実性（強さ）を考慮して決定される。

エビデンスの確実性（強さ）決定の基本原則は、その治療効果推定値に対する作成者の確信が、ある特定の推奨を支持する上で、どの程度十分かである。

GRADEでは、重大とされたアウトカムの中から、一番弱いエビデンスの確実性（強さ）をエビデンス総体の総括としてのエビデンスの確実性（強さ）として採用するとしている。

表 6-1 CQ に対するエビデンス総体の総括

A (強) :	効果の推定値が推奨を支持する適切さに強く確信がある
B (中) :	効果の推定値が推奨を支持する適切さに中程度の確信がある
C (弱) :	効果の推定値が推奨を支持する適切さに対する確信は限定的である
D (非常に弱い) :	効果の推定値が推奨を支持する適切さにほとんど確信できない

(3) 推奨の強さを決定するための評価項目について仮の判定を行う

推奨の強さを決定するための評価項目について、システムティックレビューのサマリーレポート（や、その他必要に応じて収集した資料）を基に、暫定的な判定を行う。また、その評価の結果を導いた根拠について補足事項として記述しておく。

推奨を決定する際には、以下の項目を検討する必要がある。なお、以下の項目について、「はい」の回答が多いと、推奨度が「強い」とされる可能性が高くなる。

1) アウトカム全般に関する全体的なエビデンスの確実性（強さ）

アウトカム全般の、全体的なエビデンスの確実性（強さ）が強いほど、推奨は「強い」とされる可能性が高く、逆に全体的なエビデンスの確実性（強さ）が弱いほど、推奨は「弱い」とされる可能性が高くなる。エビデンスの確実性（強さ）決定の基本原則は、その治療効果推定値に対する確信が、ある特定の推奨を支持する上でどの程度つかである。

判定結果の表記は、「はい」「いいえ」のいずれかとする。
 「はい」：明らかに当てはまる場合
 「いいえ」：「はい」以外の全て

2) 望ましい効果（益）と望ましくない効果（害と負担など）のバランス

まず、益と害（benefit and harm）を比較することから始める。重要度（下記①）により重み付けした効果の大きさ（下記②）を、益と害のそれぞれで合計し、比較、考慮する。

① 重要度（重み付け）については、CQ リスト（テンプレート【SC-4】）の重要度、および総体評価リスト（テンプレート【SR-7】）の重要性を参考にするが、推奨作成にあたってあらためてアウトカムごとの重要度を再評価・決定する。益と害のバランスを評価する際にアウトカムの重要性を評価する考え方には次のようなものがある。

- ・アウトカム 1 は、アウトカム 2 に対して何倍重要か？
- ・最も重要度の低いアウトカムに比べ、他のアウトカムは何倍重要か？
- ・アウトカム 1 を重要度 9 とした場合、アウトカム 3 の重要度はどれほどか？

② 効果の大きさは、各アウトカムにおける絶対効果〔リスク差（risk difference:RD）〕を指標として考慮すべきである。

次に、上記のようにして「益」が「害」を上回るか評価した上で、さらに「負担」も合わせて、望ましい効果（益）と望ましくない効果（害と負担）のバランスの評価を行う。副作用や有害事象といった「害」は、意図せず起きる負の事象である。それに対して、負担は意図した上で起きる負の事象であり、例えば通院や入院などの負担や、手術の切開やそれに伴う痛み、手術痕や機能喪失などのことである。

- ・定量的益と害の評価、例えば、決定木（decision tree）のような手法はバランスを考慮する際の強力な手続きのひとつになり得る。
- ・多基準決断分析（multi-criteria decision analysis）にはさまざまな方法が報告されているが（Boyd CM 2012, Mt-Isa S 2014, PROTECT^{*1}, Thokala P 2016, Wen S 2014, Marsh K 2016, Marsh K 2017, Keeney R 1993, Gail MH 1999, Dolan JG 2013），益と害のバランスが接近している、益と害のアウトカムに対する効果のトレードオフが複雑、価値観によって判断が変わり得る、直観的な判断の裏付けを取りたい、などの場合に適用を考慮する。原理はほぼ共通しており、それぞれの介入について効果推定値をスコア化し、価値観に基づいた相対的なアウトカムの重要性を数値化し、両者の積（Alper BS 2019）の総計（aggregate score）を算出する。その際に効果推定値やアウトカムの重要性の値に不確実性を設定して、確率的感度分析（確率的シミュレーション）で正味の益が設定した閾値を上回る確率を計算することも可能である。
- ・望ましい効果（益）と望ましくない効果（害や負担など）の差が大きければ大きいほど、推奨が強くなる可能性が高い。
- ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨「弱い」とされる可能性が高くなる。

判定結果の表記は、「はい」「いいえ」のいずれかとする。

「はい」：明らかに当てはまる場合

「いいえ」：「はい」以外の全て

3) 推奨の強さの評価の際に考慮すべき要因

① 患者・市民の価値観・希望、負担の確実さ（あるいは相違）

- ・価値観・希望に確実性（一致性）があるか検討する。

例）先行研究を引用する、新たにアンケート調査を行う、など

- ・ばらつきがあればあるほど、または価値観や希望における不確実性が大きければ大きいほど、推奨が「弱い」とされる可能性が高くなる。
- ・パネリストに患者・市民が入ることが望ましい。メンバーに患者・市民が属していない場合は、患者・市民がどう思うか委員が推定することになり、限界があると認識する必要がある。

* 1 治療アウトカムに関する薬剤疫学的研究ヨーロッパコンソーシアム（Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium : PROTECT）（<https://protectbenefitrisk.eu/index.html>）

② 資源利用

- ・個々の、あるいはスコープに含めた資源利用全体の介入間の差が正味の益に対して大きいか小さいかを検討する（ただし、複雑な状況を反映し、より包括的な資源利用を考慮して各介入の比較を明示的に行うためには経済学的モデリングを行う必要がある）。
- ・保険診療であればそのことを記載する。
- ・Individual perspective の場合は、患者の自己負担分を含む患者の立場から見た資源利用を推奨に考慮することができる。
- ・Population perspective の場合、推奨作成の際に（公的医療の立場などから）資源利用の介入間の差のエビデンスを考慮できる。

③ 費用対効果

Individual perspective で作成する場合は、費用対効果の観点から留意事項への記載を検討する。Individual perspective の場合にどのようなものを評価対象とするかについては、**第5章 5.2.2 節**を参照のこと。

留意事項の記載方法としては、

- ・患者にとって治療上同等であると考えられる場合には、後発品を含め最も安価な医療技術の使用を検討する。
- ・特定の使用順序を検討する（安価な医療技術から使用し、無効などの場合には高価なものに切り替える、など）。
- ・対象集団を限定することを検討する（費用対効果が良好と考えられる一部の集団に限定して使用を検討する、など）。

などが考えられる。

Population perspective の場合は、「推奨の強さを決定するための評価項目」において、費用対効果のエビデンスの評価結果を推奨作成の際に考慮できる。

(4) 推奨文草案を提示する

原則として、CQに対する答えとしての推奨文草案を作成する。パネル会議での議論の基本資料として、重要である。

ただし、草案作成時点で草案文章を示すことが推奨作成、推奨の強さの決定に関する議論を方向付けしてしまう可能性があると想定される場合、記載しない方法もある。例えば、推奨の内容として、行うこと、行わないことが定まらない場合には、記載をしないでおくことで、診療ガイドライン作成グループでのオープンな議論がしやすくなる場合もある。

推奨の記載方法は、前述の「患者 P に対して介入 I を行うことを推奨する／行わないことを推奨する」の形式を基本とする。「患者 P に対して介入 I を行うことを推奨しない」や「患者 P に対して介入 I が有効である」とは書かないことに注意が必要である。

● テンプレート

[RC-1 推奨文草案 (Individual perspective)]

[RC-2 推奨文草案 (Population perspective)]

● テンプレート 記入方法

[RC-1 推奨文草案 (Individual perspective) 記入方法]

● テンプレート 記入例

[RC-1 推奨文草案 (Individual perspective) 記入例]

6.3 ステップ3：推奨の作成、推奨の強さの判定

担当者がまとめた推奨文草案に関する資料を基に、推奨を作成し、推奨の強さを決定する。推奨は、システムティックレビューチームが作成したサマリーレポートの結果などを基にして判定される。その際、重大なアウトカムに関するエビデンスの確実性（強さ）に加えて、益と害のバランスなどを検討して決定する。さらに、患者・市民の価値観・希望、資源利用についても十分考慮することが望ましい（Jaeschke R 2008, Andrews JC 2013a, Andrews JC 2013b）。

● 手順

(1) CQに対するエビデンス総体の総括を決定する。

ステップ2で担当者がまとめたテンプレート【RC-1】（または【RC-2】）の内容などを確認し、診療ガイドライン作成グループ全体の意見として決定する。

(2) 推奨の強さの決定に影響する要因を決定する。

ステップ2で担当者がまとめたテンプレート【RC-1】（または【RC-2】）の内容などを確認し、診療ガイドライン作成グループ全体の意見として決定する。

(3) 4つの項目（エビデンスの確実性、望ましい効果と望ましくない効果のバランス、患者・市民の価値観・希望、資源利用など）を考慮した上で、推奨の向きと強さを考える際に、全ての項目のバランスを考慮する（図6-1参照）。

(参考) 全ての要素を考慮した判定の例

まず、益と害を比較することから始め、害に負担を加え、さらに資源利用を加え、最終的に、全ての要素を含めて、益と不利益（害、負担、資源利用）のバランスを考える。全ての不利益を考慮しても益が勝るとき、その介入を推奨することになる。

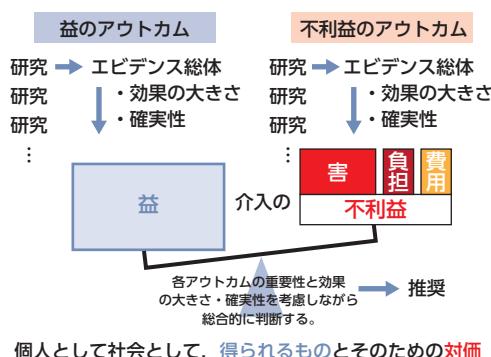


図6-1 益と不利益のバランス

(4) 推奨の強さの判定、推奨の作成

先に定めた推奨決定の方法・基準を用いて、前記結果を基に、CQに対する推奨文草案に対する推奨の強さを決定する。エビデンスの確実性（強さ）、益と害のバランスを中心に、患者・市民の価値観・希望、負担、資源利用を加味して総合的に勘案して決定する。

● 推奨決定会議のために準備する資料の作成

準備する資料リスト

- ・ CQ リスト（テンプレート【SC-4】）
- ・ 個々の研究報告の評価結果（テンプレート【SR-5】、【SR-6】）
- ・ エビデンス総体評価結果（テンプレート【SR-7】、【SR-8】）
- ・ メタアナリシス（meta-analysis、メタ解析と同義）のレポート（テンプレート【SR-10】）
- ・ 推奨文草案（テンプレート【RC-1】、【RC-2】）
- ・ EtD フレームワーク（テンプレート【RC-7】、【RC-8】）を用いる場合は準備する。
- ・ その他必要に応じた資料

● 推奨決定会議進行の具体的な方法について

(1) 推奨決定方法の種類と詳細内容

1) 修正 Delphi 法 (RAM) による合意形成方式

提案された推奨文草案に対する同意の度合いを、まったく同意しない（1点）～強く同意する（9点）で評価、集計する。投票の中央値、IQR（四分位範囲）を用いて、多くの意見が7～9点に集約されたことで、同意と判定する方法である。詳細は、原著（Fitch K 2001）を参照されたい。

2) NGT 法による合意形成方式 (expert panel) (Delbecq AL 1971)

ブレインストーミング+投票によって検討項目に対するアイデアの優先付けを行う形式である。9～12人の少数の参加者（nominal group）によって行われる。参加者が多い場合は6人以内の小グループに分かれることもある。

第1段階

1つの検討項目に対して参加者全員が一定時間熟考し、全ての意見を書き出す。参加者は、書き出した意見を一つひとつ説明する。この説明を参加者全員が順に行う。それぞれの意見に対して共感や相乗りは歓迎されるが、否定的な意見は原則禁止とする。

全ての意見が検討されたら、出された意見の中で似た意見を集約する。

第2段階

集約された意見に対して、参加者各自が点数をつける。

例 1) それぞれの意見に対して、まったく同意しない（1点）～強く同意する（9点）で記載する。

例 2) 全ての意見の中で、自分が最も同意するものを7つ選ぶ。その中で、自分が最も同意するものを7点、以下1点まで点数をつける。

第3段階

点数の分布を表にして、参加者全員にフィードバックする。フィードバックされた内容について参加者全員で討論する。

第4段階

参加者各自が再度点数をつける。最終的な分布を再度表にまとめて参加者全員にフィードバックする。

3) GRADE グリッドによる合意形成方式 (Jaeschke R 2008)

作成した推奨文草案に対して、**表6-2** を用いて投票を行う。推奨のタイプの5項目のいずれかに投票し、その得票率があらかじめ決めておいた一定の集中を得た場合に同意とする方式である（**6.1節** 参照）。

なお、Aという介入とBという介入の双方が推奨される場合は、「ある限定的な状況ではA、またある限定的な状況ではB」という選択肢もあり得る。

表 6-2 EtD フレームワーク推奨の結論テーブル

推奨のタイプ	当該介入に反対する強い推奨	当該介入に反対する条件付きの推奨	当該介入または比較対照のいずれかについての条件付きの推奨	当該介入の条件付きの推奨	当該介入の強い推奨
	○	○	○	○	○
推奨					
正当性					
サブグループに関する検討事項					
実施に関する検討事項					
監視と評価					
研究上の優先事項					

出典：Schünemann H, Brožek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook. より作成 ※EtD フレームワークより一部抜粋

(参考) GRADE グリッド

GRADE アプローチで診療ガイドライン作成パネルが投票を行う場合の合意形成方法

GRADE グリッドによる合意形成フォーム（投票用紙）				
CQ：				
推奨文草案：				
推奨の強さ	強い	弱い	弱い	強い
推奨の内容	介入支持の強い推奨	介入支持の条件付き（弱い）推奨	介入反対の条件付き（弱い）推奨	介入反対の強い推奨
推奨の表現	「実施する」ことを推奨する。	「実施する」ことを提案（条件付きで推奨）する。	「実施しない」ことを提案（条件付きで推奨）する。	「実施しない」ことを推奨する。
投票 (右のいずれかに○)				

出典：Jaeschke R, et al. (2008) Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. BMJ 337:327-330. より作成

(2) 推奨決定会議進行の具体例

1) 会議開会宣言：議長

2) 推奨度決定に関する合意形成方法についての確認：議長、作成方法論担当委員

3) 投票者、棄権者の確定

CQごとに利益相反 (conflict of interest : COI) を確認し、投票棄権者を確定し、記載する。仮に、議長に COI がある場合は、議長は当該 CQ に関するコンセンサス会議に参加せず、作成方法論担当委員、またはこれに類する委員が会議進行を担当する (COI については、[第2章](#)参照)。

4) 推奨文草案の討議

資料を基に、推奨について検討する。必要に応じて、担当者が資料の内容について説明する。

推奨決定会議では、議論の結果全員一致で同意が得られれば決定とする。一方、意見がばらつく場合は投票を行う。

5) 投票

投票は、診療ガイドライン作成パネル（診療ガイドライン作成グループ）委員が行う。診療ガイドライン作成パネルや推奨決定方法は、スコープ作成の時点で決定されるべきである。推奨決定作業は、推奨文や推奨の強さ決定に直接関係するため、その人選は公平で偏りのないことが重要で、投票は、患者などの非専門家も含めて可能な限り全員参加で行うことが勧められる。システムティックレビューチームメンバーなど、診療ガイドライン作成パネル以外の者は投票には参加しない。

投票は独立して行うことが基本である。例として、アンサーパッドを用いる、無記名投票用紙利用、インターネット投票、などがある。アンサーパッドを用いた投票は集計も早く再投票も行える利点があるが、費用がかさむ点と、投票者がその場のテンポ（流れ）に流されやすい欠点があることを理解することが必要である。また、無記名投票は周囲のさまざまな影響を受けない利点はあるが、記名投票は各パネル委員が自分の立場を明らかにし、責任をもって投票するという意味で有用な場合もある。

6) 投票による推奨の決定

1回目の投票で事前に決めておいた得票分布が得られれば決定とする。

1回目の投票で意見の集中が見られなかった場合は、再度投票が行われる。この場合、Delphi 法では、投票結果を示した上で、討論なしに再投票を繰り返す。一方、NGT 法や、修正 Delphi 法 (RAM), GRADE グリッドを用いた場合は、結果を示した上で十分な討論を行って、再投票する。

コンセンサスが得られるまで投票を繰り返すが、あらかじめ設定した回数を繰り返しても意見の集約が得られない場合、「推奨の強さは決定できない (not graded)」とし、その経過や結果の要約を解説に記載する（ただし、これはなるべく避ける）。

7) 補記：公聴会、外部評価

診療ガイドライン草案に対する公聴会での結果や外部評価委員の意見を推奨文・推奨の強さに取り入れる場合は、診療ガイドライン作成パネルによる推奨決定会議を再度開催する必要があるが、場合によってはインターネットを用いた投票なども考慮すべきである。

● テンプレート

【RC-1 推奨文草案 (Individual perspective)】

【RC-2 推奨文草案 (Population perspective)】

【RC-3 推奨の強さ決定投票用紙】

【RC-4 推奨提示】

● テンプレート 記入方法

【RC-1 推奨文草案 (Individual perspective) 記入方法】

● テンプレート 記入例

【RC-1 推奨文草案 (Individual perspective) 記入例】

【RC-3 推奨の強さ決定投票用紙 記入例】

【RC-4 推奨提示 記入例】

6.4 ベース4：解説の執筆

CQ および対応する推奨文、推奨の強さが決定されたのち、解説の文章を執筆する。

診療ガイドライン作成グループは、最終的に推奨が決定するまでの過程を診療ガイドラインの本文に記載する。システムティックレビューチームが作成したシステムティックレビューレポート（サマリーレポートともいう）のまとめやその他の資料の内容に合わせて、推奨決定の過程を詳細に記載する。

システムティックレビューレポートの中にランダム化比較試験（randomized controlled trial : RCT）のエビデンスの強さを下げた記載や観察研究（observational study）のエビデンスの強さを上げた記載がある場合などは、その理由を解説に記載する。

また、推奨とそれを支持するエビデンスの対応関係が明確になるように、システムティックレビューで用いた文献を章末または巻末に提示する。

用いる表現は明瞭にし、あいまいでないよう努める（複数の介入を比較する場合は、全てに順位付けすることに固執せず、医療現場で役立つように、柔軟な表現を心がける）。

● 手順

解説に記載すべき具体的な内容として、推奨決定に採用した各研究の強みと限界、定性的・定量的システムティックレビューの結果を「日本語として流暢な文章で記載する」。

(1) CQ が重要な臨床課題である旨の背景を記載する。

(2) エビデンス評価

アウトカムごとにシステムティックレビューを行った過程と結果（説明）を記載する。

検索：各アウトカムについて、系統的文献検索を行った結果、採用した論文数

評価：個々の論文を評価した結果

統合：エビデンス総体としての評価とそれぞれのアウトカムごとのエビデンスのまとめ。メタアナリシスを行った場合はその結果も記載し、この推奨に至った経緯を解説する。

(3) 益と害のバランス評価

システムティックレビューで得られた望ましいアウトカムと望ましくないアウトカムのバランスを比較検討した内容を記載する。

なお、CQ 設定の際に、介入の結果である益と害のアウトカムを挙げるが、それがシステムティックレビューの結果としてそのまま益と害にならない場合もある。例えば、CQ 設定時には「全生存（の改善）」を益のアウトカムと設定したが、総体評価を行ったところ死亡が増えるという結果になる可能性もある。

本マニュアルでは、総体評価した結果を「患者にとって好ましい効果」と「患者にとって好ましくない効果」に大別し、2群のバランス評価を行って記載することを提案する。

(4) 患者・市民の価値観・希望

患者・市民の価値観・希望の多様性や不確実性について検討した結果を記載する。

患者・市民の価値観・希望を調べる際には、調査研究があればその文献を用いるとともに、患者会の意見や手記など、可能な範囲で採用する。

患者（という専門家）の意見は、エビデンスとしては採用されないが、推奨決定の1因子となる。検索可能な範囲で情報を探し求める努力をする。

(5) 資源利用と費用対効果

Individual perspectiveの場合、患者の自己負担分、患者の立場から見たその他の資源利用についてのシステムティックレビューの結果に基づいて記載することが望ましいが、推奨には自己負担を含む患者の立場から見た資源利用だけが反映されることに留意する。費用対効果のシステムティックレビューおよび新規の分析を行った場合は、それら留意事項の基になったエビデンスや解釈について解説する。

Population perspectiveの場合、公的医療での資源利用と費用対効果のエビデンスの評価結果とその解釈について記載する。

Individual perspectiveまたはpopulation perspectiveのいずれの場合においても、保険診療であるか否かは、極めて重要な情報である。保険診療でない介入方法については、それを明記する。

(6) 推奨がどうしても決定できないときは、まれに「明確な推奨ができない」とする場合もある。

この場合、その経過と討論内容を本文中に記載する。

推奨に関する外部評価が寄せられた場合には、当該推奨の決定について記載する章に、外部評価と検討の内容を記載する。

推奨決定コンセンサス会議において、委員から出された反対意見の内容についても（少数意見も含めて）記載する。

(7) 今後の研究 (future research) (WHO 2014)

1) 今後の研究が推奨される場面

エビデンスに乖離があり、それによって介入による益と害のバランスがかなり不確実になってしまうほどであるとき、そのような知識の乖離は記述されなければならず、その乖離を記述（検討）するための臨床質問と記載方法についても、提案されるべきである。

具体的に今後の研究が推奨される場面としては、今回の診療ガイドライン作成において、重要と考えられた臨床課題に対してシステムティックレビューを行ったが、適切な論文が検索されなかった場合、または検索された全ての論文の質が高くなかった場合である。

その研究課題が適切であれば、未来に向けた優先研究課題と考えられ、その臨床研究の結果を基に診療ガイドラインの適正な推奨作成が可能となると期待されるものである。

つまり、今回提示された今後の研究は、臨床研究として推進され、次回改訂時には、エビデンスとして活用されることが期待される（図6-2参照）。



図 6-2 診療ガイドライン作成のサイクルにおける研究推奨 (research recommendations) の役割

出典 : National Institute for Health and Care Excellence (NICE). Research recommendations Process and methods guide, 2015 (<https://www.nice.org.uk/Media/Default/About/what-we-do/Science-policy-and-research/research-recommendation-process-methods-guide-2015.pdf>) ©Copyright National Institute for Health and Care Excellence, 2014. より作成

2) 今後の研究が推奨される場合の記載方法

- ・ 診療ガイドライン作成グループは、診療ガイドラインの本文に、この項は、未来の診療ガイドラインにおいてより適正な推奨が提案されることになると記載する。今回の診療ガイドラインでは、推奨文や推奨度作成は必須ではない。
 - ・ 診療ガイドライン作成グループは、①今後の研究が推奨される臨床疑問 (future research question) を記載し、②今回行われたシステムティックレビューの方法と結果を示し、③なぜ必要か (背景)、④今後どのような研究が必要か (可能な研究計画の概略) について可能な限り詳細に提示する必要がある。
- 上記の future research question の記載方式として、PICO (Population, Intervention, Comparator, Outcome) 形式または、EPICOT (Evidence, Population, Intervention, Comparator, Outcome, Time) 形式で作成することも、大変有用な方法である。
- ・ 今後の研究は、その CQ に対する現在の推奨の有無、または、強弱によらず、未来の研究が必要と考えられる場合に推奨されることになる。その一方で、今後の研究が複数推奨されているときは、重要度や順位付けなども考慮すべきである。

● テンプレート

【RC-5 推奨作成の経過】

【SR-15 Future Research Question】

● テンプレート 記入方法

【RC-5 推奨作成の経過 記入方法】

【SR-15 Future Research Question 記入方法】

● テンプレート 記入例

(なし)

6.5 一般向けサマリーの執筆

診療ガイドライン作成グループは、患者・市民にも理解できる平易な言葉でCQと推奨を簡単に解説する一般向けサマリーを執筆する。ここで執筆した一般向けサマリーをまとめて、一般向けの診療ガイドライン関連資料を作成することもできる。

● テンプレート

[RC-6 一般向けサマリー]

● テンプレート 記入方法

[RC-6 一般向けサマリー 記入方法]

● テンプレート 記入例

(なし)

6.6 GRADE アプローチを用いる場合の資料

これまでに示してきた方法は海外で開発されてきたさまざまな診療ガイドラインの作成方法を参考にして、日本の診療ガイドライン作成において適切と考えられる方法をまとめた提案である。したがって、診療ガイドライン作成グループが、診療ガイドライン作成において本マニュアルが参照してきたさまざまな作成方法の「GRADE の原著」に従う必要があると判断した場合には、オリジナルを参照し、オリジナルに準拠して作成を行う必要がある。

参考のために、本マニュアルを作成するにあたって参考にした方法の1つであるGRADE アプローチが用いているテンプレートを挙げておく。

(参考) GRADE アプローチで診療ガイドライン作成パネルが検討する推奨文草案

1. CQ											
2. 推奨文草案											
3. 推奨に関連する価値観や意向 (検討したアウトカム別に、一連の価値観を想定する) (例：この推奨の作成にあたっては、致死率の高い疾患における死亡の回避を重要視した。)											
4. 重大なアウトカム全般に関する全体的なエビデンスの質 (エビデンスレベル)											
5. 推奨の強さの判定 (下記の4項目について判定し、総合して判定する)	<table border="1"> <thead> <tr> <th>推奨の強さに影響する要因</th><th>判定</th></tr> </thead> <tbody> <tr> <td>アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。</td><td> <input type="checkbox"/> はい <input type="checkbox"/> いいえ </td></tr> <tr> <td>益と害のバランスが確実 (資源利用は含まず) ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。</td><td> <input type="checkbox"/> はい <input type="checkbox"/> いいえ </td></tr> <tr> <td>患者・市民の価値観や希望、負担の確実さ (あるいは相違) ・価値観や希望に確実性 (一致性) があるか？ ・逆に、ばらつきがあればあるほど、または価値観・希望や好みにおける不確実性が大きければ大きいほど、推奨度が「弱い」とされる可能性が高くなる。</td><td> <input type="checkbox"/> はい <input type="checkbox"/> いいえ </td></tr> <tr> <td>正味の益が費用や資源利用に十分見合ったものかどうか ・費用に見合った益があると判定できるか？ (費用に関する報告があれば利用する。)</td><td> <input type="checkbox"/> はい <input type="checkbox"/> いいえ </td></tr> </tbody> </table>	推奨の強さに影響する要因	判定	アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	益と害のバランスが確実 (資源利用は含まず) ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	患者・市民の価値観や希望、負担の確実さ (あるいは相違) ・価値観や希望に確実性 (一致性) があるか？ ・逆に、ばらつきがあればあるほど、または価値観・希望や好みにおける不確実性が大きければ大きいほど、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	正味の益が費用や資源利用に十分見合ったものかどうか ・費用に見合った益があると判定できるか？ (費用に関する報告があれば利用する。)	<input type="checkbox"/> はい <input type="checkbox"/> いいえ
推奨の強さに影響する要因	判定										
アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ										
益と害のバランスが確実 (資源利用は含まず) ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ										
患者・市民の価値観や希望、負担の確実さ (あるいは相違) ・価値観や希望に確実性 (一致性) があるか？ ・逆に、ばらつきがあればあるほど、または価値観・希望や好みにおける不確実性が大きければ大きいほど、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ										
正味の益が費用や資源利用に十分見合ったものかどうか ・費用に見合った益があると判定できるか？ (費用に関する報告があれば利用する。)	<input type="checkbox"/> はい <input type="checkbox"/> いいえ										

