

表1 AGREE 新旧対象表

領域	AGREE II		AGREE I		備考
	質問番号	質問	質問番号	質問	
対象と目的	1	ガイドライン全体の目的が具体的に記載されている。	1	ガイドライン全体の目的が具体的に記載されている。	同じ
	2	ガイドラインで取り扱う健康上の問題が具体的に記載されている。	2	ガイドラインで取り扱う臨床上の問題が具体的に記載されている。	マイナーチェンジ
	3	どのような集団（患者、市民など）を対象としたガイドラインであるかが具体的に記載されている。	3	どのような患者を対象としたガイドラインであるかが具体的に記載されている。	マイナーチェンジ
利害関係者の参加	4	ガイドライン作成グループには、関係する全ての専門家グループの代表者が加わっている。	4	ガイドライン作成グループには、関係する全ての専門家グループの代表者が加わっている。	同じ
	5	対象集団（患者、市民など）の価値観や好みが十分に考慮されている。	5	患者の価値観や好みが十分に考慮されている。	マイナーチェンジ
	6	ガイドラインの利用者が明確に定義されている。	6	ガイドラインの利用者が明確に定義されている。	同じ
	7		7	ガイドラインの想定する利用者で既に試行されたことがある。	削除し、19に包含
作成の厳密さ	7	エビデンスを検索するために体系的な方法が用いられている。	8	エビデンスを検索するために体系的な方法が用いられている。	同じ
	8	エビデンスの選択基準が明確に記載されている。	9	エビデンスの選択基準が明確に記載されている。	同じ
	9	エビデンスの強みと限界が明確に記載されている。			新規
	10	推奨を決定する方法が明確に記載されている。	10	推奨を決定する方法が明確に記載されている。	同じ

	11	推奨の決定にあたって、健康上の利益、副作用、リスクが考慮されている。	11	推奨の決定にあたって、健康上の利益、副作用、リスクが考慮されている。	同じ
	12	推奨とそれを支持するエビデンスとの対応関係が明確である。	12	推奨とそれを支持するエビデンスとの対応関係が明確である。	同じ
	13	ガイドラインの公表に先立って、外部審査がなされている。	13	ガイドラインの公表に先立って、外部審査がなされている。	同じ
	14	ガイドラインの改訂手続きが予定されている。	14	ガイドラインの改訂手続きが予定されている。	同じ
明確さと提示の仕方	15	推奨が具体的であり、曖昧でない。	15	推奨が具体的であり、曖昧でない。	同じ
	16	患者の状態や健康上の問題に応じて、可能な他の選択肢が明確に示されている。	16	患者の状態に応じて、可能な他の選択肢が明確に示されている。	マイナーチェンジ
	17	どれが重要な推奨か容易に見分けられる。	17	どれが重要な推奨か容易に見分けられる。	同じ
			18	利用のためのツールが用意されている。	19に移行、領域も「明確さと提示の仕方」から「適用可能性」に移行
適用可能性	18	ガイドラインは、利用の促進および障壁を記載している。	19	推奨の適用にあたって予想される制度・組織上の障壁が論じられている。	マイナーチェンジ
	19	どのように推奨を適応するかについての助言、ツールを提供している。			
	20	推奨の適応に伴う付加的な資源が考慮されている。	20	推奨の適用に伴う付加的な費用（資源）が考慮されている。	マイナーチェンジ
	21	ガイドラインにモニタリング、監査のための基準が示されている。	21	ガイドラインにモニタリング・監査のための主要な基準が示されている。	マイナーチェンジ
編集の独立性	22	資金源によりガイドラインの内容が影響されていない。	22	ガイドラインは編集に関して資金源から独立している。	マイナーチェンジ
	23	ガイドライン作成グループのメンバーの利害の衝突が記載さ	23	ガイドライン作成グループの利害の衝突が記載されている。	マイナーチェンジ

