

診療ガイドライン作成のツール

Minds2007からGRADEへ



相原守夫

Member of the GRADE Working Group

Minds 2007: 診療ガイドライン作成の手引き 2007 <http://minds4.jcqhc.or.jp/minds/glgl/glgl.pdf>

Minds: Medical Information Network Distribution Service

GRADE: Grading of Recommendations Assessment, Development, and Evaluation

医療の質の最適化や診療のばらつきの解消に国際的な注目が寄せられる中、臨床医、患者、政策決定者によるエビデンスに合致した決断を支援するための診療ガイドラインの作成が爆発的に増加している。たとえば、National Guidelines Clearinghouse をざっと検索しただけでも心不全に関するガイドラインが500件存在することがわかる。このようなガイドラインの氾濫は、診療のばらつきを解消するというよりはむしろ診療のばらつきの原因となっているのではないかと議論する声があるかもしれない。これらの様々なガイドラインの適用を試みる臨床医や政策決定者にとってこの状況をさらに複雑にしているのが、ガイドライン作成者によるエビデンスの評価と推奨作成のためのアプローチに一貫性がないことである。また、これらの様々なアプローチの使用が異なる推奨を生じさせることも明らかになっている。同じ臨床トピックを扱うガイドラインでも異なる組織により作成されたものでは推奨が異なるというその結果がガイドラインを実行しようとする人たちの混乱を招いている。

国内の診療ガイドラインの作成においても、エビデンスの質の評価方法はさまざまであり、推奨の判定も統一されていない。国内で多用されているMinds2007基準は、個別研究を評価するという簡便さを有しているものの研究デザインに依存したエビデンスレベルや推奨評価であり、今後の改訂が望まれる。

Grading of Recommendations Assessment, Development and Evaluation (GRADE) は、エビデンスの質と推奨の強さを評価するための透明性の高い方法により明確性を確保しようとする試みである^{1,2}。本稿では、国際的に普及しているGRADEシステムとMinds2007基準との違い、Minds2007を使って作成されたガイドラインにGRADEシステムを適用させる際の注意点について個人的な意見を述べたい。


1. 診療ガイドラインのためのGRADEシステム(相原 他、凸版メディア出版、2010)