「はい」の回答が多いと、推奨度が「強い」とされる可能性が高くなる。

出典：Andrews JC, et al. (2013) GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 66:726-735. By courtesy of Dr. Andrews. より作成

● EtD フレームワーク (DECIDE) ---

(1) 基本情報

1) DECIDE プロジェクトと EtD フレームワークについて

GRADE では、推奨決定の手順として DECIDE (Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence) プロジェクトによって開発された、EtD フレームワークの利用を紹介している。

2) ウェブサイト

- DECIDE プロジェクト

<http://www.decide-collaboration.eu/>

- Evidence to Decision (EtD) Frameworks

<https://ietd.epistemonikos.org/#/help/guidance>

3) Individual perspective と population perspective について (第5章参照)

推奨を決定する場合に、患者個別の立場から推奨を考える場合 (individual perspective) と、国民や社会、集団としての医療費などを考慮して推奨を考える場合 (population perspective) が考えられる。どちらの立場をとるかは、診療ガイドライン作成時のスコープによって選択されるべきである。選択した立場によって、検討項目は変化するが、必要によって、項目を付け加えることが可能であり、十分に吟味すべきである。

4) EtD フレームワークのテンプレート例 (Moberg J 2015, Alonso-Coello P 2016a, Alonso-Coello P 2016b より作成)

テンプレート 【RC-7】、【RC-8】 を参照されたい。

その他、目的に応じて異なるタイプのフレームワークが提案されている。

- Test (Schünemann 2016, Schünemann 2019)
- Coverage decisions (Alonso-Coello P 2016a)
- Health system and public health recommendations/decisions (Alonso-Coello P 2016a)

(2) EtD フレームワークは誰が作るか？

GRADE では明確に定義されていないが、システムティックレビューチームと診療ガイドライン作成パネルの合同作業となる場合が多い。最終成果物の最終責任はパネル会議が持つ。

(3) EtD フレームワーク（全体）の解説と記載方法

テンプレート【RC-7】、【RC-8】 は、EtD フレームワークに合わせ、基本のフォーマットを提示している。診療ガイドライン作成のスコープに併せて選択し、項目を追加して用いる。以下で、EtD フレームワークの使い方を紹介する (Moberg J 2015, Alonso-Coello P 2016a, Alonso-Coello P 2016b)。

〈疑問〉

- ・**臨床質問 (CQ)** : Intervention vs. Comparison は、Population に使うべきか？
- ・**集団** : 対象集団 Population
- ・**介入** : Intervention
- ・**比較対照** : Comparison
- ・**主要なアウトカム** : 複数ある場合は列挙する。
- ・**セッティング** : 本診療ガイドラインが利用される環境や背景など。
例) 医療体制の確立した地域
- ・**視点** : 本診療ガイドラインがカバーする視点、臨床場面における individual perspective もしくは population perspectiveなどを選択する。
- ・**背景** : 本課題の背景と重要臨床課題 (key clinical issues) であることを明記する。

〈評価〉 ※以下、population perspective の項目を記載。

基準 1. 問題：この問題は優先事項か？

→本 CQ が重要臨床課題であることを記述する。下記 4 項目で投票する。

- いいえ
- おそらく、いいえ
- おそらく、はい
- はい

→4 項目の中から決めきれないときは、下記 2 項目のいずれかとする。

- さまざま
- 分からない

基準 2. 望ましい効果：予期される望ましい効果はどの程度のものか？

→判定の方法は、基準 1 と同様。

基準 2～4 について、知見のまとめ (SoF:summary of findings) を「リサーチエビデンス」の欄に記載する。

基準 3. 望ましくない効果：予期される望ましくない効果はどの程度のものか？

基準 4. エビデンスの確実性：効果に関する全体的なエビデンスの確実性はどの程度か？

→ GRADE のエビデンスの確実性の評価に基づいて、下記から選択する。

- 非常に弱い
- 弱
- 中
- 強
- 採用研究なし

基準 5. 値値観：人々が主要なアウトカムをどの程度重視するかについて重要な不確実性やばらつきはあるか？

→ 下記 4 項目から選択する。

- 重要な不確実性またはばらつきあり
- 重要な不確実性またはばらつきの可能性あり
- 重要な不確実性またはばらつきはおそらくなし
- 重要な不確実性またはばらつきはなし

「リサーチエビデンス」の欄に、①関心のある主アウトカムの相対的な重要性や価値（主アウトカム、重要度、論文の確実度の表など）、②患者・市民の価値観や希望として、患者の価値観や希望に関する報告、重要とはされていない追加アウトカム、有害事象などを記載する。

基準 6. 効果のバランス：望ましい効果と望ましくない効果のバランスは介入もしくは比較対照を支持するか？

→ 選択肢から選択する。

基準 7. 必要資源量：資源利用はどの程度大きいか？

→ 選択肢から選択する。

基準 8. 必要資源量に関するエビデンスの確実性：資源利用に関するエビデンスの確実性はどの程度か？

→ 選択肢から選択する。

基準 9. 費用対効果：その介入の費用対効果は介入または比較対照を支持するか？

→ 選択肢から選択する。

基準 10. 公平性：医療上の公平性への影響はどの程度か？

→ 選択肢から選択する。

- 減る
- おそらく減る
- おそらく影響なし
- おそらく増える
- 増える

→ 意見がばらつくときは、下記の選択肢から選択する。

- さまざま
- 分からない

エビデンスが収集された場合は、「リサーチエビデンス」の欄に記載する。

パネリストが考慮すべきは次の点である。

- ・エビデンスが収集されたか。
- ・(医療上) 患まれていない人ほど、その介入からの益を得られるか。
- ・その介入は、長期間にわたり疾病負荷に関する不公平を改善するか。
- ・仮にその介入が活用されたときに、(医療上の) 不公平が増えないか？
- ・保障されるべきであるということを十分に考慮しているか？

なお、GRADEでは、メンバー構成、スコープ作成、CQ設定、システムティックレビュー、推奨作成といった段階で公平性について検討する方法を提案している。推奨作成場面では、各項目について評価を行う中で公平性の観点から検討し、また、推奨が公平性に与える影響について考慮することを提案している (Welch VA 2017a, Akl EA 2017, Welch VA 2017b, Pottie K 2017)。

基準 11. 容認性：この選択肢は重要な利害関係者にとって妥当なものか？

→選択肢から選択する。

基準 12. 実行可能性：その介入は実行可能か？

→選択肢から選択する。

エビデンスが収集された場合は、「リサーチエビデンス」の欄に記載する。

パネリストが考慮すべきは次の点である。

- ・推奨／診療ガイドラインの受け入れ可能性：投与量の変更は受け入れられるか？
- ・医療職種による特徴：知識（理解）
- ・介入の有用性
- ・インセンティブと財源：医療費支払いの構図
- ・ケア提供組織：職員、システム、設備
- ・社会的、法的、社会的因素

(4) EtD フレームワーク（判断の要約テーブル）

EtD で作成した資料を要約する。推奨決定会議で、一見して全項目の配置が分かる。

(5) EtD フレームワーク（推奨の結論テーブル）の解説と記載方法

〈推奨のタイプ〉

→選択肢から選択する。

- 当該介入に反対する強い推奨
- 当該介入に反対する条件付きの推奨
- 当該介入または比較対照のいずれかについての条件付きの推奨
- 当該介入の条件付きの推奨
- 当該介入の強い推奨

「当該介入または比較対照のいずれかについての条件付きの推奨」は対照、介入の双方が推奨される場合にのみ選択可能である。

〈結論〉

推奨： 推奨文を記載する。

正当性： 決定や推奨に至った根拠を簡潔に提示する。

サブグループに関する検討事項： 推奨作成において、検討したサブグループについて記載し、総括的な決定・

推奨を特に改変した事項についてのまとめを記載する。

実施に関する検討事項： 診療ガイドライン作成グループは、介入（オプション）の賛否について、特に受け入れ可能性や適用可能性を考慮し、あらゆる障壁への推奨の導入に対する責任を持つことに警鐘を鳴らす立場にある。

監視（モニタリング）と評価： 推奨が活用される場合に、どのような指標を用いて評価するかを提示する。

研究上の優先事項（将来的な研究課題）： 今後の臨床研究によって、不確実性を減少させられる重要な課題を検討し、提示する。

● テンプレート

【RC-7 EtD フレームワーク (Clinical recommendation: Individual perspective)】

【RC-8 EtD フレームワーク (Clinical recommendation: Population perspective)】

● テンプレート 記入方法

(なし)

● テンプレート 記入例

(なし)

6.7 その他の意思決定に関するフレームワーク

近年、新たな健康に関する意思決定に関するフレームワークとして、WHO-INTEGRATE フレームワーク（WHO-INTEGRATE Evidence framework）が提案された（Rehfuss EA 2019）。この背景として、EtD フレームワークの項目には、健康の社会的および経済的決定要因に関する十分な配慮がないことや各項目の評価法が明示的でないため、やや表面的評価となってしまうことが指摘されている。この新しいフレームワークは、主に公衆衛生的な意思決定のために策定されたものであるが、今後十分な議論がなされ、各国の診療ガイドライン作成において利用可能性について明らかになっていくことを期待したい。

また、益害バランスを定量的に評価する方法として、多基準決断分析（multi-criteria decision analysis または multiple criteria decision analysis : MCDA）や GRADE ワーキンググループ（GRADE Working Group）による方法が提案されている（Broekhuizen H 2015, Alper BS 2019）。特に、益と害の関係や価値観が複雑な場合に有用である可能性があり、今後の議論がまたれる。

6章 テンプレート

【RC-1 推奨文草案 (Individual perspective)】

1. CQ		
2. 推奨文草案		
3. 作成グループにおける、推奨に関連する価値観や意向（検討したアウトカム別に、一連の価値観を想定する）		
4. CQに対するエビデンス総体の総括（重大なアウトカム全般に関する全体的なエビデンスの強さ） A（強） B（中） C（弱） D（非常に弱い）		
5. 推奨の強さを決定するための評価項目（下記の項目について総合して判定する）		
推奨の強さの決定に影響する要因	判定	説明
アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	
益と害のバランスが確実（コストは含まず） ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	
推奨の強さに考慮すべき要因		
患者・市民の価値観・希望や好み、負担の確実さ（あるいは相違）、医療費のうち自己負担分、患者の立場から見たその他の資源利用など：		
6. 費用対効果の観点からの留意事項（費用対効果を検討した場合のみ記載するが、臨床的な推奨とは別に取り扱う）：		

明らかに当てはまる場合は「はい」とし、それ以外は、どちらともいえないを含め「いいえ」とする。

Andrews JC, et al. (2013b) GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 66:726-735. より作成

【RC-2 推奨文草案 (Population perspective)】

1. CQ		
2. 推奨文草案		
3. 作成グループにおける、推奨に関連する価値観や意向（検討したアウトカム別に、一連の価値観を想定する）		
4. CQに対するエビデンス総体の総括（重大なアウトカム全般に関する全体的なエビデンスの強さ） A（強） B（中） C（弱） D（非常に弱い）		
5. 推奨の強さを決定するための評価項目（下記の項目について総合して判定する）		
推奨の強さの決定に影響する要因	判定	説明
アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	
益と害のバランスが確実（コストは含まず） ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ	
推奨の強さに考慮すべき要因		
患者・市民の価値観や希望、負担の確実さ（あるいは相違）など		
資源利用、費用対効果		

明らかに当てはまる場合は「はい」とし、それ以外は、どちらともいえないを含め「いいえ」とする。

Andrews JC, et al. (2013b) GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 66:726-735. より作成

【RC-3 推奨の強さ決定投票用紙】

CQ :	
推奨文 :	
推奨 (いずれかに○)	<ul style="list-style-type: none">・行うことを強く推奨する・行うことを弱く推奨する（提案する）・行わないことを弱く推奨する（提案する）・行わないことを強く推奨する

[RC-4 推奨提示]

CQ :	
推奨文 :	
推奨の強さ (いずれかに○)	1 (強い) : 「実施する」, または, 「実施しない」ことを推奨する 2 (弱い) : 「実施する」, または, 「実施しない」ことを提案する
エビデンスの強さ (いずれかに○)	A (強) B (中) C (弱) D (非常に弱い)
費用対効果の観点 からの留意事項	

どうしても決定できないときは、まれに「明確な推奨ができない」とする場合もある。この場合、その経過と討論内容を本文中に記載する。

【RC-5 推奨作成の経過】



【RC-6 一般向けサマリー】



[RC-7 EtD フレームワーク (Clinical recommendation: Individual perspective)]

疑問

CQ :	
集団	
介入	
比較対照	
主要なアウトカム	
セッティング	
視点	
背景	
利益相反	

評価

基準 1. 問題 この問題は優先事項か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> いいえ <input type="radio"/> おそらく、いいえ <input type="radio"/> おそらく、はい <input type="radio"/> はい <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 2. 望ましい効果 予期される望ましい効果はどの程度のものか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> わずか <input type="radio"/> 小さい <input type="radio"/> 中 <input type="radio"/> 大きい <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 3. 望ましくない効果 予期される望ましくない効果はどの程度のものか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 大きい <input type="radio"/> 中 <input type="radio"/> 小さい <input type="radio"/> わずか <input type="radio"/> さまざま <input type="radio"/> 分からない		

基準 4. エビデンスの確実性 効果に関する全体的なエビデンスの確実性はどの程度か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 非常に弱い <input type="radio"/> 弱 <input type="radio"/> 中 <input type="radio"/> 強 <input type="radio"/> 採用研究なし		
基準 5. 価値観 人々が主要なアウトカムをどの程度重視するかについて重要な不確実性やばらつきはあるか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 重要な不確実性またはばらつきあり <input type="radio"/> 重要な不確実性またはばらつきの可能性あり <input type="radio"/> 重要な不確実性またはばらつきはおそらくなし <input type="radio"/> 重要な不確実性またはばらつきはなし		
基準 6. 効果のバランス 望ましい効果と望ましくない効果のバランスは介入もしくは比較対照を支持するか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 比較対照が優れている <input type="radio"/> 比較対照がおそらく優れている <input type="radio"/> 介入も比較対照もいずれも支持しない <input type="radio"/> おそらく介入が優れている <input type="radio"/> 介入が優れている <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 7. 費用対効果 その介入の費用対効果は介入または比較対照のどちらを支持するか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 比較対照の費用対効果がよい <input type="radio"/> 比較対照の費用対効果がおそらくよい <input type="radio"/> 介入も比較対照もいずれも支持しない <input type="radio"/> 介入の費用対効果がおそらくよい <input type="radio"/> 介入の費用対効果がよい <input type="radio"/> さまざま <input type="radio"/> 採用研究なし		

基準 8. 必要資源量 資源利用はどの程度大きいか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 大きな増加 <input type="radio"/> 中等度の増加 <input type="radio"/> 無視できるほどの増加や減少 <input type="radio"/> 中等度の減少 <input type="radio"/> 大きな減少 <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 9. 容認性 この選択肢は重要な利害関係者にとって妥当なものか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> いいえ <input type="radio"/> おそらく、いいえ <input type="radio"/> おそらく、はい <input type="radio"/> はい <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 10. 実行可能性 その介入は実行可能か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> いいえ <input type="radio"/> おそらく、いいえ <input type="radio"/> おそらく、はい <input type="radio"/> はい <input type="radio"/> さまざま <input type="radio"/> 分からない		

判断の要約

問題	判断						
	いいえ	おそらく, いいえ	おそらく, はい	はい		さまざま	分からぬ
望ましい効果	わずか	小さい	中	大きい		さまざま	分からぬ
望ましくない 効果	大きい	中	小さい	わずか		さまざま	分からぬ
エビデンスの 確実性	非常に弱い	弱	中	強			採用研究 なし
価値観	重要な不確実 性またはばら つきあり	重要な不確実 性またはばら つきの可能性 あり	重要な不確実 性またはばら つきはおそら くなし	重要な不確実 性またはばら つきはなし			
効果のバランス	比較対照が 優れている	比較対照が おそらく 優れている	介入も比較対 照もいすれも 支持しない	おそらく 介入が 優れている	介入が 優れている	さまざま	分からぬ
費用対効果	比較対照の 費用対効果が よい	比較対照の 費用対効果が おそらくよい	介入も比較対 照もいすれも 支持しない	介入の 費用対効果が おそらくよい	介入の 費用対効果が よい	さまざま	採用研究 なし
必要資源量	大きな増加	中等度の 増加	無視できる ほどの増加や 減少	中等度の 減少	大きな減少	さまざま	分からぬ
容認性	いいえ	おそらく, いいえ	おそらく, はい	はい		さまざま	分からぬ
実行可能性	いいえ	おそらく, いいえ	おそらく, はい	はい		さまざま	分からぬ

推奨のタイプ

当該介入に反対する 強い推奨	当該介入に反対する 条件付きの推奨	当該介入または比較 対照のいずれかに ついての条件付きの 推奨	当該介入の条件付き の推奨	当該介入の強い推奨
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

結論

推奨
正当性
サブグループに関する検討事項
実施に関わる検討事項
監視と評価
研究上の優先事項

出典 : Schünemann H, Brozek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook. より作成

[RC-8 EtD フレームワーク (Clinical recommendation: Population perspective)]

疑問

CQ :	
集団	
介入	
比較対照	
主要なアウトカム	
セッティング	
視点	
背景	
利益相反	

評価

基準 1. 問題 この問題は優先事項か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> いいえ <input type="radio"/> おそらく、いいえ <input type="radio"/> おそらく、はい <input type="radio"/> はい <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 2. 望ましい効果 予期される望ましい効果はどの程度のものか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> わずか <input type="radio"/> 小さい <input type="radio"/> 中 <input type="radio"/> 大きい <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 3. 望ましくない効果 予期される望ましくない効果はどの程度のものか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 大きい <input type="radio"/> 中 <input type="radio"/> 小さい <input type="radio"/> わずか <input type="radio"/> さまざま <input type="radio"/> 分からない		

基準 4. エビデンスの確実性 効果に関する全体的なエビデンスの確実性はどの程度か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 非常に弱い <input type="radio"/> 弱 <input type="radio"/> 中 <input type="radio"/> 強 <input type="radio"/> 採用研究なし		
基準 5. 価値観 人々が主要なアウトカムをどの程度重視するかについて重要な不確実性やばらつきはあるか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 重要な不確実性またはばらつきあり <input type="radio"/> 重要な不確実性またはばらつきの可能性あり <input type="radio"/> 重要な不確実性またはばらつきはおそらくなし <input type="radio"/> 重要な不確実性またはばらつきはなし		
基準 6. 効果のバランス 望ましい効果と望ましくない効果のバランスは介入もしくは比較対照を支持するか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 比較対照が優れている <input type="radio"/> 比較対照がおそらく優れている <input type="radio"/> 介入も比較対照もいずれも支持しない <input type="radio"/> おそらく介入が優れている <input type="radio"/> 介入が優れている <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 7. 必要資源量 資源利用はどの程度大きいか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 大きな増加 <input type="radio"/> 中等度の増加 <input type="radio"/> 無視できるほどの増加や減少 <input type="radio"/> 中等度の減少 <input type="radio"/> 大きな減少 <input type="radio"/> さまざま <input type="radio"/> 分からない		

基準 8. 必要資源量に関するエビデンスの確実性 資源利用に関するエビデンスの確実性はどの程度か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 非常に弱い <input type="radio"/> 弱 <input type="radio"/> 中 <input type="radio"/> 強 <input type="radio"/> 採用研究なし		
基準 9. 費用対効果 その介入の費用対効果は介入または比較対照のどちらを支持するか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 比較対照の費用対効果がよい <input type="radio"/> 比較対照の費用対効果がおそらくよい <input type="radio"/> 介入も比較対照もいずれも支持しない <input type="radio"/> 介入の費用対効果がおそらくよい <input type="radio"/> 介入の費用対効果がよい <input type="radio"/> さまざま <input type="radio"/> 採用研究なし		
基準 10. 公平性 医療上の公平性への影響はどの程度か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> 減る <input type="radio"/> おそらく減る <input type="radio"/> おそらく影響なし <input type="radio"/> おそらく増える <input type="radio"/> 増える <input type="radio"/> さまざま <input type="radio"/> 分からない		
基準 11. 容認性 この選択肢は重要な利害関係者にとって妥当なものか？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> いいえ <input type="radio"/> おそらく、いいえ <input type="radio"/> おそらく、はい <input type="radio"/> はい <input type="radio"/> さまざま <input type="radio"/> 分からない		

基準 12. 実行可能性 その介入は実行可能か？		
判断	リサーチエビデンス	追加的考察
<input type="radio"/> いいえ <input type="radio"/> おそらく、いいえ <input type="radio"/> おそらく、はい <input type="radio"/> はい <input type="radio"/> さまざま <input type="radio"/> 分からない		

判断の要約

問題	判断						
	いいえ	おそらく, いいえ	おそらく, はい	はい		さまざま	分からぬ
望ましい効果	わずか	小さい	中	大きい		さまざま	分からぬ
望ましくない 効果	大きい	中	小さい	わずか		さまざま	分からぬ
エビデンスの 確実性	非常に弱い	弱	中	強			採用研究 なし
価値観	重要な不確実 性またはばら つきあり	重要な不確実 性またはばら つきの可能性 あり	重要な不確実 性またはばら つきはおそら くなし	重要な不確実 性またはばら つきはなし			
効果のバランス	比較対照が 優れている	比較対照が おそらく 優れている	介入も比較対 照もいすれも 支持しない	おそらく 介入が 優れている	介入が 優れている	さまざま	分からぬ
必要資源量	大きな増加	中等度の 増加	無視できる ほどの 増加や減少	中等度の 減少	大きな減少	さまざま	分からぬ
必要資源量に關 するエビデンス の確実性	非常に弱い	弱	中	強			採用研究 なし
費用対効果	比較対照の 費用対効果が よい	比較対照の 費用対効果が おそらくよい	介入も比較対 照もいすれも 支持しない	介入の 費用対効果が おそらくよい	介入の 費用対効果が よい	さまざま	採用研究 なし
公平性	減る	おそらく 減る	おそらく 影響なし	おそらく 増える	増える	さまざま	分からぬ
容認性	いいえ	おそらく, いいえ	おそらく, はい	はい		さまざま	分からぬ
実行可能性	いいえ	おそらく, いいえ	おそらく, はい	はい		さまざま	分からぬ

推奨のタイプ

当該介入に反対する 強い推奨	当該介入に反対する 条件付きの推奨	当該介入または比較 対照のいずれかに ついての条件付きの 推奨	当該介入の条件付き の推奨	当該介入の強い推奨
○	○	○	○	○

結論

推奨
正当性
サブグループに関する検討事項
実施に関わる検討事項
監視と評価
研究上の優先事項

出典：Schünemann H, Brozek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook. より作成

6章 テンプレート 記入方法

【RC-1 推奨文草案 (Individual perspective) 記入方法】

1. CQ												
2. 推奨文草案												
3. 作成グループにおける、推奨に関連する価値観や意向（検討したアウトカム別に、一連の価値観を想定する） CQ に対して提示された全アウトカムの中で、推奨文草案を作成するにあたって最も重要・重大と考えられる価値を有するもの、または、全アウトカムの総まとめを記載する。 (例：この推奨の作成にあたっては、致死率の高い疾患における死亡の回避を重要視した。)												
4. CQに対するエビデンス総括（重大なアウトカム全般に関する全体的なエビデンスの強さ） A（強） B（中） C（弱） D（非常に弱い）												
5. 推奨の強さを決定するための評価項目（下記の項目について総合して判定する）												
<table border="1"><thead><tr><th>推奨の強さの決定に影響する要因</th><th>判定</th><th>説明</th></tr></thead><tbody><tr><td>アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。</td><td><input type="checkbox"/> はい <input type="checkbox"/> いいえ</td><td></td></tr><tr><td>益と害のバランスが確実（コストは含まず） ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。</td><td><input type="checkbox"/> はい <input type="checkbox"/> いいえ</td><td></td></tr><tr><td>推奨の強さに考慮すべき要因 患者・市民の価値観・希望や好み、負担の確実さ（あるいは相違）、医療費のうち自己負担分、患者の立場から見たその他の資源利用など：</td><td></td><td></td></tr></tbody></table>	推奨の強さの決定に影響する要因	判定	説明	アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ		益と害のバランスが確実（コストは含まず） ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ		推奨の強さに考慮すべき要因 患者・市民の価値観・希望や好み、負担の確実さ（あるいは相違）、医療費のうち自己負担分、患者の立場から見たその他の資源利用など：		
推奨の強さの決定に影響する要因	判定	説明										
アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ											
益と害のバランスが確実（コストは含まず） ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input type="checkbox"/> いいえ											
推奨の強さに考慮すべき要因 患者・市民の価値観・希望や好み、負担の確実さ（あるいは相違）、医療費のうち自己負担分、患者の立場から見たその他の資源利用など：												
6. 費用対効果の観点からの留意事項（費用対効果を検討した場合のみ記載するが、臨床的な推奨とは別に取り扱う）：												

明らかに当てはまる場合は「はい」とし、それ以外は、どちらともいえないを含め「いいえ」とする。

【RC-5 推奨作成の経過 記入方法】

推奨決定に採用した各研究の強みと限界、定性的・定量的システムティックレビューの結果を「日本語として流暢な文章で記載する」（詳細は **6.4 節** 参照）。

1. CQ が重要な臨床課題である旨の背景を記載する。

2. エビデンス評価

アウトカムごとにシステムティックレビューを行った過程と結果（説明）を記載する。

3. 益と害のバランス評価

システムティックレビューで得られた望ましいアウトカムと望ましくないアウトカムのバランスを比較検討した内容を記載する。

4. 患者・市民の価値観・希望

患者・市民の価値観・希望の多様性や不確実性について検討した結果を記載する。

5. 資源利用と費用対効果

個人の立場 (individual perspective) の場合、5. には患者の自己負担分、その他患者の立場から見て重要な資源利用の情報を記載する。6. には公的医療の立場から費用対効果のエビデンスの評価を行った場合は、その結果概要を記入する。