		れ、適切な対応がなされている。		
全体		このガイドラインの全体の質の 評価。		新規
評価		このガイドラインを推奨します か。	このガイドラインを推奨します か。	同じ

AGREE instrument を用いた日本語診療ガイドラインの評価

研究代表者	長谷川 友紀	東邦大学医学部社会医学講座
研究協力者	山口 直比古	東邦大学医学メディアセンター
研究協力者	平輪 麻里子	東邦大学医学メディアセンター
研究分担者	小泉 俊三	佐賀大学医学部総合診療部
研究分担者	葛西 龍樹	福島県立医科大学家庭医療学
研究分担者	平尾 智宏	香川大学医学部医療管理学
研究協力者	藤田 茂	東邦大学医学部社会医学講座
研究協力者	城川 美佳	東邦大学医学部社会医学講座
研究協力者	北澤 健文	東邦大学医学部社会医学講座
研究協力者	瀬戸 加奈子	東邦大学医学部社会医学講座

研究要旨

診療ガイドラインは、毎年約 20-30 本が作成されるなど日本においても定着しつつある。本研究では、日本において発行された診療ガイドラインを AGREE instrument 日本語版を用いて評価を行い、経年及び改訂による AGREE 評価の結果の比較を行った。診療ガイドラインは経年変化及び改訂により質が向上している傾向が示された。今後も、診療ガイドラインの評価を継続して行うとともに、評価者間で評価の分かれた項目、低い評価の項目については、原因を明らかにするとともに、作成者を対象にした支援が検討される必要がある。

った。

A. 研究目的

日本における診療ガイドラインの作成は 2000 年よりはじまり、年間約 20 本が新規もしくは改訂により作成されている。本研究では、2009 年に日本で公表された 33 診療ガイドラインを対象に、(1) AGREE instrument を用いて評価を行い、経時比較を行った。(2) AGREE I と AGREE II では、評価がそれぞれ 4 段階、7 段階と異なることから、AGREE I の各項目について 4 段階、7 段階評価を行い、両者の結果の整合性について検討を行った。(3) 昨年度開発した患者の視点からの評価項目について試行調査を行

B. 研究方法

現行の AGREE instrument を用いて 3 人の専門家に個別に 2009 年に発行された 33 診療ガイドラインの評価を実施してもらった。3 人が評価した評点の結果より、経年及び改訂前後等の推移について検討を行った。

(倫理面への配慮)

本研究は、発行された診療ガイドラインを AGREE instrument を用いて評価したものであり、倫理面の配慮が必要なデータ等は取り扱っていない。

C. 研究結果

(1) 2009年発行の33診療ガイドラインの評価

2009年は新刊及び改訂も含めて33の診療ガイドラインが発行されこれらについてAGREE instrumentを用いて評価を行った。評価結果を図1に示す。

全体の評価結果は、最近のものになるにつれて質が向上していることがうかがえた。領域別では「対象と目的」は当初から高い評価(90%以上:最高評価=100%)を受けていたが、「利害関係者の参加」「作成の厳密さ」「明確さと提示の仕方」の3領域で評価の向上(60-80%)が継続して認められた。2008年以降、「編集の独立性」が顕著に伸びており、これはCOIなどへの関心の増加が背景にあると考えられる。「適応可能性」については、上昇が認められるものの依然として改善の余地が大きい(40%以下)。

(2) AGREE Iを用いた4段階、7段階評価の比較

現在AGREE IIの開発が進められている。AGREE IIでは、質問の多くは継承されるが、重要な変化として各質問項目への回答が4段階から7段階に変更がなされる。AGREE Iの質問項目を用いて、4段階評価と7段階評価の相関をSpearmanの順位相関係数を用いて表1に示す。すべての項目において、高い相関が認められた。