<http://homepage3.nifty.com/aihara/grade-cover.html>

2. GRADEハンドブック (version 3.2 翻訳版、Free) <http://homepage3.nifty.com/aihara/GRADEproHelp.html>

エビデンスのレベル分類

-Minds2007とGRADEの違い-

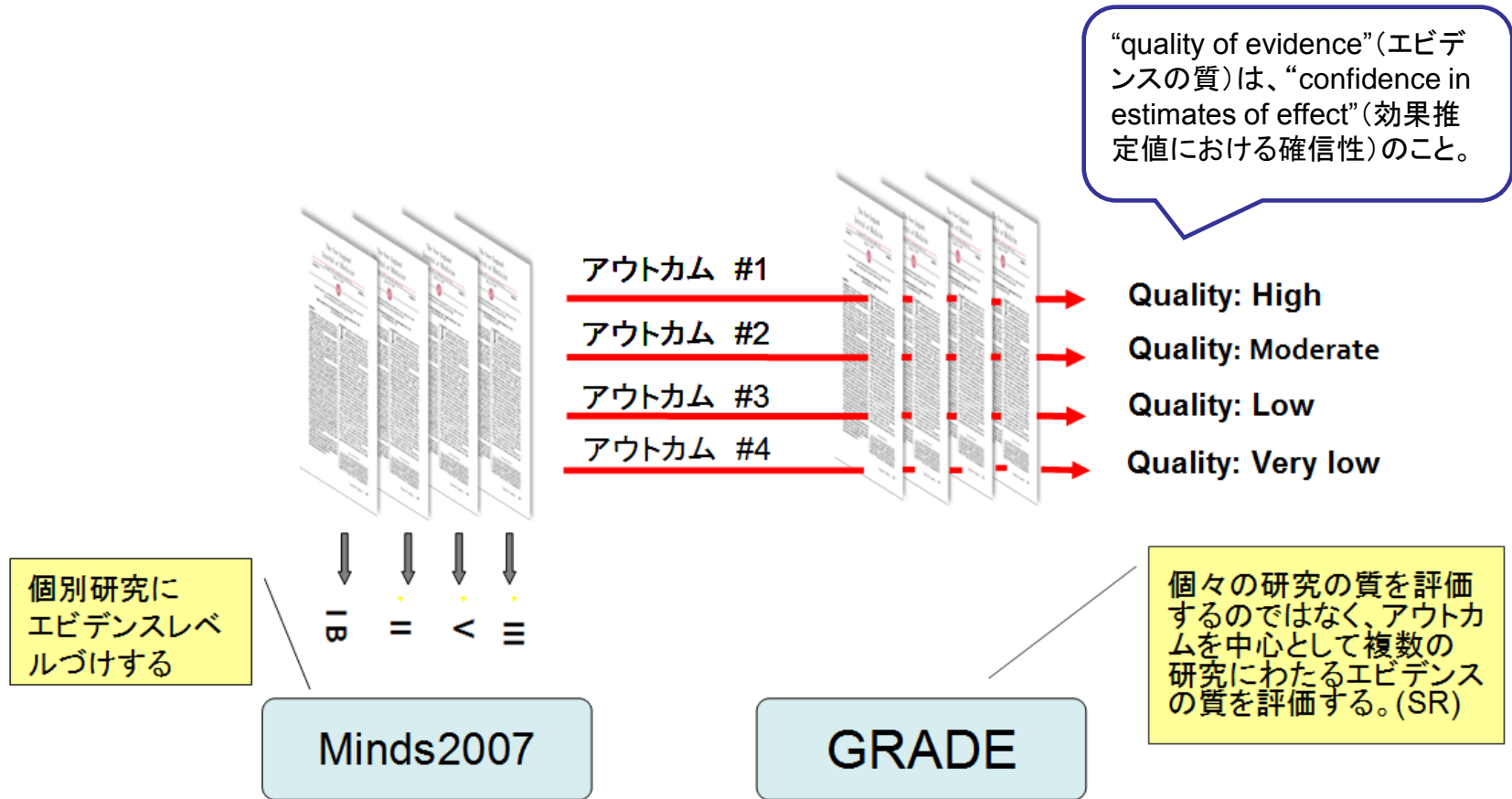
Minds 2007	GRADE
ひとつひとつの文献を研究デザインの種類によってレベル付ける(以下; 質の高いもの順)	各アウトカムについて研究横断的なシステマティック・レビュー(SR)を実施し、総体エビデンスの質を評価する
I システマティック・レビュー/RCTのメタアナリシス II ひとつ以上のランダム化比較試験による III 非ランダム化比較試験による IVa 分析疫学的研究(コホート研究) IVb 分析疫学的研究(症例対照研究、横断研究) V 記述研究(症例報告やケース・シリーズ) VI 患者データに基づかない、専門委員会や専門家個人の意見	・ランダム化比較試験* : high ・観察研究* : Low  グレードダウンの5要因、グレードアップの3要因を考慮して、最終的な総体エビデンスの質を4段階のいずれかに判定する (高、中、低、非常に低)
特定の疑問点について複数の文献(エビデンス)がある場合には、原則的には、最もレベルの高いエビデンスを採用する	各アウトカムについて総体エビデンス(body of evidence)の質を評価し、全体的なエビデンスの質の評価はGRADEルールに従う

* RCT、観察研究、いずれも単独試験を評価するものではないことに注意

GRADE はアウトカム中心である

GRADE is outcome-centric

“quality of evidence” (エビデンスの質) は、“confidence in estimates of effect” (効果推定値における確信性) のこと。