集団の立場 (population perspective) の場合、5. の最下段に資源利用および費用対効果の情報を記載する。Individual perspective または population perspective のいずれの場合においても、保険診療であるか否かは極めて重要な情報である。保険診療でない介入方法については、それを明記する。

6. 推奨がどうしても決定できないときは、まれに「明確な推奨ができない」とする場合もある。この場合、その経過と討論内容を本文中に記載する。

7. 今後の研究についても考慮する

・今後の研究を用いることが適正か検討し、提示する。

提示された今後の研究は、今後臨床研究として推進され、次回改訂時には、エビデンスとして活用されることが期待される。

・今後の研究が推奨される場合の記載方法に注意する。

【RC-6 一般向けサマリー 記入方法】

CQごとにシステムティックレビューの結果と推奨を、患者・市民にも理解できる平易な言葉で解説する。Q & Aなど、医療現場で意思決定に使用されることを想定して表現を工夫することが望ましい。

6章 テンプレート 記入例

【RC-1 推奨文草案 (Individual perspective) 記入例】

1. CQ

80歳未満の発症後3時間を超えた急性脳梗塞患者に対して、総死亡率低下、ADL保持、脳出血防止を考慮した場合、6時間以内のrt-PA投与は勧められるか？

2. 推奨文草案

80歳未満の発症後3時間を超えた急性脳梗塞患者に対して、総死亡率低下、ADL保持、脳出血防止を考慮した場合、6時間以内のrt-PA投与を行うことを提案する。

3. 作成グループにおける、推奨に関連する価値観や意向（検討したアウトカム別に、一連の価値観を想定する）

本CQに対する推奨の作成にあたっては、急性脳梗塞患者に対する死亡率の低下、ADL保持、脳出血防止を重要視した。

4. CQに対するエビデンス総括（重大なアウトカム全般に関する全体的なエビデンスの強さ）

A (強) **B (中)** C (弱) D (非常に弱い)

5. 推奨の強さを決定するための評価項目（下記の項目について総合して判定する）

推奨の強さの決定に影響する要因	判定	説明
アウトカム全般に関する全体的なエビデンスが強い ・全体的なエビデンスが強いほど推奨度は「強い」とされる可能性が高くなる。 ・逆に全体的なエビデンスが弱いほど、推奨度は「弱い」とされる可能性が高くなる。	<input checked="" type="checkbox"/> はい <input type="checkbox"/> いいえ	エビデンスの強さはB
益と害のバランスが確実（コストは含まず） ・望ましい効果と望ましくない効果の差が大きければ大きいほど、推奨度が強くなる可能性が高い。 ・正味の益が小さければ小さいほど、有害事象が大きいほど、益の確実性が減じられ、推奨度が「弱い」とされる可能性が高くなる。	<input type="checkbox"/> はい <input checked="" type="checkbox"/> いいえ	脳出血は増加するが、死亡率はほとんど差がなく、ADL保持もわずかに改善する。これは、アルテプラゼ投与によって改善する患者もいることを示している。
推奨の強さに考慮すべき要因		
患者・市民の価値観・希望や好み、負担の確実さ（あるいは相違）、医療費のうち自己負担分、患者の立場から見たその他の資源利用など： この治療に対する患者（家族）の意向は、大きくばらつくと考えられる。 薬品の単価は高額だが、入院期間延長による経費や介護費用等の増額も不明確である。		
6. 費用対効果の観点からの留意事項（費用対効果を検討した場合のみ記載するが、臨床的な推奨とは別に取り扱う）： 評価未実施		

明らかに当てはまる場合は「はい」とし、それ以外は、どちらともいえないを含め「いいえ」とする。

【RC-3 推奨の強さ決定投票用紙 記入例】

CQ :	
80歳未満の発症後3時間を超えた急性脳梗塞患者に対して、総死亡率低下、ADL保持、脳出血防止を考慮した場合、6時間以内のrt-PA投与を行うべきか？	
推奨文 :	
推奨 (いずれかに○)	<ul style="list-style-type: none">・行うことを強く推奨する<input checked="" type="radio"/> ①行うことを弱く推奨する（提案する）・行わないことを弱く推奨する（提案する）・行わないことを強く推奨する

【RC-4 推奨提示 記入例】

CQ : 80歳未満の発症後3時間を超えた急性脳梗塞患者に対して、総死亡率低下、ADL保持、脳出血防止を考慮した場合、6時間以内のrt-PA投与を行うべきか？	
推奨文 : 80歳未満の発症後3時間を超えた急性脳梗塞患者に対して、総死亡率低下、ADL保持、脳出血防止を考慮した場合、6時間以内のrt-PA投与を行うことを提案する。	
推奨の強さ (いずれかに○)	1 (強い) : 「実施する」、または、「実施しない」ことを推奨する ②(弱い) : 「実施する」、または、「実施しない」ことを提案する
エビデンスの強さ (いずれかに○)	A (強) B (中) C (弱) D (非常に弱い)
費用対効果の観点からの留意事項	評価未実施

どうしても決定できないときは、まれに「明確な推奨ができない」とする場合もある。この場合、その経過と討論内容を本文中に記載する。

6章 文献

- Akl EA, et al. (2017) GRADE equity guidelines 2: considering health equity in GRADE guideline development: equity extension of the guideline development checklist. *J Clin Epidemiol* 90:68-75. doi: [10.1016/j.jclinepi.2017.01.017](https://doi.org/10.1016/j.jclinepi.2017.01.017). PMID: [28499847](#)
- Alonso-Coello P, et al. (2016a) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 353:i2016. doi: [10.1136/bmj.i2016](https://doi.org/10.1136/bmj.i2016). PMID: [27353417](#)
- Alonso-Coello P, et al. (2016b) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ* 353:i2089. doi: [10.1136/bmj.i2089](https://doi.org/10.1136/bmj.i2089). PMID: [27365494](#)
- Alper BS, et al. (2019) Defining certainty of net benefit: a GRADE concept paper. *BMJ Open* 9:e027445. doi: [10.1136/bmjopen-2018-027445](https://doi.org/10.1136/bmjopen-2018-027445). PMID: [31167868](#)
- Andrews JC, et al. (2013a) GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 66:719-725. doi: [10.1016/j.jclinepi.2012.03.013](https://doi.org/10.1016/j.jclinepi.2012.03.013). PMID: [23312392](#).
- Andrews JC, et al. (2013b) GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 66:726-735. doi: [10.1016/j.jclinepi.2013.02.003](https://doi.org/10.1016/j.jclinepi.2013.02.003). PMID: [23570745](#).
- Boyd CM, et al. (2012) Methods for Benefit and Harm Assessment in Systematic Reviews. AHRQ Methods for Effective Health Care (book) PMID: [23326898](#)
- Broekhuizen H, et al. (2015) A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. *Pharmacoconomics* 33:445-455. doi: [10.1007/s40273-014-0251-x](https://doi.org/10.1007/s40273-014-0251-x). PMID: [25630758](#)
- Delbecq AL, et al. (1971) A group process model for problem identification and Program planning. *The Journal of Applied Behavioral Science* 7:466-492
- Dalkey N, et al. (1969) The Delphi Method, III: Use of Self Ratings to Improve Group Estimates. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2006/RM6115.pdf (参照 2021/02/25)
- Dellinger RP, et al. (2008) Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008. *Crit Care Med* 36:296-327. doi: [10.1097/01.CCM.0000298158.12101.41](https://doi.org/10.1097/01.CCM.0000298158.12101.41). PMID: [18158437](#)
- Dolan JG, et al. (2013) Patients' preferences and priorities regarding colorectal cancer screening. *Med Decis Making* 33:59-70. doi: [10.1177/0272989X12453502](https://doi.org/10.1177/0272989X12453502). PMID: [22895558](#)
- Fitch K, et al. (2001) The RAND/UCLA Appropriateness Method User's Manual. RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1269.html (参照 2021/02/25)
- Gail MH, et al. (1999) Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst* 91:1829-46. doi: [10.1093/jnci/91.21.1829](https://doi.org/10.1093/jnci/91.21.1829). PMID: [10547390](#)

- Guyatt G, et al. (2013) GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol* 66:151-157. doi: [10.1016/j.jclinepi.2012.01.006](https://doi.org/10.1016/j.jclinepi.2012.01.006). PMID: [2254203](#)
- Institute of Medicine (IOM). (2011) Clinical Practice Guidelines We Can Trust. National Academies Press
- Jaeschke R, et al. (2008) Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ* 337:327-330. doi: [10.1136/bmj.a744](https://doi.org/10.1136/bmj.a744). PMID: [18669566](#)
- Keeney R, et al. (1993) Decisions with multiple objectives: Preferences and Value Trade-Offs. Cambridge University Press
- Marsh K, et al. (2016) Multiple criteria decision analysis for health care decision making--emerging good practices: report 2 of the ISPOR MCDA emerging good practices task force. *Value Health* 19:125-37. doi: [10.1016/j.jval.2015.12.016](https://doi.org/10.1016/j.jval.2015.12.016). PMID: [27021745](#)
- Marsh K, et al. (2017) Multi-Criteria Decision Analysis to Support Healthcare Decisions. Springer
- Moberg J, et al. (2015) GRADE Evidence to Decision (EtD) Frameworks Guidance. Version 1.1 [updated May 2015], The GRADE Working Group. <https://ietd.epistemonikos.org/#/help/guidance> (参照 2021/02/25)
- Mt-Isa S, et al. (2014) Balancing benefit and risk of medicines: a systematic review and classification of available methodologies. *Pharmacoepidemiol Drug Saf* 23:667-78. doi: [10.1002/pds.3636](https://doi.org/10.1002/pds.3636). PMID: [24821575](#)
- National Institute for Health and Care Excellence (NICE). (2015) Research recommendations Process and methods guide. <https://www.nice.org.uk/Media/Default/About/what-we-do/Science-policy-and-research/research-recommendation-process-methods-guide-2015.pdf> (参照 2021/02/25)
- Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) <http://www.protectbenefitrisk.eu/index.html> (参照 2021/02/25)
- Pottie K, et al. (2017) GRADE equity guidelines 4: considering health equity in GRADE guideline development: evidence to decision process. *J Clin Epidemiol* 90:84-91. doi: [10.1016/j.jclinepi.2017.08.001](https://doi.org/10.1016/j.jclinepi.2017.08.001). PMID: [28802675](#)
- Rehfuss EA, et al. (2019) The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. *BMJ Glob Health* 4:e000844. doi: [10.1136/bmjjgh-2018-000844](https://doi.org/10.1136/bmjjgh-2018-000844). PMID: [30775012](#)
- Scottish Intercollegiate Guidelines Network (SIGN). (2011) SIGN 50: A guideline developer's handbook. Revised Edition. https://www.sign.ac.uk/media/1642/sign50_2011.pdf (参照 2021/02/25)
- Thokala P, et al. (2016) Multiple criteria decision analysis for health care decision making--an introduction: report 1 of the ISPOR MCDA emerging good practices task force. *Value Health* 19:1-13. doi: [10.1016/j.jval.2015.12.003](https://doi.org/10.1016/j.jval.2015.12.003). PMID: [26797229](#)
- Welch VA, et al. (2017a) GRADE equity guidelines 1: considering health equity in GRADE guideline development: introduction and rationale. *J Clin Epidemiol* 90:59-67. doi: [10.1016/j.jclinepi.2017.01.014](https://doi.org/10.1016/j.jclinepi.2017.01.014). PMID: [28412464](#)

Welch VA, et al. (2017b) GRADE equity guidelines 3: considering health equity in GRADE guideline development: rating the certainty of synthesized evidence. *J Clin Epidemiol* 90:76-83. doi: [10.1016/j.jclinepi.2017.01.015](https://doi.org/10.1016/j.jclinepi.2017.01.015). PMID: [28389397](https://pubmed.ncbi.nlm.nih.gov/28389397/)

Wen S, et al. (2014) Two approaches to incorporate clinical data uncertainty into multiple criteria decision analysis for benefit-risk assessment of medicinal products. *Value Health* 17:619-628. doi: [10.1016/j.jval.2014.04.008](https://doi.org/10.1016/j.jval.2014.04.008). PMID: [25128056](https://pubmed.ncbi.nlm.nih.gov/25128056/)

World Health Organization (WHO). (2014) WHO handbook for guideline development, Second ed. <https://apps.who.int/iris/handle/10665/145714> (参照 2021/02/25)

第7章

公開に向けた最終調整

Minds

7章 Key Words

◎ AGREE II (The Appraisal of Guidelines for Research & Evaluation II)
◎ AGREE Reporting Checklist
◎ AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews 2)
◎ Institute of Medicine (IOM)
◎ IOM の基準 (ガイドライン評価における)
◎ PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)
◎ RIGHT (Reporting Items for practice Guidelines in HealThcare)
◎ エビデンス総体 (body of evidence)
◎ ガイドラインサマリー
◎ 外部評価
◎ 外部評価委員会
協働意思決定 (shared decision making : SDM)
◎ クオリティーインディケータ (quality indicator : QI)
◎ クリニカルクエスチョン (clinical question : CQ)
◎ 最終調整 (finalization)
◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
◎ システマティックレビューチーム (systematic review team : SR チーム)
システムティックレビューレポート (サマリーレポートともいう)
◎ 診療アルゴリズム
◎ 診療ガイドライン (clinical practice guideline : CPG)
◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
◎ 診療ガイドライン作成事務局
◎ 診療ガイドライン統括委員会
◎ 推奨
◎ スコープ (scope)
導入 (implementation)
評価 (assessment)
普及 (dissemination)
◎ メタアナリシス (meta-analysis, メタ解析と同義)
◎ 利益相反 (conflict of interest : COI)

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】 SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】 RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】 PO : post 【第8章 診療ガイドライン公開後の取り組み】

7.0 概要

推奨作成が終了した後に、それまでに作成した原稿を取りまとめて診療ガイドライン（clinical practice guideline : CPG）草案を作成し、外部評価などの手順を踏んで最終版を確定する一連の作業を、公開に向けた最終調整（finalization）として本章でまとめます。

公開に向けた最終調整の過程で、診療ガイドラインを公開した後の対応方針の決定（**第8章**）を行い、それを草案に反映させる。草案作成までは診療ガイドライン作成グループ（guideline development group : GDG、診療ガイドライン作成パネルともいう）が担当するが、診療ガイドラインの作成経過の報告を草案に反映させた上で、最終版を承認するプロセスは診療ガイドライン統括委員会が行う。

○ ステップ

推奨作成が完了した後に公開に向けた最終調整のプロセスを開始する。診療ガイドライン作成グループが本プロセスを完了する。

- (1) 診療ガイドライン公開後の対応について協議と決定
- (2) 作成経過に関する報告事項の作成
- (3) 診療ガイドライン草案の決定
- (4) 外部評価の実施
- (5) 診療ガイドライン最終案の決定

7.1 ステップ1：診療ガイドライン公開後の対応について協議と決定

診療ガイドラインには、公開した診療ガイドラインを普及（dissemination）、導入（implementation）、そして、評価（assessment）するプロセス（図7-1）についても記載することが望ましい。具体的には、公開後の組織体制（8.1節）、導入（8.2節）、普及・活用・効果の評価（8.3節）、改訂（8.4節）について、継続的に実施する活動の内容を診療ガイドラインに記載する。



図7-1 診療ガイドラインの作成、普及、導入、評価の4段階

● 手順

- ・公開後に診療ガイドライン作成主体として実施すべき事項を整理し、それに対応できる公開後の組織体制を検討する。
- ・導入に関する事項としては、診療ガイドラインが広く活用されるための取り組みを検討することを目的として、活用の促進要因、阻害要因を整理する。さらに、阻害要因への対策として取り得る方策を検討し、実行可能性も検討して、公開後の取り組み案を決定する。

● テンプレート

(1) 公開後の組織体制

【PO-1 公開後の組織体制】

(2) 導入

【PO-2 導入】

(3) 普及・活用・効果の評価

【PO-3 普及・活用・効果の評価】

(4) 改訂

【PO-4 改訂】

● テンプレート 記入方法

(1) 公開後の組織体制 【PO-1 公開後の組織体制 記入方法】

診療ガイドライン統括委員会、診療ガイドライン作成グループなどが診療ガイドラインを公開した後にどのような形で組織を継続するかについて方針を決定して、診療ガイドラインに記載する。

(2) 導入 【PO-2 導入 記入方法】

導入とは、診療ガイドラインが、活用が想定される場で適切に活用されるように、さまざまな工夫によって適用可能性（applicability）を高めることをいう。要約版の作成、多様な情報媒体の活用、診療ガイドライン活用の促進要因と阻害要因について、診療ガイドライン作成グループ会議で検討し、対応

方針を記入する。

・要約版の作成

詳細版（full version） 診療ガイドラインを基にした日常診療で参照が可能な「実用版」、日常診療の現場で迅速に活用できるクイックレファレンスとしての「簡易版」を作成する場合は、その予定を記載する。また、患者・市民のために、診療ガイドラインを一般向けに解説する「一般向けガイドライン解説」を作成する場合は記載する。

・多様な情報媒体の活用

印刷版として書籍、小冊子など、電子版としてウェブサイトでの公開、オンラインアプリケーションによる提供など、利用媒体の予定について記載する。

・診療ガイドライン活用の促進要因と阻害要因

診療ガイドラインの導入にあたって考えられる促進要因、阻害要因を、診療ガイドライン自体の問題、提供方法の問題、利用者側の問題などに分けて検討し、対応方針を記載する。

(3) 普及・活用・効果の評価 【PO-3 普及・活用・効果の評価 記入方法】

診療ガイドラインの評価として、アンケートなどを用いた普及に関する調査、診療情報の集積による活用に関する調査、クオリティーインディケータ（quality indicator : QI）などによる効果に関する評価などを実施するかどうか、その方針を検討して記載する。

(4) 改訂 【PO-4 改訂 記入方法】

診療ガイドライン公開後の改訂に関する方針について、次の項目を中心に記載する（改訂に関する作業方法の詳細については、**第8章**参照）。

・実施時期

改訂版を発行する時期、もしくは、改訂作業を行う時期・タイミング

・実施方法

改訂について検討する方法

改訂するかどうかを判定する基準（例：エビデンスの変化、制度の変化、他）等

・実施体制

改訂についての検討を行う組織体制

● テンプレート 記入例

【PO-4 改訂 記入例】

7.2 ベース2：作成経過に関する報告事項の作成

以下のコンテンツを作成し、診療ガイドライン草案の中に含める。

① ガイドラインサマリー

診療ガイドラインが取り上げたクリニカルクエスチョン（clinical question : CQ）と推奨のリストを提示する。

② 診療アルゴリズム

疾患トピックの基本的特徴として作成した診療アルゴリズムの図を基に、診療ガイドラインで作成した推奨を反映して、診療アルゴリズムの最終版を作成する。

③ 用語

診療ガイドラインの中で使用される重要な用語について、定義を記載する。その際に、草案の中で、重要な用語が適切に統一されていることを確認する。また、診療ガイドラインの中で使用される略語について、一括して説明する。

④ 作成経過

作成経過は、以下のような項目について記載する。具体例は**テンプレート【FN-4】記入方法**を参照のこと。

- ・作成方針
- ・使用上の注意
- ・利益相反（conflict of interest : COI）
- ・作成資金
- ・組織編成
- ・作成工程

①ガイドラインサマリー、②診療アルゴリズム、③用語は診療ガイドライン作成グループが作成する。④作成経過は、診療ガイドライン統括委員会が決定した事項について記載する。その他に記載すべき項目がある場合は追加する。

● テンプレート

【FN-1 ガイドラインサマリー】

【SC-2 診療アルゴリズム】

【FN-2 重要用語の定義】

【FN-3 略語一覧】

【FN-4 作成経過】

● テンプレート 記入方法

[FN-1 ガイドラインサマリー 記入方法]

[SC-2 診療アルゴリズム 記入方法]

[FN-2 重要用語の定義 記入方法]

[FN-3 略語一覧 記入方法]

[FN-4 作成経過 記入方法]

● テンプレート 記入例

[SC-2 診療アルゴリズム 記入例]

7.3 ベース3：診療ガイドライン草案の決定

ここまでに作成したテンプレートを基に、診療ガイドライン草案を作成する。

● 手順

診療ガイドライン草案のレイアウトとテンプレートの関係は以下のとおりである。

目次

タイトルページ

- タイトル
- 作成主体（学会名等）
- 版
- 発行年月日

前付

- ガイドラインサマリー
【FN-1 ガイドラインサマリー】
- 診療アルゴリズム
【SC-2 診療アルゴリズム】
- 用語・略語一覧
【FN-2 重要用語の定義】
【FN-3 略語一覧】

(I) 作成組織・作成方針

1. 作成組織
 - 1.1. 作成主体
 - 1.2. 診療ガイドライン統括委員会
 - 1.3. 診療ガイドライン作成グループ
 - 1.4. システマティックレビューチーム
 - 1.5. 外部評価委員会
 - 1.6. 診療ガイドライン作成事務局

(以上まとめて) 【PR-6 診療ガイドライン作成組織】

2. 作成経過
 - 2.1. 作成方針
 - 2.2. 使用上の注意
 - 2.3. COI
 - 2.4. 作成資金

2.5. 組織編成

2.6. 作成工程

(以上をまとめて) **【FN-4 作成経過】**

(II) スコープ

1. 疾患トピックの基本的特徴

1.1. 臨床的特徴

1.2. 痘学的特徴

1.3. 疾患トピックの診療の全体的な流れ

(以上をまとめて) **【SC-1 疾患トピックの基本的特徴】**

2. 診療ガイドラインがカバーする内容に関する事項

3. システマティックレビューに関する事項

4. 推奨決定から公開に向けた最終調整、導入方針まで

(2, 3, 4 をまとめて) **【SC-3 スコープ】**

(III) 推奨

1. CQ1 (CQ ごとに以下の事項を記載する)

1.1. CQ 本文

1.2. 推奨本文・エビデンスの強さ・推奨の強さ

(1.1 と 1.2 をまとめて) **【RC-4 推奨提示】**

1.3. 解説

【RC-5 推奨作成の経過】

1.4. 一般向けサマリー

【RC-6 一般向けサマリー】

1.5. 引用文献

【SR-4 引用文献リスト】

2. CQ2

: (同上)

(以下、CQ の数だけ繰り返す)

(IV) 公開後の取り組み

診療ガイドラインの日常診療への導入と活用促進を図るための方策、改訂予定などを記載する。

1. 公開後の組織体制

【PO-1 公開後の組織体制】

2. 導入

【PO-2 導入】

3. 普及・活用・効果の評価

【PO-3 普及・活用・効果の評価】

4. 改訂

【PO-4 改訂】

(V) 付録

1. 参考資料

2. 作業資料

・CQ 設定表

【SC-4 CQ の設定】

・全ての文献検索データベースごとの検索式とフローチャートと文献

【SR-1 データベース検索結果】

【SR-2 文献検索フローチャート】

【SR-3 二次スクリーニング後の一覧表】

【SR-4 引用文献リスト】

・エビデンスの評価シート アウトカムごとの **STEP 1** シート

【SR-5 評価シート 介入研究】

【SR-6 評価シート 観察研究】

・エビデンスの統合シート CQ ごとの **STEP 2** シート

【SR-7 評価シート エビデンス総体】

・エビデンスのまとめ

【SR-9 定性的システムティックレビュー】

【SR-10 メタアナリシス】

【SR-11 システマティックレビューレポートのまとめ】

・外部評価まとめ

【FN-6 外部評価まとめ】

7.4 利用者視点からの再確認

診療ガイドラインは患者と医療者の双方の意思決定を支援するものであり、その内容は患者・市民にとっても、また医療者にとっても、理解しやすいものとなっている必要がある。そのため、公開に向けた最終調整にあたり、以下の点に配慮がされているかどうか確認しておくことが期待される。

① 誤解のない理解に向けて

利用者が推奨の内容を正しく理解するためには、推奨の内容が分かりやすく、論理的であることが求められる。また、診療全体におけるその推奨の位置付けに対する理解を助けるためには診療アルゴリズムなどの提示が有用と考えられる。さらに、作成者、発行者、作成年月日などの書誌情報の明記も重要である。誤解される恐れがある用語が含まれる場合には用語集を、略語が多用される場合には略語集を準備するなども有用であろう。

② 臨床現場における活用の促進

診療ガイドラインの臨床における活用促進のためには、時間の限られた状況での利用を想定した工夫が求められる。例えば、診療ガイドラインの主要なメッセージであるCQと推奨が容易に見つけられるよう、該当箇所が分かるようなクイックレファレンスガイド、CQ-推奨一覧などの提供が考えられる。

また、医療現場では、従来の書籍、論文などの媒体でなく、スマートフォンやタブレットなどの端末で閲覧されることも考慮する。現在の情報環境と利便性を鑑み、多様な利用形態に合わせて診療ガイドラインを提供することが求められる。

③ 患者・市民の活用の促進

診療ガイドラインの本来の目的は、医療者だけを対象とした資料となることではなく、患者と医療者との協働意思決定（shared decision making : SDM）を支援することにある。そのため、患者・市民は本来の診療ガイドラインの主たる利用者の一部である。診療ガイドラインの作成においては、エビデンスなどに合わせて、患者・市民の価値観・希望を考慮し、CQや推奨に反映させることで、患者・市民による活用を促進することができると考えられる。また、図を用い、分かりやすい用語で記述された患者・市民に向けた診療ガイドラインの解説を準備することも、患者・市民の診療ガイドラインに対する理解を助けることになると期待される。

7.5 ベンチマーク調査の実施

7.5.1 外部評価とは

外部評価とは、診療ガイドラインの作成中、または、診療ガイドラインを公開する前に、スコープ、CQ、システムティックレビュー、エビデンス評価、推奨、診療ガイドライン全体の草案などについて、それを作成・実行したグループ以外から評価を受けることである。

診療ガイドライン作成過程全体を通して、診療ガイドライン作成者は不偏性を保つよう努めることが求められるが、それでも作成過程、および、完成した診療ガイドラインから完全に偏りを排除することは非常に困難である。作成過程には限られた専門家が参加し、採用される考え方も限定的にならざるを得ないとされている（IOM 2011a: 131）。そこで、診療ガイドラインの不偏性をさらに高めるために、診療ガイドラインを作成したグループ以外から評価を受ける必要がある。そのような評価を外部評価と呼ぶ。

外部評価における「外部」の意味は診療ガイドラインを作成するグループの構成によって異なる。ここでは、診療ガイドライン作成の実働に携わってきた主体とは別の主体によって行われる評価を「外部」評価と呼ぶ。この中には、診療ガイドライン作成組織全体の外部から評価を受ける場合（パブリックコメント、コンサルテーションなど）、また、診療ガイドライン作成組織全体の中に、診療ガイドライン作成グループやシステムティックレビューチームとは別に評価を行う外部評価委員会（評価委員会、評価グループ、査読グループなど）を設けて評価を受ける場合などが含まれる。