(3) 患者の視点を反映した追加項目

昨年度の研究において、患者の視点を反映した質問項目として、以下の6項目を開発した。

- ・文章以外のチャートやパスなどが分かりやすく活用されている。
- ・疾患毎に適切な時期に診療ガイドラインの改訂が行われている。
- ・診療ガイドラインの作成及び評価に患者が参加している。
- ・患者の知りたい治療に関する情報が具体的に記載されている。
- ・診療ガイドラインで推奨されている治療の予後や起こりうる合併症・後遺症等のリスクに関する説明が具体的に記載されている。
- ・退院後、患者が病気と共に質の高い豊かな生活を送れるよう、日常生活上のアドバイスなどの情報源の所在が具体的に記載されている。

これらの項目の有用性を検討するために(2)と同様の調査を実施した。4段階評価と7段階評価の結果では、すべての項目において、高い相関が認められた。

D. 考察

(1) 診療ガイドラインの経年変化

日本において診療ガイドラインが定着しつつあり、またその作成手順についても浸透してきたものと考えられる。特に、AGREE instrumentのような診療ガイドラインを評価するツールを公表するこ

とにより、良質の診療ガイドラインをより効率的に作成することを円滑にしている可能性がある。

質的には全体的に向上しているものの、項目によってはまだ十分とは言えないものも見受けられる。今後、評価が低い傾向のある項目について診療ガイドラインにどのような内容を明示する必要があるのか具体的な例示を作成する等により、作成者を支援し、質を向上させるなど検討されることが望ましい。

(2) AGREE II の検証

現在開発が進められている AGREE II について、質問の評価段階の変更による影響は少ないことが示された。しかしデータの整合性については更なる検討が必要である。

(3) 患者用の視点を反映した項目

試行では、特に問題を生じることなく評価の実施が可能であり、4段階、7段階の結果は高い相関を示した。これらの項目の有用性については更なる検討が必要である。

E. 結論

診療ガイドラインは、毎年約 20-30 作成されるなど日本においても定着しつつある。診療ガイドラインは経年変化及び改訂により質が向上していることが示された。AGREE instrument の公表は、診

療ガイドラインの質向上に寄与している可能性がある。今後も、診療ガイドラインの評価を行うとともに、評価が低い傾向のある項目については、どのような内容が不足しているのかについて具体的な明示を行うなど、改善促進策が検討される必要がある。

F. 研究発表

1. 論文発表
なし
2. 学会発表
なし

G. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得
なし
2. 実用新案登録
なし
3. その他

AGREE instrument の日本語翻訳版（暫定版）は、平成 14 年度厚生労働科学研究費補助金「医療技術総合研究事業（診療ガイドラインの評価に関する研究）」（主任研究者：長谷川友紀）の一環として作成され、東邦大学医学メディアセンターより PDF ファイルにてダウンロードすることができる。

(<http://www.mnc.toho-u.ac.jp/mmc/guideline/AGREE-final.pdf>)