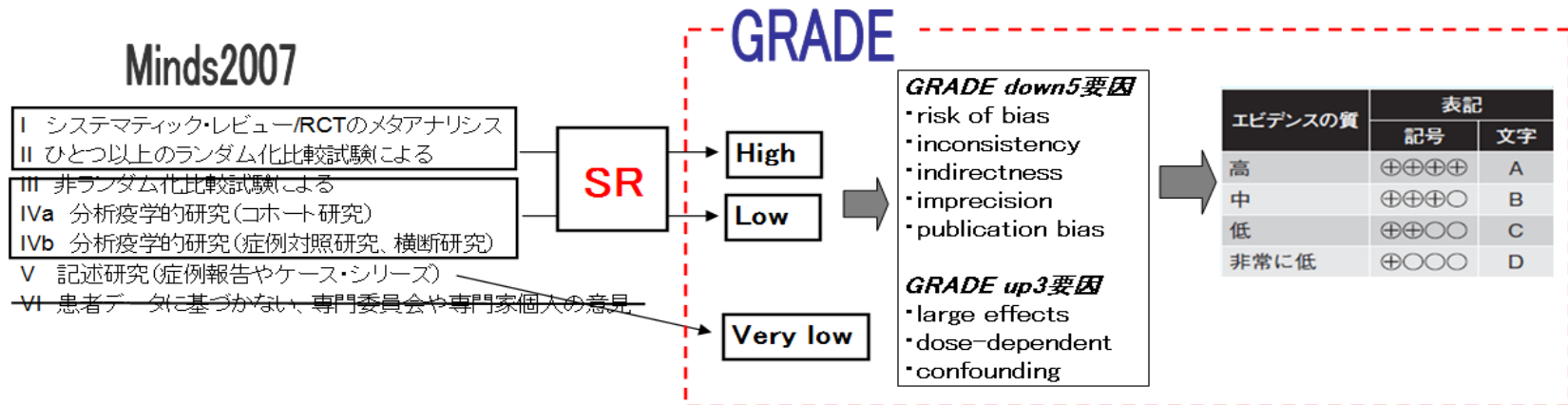


CPG作成ツール: Minds2007からGRADEへ

エビデンスの質の評価

GRADEシステムにおいては、(原則としては)個別研究を評価せず、重大なアウトカムに関するシステマティック・レビュー(SR)を利用し、エビデンスの質を評価する。個々のアウトカムのエビデンスの質を評価し、さらに(診療ガイドライン作成時は)アウトカム全般にわたるエビデンスの質をひとつだけ決定する。

(例)



エビデンスの質の分類は、高(A)、中(B)、低(C)、非常に低(D)の4つのいずれかである

エキスパートの意見はエビデンスとはいわない

推奨の強さの決定

-Minds2007とGRADEの違い-

Minds 2007	GRADE
<p>以下の要素を勘案する</p> <ol style="list-style-type: none">エビデンスのレベル <p>特定の疑問点について複数の文献(エビデンス)がある場合には、原則的には、最もレベルの高いエビデンスを採用する</p> <ol style="list-style-type: none">エビデンスの数と結論のばらつき <p>同じ結論のエビデンスが多ければ多いほど、そして結論のばらつきが小さければ小さいほど推奨は強いものとなる</p> <ol style="list-style-type: none">臨床的有効性の大きさ臨床上の適用性害やコストに関するエビデンス	<p>推奨決定のための4要因</p> <ol style="list-style-type: none">エビデンスの質 <ul style="list-style-type: none">各アウトカムが異なる方向の場合: 「重大なアウトカムに関するエビデンスの質の中で最も質が低いものを全体的なエビデンスの質とする」全てのアウトカムが同じ方向の場合: 「重大なアウトカムに関するエビデンスの質の中で最も質が高いものを全体的なエビデンスの質とする」 <ol style="list-style-type: none">利益と不利益のバランス価値観や好みコストや資源の利用

GRADEからみたMinds2007の推奨決定要因

Minds 2007	GRADE
<p>1. エビデンスのレベル</p> <p>特定の疑問点について複数の文献(エビデンス)がある場合には、原則的には、最もレベルの高いエビデンスを採用する</p>	<ul style="list-style-type: none"> ・いくつかのアウトカムが異なる方向の場合(利益と害): アウトカムに関するエビデンスの中で、最も質が低いものを採用する ・同じ方向の場合、最も質が高いものを採用する
<p>2. エビデンスの数と結論のばらつき</p> <p>同じ結論のエビデンスが多ければ多いほど、そして結論のばらつきが小さければ小さいほど推奨は強いものとなる</p>	<p>→ グレードダウン5要因のうちの、特に imprecision, inconsistencyに相当</p>
<p>3. 臨床的有効性の大きさ</p>	<p>→ グレードアップ3要因の magnitude of effectに相当</p>
<p>4. 臨床上の適用性</p> <p>(医師の能力、地域性、医療資源、保険制度)</p>	<p>→ GRADEの value and preference (一部はグレードダウンの indirectness)に相当</p>
<p>5. 害やコストに関するエビデンス</p>	<p>→ GRADEと同じである</p>