なお、診療ガイドライン公開後に作成主体とは異なる主体から評価を受けることもある。公開後の取り組みについては第8章で紹介する。

7.5.2 外部評価の計画（目的・実施数段階・評価者・ツール）

外部評価の目的を設定し、その目的に見合った実施数段階、評価者、評価方法を決定する。

診療ガイドラインの質を確かなものにするために、必要と考える外部評価を実施する段階・評価者・評価方法を決定する。

診療ガイドラインを公開する前の外部評価についてまとめると表7-1 のようになる。

表7-1 外部評価の実施段階・目的・評価者・ツール

実施段階	スコープ草案	システムティックレビュー草案	診療ガイドライン草案
目的	スコープのうち、作成方法、CQ、診療の流れ・位置付け、他の妥当性を評価する	システムティックレビューの内容（文献の包括性、エビデンスの統合・評価法、結果、他）の妥当性を評価する	診療ガイドライン全体（推奨の内容、作成方法、他）の妥当性を評価する
評価者	疾患の専門家 医療者 作成方法専門家 患者・市民、など	疾患の専門家 医療者 作成方法専門家 システムティックレビュー専門家	疾患の専門家 医療者 作成方法専門家 患者・市民、など
ツール	AGREE II, AGREE Reporting Checklist, GLIA 2.0, IOM 2011a, 他	IOM 2011b, AMSTAR 2, PRISMA, 他	AGREE II, AGREE Reporting Checklist, GLIA 2.0, IOM 2011a, 他

(1) 外部評価の目的

外部評価を行う際には、目的を明確にしておく必要がある。目的には、推奨の内容的妥当性の確認、診療ガイドラインの作成方法の確認などがある。

(2) 外部評価を実施する段階

外部評価を受けることは、作成・公開の進行を一時停止することになるため、いつでもできるものではない。ここでは、次の3つの段階で外部評価を受けることを提案したい。

1つ目はスコープの草案が完成した段階である。診療ガイドラインの企画書であるスコープはその後の作成過程全体の指針となるものであり、多様な利害関係者による評価を受けることが望ましい。

2つ目はシステムティックレビュー（systematic review：SR、系統的レビューと同義）の草案が完成した段階が挙げられる。システムティックレビューは推奨を作成する際のベースとなるものである。システムティックレビューの評価については、システムティックレビューチーム内、もしくは、診療ガイドライン作成グループによる評価を行うことで、外部評価委員による評価やパブリックコメントを実施しなくても十分な場合も多いと考えられる。しかし、さらに専門的な観点から評価を受けることが必要であると判断した場合には、専門家による外部評価を実施してもよい。

3つ目は診療ガイドライン全体の草案が完成した段階である。診療ガイドラインは、社会に広く大きな影響を与える可能性がある。そこで、診療ガイドライン全体がその利用者にとって有用なものとなっているか、さまざまな利用者の観点から評価される必要がある。

(3) 外部評価を行う評価者

外部評価を行う際には、その目的に照らして、誰から評価を受けるか明確にしておく必要がある。診療ガイドライン統括委員会が作成の方針を決める際に、外部評価を受けるにふさわしい集団が特定できている場合には、その集団の代理人を集めた外部評価委員を選定しておく。また、広く利用者全般から評価を受けることが望ましいと考えられる場合には、パブリックコメントを受ける機会を設ける。このように、外部評価の主体としては、診療ガイドライン統括委員会の決

めた外部評価委員と、パブリックコメント等の手段で評価を求める利用者一般などがある。

なお、外部評価を依頼する場合には、外部評価を行う者の COI について検討する（COI の一般的な考え方、対応方法については**第2章**参照）。

以下に外部評価の主体とその目的の例を挙げる。

1) 疾患の専門家

診療ガイドラインの取り上げるテーマによっては、関連する疾患が多岐にわたることがある。そのような場合には、関連する多様な疾患の専門家の観点を考慮することが求められる。作成過程において多様な疾患の専門家の参加が実現できなかった場合には、外部評価の評価者として関連する疾患の専門家を入れることが有用と考えられる。

2) 医療者（学会員、関連学会員を含む）

日本では診療ガイドラインは学会などを中心として作成されることが多く、作成の過程で医療者の観点が含まれていることに違いはない。しかし、実際に診療ガイドライン作成に携わる医療者の数は限られるため、多様な臨床の実情の総体を考慮することは困難である。そこで、多様な現場で診療にあたる医療者から評価を受けておくことで、現場にとってより有用な診療ガイドラインにすることができると考えられる。また、作成学会以外の、関連する領域の学会からの評価を受けておくことも有用であろう。

3) 作成方法の専門家

診療ガイドラインの作成・記載の方法が適切であるかについての評価を受けることも重要である。診療ガイドラインに詳しい専門家が作成・記載の観点からの評価者になり得ると考えられる。

4) 患者・市民

診療ガイドラインの本来の目的は患者と医療者との意思決定を支援することにある。そのため、患者やその家族、支援者や市民などの利用者としての観点を考慮して診療ガイドラインを作成することが求められる。作成過程に患者・市民メンバーを含んでいた場合でも、公開前に患者・市民から評価を受けることは多様な患者・市民の観点を加味することにつながるため、有用である。

（4）外部評価の方法・ツール

外部評価を実施する際には、草案に対して、外部評価委員や一般から自由な形式での記述的な評価を受けることが多い。記述的な方法で外部評価を実施することで、標準化されない多様な観点から意見を得ることができる。

一方、作成・記載方法の観点から評価を行うことを依頼する場合には、これまで開発されてき

た標準化された評価形式の利用を考慮する。

例えば、診療ガイドラインの作成方法や記載内容の評価としては、AGREE II (AGREE Next Steps Consortium 2009), AGREE Reporting Checklist 2016 (Brouwers MC 2016), Institute of Medicine (IOM) の基準 (IOM 2011a), RIGHT (Chen Y 2017), 記載項目の評価としてはCOGS (Shiffman RN 2003), 導入可能性の評価としてはGLIA 2.0 (Kashyap N 2011) が利用可能である。

また、診療ガイドラインを基礎付けるシステムティックレビューの項目や方法についての評価には、AMSTAR 2 (Shea BJ 2017), PRISMA (Liberati A 2009), またはIOMの基準 (IOM 2011b) が利用可能である。

診療ガイドラインの作成過程において患者・市民の視点を反映させるべきであることはいうまでもない。同様に、外部評価の段階でも患者参加は必要とされている (G-I-N PUBLIC 2015)。患者や市民が、患者・市民の観点から診療ガイドラインを評価する標準的な項目は開発されていない。ただ、標準的な評価項目とはなっていないものの、NICEが提示している、患者・支援者組織がスコープ、および、診療ガイドラインの草案に対してコメントする複数の観点 (NICE 2013) は参考になると考えられる。

ここでは診療ガイドラインを公開する前の外部評価の3つの目的・実施段階を挙げたが、それ以外の外部評価を否定するものではない。外部評価の目的の設定に応じて、必要な外部評価を実施することが望ましい。

診療ガイドライン作成グループは、システムティックレビューチームが作成したサマリーレポートの定性的システムティックレビュー、実施した場合はメタアナリシス (meta-analysis, メタ解析と同義) の結果に合わせて、外部評価の検討の結果も踏まえた推奨決定の経過を、解説として詳細に記載する。

7.5.3 外部評価の検討と評価者への返答

外部評価を受けた場合、評価内容から診療ガイドラインを修正する必要性について検討し、内容によっては評価者に返答する。そして、この一連の過程をまとめて公表することが望ましい。

外部評価を実施し、診療ガイドラインの妥当性を検討した場合には、作成過程の透明性を確保するために、診療ガイドラインの他の作成過程と同様に、外部評価の過程を公表することが重要である。また、外部評価の過程を公表するだけではなく、診療ガイドライン作成グループが外部評価に対して返答する場合には、必要な匿名化を行った上で、その返答の内容についても公表することが望ましい。

ただし、評価内容全てをまとめることは、外部評価の実情からすると現実的ではない。そのため、得られた外部評価のうち、誤字脱字の指摘、修正の必要性を指摘しないコメントといったものを除いて、診療ガイドライン作成グループが診療ガイドラインの内容に対して重大な評価内容であると判断するものに限ってまとめてもよい。NICEでは、診療ガイドライン作成グループが外部評価の内容を吟味できるように、外部評価の内容を標準的なフォーマットにまとめている (NICE 2012: 170-171)。

ここでは次のような項目で標準的な形式にまとめることを提案する。その上で、公表することが望ましい。

- ・評価受付年月日
- ・評価を受けた項目・版（スコープ、CQ、システムティックレビュー、草案、他）
- ・評価者／評価組織／グループ
- ・評価内容 ※検討箇所（節、ページ、CQ、推奨、他）を含む
- ・返答内容
- ・返答年月日

評価者に関する情報については、診療ガイドライン作成グループが公表の必要があると考え、かつ、評価者がその公表を許可する場合に限り、公表するようにする。

7.5.4 評価結果と診療ガイドライン

外部評価の結果とその過程については、診療ガイドラインの詳細版の中で記載する。

外部評価の結果とその過程については、診療ガイドラインの詳細版の中で、外部評価の節を設けて明記する。評価者の氏名・所属・役職などについては、評価者の許諾が得られた場合に限り記載する（外部評価結果の記載方法については、**テンプレート【FN-6】** 参照）。

標準的な評価ツールを用いた評価を受けた場合には、その評価ツールによる評価結果を記載する。

● 手順

- (1) 評価を受ける項目・版と目的を設定する
- (2) 評価者を特定する
- (3) 評価を受ける方法を決定する
実施時期／期間、実施方法（インターネット、会議、他）、評価項目、など
- (4) 評価を受け付ける
 - (4)-1 上記の内容を実施する ※評価結果を公表することを明記する
 - (4)-2 標準的なフォーマットでまとめる
- (5) 外部評価の結果を吟味する
 - (5)-1 評価結果から診療ガイドラインの内容を検討する
 - (5)-2 修正の必要が生じた場合は、修正を行う
- (6) 評価者へ返答を行う（7.5.3節参照）
 - (6)-1 評価結果に対する診療ガイドライン作成グループの対応について返答する
 - (6)-2 変更した場合も、変更しなかった場合も、その理由を明示する
- (7) 評価結果をまとめて、診療ガイドラインの中で記載する

● テンプレート

【FN-5 外部評価返答リスト】

【FN-6 外部評価まとめ】

● テンプレート 記入方法

【FN-5 外部評価返答リスト 記入方法】

外部評価の概要について、診療ガイドライン作成グループで検討した結果をまとめる。外部評価のID、評価を受け付けた年月日、評価を受けた項目・版、評価者情報、評価内容、評価に対する返答内容、返答を確定させた年月日を記載する。

【FN-6 外部評価まとめ 記入方法】

外部評価を受けた場合、その結果をその目的ごとにまとめる。外部評価の目的、方法（項目・版、期間、評価者、ツール、他）、経過、結果を記載する。

● テンプレート 記入例

【FN-5 外部評価返答リスト 記入例】

【FN-6 外部評価まとめ 記入例】

7.6 診療ガイドライン最終案の決定

外部評価で得られた意見を吟味して、必要な修正を加えた上で、診療ガイドライン最終案を決定する。

● 手順

外部評価の終了後に診療ガイドライン作成グループ会議を開催して、修正の要否を検討し、必要な修正を加えて最終案を作成する。

7章 テンプレート

【FN-1 ガイドラインサマリー】

ガイドラインサマリー

【FN-2 重要用語の定義】

用語名	解説

[FN-3 略語一覧]

略語名	正式名称

【FN-4 作成経過】

項目	本文
作成方針	
使用上の注意	
COI	
作成資金	
組織編成	診療ガイドライン統括委員会： ----- 診療ガイドライン作成グループ： ----- システムティックレビューチーム： ----- -----
作成工程	準備： ----- スコープ： ----- システムティックレビュー： ----- 推奨作成： ----- 公開に向けた最終調整： ----- 公開： -----

[FN-5 外部評価返答リスト]

【FN-6 外部評価まとめ】

目的	
方法	項目・版： 期間： 評価者： ツール： その他：
経過	
結果	

7章 テンプレート 記入方法

【FN-1 ガイドラインサマリー 記入方法】

ガイドラインサマリー

診療ガイドラインで取り上げた CQ と推奨の一覧を提示する。

【FN-2 重要用語の定義 記入方法】

用語名	解説
	診療ガイドラインの中で使用される重要な用語の解説を提示する。専門医でなくても診療ガイドラインが理解できるように配慮することが望ましい。

[FN-3 略語一覧 記入方法]

略語名	正式名称
	診療ガイドラインの中で頻繁に使用される略語について一括して正式名称を提示する。

【FN-4 作成経過 記入方法】

項目	本文
作成方針	診療ガイドライン作成にあたって重視した全体的な方針を記載する。
使用上の注意	診療ガイドライン利用者が、利用にあたって注意すべき点を記載する。
COI	COIについて記載する。
作成資金	作成資金の出どころを記載する。
組織編成	<p>診療ガイドライン統括委員会：</p> <p>-----</p> <p>編成の方針、過程、結果について記載する。</p> <p>診療ガイドライン作成グループ：</p> <p>-----</p> <p>編成の方針、過程、結果について記載する。</p> <p>システムティックレビューチーム：</p> <p>-----</p> <p>編成の方針、過程、結果について記載する。</p> <p>(他の委員会を組織した場合には追加する)</p> <p>-----</p> <p>編成の方針、過程、結果について記載する。</p>
作成工程	<p>準備：</p> <p>-----</p> <p>委員会の開催など、実施事項、実施結果、問題点等について記載する。</p> <p>スコープ：</p> <p>-----</p> <p>委員会の開催など、実施事項、実施結果、問題点等について記載する。</p> <p>システムティックレビュー：</p> <p>-----</p> <p>委員会の開催など、実施事項、実施結果、問題点等について記載する。</p> <p>推奨作成：</p> <p>-----</p> <p>委員会の開催など、実施事項、実施結果、問題点等について記載する。</p> <p>公開に向けた最終調整：</p> <p>-----</p> <p>委員会の開催など、実施事項、実施結果、問題点等について記載する。</p> <p>公開：</p> <p>-----</p> <p>委員会の開催など、実施事項、実施結果、問題点等について記載する。</p>

【FN-5 外部評価返答リスト 記入方法】

番号	評価			年月日	返答
	年月日	項目・版	評価者		
					<p>評価した年月日を記載する。</p> <p>評価内容を記載する。 ※重要な検討事案だけをまとめてよい。</p> <p>返答内容を記載する。</p> <p>評価を受け付けた項目・版を記載する。 例) •スコープ •システムティックレビュー・レポート •草案</p>

【FN-6 外部評価まとめ 記入方法】

目的	外部評価を行う目的を記載する。
方法	<p>項目・版：</p> <p>期間：</p> <p>評価者：</p> <p>ツール：</p> <p>その他：</p>
経過	外部評価を実施した経過について記載する。 ※実施したことを日付順にまとめる。
結果	外部評価の結果を記載する。 ※外部評価の内容、評価への返答をまとめる。 ※まとめ表を抜粋して記載してもよい。 ※標準的なツールを用いて評価した場合は、その結果を記載する。

7章 テンプレート 記入例

【FN-5 外部評価返答リスト 記入例】

【FN-6 外部評価まとめ 記入例】

目的	草案段階の診療ガイドラインの作成方法・記載項目の妥当性を確認すること。
方法	<p>項目・版： 草案</p> <p>期間： 2013年12月15日～2014年1月14日</p> <p>評価者： 診療ガイドラインの専門家、臨床の専門家の6名から構成される外部評価委員会（代表：○○△△）。外部評価委員会のメンバーは、統括委員会が選定した。</p> <p>ツール： AGREE IIを用いた。</p> <p>その他： 外部評価委員会のコメントの後に1ヵ月の修正期間を見込んだ。</p>
経過	<p>作成組織の立ち上げ時に外部評価委員会を設置。</p> <p>2013年11月30日に草案完成。</p> <p>2013年12月15日～2014年1月14日までの間に外部評価委員会に外部評価を依頼。外部評価委員会はAGREE IIを用いて個別に評価を行い、2014年1月6日に評価会議を実施し、評価結果を共有し、再度個別の評価を行った。再度行った評価の結果を1月10日までに評価委員会代表に提出し、結果を集約した。</p> <p>2014年1月31日に診療ガイドライン作成グループは会議を行い、外部評価の結果を吟味した。</p> <p>2014年2月1日～2014年3月1日まで、診療ガイドライン作成グループは、外部評価の結果を踏まえた修正を行った。2014年3月5日、診療ガイドライン作成グループは会議を行い、外部評価の結果への返答について検討した。</p>
結果	<p>全体としておおむね良好であるとの評価結果を得た。</p> <p>次の点について改善が提案された。</p> <ul style="list-style-type: none"> ①「作成の厳密さ」について、エビデンス総体を評価した枠組みの提示が不十分である。 ②「提示の明確さ」について、個々の推奨の関連が分かりにくい。 ③「適用可能性」について、コストに関する検討・記載が不十分である。 ④「編集の独立性」について、個人のCOIに対する記載がない。 <p>以上の提案を受けて、次のように修正した。</p> <ul style="list-style-type: none"> ①→エビデンス総体を評価した際の評価シートを提示した。 ②→臨床重要課題に対応する推奨のフローチャートを提示した。 ③→推奨ごとにコストに関して検討を行い、記載できる内容は追記した。 ④→個人のCOIについては、学会の利益相反委員会で管理していること、そして、委員会の判断結果について追記した。

7章 文献

- AGREE Next Steps Consortium. (2009) The AGREE II Instrument [Electronic version]. <http://www.agreertrust.org> (参照 2013/02/27)
- Brouwers MC, et al. (2016) The AGREE reporting checklist: a tool to improve reporting of clinical practice guidelines. BMJ 352:i1152. doi: [10.1136/bmj.i1152](https://doi.org/10.1136/bmj.i1152). PMID: [26957104](#)
- Chen Y, et al. (2017) A reporting tool for practice guidelines in health care: the RIGHT statement. Ann Intern Med 166:128-132. doi: [10.7326/M16-1565](https://doi.org/10.7326/M16-1565). PMID: [27893062](#)
- Guidelines International Patient and Public Involvement Working Group (G-I-N PUBLIC). (2015) G-I-N PUBLIC Toolkit: Patient and Public Involvement in Guidelines. https://g-i-n.net/document-store/resources/toolkit/gin-public-toolkit-final/at_download/file (参照 2021/02/22)
- Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press
- Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for Systematic Reviews. National Academies Press
- Kashyap N, et al. (2011) GLIA: GuideLine Implementability Appraisal v. 2.0. Yale Center for Medical Informatics. http://nutmeg.med.yale.edu/glia/doc/GLIA_v2.pdf (参照 2021/02/22)
- Liberati A, et al. (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 6:e1000100. doi: [10.1371/journal.pmed.1000100](https://doi.org/10.1371/journal.pmed.1000100). PMID: [19621070](#)
- National Institute for Health and Clinical Excellence (NICE). (2012) The Guidelines Manual: Process and Methods. <https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933> (参照 2021/02/22)
- National Institute for Health and Care Excellence (NICE). (2013) Factsheet 2: How organizations representing patients and carers can get involved. <https://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Factsheet-2-contribute-to-developing-clinical-guidelines.pdf> (参照 2021/02/22)
- Shea BJ, et al. (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Medical Res Methodol 7:10. doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10). PMID: [17302989](#)
- Shea BJ, et al. (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 358:j4008. doi: [10.1136/bmj.j4008](https://doi.org/10.1136/bmj.j4008). PMID: [28935701](#)
- Shiffman RN, et al. (2003) Standardized reporting of clinical practice guidelines: a proposal from the conference on guideline standardization. Ann Intern Med 139:493-498. doi: [10.7326/0003-4819-139-6-200309160-00013](https://doi.org/10.7326/0003-4819-139-6-200309160-00013). PMID: [13679327](#)

第8章

診療ガイドライン 公開後の取り組み

Minds

8章 Key Words

- ◎ National Institute for Health and Care Excellence (NICE)
- ◎ 益と害 (benefit and harm)
- ◎ 益と害のバランス
- ◎ エビデンスの確実性
- ◎ エビデンス総体 (body of evidence)
- ◎ 外部評価
- ◎ 外部評価委員会
- ◎ 患者アウトカム
- ◎ クオリティーインディケータ (quality indicator : QI)
- ◎ クリニカルクエスチョン (clinical question : CQ)
- ◎ システマティックレビュー (systematic review : SR, 系統的レビューと同義)
- ◎ システマティックレビューチーム (systematic review team : SR チーム)
- ◎ 診療アルゴリズム
- ◎ 診療ガイドライン (clinical practice guideline : CPG)
- ◎ 診療ガイドライン作成グループ (guideline development group : GDG, 診療ガイドライン作成パネルともいう)
- ◎ 診療ガイドライン作成事務局
- ◎ 診療ガイドライン統括委員会
- ◎ 推奨
- ◎ スコープ (scope)
 - 導入 (implementation)
 - 評価 (assessment)
 - 普及 (dissemination)
- ◎ 利益相反 (conflict of interest : COI)

◎：巻末「重要用語集」の収録語

テンプレートの頭文字の略称

PR : preparation 【第2章 準備】 SC : scope 【第3章 スコープ】

SR : systematic review 【第4章 システマティックレビュー】 EC : economic 【第5章 医療経済評価】

RC : recommendation 【第6章 推奨】 FN : finalization 【第7章 公開に向けた最終調整】

PO : post 【第8章 診療ガイドライン公開後の取り組み】

8.0 概要

本マニュアルは、診療ガイドライン（clinical practice guideline : CPG）の作成方法を提案することを中心に解説してきたが、診療ガイドラインには、公開した診療ガイドラインを普及（dissemination）、導入（implementation）、そして、評価（assessment）するプロセス（図8-1）についても記載することが望ましい。



図8-1 診療ガイドラインの作成、普及、導入、評価の4段階

以下、本文で紹介する項目について、継続的に実施する活動の内容を診療ガイドラインに記載することが望ましい。

8.1 公開後の組織体制

診療ガイドラインを公開した後も、診療ガイドラインの導入促進、利用者からのフィードバックの受け付け、普及・活用・効果の評価、そして、診療ガイドラインの推奨に影響を及ぼす新たな研究を踏まえたエビデンスの検討など、診療ガイドライン統括委員会、診療ガイドライン作成グループに求められる作業が多い。したがって、診療ガイドライン統括委員会、診療ガイドライン作成グループは、診療ガイドライン公開後も、次期の診療ガイドライン統括委員会、診療ガイドライン作成グループが新たに編成されるまでは活動を継続することが望ましい。

● テンプレート

【PO-1 公開後の組織体制】

● テンプレート 記入方法

【PO-1 公開後の組織体制 記入方法】

● テンプレート 記入例

(なし)

8.2 導入

導入とは、診療ガイドラインが、活用が想定される場で適切に活用されるように、さまざまな工夫によって適用可能性（applicability）を高めることをいう。

8.2.1 要約版の作成

詳細版を基に日常診療で参照が可能な「実用版」を、さらに、日常診療の現場で迅速に活用できるクイックレファレンスとしての「簡易版」を作成することによって、活用の促進を図る。実用版、簡易版が作成された場合には、詳細版は必要に応じて適宜参照する位置付けとなる。また、患者・市民のために、診療ガイドラインを一般向けに解説する「一般向けガイドライン解説」の作成も強く望まれる。詳細版、実用版、簡易版、一般向けの掲載内容案を表8-1に示す。

表8-1 詳細版、実用版、簡易版、一般向けの掲載内容案

項目	詳細版	実用版	簡易版	一般向け
タイトルページ				
タイトル	○	○	○	○
作成主体	○	○	○	○
版	○	○	○	○
発行年月日	○	○	○	○
前付				
ガイドラインサマリー	○	○	○	○
診療アルゴリズム	○	○	○	○
用語・略語一覧	○	○	○	○
(I) 作成組織・作成方針				
1 作成組織				
1.1 作成主体	○	○		
1.2 診療ガイドライン統括委員会	○	○		
1.3 診療ガイドライン作成グループ	○	○		
1.4 システマティックレビューチーム	○	○		
1.5 外部評価委員会	○	○		
1.6 診療ガイドライン作成事務局	○	○		
2 作成経過				
2.1 作成方針	○	○		○
2.2 使用上の注意	○	○	○	○
2.3 利益相反 (COI)	○	○		○

2.4 作成資金	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
2.5 組織編成	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
2.6 作成工程	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
(II) スコープ				
1 疾患トピックの基本的特徴				
1.1 臨床的特徴	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
1.2 痘学的特徴	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
1.3 疾患トピックの全体的な流れ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
(図) 診療アルゴリズム	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2 診療ガイドラインがカバーする内容に関する事項	<input type="radio"/>	<input type="radio"/>		
3 システマティックレビューに関する事項	<input type="radio"/>	<input type="radio"/>		
4 推奨作成から公開に向けた最終調整、導入方針まで	<input type="radio"/>	<input type="radio"/>		
(III) 推奨				
1 CQ1				
1.1 CQ 本文	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.2 推奨本文・エビデンスの強さ・推奨の強さ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.3 解説	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
1.4 一般向けサマリー	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>
1.5 引用文献	<input type="radio"/>	<input type="radio"/>		
2 CQ2				
: : (同上)				
(以下、CQ の数だけ繰り返す)				
(IV) 公開後の取り組み				
1 公開後の組織体制	<input type="radio"/>	<input type="radio"/>		
2 導入	<input type="radio"/>	<input type="radio"/>		
3 普及・活用・効果の評価	<input type="radio"/>	<input type="radio"/>		
4 改訂	<input type="radio"/>	<input type="radio"/>		
(V) 付録				
1 参考資料	<input type="radio"/>			
2 作業資料				
・ CQ 設定テンプレート	<input type="radio"/>			
・ 全ての文献検索データベースごとの検索式とフローチャートと文献	<input type="radio"/>			

・ エビデンス評価シート アウトカムごとの STEP 1 シート	○			
・ エビデンス評価シート CQ ごとの STEP 2 シート	○			
・ エビデンスのまとめ	○			
・ 外部評価まとめ	○			