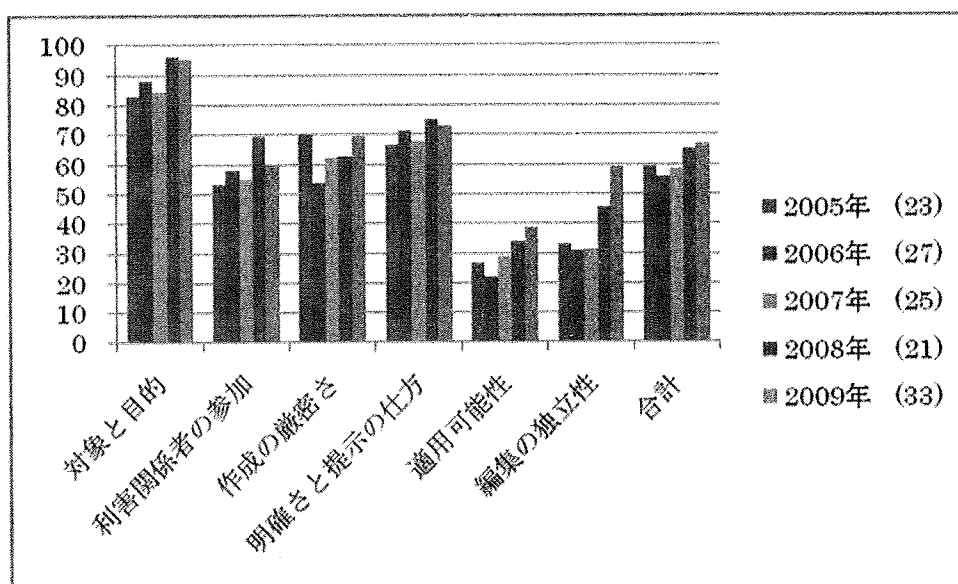


図1 AGREE instrument を用いた診療ガイドラインの評価

表1 AGREE I を用いた 4 段階評価と 7 段階評価の比較

	Spearman の順位相 関係数	p
<u>対象と目的</u>		
1 ガイドライン全体の目的が具体的に記載されている。	0.901	**
2 ガイドラインで取り扱う臨床上的の問題が具体的に記載されている。	0.935	**
3 どのような患者を対象としたガイドラインであるかが具体的に記載されている。	0.951	**
<u>利害関係者の参加</u>		
4 ガイドライン作成グループには、関係する全ての専門家グループの代表者が加わっている。	0.912	**
5 患者の価値観や好みが十分に考慮されている。	0.981	**
6 ガイドラインの利用者が明確に定義されている。	0.948	**
7 ガイドラインの想定する利用者で既に試行されたことがある。	0.990	**
<u>作成の厳密さ</u>		
8 エビデンスを検索するために系統的な方法が用いられている。	0.982	**
9 エビデンスの選択基準が明確に記載されている。	0.954	**

10	推奨を決定する方法が明確に記載されている。	0.915	**
11	推奨の決定にあたって、健康上の利益、副作用、リスクが考慮されている。	0.969	**
12	推奨とそれを支持するエビデンスとの対応関係が明確である。	0.974	**
13	ガイドラインの公表に先立って、外部審査がなされている。	0.950	**
14	ガイドラインの改訂手続きが予定されている。	0.946	**

明確さと提示の仕方

15	推奨が具体的であり、曖昧でない。	0.877	**
16	患者の状態に応じて、可能な他の選択肢が明確に示されている。	0.957	**
17	どれが重要な推奨か容易に見分けられる。	0.974	**
18	利用のためのツールが用意されている。	0.992	**

適用可能性

19	推奨の適用にあたって予想される制度・組織上の障碍が論じられている。	0.993	**
20	推奨の適用に伴う付加的な費用（資源）が考慮されている。	0.979	**
21	ガイドラインにモニタリング・監査のための主要な基準が示されている。	0.929	**

編集の独立性

22	ガイドラインは編集に関して資金源から独立している。	0.940	**
23	ガイドライン作成グループの利害の衝突が記載されている。	0.964	**

患者の視点

1	文章以外のチャートやパスなどが分かりやすく活用されている。	0.946	**
2	疾患毎に適切な時期に診療ガイドラインの改訂が行われている。	0.978	**
3	診療ガイドラインの作成及び評価に患者が参加している。	0.999	**
4	患者の知りたい治療に関する情報が具体的に記載されている。	0.855	**
5	診療ガイドラインで推奨されている治療の予後や起こりうる合併症・後遺症等のリスクに関する説明が具体的に記載されている。	0.961	**
6	退院後、患者が病気と共に質の高い豊かな生活を送れるよう、日常生活上のアドバイスなどの情報源の所在が具体的に記載されている。	0.988	**

** p<0.001

2009年発行診療ガイドライン改訂版の主な変更点に関する研究

研究代表者	長谷川 友紀	東邦大学医学部社会医学講座
研究協力者	山口 直比古	東邦大学医学メディアセンター
研究協力者	平輪 麻里子	東邦大学医学メディアセンター
研究協力者	瀬戸 加奈子	東邦大学医学部社会医学講座