このように、利益と不利益(害)を考慮したアウトカム全般に関するエビデンスの考え方の違いがあるだけで、基本的にはMinds2007とGRADEは大きな違いはない

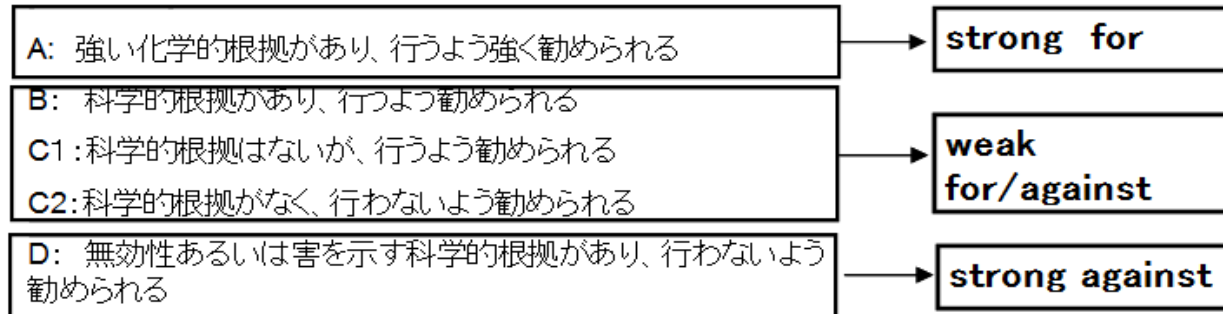
CPG作成ツール: Minds 2007からGRADEへ

推奨グレード

GRADEシステムにおいては
エビデンスの質は推奨グレードの評価要因の1つにすぎない
(エビデンスの質と推奨を切り離しているのがGRADEの特徴である)

推奨度	表記	
	記号	数字
強い推奨	↑ ↑あるいは ↓ ↓	1
弱い推奨	↑ ? あるいは ↓ ?	2

Minds2007



- ・利益／不利益バランスによっては、非常に弱いエビデンスの質に基づく強い推奨もありえる
- ・推奨の分類としては、強さ(強い、弱い)と方向性(～を使用する、使用しない)を組み合わせた4種類であり、基本的な違いはない

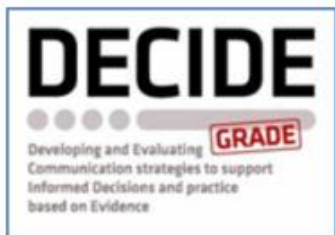
まとめ

Minds2007基準には個別研究のエビデンスを評価するという簡便さがある。しかし、WHO, Cochrane Collaboration, UpToDate, Clinical Evidence, CDC, IDSA, ACCPなどがGRADEアプローチを採用し、GINとGRADEがガイドライン作成・普及に協調して活動している現状では、body of evidenceの質評価と利益・不利益バランスによる推奨決定というGRADEアプローチを国内基準に取り込むことが望ましい。

Minds2007基準によるガイドラインにGRADEを取り込む際は、

1. 既存のSRを利用するか、もしくはアウトカムを主体としたSRを作成し、エビデンステーブル(evidence profileあるいはSummary of findingテーブル)を準備する。
2. 各アウトカムの総体エビデンスを評価し、さらにアウトカム全般にわたる全体的なエビデンスの質(overall quality of evidence across outcomes)をひとつ決定する。
3. 推奨の強さの判定には、これまでと同様に害やコストに関するエビデンスを考慮するが、推奨の強さとエビデンスのレベル評価を分離する必要がある。
4. 推奨の強さは2種類(強い、弱い)、エビデンスのレベルは4段階(高、中、低、非常に低)が基本である。
5. エビデンスの質や推奨の強さの決定のプロセスを透明化する。

GRADEガイドライン作成・・・次は？



DECIDE研究プロジェクトの発足

[Developing and Evaluating Communication Strategies to Support Informed Decisions and Practices Based on Evidence.](http://www.decide-collaboration.eu/welcome)

計画目標:

「対象を絞ったガイドライン配布のための方法を開発・評価するために、GRADE Working Groupの取り組みを土台とし、エビデンスに基づく推奨の配布を改善する」

方法:

GRADEは、エビデンスの質と推奨の強さを評価し伝達するための系統的アプローチである。GRADEは、そもそも他の評価システムの弱点に対処するために開発されたものだが、現在では国際的に広く使用されている。GRADE Working Groupのメンバーにより結成されるDECIDE協会 (DECIDE condortium) は、このアプローチをさらに発展させることで、診療の行方を左右する主な利害関係者 (医療専門家、政策決定者と管理責任者、患者と一般市民) に対し、エビデンスに基づく推奨を効果的に配布することを確実にする。われわれは、諮問グループ、審議、ユーザー調査を通じて利害関係者からのインプットを集める。この取り組みは、欧州の幅広い医療システムにわたって行われる。開発された対象を絞った配布戦略はランダム化試験で評価され、DECIDEパートナーやわれわれが支持する他のガイドライン作成者らにより開発された実際のガイドラインを使用して改良・適用・評価される。

期待される結果:

多様なセッティングで厳密に評価された推奨配布戦略は、研究の診療への反映を支援し、実際の医療システムに適用される。

<http://www.decide-collaboration.eu/welcome>を翻訳