8.2.2 多様な情報媒体の活用

提供のメディアを有効活用することも、導入による適用可能性の促進にとって重要である。**表 8-2** に印刷版と電子版の比較を挙げる。

表 8-2 診療ガイドライン公開方法の比較

	印刷版	電子版
対応可能な文章量	学術雑誌：～10 ページ程度 書籍：～300 ページ程度 小冊子：～100 ページ程度	制限なし
診療現場での携行	学術雑誌、書籍：図書館、医局、外来などで閲覧 小冊子：ベッドサイドへの携行が可能	固定端末：図書館、医局などに設置できる。 ベッドサイドへの携行は不適 ノート PC 端末：診療現場である程度の移動は可能 携帯端末：診療現場のあらゆる場所への携行が可能
必要な情報の検索	目次、索引による検索	多様な検索機能が利用可能
幅広い利用者への対応	学術雑誌：学会員に限られる場合あり 書籍：購入が必要 小冊子：配布による	オンライン情報提供：回線接続が必要 オフラインアプリ：参照時にはオンライン接続は不要
公開に必要な経費	学術雑誌：廉価で対応可能 書籍：有償販売により経費調達が可能 小冊子：無償提供の場合は作成主体の負担となる	Minds ガイドラインライブラリ：無料 学会などのウェブサイト：ページ作成費用は作成主体の負担

8.2.3 診療ガイドライン活用の促進要因と阻害要因

診療ガイドラインの導入を考える際には、診療ガイドライン自体の問題、提供方法の問題、利用者側の問題などを考える必要がある。

・ 診療ガイドライン自体の問題

診療ガイドラインが日常診療で活用されるかどうかを左右する最も重要な要素は、診療ガイドラインが信頼できるという認知を利用者側から得られることである。本マニュアルでは、システムティックレビュー（systematic review : SR、系統的レビューと同義）によるエビデンス総体（body of evidence）の評価と統合、エビデンスとともに患者・市民の価値観・希望、好みの多様性、経済

的視点も考慮した推奨の決定など、『Minds 診療ガイドライン作成の手引き 2007』では必ずしも明確に示されていなかった作成方法の詳細を提示した。

診療ガイドラインの体裁は、当初は教科書と同様に章節項による構成が少なくなかったが、手引き 2007 で推奨したこともあり、クリニカルクエスチョン（clinical question : CQ）単位とした構成が着実に増えつつあり、意思決定支援に直結した構成となってきた。ただし、CQ の中には、「～の疫学的特徴はなにか」など、診療上の意思決定とは無関係の質問も少なくないため改善が求められる。

・提供方法の問題

利用者が診療ガイドラインを必要とする場面で、必要とする情報を確實に提供できることが、診療ガイドラインの活用を促進する上で極めて重要である。診療ガイドラインは、参照の目的、必要な情報によって、詳細版、実用版、簡易版の 3 種類を作成することが望ましいと述べた。また、詳細版、実用版、簡易版の参照の場面を考えて、印刷版、電子版の中で適切なメディアを選択することが重要である。

・利用者側の問題

患者と医療者による意思決定を支援するのが診療ガイドラインの目的であるから、患者、医療者が活用の主役である。ただし、わが国では診療ガイドラインの順守は義務ではないため、利用者である患者と医療者が、自ら診療ガイドラインの提示する推奨を吟味して、その適用の可否を自主的に判断することになる。したがって、利用者である患者と医療者の診療ガイドラインに対する認知と態度が、診療ガイドラインの活用を左右する大きな要因となる。患者、医療者の双方に診療ガイドライン活用の具体的方法と注意点を伝える努力が求められる。

● テンプレート

[PO-2 導入]

● テンプレート 記入方法

[PO-2 導入 記入方法]

● テンプレート 記入例

(なし)

8.3 普及・活用・効果の評価

診療ガイドライン公開後は、診療ガイドラインの普及、活用、そして活用による効果について評価する必要がある。これらは、診療ガイドラインの改訂に向けた重要な情報となる。

まずは、診療ガイドラインの普及、活用に関する評価がある。これには、アンケートなどの手法を用いることが考えられる。また、診療ガイドラインの活用に関しては、診療に関する情報を集積することにより、該当患者に対する実施率などの検討が可能になる。

また、診療ガイドラインの評価としては、何よりも診療ガイドラインの導入による患者アウトカムの改善を評価すべきであり、クオリティーアンディケータ（quality indicator : QI）などによる評価が考えられる。QIでは、診療ガイドラインの推奨に基づいて診療が行われる割合、代理指標（surrogates）の改善の評価などが取り上げられる。また、診療ガイドラインに対する患者と医療者の満足度の評価も重要である。

これらの評価と並行して、発行した診療ガイドラインの内容、作成方法の観点からの評価も重要である。外部組織に評価を依頼する、もしくは、一般の利用者からフィードバックを受け付けるなどを通して、診療ガイドラインが持っている課題を抽出しておくことは、改訂に向けた取り組みを計画する上で重要になると期待される。

● テンプレート

【PO-3 普及・活用・効果の評価】

● テンプレート 記入方法

【PO-3 普及・活用・効果の評価 記入方法】

● テンプレート 記入例

(なし)

8.4 改訂

8.4.1 改訂の意義

診療ガイドラインが常に患者と臨床家の意思決定に役立つ資料であるために、 診療ガイドライン作成グループは、 作成した診療ガイドラインの改訂を検討する必要がある。

診療ガイドラインは、 作成時点におけるエビデンス、 値値観、 社会環境に基づいて作成されている。しかし、 その作成時点での推奨の妥当性やエビデンスの確実性などが、 その後も変化せずに持続するわけではない。50 %の診療ガイドラインが5.8年で「時代遅れ」になるとされている (Shekelle PG 2001)。

また、 診療ガイドラインの推奨はエビデンスのシステムティックレビューに基づいて作成されるが、 システムティックレビューの約半数は5.5年でその結論の変更が必要となるとされている (Shojania KG 2007)。さらに、 推奨を作成する際には、 エビデンスのシステムティックレビューだけではなく、 患者・市民の価値観・希望、 益と害 (benefit and harm) のバランス、 コストについて検討するが、 これらの要素は、 医療を取り巻く社会情勢の推移によってその内容が変化し得る。このように、 診療ガイドラインを基礎づけている項目には「寿命」がある。

そこで、 診療ガイドライン作成グループは、 診療ガイドラインを常に患者と臨床家の意思決定に役立つ資料とするために、 作成した診療ガイドラインの改訂を検討する必要がある。

しかし、 どのような期間、 どのような方法で改訂を行うかについては議論の余地があり、 厳格な基準がないのが現状である (IOM 2011: 134, Alonso-Coello P 2011)。以下は、 改訂の方法についての提案である。

8.4.2 改訂のタイミング

改訂は3～5年ごとを目安とする。

しかし、 診療ガイドラインをより厳密に最新の状態にしておくために、 期限を設定した改訂にあわせて、 研究報告の定期的な検索、 社会環境の変化の監視、 先行版の発行後の外部評価などに基づいた改訂の必要性を検討していくことが望ましい。

前述のとおり、 診療ガイドライン、 システムティックレビューの寿命は5年程度と考えられている。したがって、 診療ガイドラインを改訂する期間は、 5年以内が目安となるといえるだろう。

しかし、 論文としての結果の報告数は年々増加しているため、 現在では5年より早く改訂することが望ましいと考えられる。システムティックレビューに関して、 Cochrane 共同計画では2年ごとの改訂を推奨している (Higgins JPT 2011)。

診療ガイドラインのハンドブックでは、 3年ごと (NICE 2012: 189), 2～5年 (WHO 2012: 52) など、 さまざまな期間が示されているが、 ここでは3～5年を目安として提案する。

しかし、 実際にはテーマや領域などによって適切な改訂の期間は異なると考えられる。Shojania KG

ら（2007）は、循環器領域とそれ以外で、システムティックレビューが有効である期間が異なることを明らかにしている。また、社会環境、価値観も3～5年ごとに変化するわけではない。このようなことから、改訂について、先のように一律の改訂期間を設定することは現実的ではないといえるかもしれない。

そこで、診療ガイドラインをより厳密に最新の状態にしておくために、期限を設定した改訂にあわせて、研究報告の定期的な検索、社会環境の変化の監視、先行版の発行後の外部評価などに基づき改訂の必要性を検討していくことが望ましい。

8.4.3 改訂の種類

診療ガイドラインの改訂を検討する場合には、改訂するかどうかだけでなく、どの程度の改訂が必要になるかを決定する必要がある。

診療ガイドラインの改訂の種類としては、全面改訂、部分改訂、追加、などがある。

改訂について検討する際には、改訂が必要であると判断される場合と、改訂は必要ないと判断される場合がある。また、特定の推奨のみ改訂が必要となる場合もあるだろうし、さらには、当初は検討されていなかったトピックについての推奨を検討することが必要になる場合も想定される。部分的な修正や加筆修正が必要となる場合、もしくは、あらかじめ定めた改訂の時期に達したにもかかわらず改訂の必要がないと判断される場合に、一律に診療ガイドライン全体を再度同様に作成することは、資源的にも時間的にも負担が大きい。そのため、診療ガイドラインの改訂を検討する場合には、改訂するかどうかだけでなく、どの程度の改訂が必要になるかを決定する必要がある。

診療ガイドラインの改訂の種類としては次のようなものが考えられる。

① 全面改訂

診療ガイドラインで取り上げるトピックの位置付けが大きく変化し、重要臨床課題、CQから再度検討する必要が生じた場合には、診療ガイドライン全体の改訂である「全面改訂」を行う。その場合は、スコープ（scope）作成から外部評価までの診療ガイドライン作成の全過程を再度実施する。

② 部分改訂

診療ガイドラインで取り上げるトピックの位置付けが大きく変化してはいないが、新しく考慮すべき報告が発表された、新たなケアの選択肢が制度的に承認されたなど、重要臨床課題の位置付けを再度検討する必要が生じた場合には、診療ガイドラインを構成するシステムティックレビュー、推奨を部分的に再検討する「部分改訂」を行う。その場合は、重要臨床課題の位置付けを明確にし、関連する部分（例えば、CQ、システムティックレビュー、推奨、など）の作成過程を再度実施する。

③ 追加

新たな検査・治療法が登場したなど、先行する版では取り上げられていないかった重要臨床課題が生じた場合、その課題に限って対応する追加的なシステムティックレビュー、推奨を作成する「追加」を行う。その場合、新しく設定した重要臨床課題の診療ガイドラインの中における位置付けを検討し、全面改訂、部分改訂の必要性についても考慮する。

④ 継続

あらかじめ設定した期間に達し、改訂について検討を行う際に、文献などのレビューはしたが、改訂する必要がないと判断した場合には、現在の診療ガイドラインの内容をそのまま維持する「継続」という選択肢を考慮する。この場合、そのような検討・判断を行った日時とその事由について診療ガイドライン中に追記する。

⑤ 取り下げ

診療ガイドラインが対象とするトピックより包括的なトピックの診療ガイドラインが作成された、または、トピック自体が社会的な意義を失ったなど、作成された診療ガイドラインを公開しておく意義がなくなったと判断する場合は、利用者の混乱を避けるためにも、当該診療ガイドラインの流通・掲載を停止する「取り下げ」を行う。この場合、取り下げた日時とその事由を関連する学会・研究会などのホームページで広く伝達するように努める。これらの改訂のほかに、誤字・脱字を正したり、文言の意図を正確に伝えるように書き換えたりする「修正」なども考慮すべきである。

8.4.4 改訂の検討方法

継続的な改訂の検討を行う際には、継続的なエビデンスのモニタリング、外部評価の検討、診療ガイドライン作成グループ内での情報交換などを活用する。

診療ガイドライン作成グループは診療ガイドライン発行前に、診療ガイドライン発行後の組織体制、改訂の方針について検討し、診療ガイドラインの中に記載する。診療ガイドライン発行後には継続的に診療ガイドラインの改訂の検討を行う。

継続的なエビデンスのモニタリングとしては、診療ガイドライン作成時に作成した検索式を定期的に検索するとよい。また、データベースの全てをモニタリングするのはコストがかかるため、主要5誌 (*New England Journal of Medicine, Lancet, Annals of Internal Medicine, BMJ, JAMA*)、および、トピックに関連した主要誌のみに限定した検索を行うという方法も報告されている (Shekelle PG 2001)。

継続的に改訂を検討する場合には、Shekelle PG ら (2001) が挙げた、改訂が必要となる 6 つの状況が参考になるだろう。

- ① 介入の益と害に関するエビデンスの変化
- ② 重要なアウトカムの変化
- ③ 可能な介入の変化
- ④ 現在行われている診療が最適であることを示すエビデンスの変化
- ⑤ アウトカムに対する価値観の変化
- ⑥ ヘルスケアに対する利用可能な資源の変化

改訂の必要性についての検討には **(参考資料)** 「改訂検討シート」を利用してもよい。

診療ガイドライン作成グループは、継続的な検討に基づき、診療ガイドラインの部分的改訂、追加、継続、取り下げを実施してもよい。診療ガイドライン作成グループは、全面改訂の必要がある場合には診療ガイドライン統括委員会に改訂の必要性について報告し、診療ガイドライン統括委員会は改訂の方針について検討する。

(参考資料) 改訂検討シート

重要臨床課題	
検討項目	<p>次の項目に変化があるか？</p> <p><input type="checkbox"/> 介入の益と害に関するエビデンス <input type="checkbox"/> 重要なアウトカム <input type="checkbox"/> 可能な介入の変化 <input type="checkbox"/> 現在行われている診療が最適であることを示すエビデンス <input type="checkbox"/> アウトカムに対する価値観 <input type="checkbox"/> ヘルスケアに対する利用可能な資源</p> <p>※上の項目にチェックをつけた場合には、根拠を以下に記載。</p>
改訂の要否	<p><input type="checkbox"/> 改訂の必要あり <input type="checkbox"/> 改訂の必要なし</p> <p>※根拠を以下に記載。</p>

● テンプレート

[PO-4 改訂]

● テンプレート 記入方法

[PO-4 改訂 記入方法]

● テンプレート 記入例

[PO-4 改訂 記入例]

8.5 公開後の取り組みの決定プロセス

公開後に取り組むべき事項は、診療ガイドライン作成グループが協議の上決定する。診療ガイドライン作成グループによって決定された方針は、診療ガイドラインの草案に記載し、外部評価を経て、診療ガイドライン統括委員会による承認を得る。

8章 テンプレート

【PO-1 公開後の組織体制】

組織名称	公開後の対応
診療ガイドライン 統括委員会	
診療ガイドライン 作成グループ	
システムティック レビューチーム	

【PO-2 導入】

要約版の作成
多様な情報媒体の活用
診療ガイドライン活用の促進要因と阻害要因

【PO-3 普及・活用・効果の評価】

評価方法	具体的方針

【PO-4 改訂】

項目	方針
実施時期	
実施方法	
実施体制	

8章 テンプレート 記入方法

【PO-1 公開後の組織体制 記入方法】

組織名称	公開後の対応
診療ガイドライン 統括委員会	継続、再編成、解散などの対応を記載する。
診療ガイドライン 作成グループ	継続、再編成、解散などの対応を記載する。
システムティック レビューチーム	継続、再編成、解散などの対応を記載する。

【PO-2 導入 記入方法】

要約版の作成
詳細版を基に、実用版、簡易版、一般向け解説などを作成する場合には、計画を記載する。
多様な情報媒体の活用
印刷版として学術雑誌、書籍、冊子など、電子版としてインターネットによる情報提供、携帯端末用の情報提供など、情報提供方法の計画を記載する。
診療ガイドライン活用の促進要因と阻害要因
診療ガイドライン自体、利用者への対応、提供方法などについて、診療ガイドライン活用の促進要因と阻害要因の分析結果を記載する。

【PO-3 普及・活用・効果の評価 記入方法】

評価方法	具体的方針
方法の名称を記載する。	方法の概要、対象となる医療機関、実施期間などについて具体的に記載する。

【PO-4 改訂 記入方法】

項目	方針
実施時期	改訂版を発行する時期、もしくは、改訂作業を行う時期・タイミングを記載する。
実施方法	改訂について検討する方法を記載する。 改訂するかどうかを判定する基準（例：エビデンスの変化、制度の変化、他）など
実施体制	改訂についての検討を行う組織体制を記載する。

8章 テンプレート 記入例

【PO-4 改訂 記入例】

項目	方針
実施時期	5年後に改訂版を発行することを目指し、3年後に改訂検討会議を行う。
実施方法	エビデンス、制度の変更にあわせて、全面改訂ないし部分改訂を行う。緊急の必要がある場合には追加を行い、学会ホームページで公表する。
実施体制	診療ガイドライン作成グループの一部を診療ガイドライン検討グループとして、常設の改訂検討グループとする。

8章 文献

- Alonso-Coello P, et al. (2011) The updating of clinical practice guidelines: insights from an international survey. *Implement Sci.* 6:107. doi: [10.1186/1748-5908-6-107](https://doi.org/10.1186/1748-5908-6-107). PMID: [21914177](#)
- Higgins JPT, et al. (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. <http://handbook-5-1.cochrane.org/> (参照 2021/02/22)
- Institute of Medicine (IOM). (2011) Clinical Practice Guidelines We Can Trust. National Academies Press
- National Institute for Health and Clinical Excellence (NICE). (2012) The Guidelines Manual: Process and Methods. <https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933> (参照 2021/02/22)
- Shekelle PG, et al. (2001) Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA* 286:1461-1467. doi: [10.1001/jama.286.12.1461](https://doi.org/10.1001/jama.286.12.1461). PMID: [11572738](#)
- Shojania KG, et al. (2007) How quickly do systematic reviews go out of date? a survival analysis. *Ann Intern Med* 147:224-233. doi: [10.7326/0003-4819-147-4-200708210-00179](https://doi.org/10.7326/0003-4819-147-4-200708210-00179). PMID: [17638714](#)
- World Health Organization (WHO). (2012) WHO Handbook for Guideline Development. World Health Organization.

重要用語集

Minds

英文

AGREE II (The Appraisal of Guidelines for Research & Evaluation II)

診療ガイドラインの作成方法および作成過程に焦点を当てて、 診療ガイドラインの質を評価するためのツール。AGREE の改訂版で、 2009 年に開発された。

AGREE Reporting Checklist

診療ガイドラインの作成方法や記載内容を評価するためのツール。AGREE II に準拠しており、 6 つのドメイン、 23 項目からなる。

AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews 2)

システムティックレビューの方法論的な質を評価するために 2017 年に作成された測定ツール。AMSTAR (A Measurement Tool to Assess Systematic Reviews, 2007 年開発) の改訂版。

CINAHL (Cumulative Index to Nursing and Allied Health Literature)

看護学領域の文献データベース。EBSCO 社が提供している。

COGS (The Conference on Guideline Standardization Statement)

診療ガイドラインの記載項目の評価ツール。米国医療研究・品質庁 (Agency for Healthcare Research and Quality : AHRQ) の資金提供を受け、 2002 年に開催された会議でまとめられた。

COI →利益相反

Consensus Development Conference

推奨を作成する方法のひとつ。検討すべき事項について、パネル全体が参加する会議で互いに許容可能なコンセンサスを作る義務が負わされる。文献レビュー、プレゼンテーション、聴衆による討議を経て、報告書に対する参加者の合意をもって終了する。

CPG, CPGs →診療ガイドライン

CQ →クリニカルクエスチョン

Delphi (デルファイ) 法

推奨を決定する方法のひとつ。回答者グループに対し、あるテーマについて投票を実施し、その結果をフィードバックすることを数回繰り返すことで、回答者グループの意見を集約・収束させる技法。一部の意見に引きずられないようするため無記名で行われる。途中でパネル討議を行わないものを Delphi 法、討議を行う機会を設けるものを修正 Delphi 法という。

EBM (evidence-based medicine)

エビデンスに基づく医療のこと。最善の科学的根拠／実証報告を基に、医療者の技能および患者の価値観や希望を考え合わせて、よりよい医療を目指す医療の実践様式のこと。

Embase

世界的な医学薬学領域の文献データベース。Elsevier 社提供。

EtD フレームワーク (evidence to decisions framework)

推奨決定の手順。EU の 5 カ年プロジェクト、DECIDE (Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence) により開発された。

GDG →診療ガイドライン作成グループ

GLIA (GuideLine Implementability Appraisal)

診療ガイドラインの導入可能性の評価ツール。2005 年に開発された。

GPS (good practice statement)

Best practice statement と同義。診療上の重要度の高い医療行為について、新たにシステムティックレビューを行わなくとも、明確な理論的根拠や大きな正味の益があると診療ガイドライン作成グループが判断した医療行為を提示するもの。

GRADE (The Grading of Recommendations Assessment, Development and Evaluation)

システムティックレビューや診療ガイドライン作成の際に、エビデンスの質評価や推奨度を明確なプロセスで決定するためのシステムとして国際的な任意団体が発表した方法。現在、WHO や英国 NICE などにおける診療ガイドライン開発で採用されているアプローチ。

GRADEpro GDT

GRADE アプローチに沿った診療ガイドライン作成をウェブ上で行うためのクラウドサービス。

GRADE グリッド (GRADE grid)

推奨を決定する方法のひとつ。検討すべき事項について、診療ガイドライン作成委員全員が参加する会議で、それぞれの考えを発表し、その結果を基に投票を行う。投票項目は「強く推奨する」、「弱く推奨する」の 2 段階と、「実施することを推奨する」か「実施しないことを推奨する」の 2 つの方向性で分けた 4 つの分類となる。

HR →ハザード比

ICER →増分費用効果比

Individual perspective (個人の立場)

医療経済評価において、個々の患者の立場からの視点を示す。診療ガイドライン作成における視点のひとつ。

⇒ Population perspective (集団の立場)

Institute of Medicine (IOM)

1970年に設立された米国科学アカデミーの一部門。2016年にNational Academy of Medicine (NAM)と改称している。米国の政府機関に対して健康や医療に関する提言を行っているNational Academy of Sciencesの医学・医療部門。

International Guideline Library

Guidelines International Network (G-I-N)が提供する診療ガイドラインデータベース。世界各国の診療ガイドラインが登録されている。

IOMのSRの基準（ガイドライン評価における）

Institute of Medicineが示した診療ガイドライン評価の基準。診療ガイドラインのシステムティックレビュー(SR)の項目や方法の評価基準。

IOMの基準（ガイドライン評価における）

Institute of Medicineが示した診療ガイドライン評価の基準。診療ガイドラインの作成方法や記載内容の評価基準。

ITT (intention-to-treat) 解析

初期治療方針に基づく分析のこと。ランダム化割り付け試験における統計解析において、脱落例やプロトコル逸脱があったとしても、これらも含めて割り付けどおりに介入群と対照群とを取り扱い解析を行うこと。ITTを行う理由は、脱落例やプロトコル逸脱を除外すれば症例減少バイアスになる可能性があること、実際の介入においてもプロトコル逸脱は発生するため実行可能性の指標になること、ランダム割り付けが保証されることなどがある。

I^2 統計値

メタアナリシスにおいて、研究間の異質性を判断する指標。0から100%の値で表され、大きいほど異質性があることを意味する。

JMEDPlus

日本で発行された医学、歯学、薬学、看護学、生物科学、獣医学などに関する文献データベースで、

国立研究開発法人科学技術振興機構（JST）が作成している。

MCDA →多基準決断分析

MEDLINE

医学を中心とする生命科学の文献情報を収集したインターネット上のデータベース。米国国立医学図書館（NLM）が開発したもの。

MeSH (Medical Subject Headings)

MEDLINEにおける文献管理・検索を目的として、米国国立医学図書館（NLM）が開発したシソーラス。生命科学の用語集。

Minds (Medical Information Network Distribution Service)

日本の診療ガイドラインセンター。日本国内で発行された診療ガイドラインを評価選定し、その書誌情報および診療ガイドライン本文をウェブサイト（Minds ガイドラインライブラリ）にて公開している。また、診療ガイドラインの作成支援、普及啓発、市民・患者参画を推進している。

National Guideline Clearinghouse (NGC)

米国の診療ガイドラインセンター。2018年に予算が打ち切られ、現在は活動を停止している。

National Institute for Health and Care Excellence (NICE)

英国のイングランドとウェールズにおける診療ガイドラインセンター。

NGT 法 (Nominal Group Technique)

推薦を決定する方法のひとつ。検討すべき事項について、パネル全体が参加する会議でそれぞれの考えを発表し、そこで出てきたことの重要性をパネルが個別に評価する。その結果を集計して、コンセンサスを形成する。

NNT (number needed to treat)

治療必要人数のこと。一人のイベント発生を抑制するのに何人の治療が必要かを示す指標。NNTが小さいほど治療の効果は大きい。絶対リスク減少（率差あるいはリスク差）の絶対値の逆数。

NMA →ネットワークメタアナリシス

non-RCT →非ランダム化比較試験

NRSI (non-randomized studies of healthcare interventions)

非ランダム化ヘルスケア介入研究のこと。

OR →オッズ比**PICO**

患者の臨床問題や疑問点を整理する枠組み。PはPatients（患者）、Problem（問題）、Population（対象者）のことであり、介入を受ける対象を示す。IはInterventions（介入）、すなわち介入の選択肢のことである。CはComparisons（比較対照）、Controls（対照）、Comparators（対照）のことであり、Iと比較したい介入を示す。OはOutcomes（アウトカム）、すなわち介入を行った結果として起こり得るアウトカム事象（転帰事象）のことである。これらに、TimingとSettingが追加されることもある（PICOTS）。

Population perspective（集団の立場）

医療経済評価において、公的医療や社会など集団を対象とする立場からの視点を示す。診療ガイドライン作成における視点のひとつ。

⇒ Individual perspective（個人の立場）

PPI →患者・市民参画**PRESS (Peer Review of Electronic Search Strategies)**

システムティックレビューの検索式査読のためのガイドライン。

PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)

システムティックレビューおよびメタアナリシスについて、質の向上を図るため、報告すべき項目を明示した声明。

PROSPERO (International Prospective Register of Systematic Reviews)

システムティックレビューを実施する際の事前登録システム。これにより重複を回避し、計画されたものと完遂されたものを比較することにより出版（報告）バイアスを減らすことが期待されている。英国NIHR（National Institute for Health Research）の資金により、ヨーク大学Centre for Reviews and Dissemination（CRD）が運営している。

PsycINFO

米国心理学会が作成している心理学領域の文献データベース。

PubMed

米国国立医学図書館 (National Library of Medicine : NLM) が運営し公開している、医学領域の文献データベース MEDLINE を中心とした生命科学、生物医学文献検索エンジン。

QI → クオリティーアインディケータ

Q 統計値

メタアナリシスにおいて、研究間の統計学的異質性を判断する目的で用いられる指標で、Cochran Q 統計量とも表記される。

Rayyan

文献検索の結果から文献選定作業を共同で実施することを実現したクラウドサービス。カタールのコンピューティング研究所 (Qatar Computing Research Institute : QCRI) が開発し、無料で利用可能。

RCT → ランダム化比較試験

RD → リスク差

RevMan (Review Manager)

Cochrane 共同計画より提供されている、システムティックレビューを効率よく実施するためのソフトウェア。メタアナリシスのツールも付属している。

RIGHT (Reporting Items for practice Guidelines in HealThcare)

診療ガイドラインの報告に関するツール。チェックリストの形で報告すべき項目が記載されている。

ROBINS-I (Risk of Bias in Non-randomized Studies for Interventions)