研究要旨

2009年に刊行された11診療ガイドラインについて改訂前後の比較検討を行った。改訂版では、①新たなエビデンスが加えられた、②新たに使用可能（保険適用）となった薬剤等についての使用指針が加えられた、③章立てやその順番などが整理された、④ページ数増加の傾向が見られた、⑤新たなエビデンスにより、推奨度の変更が、若干ではあるが見られた、⑥今後の改訂についても予定されている、等の特徴が認められた。診療ガイドラインが最新の知見を反映した形でタイムリーに改訂されるためには、作成組織、その支援体制を含めた検討が、今後必要であると考えられる。

A. 研究目的

診療ガイドラインは3年～5年毎に改訂されることが望ましく、昨年あたりから、各学会作成の診療ガイドラインは改訂の時期に入ってきている。新旧診療ガイドラインの比較検討を行い、おもな改訂点とその理由を明らかにした。

B. 研究方法

2009年に刊行された診療ガイドラインで、今回AGREEによる評価の対象とした34点のガイドラインのうち、15点が改訂版であった。この中から、旧版との比較ができた11点について検討した。なお、改訂の要点についての報告が、学会誌などに報告されているので、それら

の文献も参照した。

C. 研究結果

改訂版の診療ガイドラインに見られる共通の特徴は以下のとおりであった。

- ・追加の文献検索による新たなエビデンスが加えられた
- ・新たに使用可能（保険適用）となった薬剤等についての使用指針が加えられた
- ・項目や並び順などの見直しが行われ、章立てやその順番などが整理された
- ・冊子としての体裁にはほとんど変化はないが、ページ数増加の傾向が見られた
- ・文献検索による新たなエビデンスにより、推奨度の変更が、若干ではあるが見られた
- ・今後の改訂についても予定されている

以下、各ガイドラインの改訂内容の特徴などについて報告する。

高血圧治療ガイドライン		
版表示	2004年版	2009年版
判型、頁	A4版 131頁	A4版 192頁
章、CQの数	11、CQ形式ではない	12、CQ形式ではない
推奨度の変更	推奨ではない	推奨ではない
主な変更点		<p>日本で CASE-J など複数の大規模臨床試験が実施されたのを受けて</p> <ul style="list-style-type: none"> ・「女性の高血圧」の章を新設 ・高血圧分類を軽症・中等症・重症より、I度・II度・III度変更 ・嚴重な高血圧目標を 130/80mmHg とする ・高齢者は 140/90mmHg とする ・家庭血圧の重要性 ・RA系阻害薬を第一選択薬、利尿剤、Ca拮抗薬を第二選択薬に ・複数の降圧剤の併用

小児気管支喘息治療・管理ガイドライン		
版表示	2002年版	2009年版
判型、頁	B5版 223頁	B5版 230頁
章、CQの数	16、CQ形式ではない	9、CQ形式ではない
推奨度の変更		<ul style="list-style-type: none"> ・章立てを大幅に変更した ・引用文献にエビデンスレベル（A-D）を付記した ・治療ステップを重症度ではなく、治療内容の強弱に沿ったものとし、国際ガイドラインの代表であるGINAに倣った ・治療にあたっては、無症状をコントロールとした
主な変更点		<ul style="list-style-type: none"> ・高齢者喘息の項目を充実させた ・新薬についての記述を加えた

COPD（慢性閉塞性肺疾患）診断と治療のためのガイドライン		
版表示	第2版	第3版
判型、頁	A4版 137頁	A4版 173頁
章、CQの数	4、CQではなくPOINT	5、CQではなくPoints
推奨度の変更	推奨はPOINT	推奨はPOINTS
主な変更点		<ul style="list-style-type: none"> ・疾患の重症度分類について、呼吸機能の障害程度のみで分類した病期と、症状や全身状態を加味してとらえる重症度の考え方を明示 ・引用文献にエビデンスレベル（A-D）を付記した ・65歳以上のCOPD患者には、肺炎球菌ワクチンの接種が推奨、%FEV40%未満の患者にも推奨 ・推奨度が下がっている部分あり