介入に関する観察研究のバイアスリスク評価ツール。2016 年に発表された。

RR → リスク比

SMD → 標準化平均値差

SR → システマティックレビュー

SoF 表 (summary of findings table)

GRADE ワーキンググループ (GRADE Working Group) から提案された、システムティックレビューに対応した結果のまとめ表。

The Cochrane Library

Cochrane 共同計画が提供する複数の医学文献データベース。CDSR (The Cochrane Database of Systematic Reviews) は、The Cochrane Library の中核データベース。

The Scottish Intercollegiate Guidelines Network (SIGN)

英国スコットランドにおける診療ガイドラインセンター。

WHO-INTEGRATE フレームワーク (WHO-INTEGRATE Evidence framework)

WHO の規範、価値観から生み出された、健康に関する意思決定のフレームワーク。公衆衛生的な意思決定のために策定されたもので、健康上の益と害のバランス、人権と社会文化的受容可能性、健康の公平・平等・非差別、社会的意義、財政面、経済的な観点からの検討、実行可能性と医療システムの考慮という 6 つの基準から構成される。

和文

アウトカムの重要性

クリニックエスチョン (CQ) の作成時に、複数のアウトカムの中から介入を受ける患者にとって、どの程度重要と考えられるかを評価すること。「重大」、「重要」、「重要でない」の 3 つに分類する。

一次研究

患者や健康者などを対象に実施した研究。ランダム化比較試験、コホート研究、症例対照研究、横断研究などが含まれる。

医中誌 Web

医学中央雑誌刊行会が刊行しているわが国で発表された医学文献のデータベース。

益と害 (benefit and harm)

益とは、介入によってもたらされると期待される望ましい効果のこと。害とは、介入によってもたらされる有害事象のこと。

益と害のバランス

益（介入によってもたらされると期待される望ましい効果）と、害（介入によってもたらされる有害事象）のバランスのこと。

エビデンス総体 (body of evidence)

研究論文などのエビデンスを系統的な方法で収集し、採用されたエビデンスの全体を評価し統合したもの

の。介入とアウトカムの組み合わせごとにまとめられる。

エビデンス総体の質的統合

アウトカム単位で作成したエビデンス総体を、質的に統合したもの。

エビデンス総体の総括（overall evidence）

アウトカムごとに評価されたエビデンス総体の、エビデンスの確実性（強さ）を統合したもの。

エビデンス総体の量的統合

アウトカム単位で作成したエビデンス総体に対してメタアナリシスを行い、量的に統合したもの。

エビデンスの確実性

広義のエビデンスに対する評価のこと。

エビデンスの確実性（質）

システムティックレビューにおける効果指標の確実性に対する確信のこと。エビデンスの質ともいう。

エビデンスの確実性（強さ）

推奨作成における推奨を支持する強さに対する確信のこと。エビデンスの強さともいう。

横断研究（cross-sectional study）

要因曝露と疾病／病態の有無を同時に調査することで、これらの関連を明らかにする研究方法である。要因曝露と疾病／病態発生との関連を比較的容易に明らかにできるが、因果関係を見極めるのが難しいことが欠点である。また、有病率などを明らかにする目的で行われる有病調査も含まれるが、この場合は非比較研究としての横断研究となる。診断法の研究では診断法の実施という介入が行われ、その結果と同時点での疾病／病態の有無の関係が解析される横断研究が主体となる。

オッズ比（odds ratio : OR）

オッズとは見込みのことであり、ある事象が起きる確率 p の起きない確率 $(1 - p)$ に対する比である。オッズが 1 より大きいほどその事象の起きる確率が高く、1 より小さいほど低くなる。オッズ比とは 2 つのオッズの比のことである。症例対照研究や横断研究で用いることが多い。確率 p が小さい場合、リスク比の近似として用いられる。

ガイドラインサマリー

診療ガイドラインが取り上げたクリニカルクエスチョン（CQ）と推奨をリストにまとめたもの。

介入研究 (interventional study)

研究者の介入によって要因曝露を変化させることで、疾病発生の状況が変化するかを調べる研究方法。観察研究では完全に制御できない交絡因子が存在するが、介入研究ではランダム割り付けを実施することで測定できない因子も含め、介入以外の因子のバランスをとることが可能になるため、バイアスを減らすことができる。

外部評価

診療ガイドラインの作成に携わってきた者とは別の者により行われる診療ガイドラインの評価。スコープ、クリニカルクエスチョン (CQ)、システムティックレビュー、エビデンス評価、推奨、診療ガイドライン全体が対象となる。

外部評価委員会

診療ガイドライン草案を第三者の立場で評価して、診療ガイドライン草案の改善のための助言を行う委員会のこと。

観察研究 (observational study)

要因曝露を人為的に介入せずに疾病発生などの状況を観察する研究方法。仮説の設定を目的とした記述疫学研究と、仮説の検証を目的とした分析疫学研究に分けることができる。分析疫学研究はさらに、生態学的研究、横断研究、コホート研究、症例対照研究などに分類される。

患者アウトカム

医療行為によって患者に生じる結果の全体のこと。死亡の回避などの効果（益）のみでなく、医療行為によって引き起こされる害（有害事象）も含まれる。

患者・市民参画 (patient and public involvement : PPI)

診療ガイドライン作成において、患者・市民の考えを十分に理解して、取り入れること。診療ガイドライン作成における患者・市民参画は、国際的に広く推進されており、必須のものとなりつつある。

患者・市民の価値観・希望

診療ガイドライン作成において、患者や市民の考えを取り入れる際の基準。クリニカルクエスチョン (CQ) や推奨を作成する際の観点のひとつ。

感度分析 (sensitivity analysis)

前提となる条件や仮定を変えた場合に、結果がどのように変化するのかを確認する分析。

クオリティーアインディケータ (quality indicator : QI)

医療の質を示す指標。QI の改善を促すことで医療の質の向上を図る。

クリニカルクエスチョン (clinical question : CQ)

CQのこと。診療ガイドラインで取り上げることが決まった重要臨床課題 (key clinical issues) に基づいて、疑問の構成要素を抽出したもの。患者アウトカムを左右する意思決定のポイントに設定される。

グレイリテラチャ (grey literature)

灰色文献のこと。一般的な出版が行われない文書。政府機関や学術機関、企業などが発行する文書。

傾向スコア分析

観察研究における解析方法。傾向スコアを算出して対象をマッチングすることで、選択バイアスを制御する統計解析の手法。実臨床の条件下での効果を明らかにすることが期待されている。

検出バイアス (detection bias)

比較される群間でアウトカム測定に系統的な差がある場合に生じるバイアス。

交絡因子 (confounding factor)

要因曝露や介入と疾病／病態発生の観察に影響を与え、真の関係とは異なった観察結果をもたらす第三の因子のこと。疫学では、性別や年齢は交絡因子として取り扱う。交絡因子を制御するためには、研究計画段階ではランダム化 (randomization), 限定 (restriction), マッチング (matching) を行う。統計解析段階では層化 (stratification), 数学的モデリング (mathematical modeling) を行う。交絡因子が制御されないと、介入／要因曝露とアウトカムの関係が見かけ上のものだけである可能性を否定できなくなる。交絡因子は要因曝露や介入とアウトカムの両方に関係がある因子である。

固定効果モデル (メタアナリシスにおける)

分散の逆数で重み付けしたモデル。統合の対象となった研究以外の研究は想定せずに、効果指標の平均値を求める。

コホート研究 (cohort study)

コホートとは、一定期間にわたって追跡される集団という意味である。コホート研究は、曝露群と非曝露群を設定して追跡調査を行い疾病発生や死亡を把握することで、介入／要因曝露と、疾病／病態発生（アウトカム）の関連を明らかにする研究方法である。観察の時間の流れは順行的となる。相対危険、寄与危険、累積罹患率などの疫学指標を明らかにできる。

今後の研究 (future research)

未来に向けた優先研究課題。

コンシールメント (concealment)

割り付け方法の隠蔽のこと。割り付け方法を対象者にも医療者にも知らせず、予見を可能な限り防ぐことにより選択バイアスを制御する方法。コンシールメントが保たれない場合は、意識的・無意識的に選択バイアスが発生する可能性がある。

最終調整 (finalization)

推奨の作成後、草案としてまとめた診療ガイドラインを外部評価などの手順を経て最終版として確定する一連の作業。

サンプルサイズ (sample size)

研究対象者の数のこと。臨床研究におけるサンプルサイズは、多すぎても少なすぎても問題が生じる。特にサンプルサイズが必要な数に足りない場合で統計学的に有意差が認められなかった場合に、真に差がないのか、サンプルサイズが不足するために有意差が観察されなかったのか（検出力不足）が判別できなくなる。

システムティックレビュー (systematic review : SR)

系統的レビューと同義。学術文献を系統的に検索・収集し、類似した研究を一定の基準で選択・評価した上で、明確で科学的な手法を用いてまとめる研究、またはその成果物のこと。定性的システムティックレビューでは、さまざまなバイアスを評価して、介入／要因曝露の効果への影響を推定する。また、統計学的な手法を用いて効果指標が定量的に統合されるメタアナリシスが行われる。システムティックレビューは、メタアナリシスの実施を含まないものを示す場合もある。

システムティックレビューチーム (systematic review team : SR チーム)

診療ガイドライン作成グループが作成したスコープに基づいて、システムティックレビューを担当するグループ。診療ガイドライン作成グループとは独立したチームとして、診療ガイドライン統括委員会により編成される。

システムティックレビューの評価

システムティックレビューの項目や方法の評価には、AMSTAR 2 (Shea BJ 2017), PRISMA (Liberati A 2009), IOM のシステムティックレビューの基準 (IOM 2011) が利用できる。

Shea BJ, et al. (2017) BMJ 358:j4008.

Liberati A, et al. (2009) PLoS Med 6:e1000100.

Institute of Medicine. (2011) Finding What Works in Health Care: Standards for systematic reviews. National Academies Press

システムティックレビュー論文

システムティックレビューを行った後に、学術論文としてまとめたもの。

実行バイアス (performance bias)

比較される群間で、介入・ケアの実行に系統的な差がある場合に生じるバイアス。

修正 Delphi 法

推奨を決定する方法のひとつ。検討すべき事項について、まずは専門家が個別に評価を行い（第1ラウンド）、その評価結果を資料とした会議における検討後に、再度個別に評価を行う（第2ラウンド）。第2ラウンドの結果として得られた中央値を基に、推奨についてのコンセンサスを決定する。

重要臨床課題（key clinical issues）

診療ガイドラインが取り上げる臨床上の課題のこと。クリニカルクエスチョン（CQ）の基になる。重要な臨床課題は現時点の診療プロセスにおいて複数の選択肢が存在し、これらの益と害のバランスが不明瞭な臨床上の課題を取り上げる。

出版（報告）バイアス（publication bias, reporting bias）

研究が選択的に出版（報告）されることによって発生するバイアス。益と害の効果が系統的に過小評価または過大評価される。一般的に統計学的有意差の観察されない研究、仮説に対して否定的な研究は発表されない傾向にある。

症例減少バイアス（attrition bias）

比較される群間で解析対象となる症例の減少に系統的な差がある場合に生じるバイアス。

症例集積（case series）

共通の特性を持った一連の患者についての分析。比較対照群は置かない。要因曝露と疾病発生の関係を検討することは難しい。

症例対照研究（case-control study）

すでに疾病／病態が生起している症例群とそうでない対照群を設定して、それぞれの群における過去の介入／要因曝露状況を比較することで、介入／要因曝露と疾病／病態発生の関連を明らかにする研究方法である。観察の時間の流れは逆行的である。まれな疾病や時間経過の長い因果推論に威力を発揮しやすい一方、過去のデータ測定のあいまいさやバイアスの影響が入りやすいことが欠点である。

症例報告（case reports）

1例または数例の患者についての詳細な報告。疾病の頻度やリスクについて推定することはできないが、仮説を設定する助けになるとともに、まれな疾病／病態への対処の参考になる。

診断精度（diagnostic test accuracy : DTA）研究

診断法の感度・特異度を明らかにする研究。診断法に関する研究の多くは診断能を明らかにする目的で行われている。

診療アルゴリズム

診療の流れを図示して、クリニカルクエスチョン（CQ）の位置付けを明示したもの。

診療ガイドライン（clinical practice guideline : CPG）

健康に関する重要な課題について、医療利用者と医療提供者の意思決定を支援するために、システムティックレビューによりエビデンス総体を評価し、益と害のバランスを勘案して、最適と考えられる推奨を提示する文書。

診療ガイドライン作成グループ（guideline development group : GDG）

診療ガイドライン作成パネルともいう。診療ガイドラインの作成において、スコープを作成し、クリニカルクエスチョン（CQ）を決定し、システムティックレビューの結果を受けて推奨を作成し、診療ガイドライン草案を作成するグループ。

診療ガイドライン作成事務局

診療ガイドライン作成の進行管理、診療ガイドライン作成組織のメンバー間の連絡、作成会議の日程調整、会議室の確保、文献収集などの事務作業、診療ガイドライン作成資金の管理などを担当する事務局。

診療ガイドラインデータベース

国内・海外の診療ガイドラインを検索するためのデータベース。ECRI Guidelines Trust®（米国）、NICE Evidence Search（英国）、International Guideline Library（国際組織）、東邦大学・医中誌診療ガイドライン情報データベース（日本）、Minds ガイドラインライブラリ（日本）などがある。

診療ガイドライン統括委員会

診療ガイドラインの作成主体（学会など）の下で、診療ガイドライン作成に関わる委員会の設置、予算の決定など、意思決定を担う委員会のこと。診療ガイドライン作成全体を統括する。

診療ガイドラインのためのシステムティックレビュー

診療ガイドライン作成のためのシステムティックレビューは推奨と関連したエビデンスの確実性（強さ）を明らかにすることを目的にする。通常のシステムティックレビューはエビデンスの確実性（質）を明らかにすることを目的としている点で異なる。

診療ガイドラインの適用（guideline adaptation）

新しい診療ガイドラインを作成する際に、他の組織で作成された既存の診療ガイドラインを新しい対象に合わせて利用すること。

診療ガイドラインの評価ツール

診療ガイドラインの作成方法や形式の妥当性の評価は AGREE II, IOM の基準, RIGHT, 記載項目の評価は COGS, 導入可能性の評価は GLIA が利用できる。

診療ガイドライン作成パネル → 診療ガイドライン作成グループ^o

推奨

エビデンス, 益と害, 患者・市民の価値観・希望, 費用などの評価に基づき臨床における意思決定を支援する文章のこと。診断, 治療, 予防などのための選択肢について推奨が作成される。

スコーピングサーチ

スコープを組み立てるための予備的な文献検索。どのような重要臨床課題があるかを探ること。海外で公表された診療ガイドライン, システマティックレビュー論文, ランダム化比較試験 (RCT) などが検索の対象となる。

スコープ (scope)

診療ガイドラインの作成にあたり, 診療ガイドラインがカバーする内容に関する事項, システマティックレビューに関する事項, 推奨作成から最終調整, 公開に関する事項などを明確にするための文書。診療ガイドライン作成の企画書ともいえる文書。

前景疑問 (foreground question)

現時点の診療プロセスにおいて複数の選択肢が存在し, 益と害のバランスが不明瞭な臨床上の課題など, 推奨を出すことで患者アウトカムの改善が期待できる疑問のこと。重要臨床課題およびクリニカルクエスチョン (CQ) として取り上げる。

選択バイアス (selection bias)

研究対象の選択の偏りにより生じるバイアス。

早期試験中止バイアス

中間解析が計画された研究デザインではない, または適切に計画された研究 (adaptive study design : 適応的研究デザイン) ではないにもかかわらず計画されたサンプルサイズを満たす前に効果が証明されたとして中止されたなど, 臨床試験において効果が過大評価されるバイアス。

増分費用効果比 (incremental cost-effectiveness ratio : ICER)

もしくは増分費用対効果。「費用の增加分」を「効果の增加分」で除して求められる医療経済評価の指標。費用と効果を統合した指標。

多基準決断分析（multi-criteria decision analysis, multiple criteria decision analysis : MCDA）

複数の（相反する）基準を評価し、決断へと結びつける手法のこと。診療ガイドライン作成の文脈では、益と害のバランスを定量的に評価すること。

単一群試験（single arm trial）

介入群のみで、比較対照群を置かない臨床試験。期待される効果に一定の閾値を設定して、それを超えることを証明するようにデザインされる。

二次研究

すでに発表されている論文データを対象に実施する研究。システムティックレビューやメタアナリシスが含まれる。

ネットワークメタアナリシス（network meta-analysis : NMA）

Multi-treatment meta-analysis、または mixed treatment comparisons (MTC) ともいう。3つ以上の介入・要因曝露の効果を直接的な比較だけでなく、間接的な比較も含めてペア比較や順位付けを可能にするメタアナリシス。

バイアスリスク（risk of bias）

バイアス（系統的偏り）が研究結果に入り込むリスクのこと。研究の内的妥当性を評価する際に考慮する。

背景疑問（background question）

疾患の罹患率、症状、発症経過など、疾患トピックの背景となるような情報に関する疑問のこと。重要な臨床課題およびクリニカルクエスチョン (CQ) として取り上げる必要はなく、疾患トピックの基本的特徴に記載すればよい。

ハザード比（hazard ratio : HR）

ハザードは時間 t における瞬間的な死亡確率を示す。ハザード比は2群のハザードの比のこと。リスクが何倍になるかを示している。広義の相対危険の一種。生存時間分析で用いられることが多い。

非一貫性（inconsistency）

研究報告によって治療効果の推定値が異なる（ばらつきが存在する）こと。

非直接性（indirectness）

一般的に外的妥当性（external validity）、一般化可能性（generalizability）、適用可能性（applicability）と同義。本マニュアルでは、現在考えているクリニカルクエスチョン (CQ) と評価される論文との間の臨床状況・集団・条件の相違を示す。

標準化平均値差（standardized mean difference : SMD）

標準化した平均値の差のこと。Hedge's unbiased estimator, Cohen's d , Hedge's g , Glass's Δ などが用いられる。

非ランダム化比較試験（non-randomized controlled trial : non-RCT）

非無作為化比較試験のこと。ランダム化（無作為化）を行わず、任意の群に対象者を登録して介入した研究のこと。準ランダム化比較試験ともいわれる。選択バイアスが大きくなりやすい。

ファンネルプロット（funnel plot）

システムティックレビューやメタアナリシスを実施する際に、収集した文献に出版（報告）バイアスが存在しないかを確認するために用いる。縦軸には精度（標本分散の逆数またはサンプルサイズ）、横軸には効果の大きさ（相対危険やオッズ比）をとり、個々の研究をプロットする。出版（報告）バイアスが存在しない場合は、プロットは漏斗が逆さまになった形状を示す。出版（報告）バイアスが存在する場合は、グラフ下側の片方が空白になる。

フォレストプロット（forest plot）

メタアナリシスの結果を示す図のこと。各研究の効果指標値と信頼区間は、四角と線で示す。四角（サンプルサイズ）が大きく、線（信頼区間）が短いほど信頼性が高い研究である。メタアナリシスにより得られた統合値と信頼区間はひし形で表す。ひし形の中心は統合された相対危険やオッズ比、ひし形の幅は95%信頼区間を示す。ひし形が1をまたがない場合、有意差ありと判定できることが多い。

不精確性（imprecision）

サンプルサイズやイベント数が少ないため、効果推定値の信頼区間の幅が広いこと。

分析的枠組み（analytic framework）

臨床課題を解析するための枠組み。重要臨床課題（key clinical issues）からクリニカルクエスチョン（CQ）を作成する際に、重要臨床課題とPICOの臨床的位置付けを具体的に図示するもの。

ペイシェント・クエスチョン

診療ガイドラインの作成にあたり、患者や市民から寄せられた臨床質問のこと。

メタアナリシス（meta-analysis）

メタ解析と同義。複数の研究結果について統計的方法を用いることで、効果指標の統合値と信頼区間を算出する方法。定量的システムティックレビューに用いられる。これらの研究結果を定量的に評価する際に実施する。結果はフォレストプロットとして表示する。出版（報告）バイアスの評価には、ファンネルプロットを表示して検討する。

盲検化 (blinding)

マスキング (masking) ともいう。情報バイアスの制御方法。対象者が介入群または対照群のどちらの群に割り付けられたのか、分からないように隠すこと。対象者がどちらの群に入ったのか分かってしまうと、判断、行動、心理に影響を与え、結果にも影響する可能性があるため実施する。患者に加え、医療者も盲検化の対象にする場合は二重盲検化という。アウトカム測定者も盲検化の対象である。質の高い研究では、二重盲検以上が要求される。

用量 - 反応勾配 (dose-dependent gradient)

用量反応関係 (dose-response relationship) ともいう。要因曝露が増加するほど、相対危険が大きくなること、あるいはその逆のこと。因果関係の有無を判断する際の基準のひとつで、観察研究のエビデンス評価の一項目として位置付けられている。

ランダム化 (randomization)

ランダム割り付け (random allocation) とほぼ同義。選択バイアスや交絡バイアスの制御方法。コンピューター生成乱数配列や乱数表などを用いて、全ての対象者が同じ確率で介入群または対照群になるかを決定すること。

ランダム化比較試験 (randomized controlled trial : RCT)

介入群と対照群にランダム割り付けを行い、介入の実施後、アウトカムを観察することで、介入群と対照群を比較する研究方法。介入の効果を明らかにする上では最も優れた研究方法である。一次研究においては、エビデンスの強さは最も高いと見なされているが、実行可能性は限定されており、全ての臨床的状況で実施できるわけではない。

ランダム効果モデル（メタアナリシスにおける）

研究間の分散を加算して重み付けしたモデル。母集団として統合の対象となった研究以外の研究が想定され、統合の対象となった研究は母集団からランダムに抜き出されたものとして取り扱われている。

利益相反 (conflict of interest : COI)

教育・研究に携わる専門家としての社会的責任と産学連携の活動に伴い生じる利益などが衝突・相反する状態のこと。

リスク差 (risk difference : RD)

2つのリスクの差、すなわちイベント発生率の差。絶対リスク減少、率差とも呼ばれる。

リスク比 (risk ratio : RR)

2つのリスクの比、すなわちイベント発生率の比。狭義の相対危険 (relative risk) とも呼ばれる。(1 - リスク比) は相対危険度減少率 (relative risk reduction) と呼ばれる。

文献一覧

Minds

1 章 文献

Institute of Medicine (IOM). (2011) Clinical Practice Guidelines We Can Trust. National Academies Press

2章 文献

- AGREE Next Steps Consortium. (2009) The AGREE II Instrument [Electronic version]. <http://www.agreertrust.org> (参照 2013/02/28)
- AGREE Next Steps Consortium. (2017) The AGREE II Instrument [Electronic version]. <http://www.agreertrust.org> (参照 2019/10/10)
- Guidelines International Network (G-I-N). (2015) G-I-N Public Toolkit: Patient and Public Involvement in Guidelines. <https://g-i-n.net/document-store/working-groups-documents/g-i-n-public/toolkit/toolkit-2015> (参照 2021/02/24)
- Guyatt G, et al. (2010) The vexing problem of guidelines and conflict of interest: a potential solution. Ann Intern Med 152:738-741. doi: <10.7326/0003-4819-152-11-201006010-00254>. PMID: <20479011>
- International Committee of Medical Journal Editors. (2019) Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals (Updated December 2019) <http://www.icmje.org/> (参照 2021/03/03)
- Institute of Medicine (IOM). (2009) Conflict of Interest in Medical Research, Education, and Practice. National Academies Press
- Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press
- Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for Systematic Reviews. National Academies Press
- Melissa J, et al. (2017) Framework for enhancing clinical practice guidelines through continuous patient engagement. Health Expect 20:3-10. doi: <10.1111/hex.12467>. PMID: <27115476>
- Minds 患者・市民支援部会. (2017) 「診療ガイドライン作成への患者・市民の参加」の基本的な考え方. <https://minds.jcqhc.or.jp/docs/minds/guideline/pdf/Proposal3.pdf> (参照 2021/02/24)
- National Institute for Health and Care Excellence (NICE). (2015) NICE's approach to public involvement in guidance and standards: a practical guide. <https://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/Public-involvement-programme/PIP-process-guide-apr-2015.pdf> (参照 2021/02/24)
- Schünemann HJ, et al. (2009) An official American Thoracic Society Policy statement: managing conflict of interest in professional societies. Am J Respir Crit Care Med 180:564-580. doi: <10.1164/rccm.200901-0126ST>. PMID: <19734351>
- Schünemann HJ, et al. (2015) Guidelines international network: principles for disclosure of interests and management of conflicts in guidelines. Ann Intern Med 163:548-553. doi: <10.7326/M14-1885>. PMID: <26436619>
- Scottish Intercollegiate Guidelines Network (SIGN). (2019) SIGN 100: A handbook for patient and carer representatives. <https://www.sign.ac.uk/assets/sign100.pdf> (参照 2021/02/24)

浅見万里子 他. (2007) 診療ガイドライン作成過程への患者・支援者参画のためのガイドライン “Patient Involvement Guidelines” (略称：PIGL) . <http://www.kanjyakai.net/manage/wp-content/themes/kanjakai/docs/guideline01.pdf> (参照 2021/02/24)

厚生労働省. (2018) 厚生労働科学研究における利益相反の管理に関する指針 . <https://www.mhlw.go.jp/stf/seisaku/jouhou-10600000-Daijinkanboukouseikagakuka/0000152586.pdf> (参照 2019/10/10)

全国医学部長病院長会議. (2018) 医学系研究機関における組織 COI 管理ガイドンス . https://www.ajmc.jp/pdf/20190425_01.pdf (参照 2021/03/03)

日本医学会利益相反委員会. (2017) 診療ガイドライン策定参加資格基準ガイドンス . http://jams.med.or.jp/guideline/clinical_guidance.pdf (参照 2021/03/03)

日本医学会利益相反委員会. (2020) COI 管理ガイドンス . http://jams.med.or.jp/guideline/coi_guidelines_2020.pdf (参照 2021/03/03)

3章 文献

- Agency for Healthcare Research and Quality (AHRQ). (2014) Methods Guide for Effectiveness and Comparative Effectiveness Reviews. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/cer-methods-guide_overview.pdf (参照 2021/01/22)
- Alonso-Coello P, et al. (2016a) GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 353:i2016. doi: [10.1136/bmj.i2016](https://doi.org/10.1136/bmj.i2016). PMID: [27353417](#)
- Alonso-Coello P, et al. (2016b) GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ 353:i2089. doi: [10.1136/bmj.i2089](https://doi.org/10.1136/bmj.i2089). PMID: [27365494](#)
- Dalkey N, et al. (1969) The Delphi Method, III: Use of Self Ratings to Improve Group Estimates. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2006/RM6115.pdf (参照 2021/01/22)
- GRADE Working Group. (2015) Evidence to Decision (EtD) Frameworks Guidance. https://s3.amazonaws.com/ietd_pdf/EtD+guidance+updated+2015+05+19.pdf (参照 2021/01/22)
- Guyatt GH, et al. (2015) Guideline panels should not GRADE good practice statements. J Clin Epidemiol 68:597-600. doi: [10.1016/j.jclinepi.2014.12.011](https://doi.org/10.1016/j.jclinepi.2014.12.011). PMID: [25660962](#)
- Guyatt GH, et al. (2016) Guideline panels should seldom make good practice statements: guidance from the GRADE working group. J Clin Epidemiol 80:3-7. doi: [10.1016/j.jclinepi.2016.07.006](https://doi.org/10.1016/j.jclinepi.2016.07.006). PMID: [27452192](#)
- Harris RP, et al. (2001) Current methods of the US preventive services task force: a review of the process. Am J Prev Med 20:21-35. doi: [10.1016/s0749-3797\(01\)00261-6](https://doi.org/10.1016/s0749-3797(01)00261-6). PMID: [11306229](#)
- National Institute for Health and Care Excellence (NICE). (2014) Developing NICE guidelines: the manual. <https://www.nice.org.uk/process/pmg20/chapter/developing-review-questions-and-planning-the-evidence-review> (参照 2021/01/22)
- Schünemann H, et al. (2013) 3. Selecting and rating the importance of outcomes. GRADE Handbook. <https://gdt.gradepro.org/app/handbook/handbook.html> (参照 2021/01/22)
- Straus SE, et al. (2010) Evidence-Based Medicine: How to Practice and Teach It. Fourth ed. Churchill Livingstone
- World Health Organization (WHO). (2014) 14. Strong recommendations when the evidence is low quality. WHO handbook for guideline development. Second ed. World Health Organization.
- Williamson PR, et al. (2012) Developing core outcome sets for clinical trials: issues to consider. Trials 13:132. doi: [10.1186/1745-6215-13-132](https://doi.org/10.1186/1745-6215-13-132). PMID: [22867278](#)

4章 文献

- Aberegg SK, et al. (2010) Delta inflation: a bias in the design of randomized controlled trials in critical care medicine. Crit Care 14:R77. doi: [10.1186/cc8990](https://doi.org/10.1186/cc8990). PMID: [20429873](#)
- Agency for Healthcare Research and Quality (AHRQ). (2012) Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(12)-EHC063-EF. <https://effectivehealthcare.ahrq.gov/> (参照 2021/02/24)
- Als-Nielsen B, et al. (2003) Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 290:921-928. doi: [10.1001/jama.290.7.921](https://doi.org/10.1001/jama.290.7.921). PMID: [12928469](#)
- Arber M, et al. (2017) Which information sources should be used to identify studies for systematic reviews of economic evaluations in healthcare? Value in Health 20:PA738.
- Armijo-Olivo S, et al. (2015) What is the influence of randomisation sequence generation and allocation concealment on treatment effects of physical therapy trials? A meta-epidemiological study. BMJ Open 5:e008562. doi: [10.1136/bmjopen-2015-008562](https://doi.org/10.1136/bmjopen-2015-008562). PMID: [26338841](#)
- Armijo-Olivo S, et al. (2017) Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. Am J Phys Med Rehabil 96:34-44. doi: [10.1097/PHM.0000000000000521](https://doi.org/10.1097/PHM.0000000000000521). PMID: [27149591](#)
- Atkins D, et al. (2011) Assessing the Applicability of Studies When Comparing Medical Interventions. Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews. [AHRQ Publication No. 11-EHC019-EF. https://effectivehealthcare.ahrq.gov/](https://effectivehealthcare.ahrq.gov/) (参照 2021/02/24)
- Bafeta A, et al. (2012) Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. BMJ 344:e813. doi: [10.1136/bmj.e813](https://doi.org/10.1136/bmj.e813). PMID: [22334559](#)
- Balshem H, et al. (2011) GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 64:401-406. doi: [10.1016/j.jclinepi.2010.07.015](https://doi.org/10.1016/j.jclinepi.2010.07.015). PMID: [21208779](#)
- Bassler D, et al. (2013) Reflections on meta-analyses involving trials stopped early for benefit: is there a problem and if so, what is it? Stat Methods Med Res 22:159-168. doi: [10.1177/0962280211432211](https://doi.org/10.1177/0962280211432211). Epub 2011 Dec 13. PMID: [22170891](#)
- Berkman ND, et al. (2014) The Empirical Evidence of Bias in Trials Measuring Treatment Differences. Agency for Healthcare Research and Quality. PMID: [25392898](#)
- Bialy L, et al. (2014) A meta-epidemiological study to examine the association between bias and treatment effects in neonatal trials. Evid Based Child Health 9:1052-1059. doi: [10.1002/ebch.1985](https://doi.org/10.1002/ebch.1985). PMID: [25504975](#)
- Brignardello-Petersen R, et al. (2019) GRADE approach to rate the certainty from a network meta-analysis: addressing incoherence. J Clin Epidemiol 108:77-85. doi: [10.1016/j.jclinepi.2018.11.025](https://doi.org/10.1016/j.jclinepi.2018.11.025). Epub 2018 Dec 5. PMID: [30529648](#)
- Britton A, et al. (1999) Threats to applicability of randomised trials: exclusions and selective participation.

- J Health Serv Res Policy 4:112-121. doi: [10.1177/135581969900400210](https://doi.org/10.1177/135581969900400210). PMID: [10387403](#)
- Cancer Research and Biostatistics (CRAB). Statistical Tools. <https://stattools.crab.org/> (参照 2021/02/24)
- Carrasco-Labra A, et al. (2015) Comparison between the standard and a new alternative format of the Summary-of-Findings tables in Cochrane review users: study protocol for a randomized controlled trial. Trials 16:164. doi: [10.1186/s13063-015-0649-6](https://doi.org/10.1186/s13063-015-0649-6). PMID: [25873338](#)
- Carrasco-Labra A, et al. (2016) Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. J Clin Epidemiol 74:7-18. doi: [10.1016/j.jclinepi.2015.12.007](https://doi.org/10.1016/j.jclinepi.2015.12.007). Epub 2016 Jan 11. PMID: [26791430](#)
- Centre for Clinical Research and Biostatistics (CCRB). (2017) Sample Size Calculation. <http://www2.ccrb.cuhk.edu.hk/web/> (参照 2021/02/24)
- Chalmers TC, et al. (1983) Bias in treatment assignment in controlled clinical trials. N Engl J Med 309:1358-1361. doi: [10.1056/NEJM198312013092204](https://doi.org/10.1056/NEJM198312013092204). PMID: [6633598](#)
- Chan AW, et al. (2004a) Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 291:2457-2465. doi: [10.1001/jama.291.20.2457](https://doi.org/10.1001/jama.291.20.2457). PMID: [15161896](#)
- Chan AW, et al. (2004b) Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 171:735-740. doi: [10.1503/cmaj.1041086](https://doi.org/10.1503/cmaj.1041086). PMID: [15451835](#)
- Chang SM, et al. (2012) Methods Guide for Medical Test Reviews [Internet]. Agency for Healthcare Research and Quality (US) PMID: 22834019. <https://www.ncbi.nlm.nih.gov/books/NBK98241/> (参照 2021/02/24)
- Chess LE, et al. (2013) Risk of bias of randomized controlled trials published in orthopaedic journals. BMC Med Res Methodol 13:76. doi: [10.1186/1471-2288-13-76](https://doi.org/10.1186/1471-2288-13-76). PMID: [23758875](#)
- Clark T, et al. (2013) Sample size determinations in original research protocols for randomised clinical trials submitted to UK research ethics committees: review. BMJ 346:f1135. doi: [10.1136/bmj.f1135](https://doi.org/10.1136/bmj.f1135). PMID: [23518273](#)
- Dechartres A, et al. (2011) Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. J Clin Epidemiol 64:136-144. doi: [10.1016/j.jclinepi.2010.04.015](https://doi.org/10.1016/j.jclinepi.2010.04.015). PMID: [20705426](#)
- Deeks JJ, et al. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration. <http://methods.cochrane.org/sdt/handbook-dta-reviews> (参照 2021/02/24)
- Dias S, et al. TSDs completed or currently in progress. <http://nicedsu.org.uk/technical-support-documents/technical-support-documents/> (参照 2021/02/24)
- Dias S, et al. (2010) Checking consistency in mixed treatment comparison meta-analysis. Stat Med 29:932-944. doi: [10.1002/sim.3767](https://doi.org/10.1002/sim.3767). PMID: [20213715](#)
- Dias S, et al. (2011) NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. <http://www.nicedsu.org.uk> (参照 2021/02/24)

- Dias S, et al. (2018) Network Meta-Analysis for Decision Making. John Wiley & Sons Ltd.
- Driessen E, et al. (2015) Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of US national institutes of health-funded trials. PLoS One 10:e0137864. doi: [10.1371/journal.pone.0137864](https://doi.org/10.1371/journal.pone.0137864). PMID: [26422604](#)
- Dwan K, et al. (2010) Assessing the potential for outcome reporting bias in a review: a tutorial. Trials 11:52. doi: [10.1186/1745-6215-11-52](https://doi.org/10.1186/1745-6215-11-52). PMID: [20462436](#)
- Fenwick J, et al. (2008) The effect of bias on the magnitude of clinical outcomes in periodontology: a pilot study. J Clin Periodontol 35:775-782. doi: [10.1111/j.1600-051X.2008.01291.x](https://doi.org/10.1111/j.1600-051X.2008.01291.x). PMID: [18840153](#)
- Gagnier JJ, et al. (2012) Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. BMC Med Res Methodol 12:111. doi: [10.1186/1471-2288-12-111](https://doi.org/10.1186/1471-2288-12-111). PMID: [22846171](#)
- Gordon M, et al. (2020) forestplot: Advanced Forest Plot Using ‘grid’ Graphics. <https://cran.r-project.org/web/packages/forestplot/index.html> (参照 2021/02/24)
- GRADEpro GDT. (2021) <https://gradepro.org/> (参照 2021/02/24)
- Gurusamy KS, et al. (2009) Assessment of risk of bias in randomized clinical trials in surgery. Br J Surg 96:342-349. doi: [10.1002/bjs.6558](https://doi.org/10.1002/bjs.6558). PMID: [19283747](#)
- Guyatt G, et al. (2011a) GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 64:383-394. doi: [10.1016/j.jclinepi.2010.04.026](https://doi.org/10.1016/j.jclinepi.2010.04.026). PMID: [21195583](#)
- Guyatt GH, et al. (2011b) GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 64:395-400. doi: [10.1016/j.jclinepi.2010.09.012](https://doi.org/10.1016/j.jclinepi.2010.09.012). PMID: [21194891](#)
- Guyatt GH, et al. (2011c) GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 64:407-415. doi: [10.1016/j.jclinepi.2010.07.017](https://doi.org/10.1016/j.jclinepi.2010.07.017). PMID: [21247734](#)
- Guyatt GH, et al. (2011d) GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol 64:1277-1282. doi: [10.1016/j.jclinepi.2011.01.011](https://doi.org/10.1016/j.jclinepi.2011.01.011). PMID: [21802904](#)
- Guyatt GH, et al. (2011e) GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 64:1283-1293. doi: [10.1016/j.jclinepi.2011.01.012](https://doi.org/10.1016/j.jclinepi.2011.01.012). PMID: [21839614](#)
- Guyatt GH, et al. (2011f) GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol 64:1294-1302. doi: [10.1016/j.jclinepi.2011.03.017](https://doi.org/10.1016/j.jclinepi.2011.03.017). PMID: [21803546](#)
- Guyatt GH, et al. (2011g) GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol 64:1303-1310. doi: [10.1016/j.jclinepi.2011.04.014](https://doi.org/10.1016/j.jclinepi.2011.04.014). PMID: [21802903](#)
- Guyatt GH, et al. (2011h) GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 64:1311-1316. doi: [10.1016/j.jclinepi.2011.06.004](https://doi.org/10.1016/j.jclinepi.2011.06.004). PMID: [21802902](#)
- Guyatt GH, et al. (2012) GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. J Clin Epidemiol 66:158-172. doi: [10.1016/j.jclinepi.2012.01.012](https://doi.org/10.1016/j.jclinepi.2012.01.012). PMID: [22609141](#)
- Guyatt G, et al. (2013a) GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol 66:151-157. doi: [10.1016/j.jclinepi.2012.01.006](https://doi.org/10.1016/j.jclinepi.2012.01.006). PMID: [22542023](#)
- Guyatt GH, et al. (2013b) GRADE guidelines: 13. Preparing summary of findings tables and evidence

- profiles-continuous outcomes. J Clin Epidemiol 66:173-183. doi: [10.1016/j.jclinepi.2012.08.001](https://doi.org/10.1016/j.jclinepi.2012.08.001). PMID: [23116689](#)
- Hahn S, et al. (2000) Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. Stat Med 19:3325-3336. doi: [10.1002/1097-0258\(20001230\)19:24<3325::aid-sim827>3.0.co;2-d](https://doi.org/10.1002/1097-0258(20001230)19:24<3325::aid-sim827>3.0.co;2-d). PMID: [11122498](#)
- Hamm MP, et al. (2010) A descriptive analysis of a representative sample of pediatric randomized controlled trials published in 2007. BMC Pediatr 10:96. doi: [10.1186/1471-2431-10-96](https://doi.org/10.1186/1471-2431-10-96). PMID: [21176224](#)
- Hartling L, et al. (2011) Applying the risk of bias tool in a systematic review of combination long-acting beta-agonists and inhaled corticosteroids for persistent asthma. PLoS One 6:e17242. doi: [10.1371/journal.pone.0017242](https://doi.org/10.1371/journal.pone.0017242). PMID: [21390219](#)
- Hartling L, et al. (2013) Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. J Clin Epidemiol 66:973-981. doi: [10.1016/j.jclinepi.2012.07.005](https://doi.org/10.1016/j.jclinepi.2012.07.005). PMID: [22981249](#)
- Hawkins N, et al. (2016) 'Arm-based' parameterization for network meta-analysis. Res Synth Methods 7:306-313. doi: [10.1002/jrsm.1187](https://doi.org/10.1002/jrsm.1187). PMID: [26610409](#)
- Haynes RB, et al. (2004) Optimal search strategies for retrieving scientifically strong studies of diagnosis from MEDLINE: analytical survey. BMJ 328:1040. doi: [10.1136/bmj.38068.557998.EE](https://doi.org/10.1136/bmj.38068.557998.EE). PMID: [15073027](#).
- Hedin RJ, et al. (2016) Publication bias and nonreporting found in majority of systematic reviews and meta-analyses in anesthesiology journals. Anesth Analg 123:1018-1025. doi: [10.1213/ANE.0000000000001452](https://doi.org/10.1213/ANE.0000000000001452). PMID: [27537925](#)
- Herbison P, et al. (2011) Different methods of allocation to groups in randomized trials are associated with different levels of bias. A meta-epidemiological study. J Clin Epidemiol 64:1070-1075. doi: [10.1016/j.jclinepi.2010.12.018](https://doi.org/10.1016/j.jclinepi.2010.12.018). PMID: [21474279](#)
- Higgins JP, et al. (2011a) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343:d5928. doi: [10.1136/bmj.d5928](https://doi.org/10.1136/bmj.d5928). PMID: [22008217](#)
- Higgins JPT, et al. (2011b) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. <http://handbook-5-1.cochrane.org/> (参照 2021/02/24)
- Higgins JPT, et al. (2019) RoB 2 tool A revised Cochrane risk of bias tool for randomized trials. <https://www.riskofbias.info/welcome/rob-2-0-tool> (参照 2021/02/24)
- Hong H, et al. (2015) Incorporation of individual-patient data in network meta-analysis for multiple continuous endpoints, with application to diabetes treatment. Stat Med 34:2794-2819. doi: [10.1002/sim.6519](https://doi.org/10.1002/sim.6519). PMID: [25924975](#)
- Hong H, et al. (2016) A Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons. Res Synth Methods 7:6-22. doi: [10.1002/jrsm.1153](https://doi.org/10.1002/jrsm.1153). PMID: [26536149](#)
- Howard BE, et al. (2016) SWIFT-review: a text-mining workbench for systematic review. Syst Rev 5:87. doi: [10.1186/s13643-016-0263-z](https://doi.org/10.1186/s13643-016-0263-z). PMID: [27216467](#)
- Hróbjartsson A, et al. (2012) Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. BMJ 344:e1119. doi: [10.1136/bmj.e1119](https://doi.org/10.1136/bmj.e1119)

[e1119](#). PMID: [22371859](#)

Hughes MD, et al. (1992) The impact of stopping rules on heterogeneity of results in overviews of clinical trials. *Biometrics* 48:41-53. PMID: [1581492](#)

Hultcrantz M, et al. (2017) The GRADE working group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 87:4-13. doi: [10.1016/j.jclinepi.2017.05.006](#). PMID: [28529184](#)

Hutton B, et al. (2015) The PRISMA extension statement. *Ann Intern Med* 163:566-567. doi: [10.7326/L15-5144-2](#). PMID: [26436629](#)

Inaba Y, et al. (2009) Impact of single or multicentre study design on the results of trials examining the efficacy of adjunctive devices to prevent distal embolisation during acute myocardial infarction. *EuroIntervention* 5:375-383. doi: 10.4244/v5i3a59. PMID: [19736164](#)

Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press

Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for systematic reviews. National Academies Press

Ioannidis JP. (2011) Excess significance bias in the literature on brain volume abnormalities. *Arch Gen Psychiatry* 68:773-780. doi: [10.1001/archgenpsychiatry.2011.28](#). PMID: [21464342](#)

Jüni P, et al. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 282:1054-1060. doi: [10.1001/jama.282.11.1054](#). PMID: [10493204](#)

Kirkham JJ, et al. (2010) The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 340:c365. doi: [10.1136/bmj.c365](#). PMID: [20156912](#)

Kjaergard LL, et al. (2001) Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 135:982-989. doi: [10.7326/0003-4819-135-11-200112040-00010](#). PMID: [11730399](#)

Knudsen AB, et al. (2016) Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the US preventive services task force. *JAMA* 315:2595-2609. doi: [10.1001/jama.2016.6828](#). PMID: [27305518](#)

Krahn U, et al. (2013) A graphical tool for locating inconsistency in network meta-analyses. *BMC Med Res Methodol* 13:35. doi: [10.1186/1471-2288-13-35](#). PMID: [23496991](#)

Krogh J, et al. (2015) DEPERROR: Risks of systematic errors in drug and non-drug randomized clinical trials assessing intervention effects in patients with unipolar depression. *J Affect Disord* 179:121-127. doi: [10.1016/j.jad.2015.03.042](#). PMID: [25863907](#)

Kunz R, et al. (2007) Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev* MR000012. doi: [10.1002/14651858.MR000012.pub2](#). PMID: [17443633](#)

Kyzas PA, et al. (2005) Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst* 97:1043-1055. doi: [10.1093/jnci/dji184](#). PMID: [16030302](#)

Lee A, et al. (2006) Publication bias affected the estimate of postoperative nausea in an acupoint stimulation systematic review. *J Clin Epidemiol* 59:980-983. doi: [10.1016/j.jclinepi.2006.02.003](#). PMID: [16895822](#)

- Lefebvre C, et al. (2020) Chapter 4: Searching for and selecting studies. In: Higgins JPT, et al. editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.1. <https://training.cochrane.org/handbook/current> (参照 2021/02/24)
- Lin L, et al. (2017) Performing arm-based network meta-analysis in R with the pcnetmeta package. *J Stat Softw* 80. doi: [10.18637/jss.v080.i05](https://doi.org/10.18637/jss.v080.i05). PMID: [28883783](https://pubmed.ncbi.nlm.nih.gov/28883783/) [ノックページ：<https://cran.r-project.org/web/packages/pcnetmeta/>] (参照 2021/02/24)
- Linde K, et al. (1999) Impact of study quality on outcome in placebo-controlled trials of homeopathy. *J Clin Epidemiol* 52:631-636. doi: [10.1016/s0895-4356\(99\)00048-7](https://doi.org/10.1016/s0895-4356(99)00048-7). PMID: [10391656](https://pubmed.ncbi.nlm.nih.gov/10391656/)
- Liu CJ, et al. (2011) Do unblinded assessors bias muscle strength outcomes in randomized controlled trials of progressive resistance strength training in older adults? *Am J Phys Med Rehabil* 90:190-196. doi: [10.1097/PHM.0b013e31820174b3](https://doi.org/10.1097/phm.0b013e31820174b3). PMID: [21173683](https://pubmed.ncbi.nlm.nih.gov/21173683/)
- Marshall IJ, et al. (2016) RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials. *J Am Med Inform Assoc* 23:193-201. doi: [10.1093/jamia/ocv044](https://doi.org/10.1093/jamia/ocv044). PMID: [26104742](https://pubmed.ncbi.nlm.nih.gov/26104742/)
- McDonagh M, et al. (2013) Avoiding Bias in Selecting Studies. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Oregon Health & Science University, McMaster University, and Southern California Evidencebased Practice Centers under Contract No. 290-2007-10057-I.) AHRQ Publication No. 13- EHC045-EF. Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/methods-guidance-bias_methods.pdf (参照 2021/02/24)
- McGowan J, et al. (2016) PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E). Ottawa: CADTH. https://www.cadth.ca/sites/default/files/pdf/CP0015_PRESS_Update_Report_2016.pdf (参照 2021/03/08)
- McGuire W, et al. (2010) The effect of fluconazole prophylaxis in very low birthweight infants is overestimated by before-after studies. *Arch Dis Child Fetal Neonatal Ed* 95:F182-187. doi: [10.1136/adc.2009.160796](https://doi.org/10.1136/adc.2009.160796). PMID: [19822526](https://pubmed.ncbi.nlm.nih.gov/19822526/)
- McInnes MDF, et al. (2018) Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 319:388-396. doi: [10.1001/jama.2017.19163](https://doi.org/10.1001/jama.2017.19163). PMID: [29362800](https://pubmed.ncbi.nlm.nih.gov/29362800/)
- McLeod BD, et al. (2004) Using dissertations to examine potential bias in child and adolescent clinical trials. *J Consult Clin Psychol* 72:235-251. doi: [10.1037/0022-006X.72.2.235](https://doi.org/10.1037/0022-006X.72.2.235). PMID: [15065958](https://pubmed.ncbi.nlm.nih.gov/15065958/)
- Mhaskar R, et al. (2012) Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. *J Clin Epidemiol* 65:602-609. doi: [10.1016/j.jclinepi.2011.10.016](https://doi.org/10.1016/j.jclinepi.2011.10.016). PMID: [22424985](https://pubmed.ncbi.nlm.nih.gov/22424985/)
- Millard LA, et al. (2016) Machine learning to assist risk-of-bias assessments in systematic reviews. *Int J Epidemiol* 45:266-277. doi: [10.1093/ije/dyv306](https://doi.org/10.1093/ije/dyv306). PMID: [26659355](https://pubmed.ncbi.nlm.nih.gov/26659355/)
- Mills EJ, et al. (2013) Demystifying trial networks and network meta-analysis. *BMJ* 346:f2914. doi: [10.1136/bmj.f2914](https://doi.org/10.1136/bmj.f2914). PMID: [23674332](https://pubmed.ncbi.nlm.nih.gov/23674332/)
- Morton SC, et al. (2018) Quantitative Synthesis-An Update. Methods Guide for Comparative Effectiveness

- Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 18-EHC007- EF. Agency for Healthcare Research and Quality. Posted final reports are located on the Effective Health Care Program search page. https://doi.org/10.23970/AHRQE_PCMETHGUIDE3 (参照 2021/02/24)
- Niemeyer H, et al. (2012) Publication bias in meta-analyses of the efficacy of psychotherapeutic interventions for schizophrenia. *Schizophr Res* 138:103-112. doi: [10.1016/j.schres.2012.03.023](https://doi.org/10.1016/j.schres.2012.03.023). PMID: [22484024](https://pubmed.ncbi.nlm.nih.gov/22484024/)
- Norris SL, et al. (2012) Selective outcome reporting as a source of bias in reviews of comparative effectiveness. Agency for healthcare research and quality. PMID: [22993870](https://pubmed.ncbi.nlm.nih.gov/22993870/)
- Nüesch E, et al. (2009a) The importance of allocation concealment and patient blinding in osteoarthritis trials: a meta-epidemiologic study. *Arthritis Rheum* 61:1633-1641. doi: [10.1002/art.24894](https://doi.org/10.1002/art.24894). PMID: [19950329](https://pubmed.ncbi.nlm.nih.gov/19950329/)
- Nüesch E, et al. (2009b) The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 339:b3244. doi: [10.1136/bmj.b3244](https://doi.org/10.1136/bmj.b3244). PMID: [19736281](https://pubmed.ncbi.nlm.nih.gov/19736281/)
- O'Mara-Eves A, et al. (2015) Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Syst Rev* 4:5. doi: [10.1186/2046-4053-4-5](https://doi.org/10.1186/2046-4053-4-5). PMID: [25588314](https://pubmed.ncbi.nlm.nih.gov/25588314/)
- Ouzzani M, et al. (2016) Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 5:210. doi: [10.1186/s13643-016-0384-4](https://doi.org/10.1186/s13643-016-0384-4). PMID: [27919275](https://pubmed.ncbi.nlm.nih.gov/27919275/). <https://rayyan.qcri.org/welcome> (解説動画：<https://www.youtube.com/watch?v=irAOQgzFMs4>) (参照 2021/02/24)
- Peters JL, et al. (2007) Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Stat Med* 26:4544-4562. doi: [10.1002/sim.2889](https://doi.org/10.1002/sim.2889). PMID: [17476644](https://pubmed.ncbi.nlm.nih.gov/17476644/)
- Pildal J, et al. (2007) Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 36:847-857. doi: [10.1093/ije/dym087](https://doi.org/10.1093/ije/dym087). PMID: [17517809](https://pubmed.ncbi.nlm.nih.gov/17517809/)
- Plummer M. (2017) JAGS. <http://mcmc-jags.sourceforge.net/> (参照 2021/02/24)
- Plummer M, et al. (2019) jags: Bayesian Graphical Models using MCMC. <https://cran.r-project.org/web/packages/rjags/> (参照 2021/02/24)
- pmSearch (2019) <https://stat.zanet.biz/sr/pmsearch.htm> (参照 2021/02/24)
- Poorolajal J, et al. (2010) Capture-recapture method for assessing publication bias. *J Res Med Sci* 15:107-115. PMID: [21526067](https://pubmed.ncbi.nlm.nih.gov/21526067/)
- Puhan MA, et al. (2014) A GRADE working group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 349:g5630. doi: [10.1136/bmj.g5630](https://doi.org/10.1136/bmj.g5630). PMID: [25252733](https://pubmed.ncbi.nlm.nih.gov/25252733/)
- Riley RD, et al. (2004) Sensitivity analyses allowed more appropriate and reliable meta-analysis conclusions for multiple outcomes when missing data was present. *J Clin Epidemiol* 57:911-924. doi: [10.1016/j.jclinepi.2004.01.018](https://doi.org/10.1016/j.jclinepi.2004.01.018). PMID: [15504634](https://pubmed.ncbi.nlm.nih.gov/15504634/)
- Roest AM, et al. (2015) Reporting bias in clinical trials investigating the efficacy of second-generation antidepressants in the treatment of anxiety disorders: a report of 2 meta-analyses. *JAMA Psychiatry* 72:500-510. doi: [10.1001/jamapsychiatry.2015.15](https://doi.org/10.1001/jamapsychiatry.2015.15). PMID: [25806940](https://pubmed.ncbi.nlm.nih.gov/25806940/)
- Rücker G, et al. (2019) Network meta-analysis of multicomponent interventions. *Biometrical Journal*