急性膵炎診療ガイドライン		
版表示	第2版	2010（第3版）
判型、頁	A4版 131頁	A4版 160頁
章、CQの数	8、65	10、74
推奨度の変更		CQ43（旧版CQ37）で後半が推奨度AからBへ下げられている、などがある。根拠となるRCTの引用が旧版の15件より20件へと増加して引用されている。
主な変更点		<ul style="list-style-type: none"> ・重症度分類に中等症を用いず、軽症と重症のみとした ・ERCP後膵炎予防のためのステンドの留置、NSAIDsの予防的投与を推奨 ・感染性膵炎壊死はインターベンション治療（手術、IVRなど）適応

大腸癌治療ガイドライン 医師用		
版表示	2005年版	2009年版
判型、頁	B4版 54頁	B4版 86頁
章、CQの数	2、CQの形式ではない	2、CQ19
推奨度の変更	記述的で推奨度はない	推奨度（A-D）を設定
主な変更点		<ul style="list-style-type: none"> ・クリニカルクエスションの方式を採用、推奨度も設定 ・各論とCQとを分けた ・文献のエビデンスレベルを意識して文献検索を行った ・内視鏡治療後の追加手術について記載 ・腹腔鏡下手術は適用を拡大 ・化学療法のレジメンを多様化 ・術後サーベイランスはCTが望ましいとした

科学的根拠に基づく肝癌診療ガイドライン		
版表示	2005年版	2009年版
判型、頁	B5版 169頁	B5版 179頁
章、CQの数	6、CQ58	6、CQ51
推奨度の変更		推奨度のグレードは同一
主な変更点		<ul style="list-style-type: none"> ・改訂のないCQは2件のみで、47CQを改訂、7CQを新設し、全体を51CQに統合整理した ・複数のRCTにより、RFAをPEIより優れたものとして推奨 ・Sorafenibの有効性についてのRCTが出、また保険適用となったため、治療薬として推奨

鼻アレルギー診療ガイドライン		
版表示	2005年版（改訂5版）	2009年版（改訂第6版）
判型、頁	B5版 103頁	B5版 115頁
章、CQの数	6、CQは無し	6、CQは無し
推奨度の変更	推奨度を設定していない	推奨度を設定していない

主な変更点		<ul style="list-style-type: none"> ・第5章治療に舌下免疫療法が追加された ・第6章その他に専門医への紹介が追加された
-------	--	------------------------------------------------------------------------------------------------------

科学的根拠に基づく膀胱診療ガイドライン		
版表示	2006年版	2009年版
判型、頁	B5版 100頁	B5版 151頁
章、CQの数	5, 22CQ	5, 25CQ
推奨度の変更		<ul style="list-style-type: none"> ・2項目で推奨度があげられた 1 局所切除不能癌に対する化学療法単独治療が推奨度CからBへ 2 術後補助化学療法に塩酸ゲムシタビンの利用がCからBへ
主な変更点		<ul style="list-style-type: none"> ・推奨度CをC1、C2に分け、Mindsのグレードを採用した ・CQの数が増加し、部分的に文言を修正した

子宮体がん治療ガイドライン		
版表示	2006年版	2009年版
判型、頁	B5版 127頁	B5版 183頁
章、CQの数	8, 41CQ	10, 51CQ
推奨度の変更		Mindsの推奨度を採用し、CをC1とC2とした。これにより、旧版の推奨グレードDがC2となったものがある
主な変更点		<ul style="list-style-type: none"> ・2章追加し、CQの数も増加。 追加したのは「肉腫」など ・名称を「子宮体癌」より「子宮体がん」とした。これは間葉系悪性腫瘍も対象としたため。 ・診断、治療のアルゴリズムを示すフローチャートを新設した