- 62:808-821. doi: [10.1002/bimj.201800167](https://doi.org/10.1002/bimj.201800167). PMID: 31021449 [R のパッケージ : <https://cran.r-project.org/web/packages/netmeta/> (参照 2021/03/08)]
- Salanti G, et al. Multiple-Treatments Meta-Analysis. <http://www.mtm.uoi.gr/> (参照 2021/02/24)
- Salanti G, et al. (2011) Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 64:163-171. doi: [10.1016/j.jclinepi.2010.03.016](https://doi.org/10.1016/j.jclinepi.2010.03.016). PMID: 20688472
- Savovi J, et al. (2012) Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess* 16:1-82. doi: [10.3310/hta16350](https://doi.org/10.3310/hta16350). PMID: 22989478
- Schou IM, et al. (2013) Meta-analysis of clinical trials with early stopping: an investigation of potential bias. *Stat Med* 32:4859-4874. doi: [10.1002/sim.5893](https://doi.org/10.1002/sim.5893). PMID: 23824994
- Schulz KF, et al. (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 273:408-412. doi: [10.1001/jama.273.5.408](https://doi.org/10.1001/jama.273.5.408). PMID: 7823387
- Schünemann H, et al. (2013) GRADE Handbook. <http://gdt.guidelinedevelopment.org/app/handbook/handbook.html> (参照 2021/02/24)
- Schünemann HJ (2016) Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision? *J Clin Epidemiol* 75:6-15. doi: [10.1016/j.jclinepi.2016.03.018](https://doi.org/10.1016/j.jclinepi.2016.03.018). PMID: 27063205
- Schünemann HJ, et al. (2019a) GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol* 111:105-114. doi: [10.1016/j.jclinepi.2018.01.012](https://doi.org/10.1016/j.jclinepi.2018.01.012). PMID: 29432858
- Schünemann HJ, et al. (2019b) GRADE guidelines: 22. The GRADE approach for tests and strategies—from test accuracy to patient-important outcomes and recommendations. *J Clin Epidemiol* 111:69-82. doi: [10.1016/j.jclinepi.2019.02.003](https://doi.org/10.1016/j.jclinepi.2019.02.003). PMID: 30738926
- Schwarzer G, et al. (2010) Empirical evaluation suggests Copas selection model preferable to trim-and-fill method for selection bias in meta-analysis. *J Clin Epidemiol* 63:282-288. doi: [10.1016/j.jclinepi.2009.05.008](https://doi.org/10.1016/j.jclinepi.2009.05.008). PMID: 19836925
- Selva A, et al. (2017) Development and use of a content search strategy for retrieving studies on patients' views and preferences. *Health Qual Life Outcomes* 15:126. doi: [10.1186/s12955-017-0698-5](https://doi.org/10.1186/s12955-017-0698-5). PMID: 28851437
- Shea BJ, et al. (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 7:10. doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10). PMID: 17302989
- Shea BJ, et al. (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomized or non-randomised studies of healthcare interventions, or both. *BMJ* 358:j4008. doi: [10.1136/bmj.j4008](https://doi.org/10.1136/bmj.j4008). PMID: 28935701
- Shemilt I, et al. (2019) Chapter 20: Economic evidence. In: Higgins JPT, et al. editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0*. <https://training.cochrane.org/handbook/archive/>

v6 (参照 2021/02/24)

Song F, et al. (2000) Publication and related biases. *Health Technol Assess* 4:1-115. PMID: [10932019](#)

Song F, et al. (2010) Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess* 14:iii, ix-xi, 1-193. doi: [10.3310/hta14080](#). PMID: [20181324](#)

Sox HC, et al. (2013) Medical decision making, 2nd ed. Wiley-Blackwell

Sterne JA, et al. (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 355:i4919. doi: [10.1136/bmj.i4919](#). PMID: [27733354](#). <https://www.riskofbias.info/welcome/home/current-version-of-robins-i> (参照 2021/02/24)

Sterne JA, et al. (2001) Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 323:101-105. doi: [10.1136/bmj.323.7304.101](#). PMID: [11451790](#)

Sterne JAC, et al. (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 366: l4898. doi: [10.1136/bmj.l4898](#). PMID: [31462531](#)

Tierney JF, et al. (2007) Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 8:16. doi: [10.1186/1745-6215-8-16](#). PMID: [17555582](#)

Turner RM, et al. (2009) Bias modelling in evidence synthesis. *J R Stat Soc Ser A Stat Soc* 172:21-47. doi: [10.1111/j.1467-985X.2008.00547.x](#). PMID: [19381328](#)

Unverzagt S, et al. (2013) Single-center trials tend to provide larger treatment effects than multicenter trials: a systematic review. *J Clin Epidemiol* 66:1271-1280. doi: [10.1016/j.jclinepi.2013.05.016](#). PMID: [23972520](#)

Vale CL, et al. (2013) Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. *BMJ* 346:f1798. doi: [10.1136/bmj.f1798](#). PMID: [23610376](#)

van Aert RC, et al. (2016) Conducting meta-analyses based on p values: reservations and recommendations for applying p-uniform and p-curve. *Perspect Psychol Sci* 11:713-729. doi: [10.1177/1745691616650874](#). PMID: [27694466](#)

van Hoorn R, et al. (2016) The development of PubMed search strategies for patient preferences for treatment outcomes. *BMC Med Res Methodol* 16:88. doi: [10.1186/s12874-016-0192-5](#). PMID: [27473226](#)

van Tulder MW, et al. (2009) Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine (Phila Pa 1976)* 34:1685-1692. doi: [10.1097/BRS.0b013e3181ab6a78](#). PMID: [19770609](#)

van Valkenhoef G, et al. gemtc. <https://cran.r-project.org/web/packages/gemtc/> (参照 2021/02/24)

van Valkenhoef G, et al. (2016) Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis. *Res Synth Methods* 7:80-93. doi: [10.1002/jrsm.1167](#). PMID: [26461181](#)

Viechtbauer W. (2020) metafor: Meta-Analysis Package for R. <https://cran.r-project.org/web/packages/metafor/index.html> (参照 2021/02/24)

Viswanathan M, et al. (2012) Development of the RTI item bank on risk of bias and precision of observational studies. *J Clin Epidemiol* 65:163-178. doi: [10.1016/j.jclinepi.2011.05.008](#). PMID: [21959223](#)

- Viswanathan M, et al. (2017) Assessing the Risk of Bias in Systematic Reviews of Health Care Interventions. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 17(18)-EHC036-EF. Rockville, MD: Agency for Healthcare Research and Quality. Posted final reports are located on the Effective Health Care Program search page. <https://doi.org/10.23970/AHQEPCMETHODGUIDE2> (参照 2021/02/24)
- Watzlawick R, et al. (2014) Effect and reporting bias of RhoA/ROCK-blockade intervention on locomotor recovery after spinal cord injury: a systematic review and meta-analysis. JAMA Neurol 71:91-99. doi: [10.1001/jamaneurol.2013.4684](https://doi.org/10.1001/jamaneurol.2013.4684). PMID: [24297045](#)
- Wessels M, et al. (2016) How to identify existing literature on patients' knowledge, views, and values: the development of a validated search filter. J Med Libr Assoc 104:320-324. doi: [10.3163/1536-5050.104.4.014](https://doi.org/10.3163/1536-5050.104.4.014). PMID: [27822157](#)
- White CM, et al. (2009) Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. In: Agency for Healthcare Research and Quality (AHRQ). Methods Guide for Comparative Effectiveness Reviews. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/methods-guidance-de-novo-processes_methods.pdf (参照 2021/02/24)
- White IR, et al. (2019) A comparison of arm-based and contrast-based models for network meta-analysis. Stat Med 38:5197-5213. doi: [10.1002/sim.8360](https://doi.org/10.1002/sim.8360). PMID: [31583750](#)
- Whiting PF, et al. (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529-536. doi: [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009). PMID: [22007046](#)
- World Health Organization (WHO). (2019) WHO Handbook for Guideline Development. 2nd edition. 17: Developing guideline recommendations for tests or diagnostic tools. <https://apps.who.int/iris/bitstream/handle/10665/145714/9789241548960-chap17-eng.pdf?sequence=20&isAllowed=y> (参照 2021/02/24)
- Wood L, et al. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 336:601-605. doi: [10.1136/bmj.39465.451748.AD](https://doi.org/10.1136/bmj.39465.451748.AD). PMID: [18316340](#)
- Yepes-Nuñez JJ, et al. (2019) Development of the summary of findings table for network meta-analysis. J Clin Epidemiol 115:1-13. doi: [10.1016/j.jclinepi.2019.04.018](https://doi.org/10.1016/j.jclinepi.2019.04.018). PMID: [31055177](#)
- Zhang J, et al. (2014) Network meta-analysis of randomized clinical trials: reporting the proper summaries. Clin Trials 11:246-262. doi: [10.1177/1740774513498322](https://doi.org/10.1177/1740774513498322). PMID: [24096635](#)
- Zhang J, et al. (2019) Bayesian hierarchical methods for meta-analysis combining randomized-controlled and single-arm studies. Stat Methods Med Res 28:1293-1310. doi: [10.1177/0962280218754928](https://doi.org/10.1177/0962280218754928). PMID: [29433407](#)
- Zhang Y, et al. (2017) Using patient values and preferences to inform the importance of health outcomes in practice guideline development following the GRADE approach. Health Qual Life Outcomes 15:52. doi: [10.1186/s12955-017-0621-0](https://doi.org/10.1186/s12955-017-0621-0). PMID: [28460638](#)
- Zhang E, et al. (2017) TrialSize. <https://cran.r-project.org/web/packages/TrialSize/index.html> (参照

2021/02/24)

上岡洋晴 他. (2018) 「AMSTAR 2：ヘルスケア介入のランダム化または非ランダム化研究あるいは両方を含むシステムティック・レビューのための厳密な評価ツール」の紹介と解. 薬理と治療 46:111-113.

河合富士美 他. (2017) システマティックレビュー検索式査読のためのガイドライン：PRESS Peer Review of Electronic Search Strategies. 医学図書館 64:75-78.

森實敏夫 他. (2015) 特別寄稿 5 診断に関する診療ガイドライン (CPG) の作成. https://minds.jcqhc.or.jp/s/guidance_special_articles5_1 (参照 2021/1/28)

5章 文献

Alonso-Coello P, et al. (2016) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 353:i2016. doi: [10.1136/bmj.i2016](https://doi.org/10.1136/bmj.i2016). PMID: [27353417](#)

Ara R, et al. (2017) The identification, review and synthesis of health state utility values from the literature. Pharamcoeconomics 35:43-55. doi: [10.1007/s40273-017-0547-8](https://doi.org/10.1007/s40273-017-0547-8). PMID: [29052156](#)

Arber M, et al. (2016) Performance of search filters to identify health state utility studies. Value in Health 19:A390-A391.

Arber M, et al. (2017) Performance of ovid MEDLINE search filters to identify health state utility studies. Int J Technol Assess Health Care 33:472-480. doi: [10.1017/S0266462317000897](https://doi.org/10.1017/S0266462317000897). PMID: [29065942](#)

Arber, M. et al. (2018) Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care 34:547-554. doi: [10.1017/S0266462318000636](https://doi.org/10.1017/S0266462318000636). PMID: [30442221](#)

Arber M, et al. (2018) Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care 34:547-554. doi: [10.1017/S0266462318000636](https://doi.org/10.1017/S0266462318000636). PMID: [30442221](#)

Brunetti M, et al. (2013) GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol 66:140-150. doi: [10.1016/j.jclinepi.2012.04.012](https://doi.org/10.1016/j.jclinepi.2012.04.012). PMID: [22863410](#)

Drummond MF, et al. (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ 313:275-283. doi: [10.1136/bmj.313.7052.275](https://doi.org/10.1136/bmj.313.7052.275). PMID: [8704542](#)

Drummond MF, et al. (2015) Methods for Economic Evaluation of Health Care Programmes. Fourth ed. Oxford University Press

European network for Health Technology Assessment. (2020) METHODOLOGICAL GUIDELINES: Practical considerations when critically assessing economic evaluations. https://eunethta.eu/wp-content/uploads/2020/03/EUNetHTA-JA3WP6B2-5-Guidance-Critical-Assessment-EE_v1-0.pdf (参照 2020/10/25)

Frick K, et al. (2012) Usefulness of Economic Evaluation Data in Systematic Reviews of Evidence [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US) <https://www.ncbi.nlm.nih.gov/books/NBK114533/> (参照 2020/10/15)

Glanville J, et al. (2009) How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. Int J Technol Assess Health Care 25:522-529. doi: [10.1017/S0266462309990523](https://doi.org/10.1017/S0266462309990523). PMID: [19845982](#)

Husereau D, et al. (2013) Consolidated Health Economic Evaluation Reporting Standards (CHEERS)-- explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 16:231-250. doi: [10.1016/j.jval.2013.02.002](https://doi.org/10.1016/j.jval.2013.02.002). PMID: 23538175

ISSG Search Filters Resource. <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home> (参照 2020/10/25)

Luce BR, et al. (1996) Estimating costs in cost-effectiveness analysis. In: Gold MR, et al. editors. Cost-effectiveness in health and medicine. Oxford University Press

McKinlay RJ, et al. (2006) Optimal search strategies for detecting cost and economic studies in EMBASE. BMC Health Serv Res 6:67. doi: [10.1186/1472-6963-6-67](https://doi.org/10.1186/1472-6963-6-67). PMID: [16756662](#)

National Collaborating Centre for Mental Health (NCCMH). (2014) Antenatal and postnatal mental health (update): clinical management and service guidance. <https://www.nice.org.uk/guidance/cg192/evidence> (参照 2020/10/25)

National Institute for Health and Care Excellence (NICE). (2020) Developing NICE guidelines: the manual. [PMG20] <https://www.nice.org.uk/process/pmg20/> (参照 2020/10/15)

Paisley S. (2016) Identification of evidence for key parameters in decision-analytic models of cost effectiveness: a description of sources and a recommended minimum search requirement. Pharmacoeconomics 34:597-608. doi: [10.1007/s40273-015-0372-x](https://doi.org/10.1007/s40273-015-0372-x). PMID: [26861793](#)

Papaioannou D, et al. (2010) NICE DSU Technical support document 9: the identification, review and synthesis of health state utility values from the literature. <https://pubmed.ncbi.nlm.nih.gov/28481496/> (参照 2020/10/15)

Philips Z, et al. (2006) Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics 24:355-371. doi: [10.2165/00019053-200624040-00006](https://doi.org/10.2165/00019053-200624040-00006). PMID: [16605282](#)

Shemilt I, et al. (2019) Chapter 20: Economic evidence. In: Higgins JPT, et al. editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). <https://www.training.cochrane.org/handbook> (参照 2020/10/15)

ScHARRHUD (Health Utilities Database) <http://www.scharrhud.org/> (参照 2021/03/06)

Thielen FW, et al. (2016) How to prepare a systematic review of economic evaluations for clinical practice guidelines: database selection and search strategy development (part 2/3). Expert Rev Pharmacoecon Outcomes Res 16:705-721. doi: [10.1080/14737167.2016.1246962](https://doi.org/10.1080/14737167.2016.1246962). PMID: [27805466](#)

van Mastrigt GA, et al. (2016) How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3). Expert Rev Pharmacoecon Outcomes Res 16:689-704. doi: [10.1080/14737167.2016.1246960](https://doi.org/10.1080/14737167.2016.1246960). PMID: [27805469](#)

Wilczynski NL, et al. (2004) Optimal search strategies for detecting health services research studies in MEDLINE. CMAJ 171:1179-1185. doi: [10.1503/cmaj.1040512](https://doi.org/10.1503/cmaj.1040512). PMID: [15534310](#)

Wijnen B, et al. (2016) How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: data extraction, risk of bias, and transferability (part 3/3). Expert Rev Pharmacoecon Outcomes Res 16:723-732. doi: [10.1080/14737167.2016.1246961](https://doi.org/10.1080/14737167.2016.1246961). PMID: [27762640](#)

6章 文献

- Akl EA, et al. (2017) GRADE equity guidelines 2: considering health equity in GRADE guideline development: equity extension of the guideline development checklist. *J Clin Epidemiol* 90:68-75. doi: [10.1016/j.jclinepi.2017.01.017](https://doi.org/10.1016/j.jclinepi.2017.01.017). PMID: [28499847](#)
- Alonso-Coello P, et al. (2016a) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 353:i2016. doi: [10.1136/bmj.i2016](https://doi.org/10.1136/bmj.i2016). PMID: [27353417](#)
- Alonso-Coello P, et al. (2016b) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ* 353:i2089. doi: [10.1136/bmj.i2089](https://doi.org/10.1136/bmj.i2089). PMID: [27365494](#)
- Alper BS, et al. (2019) Defining certainty of net benefit: a GRADE concept paper. *BMJ Open* 9:e027445. doi: [10.1136/bmjopen-2018-027445](https://doi.org/10.1136/bmjopen-2018-027445). PMID: [31167868](#)
- Andrews JC, et al. (2013a) GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 66:719-725. doi: [10.1016/j.jclinepi.2012.03.013](https://doi.org/10.1016/j.jclinepi.2012.03.013). PMID: [23312392](#).
- Andrews JC, et al. (2013b) GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 66:726-735. doi: [10.1016/j.jclinepi.2013.02.003](https://doi.org/10.1016/j.jclinepi.2013.02.003). PMID: [23570745](#).
- Boyd CM, et al. (2012) Methods for Benefit and Harm Assessment in Systematic Reviews. AHRQ Methods for Effective Health Care (book) PMID: [23326898](#)
- Broekhuizen H, et al. (2015) A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. *Pharmacoeconomics* 33:445-455. doi: [10.1007/s40273-014-0251-x](https://doi.org/10.1007/s40273-014-0251-x). PMID: [25630758](#)
- Delbecq AL, et al. (1971) A group process model for problem identification and Program planning. *The Journal of Applied Behavioral Science* 7:466-492
- Dalkey N, et al. (1969) The Delphi Method, III: Use of Self Ratings to Improve Group Estimates. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2006/RM6115.pdf (参照 2021/02/25)
- Dellinger RP, et al. (2008) Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008. *Crit Care Med* 36:296-327. doi: [10.1097/01.CCM.0000298158.12101.41](https://doi.org/10.1097/01.CCM.0000298158.12101.41). PMID: [18158437](#)
- Dolan JG, et al. (2013) Patients' preferences and priorities regarding colorectal cancer screening. *Med Decis Making* 33:59-70. doi: [10.1177/0272989X12453502](https://doi.org/10.1177/0272989X12453502). PMID: [22895558](#)
- Fitch K, et al. (2001) The RAND/UCLA Appropriateness Method User's Manual. RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1269.html (参照 2021/02/25)
- Gail MH, et al. (1999) Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst* 91:1829-46. doi: [10.1093/jnci/91.21.1829](https://doi.org/10.1093/jnci/91.21.1829). PMID: [10547390](#)

- Guyatt G, et al. (2013) GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol* 66:151-157. doi: [10.1016/j.jclinepi.2012.01.006](https://doi.org/10.1016/j.jclinepi.2012.01.006). PMID: [2254203](#)
- Institute of Medicine (IOM). (2011) Clinical Practice Guidelines We Can Trust. National Academies Press
- Jaeschke R, et al. (2008) Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ* 337:327-330. doi: [10.1136/bmj.a744](https://doi.org/10.1136/bmj.a744). PMID: [18669566](#)
- Keeney R, et al. (1993) Decisions with multiple objectives: Preferences and Value Trade-Offs. Cambridge University Press
- Marsh K, et al. (2016) Multiple criteria decision analysis for health care decision making--emerging good practices: report 2 of the ISPOR MCDA emerging good practices task force. *Value Health* 19:125-37. doi: [10.1016/j.jval.2015.12.016](https://doi.org/10.1016/j.jval.2015.12.016). PMID: [27021745](#)
- Marsh K, et al. (2017) Multi-Criteria Decision Analysis to Support Healthcare Decisions. Springer
- Moberg J, et al. (2015) GRADE Evidence to Decision (EtD) Frameworks Guidance. Version 1.1 [updated May 2015], The GRADE Working Group. <https://ietd.epistemonikos.org/#/help/guidance> (参照 2021/02/25)
- Mt-Isa S, et al. (2014) Balancing benefit and risk of medicines: a systematic review and classification of available methodologies. *Pharmacoepidemiol Drug Saf* 23:667-78. doi: [10.1002/pds.3636](https://doi.org/10.1002/pds.3636). PMID: [24821575](#)
- National Institute for Health and Care Excellence (NICE). (2015) Research recommendations Process and methods guide. <https://www.nice.org.uk/Media/Default/About/what-we-do/Science-policy-and-research/research-recommendation-process-methods-guide-2015.pdf> (参照 2021/02/25)
- Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) <http://www.protectbenefitrisk.eu/index.html> (参照 2021/02/25)
- Pottie K, et al. (2017) GRADE equity guidelines 4: considering health equity in GRADE guideline development: evidence to decision process. *J Clin Epidemiol* 90:84-91. doi: [10.1016/j.jclinepi.2017.08.001](https://doi.org/10.1016/j.jclinepi.2017.08.001). PMID: [28802675](#)
- Rehfuss EA, et al. (2019) The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. *BMJ Glob Health* 4:e000844. doi: [10.1136/bmjjgh-2018-000844](https://doi.org/10.1136/bmjjgh-2018-000844). PMID: [30775012](#)
- Scottish Intercollegiate Guidelines Network (SIGN). (2011) SIGN 50: A guideline developer's handbook. Revised Edition. https://www.sign.ac.uk/media/1642/sign50_2011.pdf (参照 2021/02/25)
- Thokala P, et al. (2016) Multiple criteria decision analysis for health care decision making--an introduction: report 1 of the ISPOR MCDA emerging good practices task force. *Value Health* 19:1-13. doi: [10.1016/j.jval.2015.12.003](https://doi.org/10.1016/j.jval.2015.12.003). PMID: [26797229](#)
- Welch VA, et al. (2017a) GRADE equity guidelines 1: considering health equity in GRADE guideline development: introduction and rationale. *J Clin Epidemiol* 90:59-67. doi: [10.1016/j.jclinepi.2017.01.014](https://doi.org/10.1016/j.jclinepi.2017.01.014). PMID: [28412464](#)

7章 文献

- AGREE Next Steps Consortium. (2009) The AGREE II Instrument [Electronic version]. <http://www.agreertrust.org> (参照 2013/02/27)
- Brouwers MC, et al. (2016) The AGREE reporting checklist: a tool to improve reporting of clinical practice guidelines. BMJ 352:i1152. doi: [10.1136/bmj.i1152](https://doi.org/10.1136/bmj.i1152). PMID: [26957104](#)
- Chen Y, et al. (2017) A reporting tool for practice guidelines in health care: the RIGHT statement. Ann Intern Med 166:128-132. doi: [10.7326/M16-1565](https://doi.org/10.7326/M16-1565). PMID: [27893062](#)
- Guidelines International Patient and Public Involvement Working Group (G-I-N PUBLIC). (2015) G-I-N PUBLIC Toolkit: Patient and Public Involvement in Guidelines. https://g-i-n.net/document-store/resources/toolkit/gin-public-toolkit-final/at_download/file (参照 2021/02/22)
- Institute of Medicine (IOM). (2011a) Clinical Practice Guidelines We Can Trust. National Academies Press
- Institute of Medicine (IOM). (2011b) Finding What Works in Health Care: Standards for Systematic Reviews. National Academies Press
- Kashyap N, et al. (2011) GLIA: GuideLine Implementability Appraisal v. 2.0. Yale Center for Medical Informatics. http://nutmeg.med.yale.edu/glia/doc/GLIA_v2.pdf (参照 2021/02/22)
- Liberati A, et al. (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 6:e1000100. doi: [10.1371/journal.pmed.1000100](https://doi.org/10.1371/journal.pmed.1000100). PMID: [19621070](#)
- National Institute for Health and Clinical Excellence (NICE). (2012) The Guidelines Manual: Process and Methods. <https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933> (参照 2021/02/22)
- National Institute for Health and Care Excellence (NICE). (2013) Factsheet 2: How organizations representing patients and carers can get involved. <https://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Factsheet-2-contribute-to-developing-clinical-guidelines.pdf> (参照 2021/02/22)
- Shea BJ, et al. (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Medical Res Methodol 7:10. doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10). PMID: [17302989](#)
- Shea BJ, et al. (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 358:j4008. doi: [10.1136/bmj.j4008](https://doi.org/10.1136/bmj.j4008). PMID: [28935701](#)
- Shiffman RN, et al. (2003) Standardized reporting of clinical practice guidelines: a proposal from the conference on guideline standardization. Ann Intern Med 139:493-498. doi: [10.7326/0003-4819-139-6-200309160-00013](https://doi.org/10.7326/0003-4819-139-6-200309160-00013). PMID: [13679327](#)

8章 文献

- Alonso-Coello P, et al. (2011) The updating of clinical practice guidelines: insights from an international survey. *Implement Sci.* 6:107. doi: [10.1186/1748-5908-6-107](https://doi.org/10.1186/1748-5908-6-107). PMID: [21914177](#)
- Higgins JPT, et al. (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. <http://handbook-5-1.cochrane.org/> (参照 2021/02/22)
- Institute of Medicine (IOM). (2011) Clinical Practice Guidelines We Can Trust. National Academies Press
- National Institute for Health and Clinical Excellence (NICE). (2012) The Guidelines Manual: Process and Methods. <https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933> (参照 2021/02/22)
- Shekelle PG, et al. (2001) Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA* 286:1461-1467. doi: [10.1001/jama.286.12.1461](https://doi.org/10.1001/jama.286.12.1461). PMID: [11572738](#)
- Shojania KG, et al. (2007) How quickly do systematic reviews go out of date? a survival analysis. *Ann Intern Med* 147:224-233. doi: [10.7326/0003-4819-147-4-200708210-00179](https://doi.org/10.7326/0003-4819-147-4-200708210-00179). PMID: [17638714](#)
- World Health Organization (WHO). (2012) WHO Handbook for Guideline Development. World Health Organization.