小児中耳炎診療ガイドライン		
版表示	2006年版	2009年版
判型、頁	A4版 59頁	A4版 75頁及び図表
章、CQの数	20、11CQ	20、12CQ
推奨度の変更	推奨度はAHRQを使用	推奨度表は変更なし
主な変更点		<ul style="list-style-type: none"> ・重症度スコアを一部修正1 ・文献検索によりエビデンスレベルの高い文献を追加

脳卒中治療ガイドライン		
版表示	2004	2009
判型、頁	A4版 234頁	A4版 358頁
章、CQの数	5、CQ形式ではないが推奨あり	7、CQ形式ではないが推奨あり
推奨度の変更	推奨度あり	推奨度表の文言に修正あり
主な変更点		<ul style="list-style-type: none"> ・脳梗塞にTIA、無症候性脳血管障害を追加 ・メタボリックシンドロームや睡眠時無呼吸症候群についても記載 ・薬物療法で、tPA やクロピドグレル、シロスタゾール（保険適用された）を推奨（推奨度に変更は無い）

D. 考察 と E. 結論

AGREE による領域別、総合評価とは別に個々の診療ガイドラインの新旧版の比較検討を行った。診療ガイドラインの望ましい改訂時期については、新たな知見が得られた、(保険収載など)制度変更が行われた時期とされるが、これは対象疾患により異なり一律に規定することは困難である。例えば、悪性腫瘍の化学療法はその開発スピードが早く、最新の情報入手、治療機会を得るために治験に参加することを推奨する診療ガイドラインもある。また、最近では診療ガイドラインの多くは関連学会によ

り作成されるが、診療ガイドラインに関わる常設委員会を有するか否か、診療ガイドライン作成・維持に関わる費用をどのような形で賄うかにより機動的に改訂作業を行うことができるか否かは変わってくる。一般には、常設委員会を有している場合、当該診療ガイドライン改訂の経費がその販売により賄われる場合には、患者数の多い疾患ほど、改訂作業は機動的に実施することが可能になる。

また、新しい知見を反映した診療ガイドラインが実効性を有するには、その内容の周知方法についても検討される必要がある。

診療ガイドラインの重要性を鑑みると、作成

組織、その支援体制を含めた検討が、今後必要
であると考えられる。

F. 研究発表

1. 論文発表
なし
2. 学会発表
なし

G. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

診療ガイドラインの GRADE システムについて

分担研究者 平尾 智広 香川大学医学部医療管理学

研究要旨

現在、多くの診療ガイドラインが整備されているが、提示されるエビデンスの質と推奨の強さについて、多数の判断基準が存在し、一貫性があるとはいえない。またレベルやグレードの標記方法もガイドライン作成者に委ねられており、混乱の要因となっている。GRADE (Grading of Recommendations Assessment, Development and Evaluation) システムは、このような状況を整理し改善するために開発された。Cochrane collaboration、WHO 等の有力機関で採用されており、診療ガイドライン作成の標準的手法となりつつある。我が国における GRADE システムの導入は他の先進国に比べて遅れており、普及を促進すべきと考える。しかし導入について問題点や限界が指摘されており、トレーニングや普及・啓発を行う体制整備が必要である。

A. 研究目的

多くの診療ガイドラインが整備されており、根拠に基づいたエビデンスと推奨が提示されている。臨床家にとって大変有用なことであるが、提示される“エビデンスの質”と“推奨の強さ”について、多数の判断基準が存在し、一貫性があるとはいえない。またレベルやグレードの標記方法もガイドライン作成者に委ねられており、混乱の要因となっている。

GRADE (Grading of Recommendations Assessment, Development and Evaluation) システムは、このような状況を整理し改善するために、Gordon H Guyatt、Andrew D Oxmanらが中心となって開発されたものである。Cochrane collaboration、WHO、AHRQ等の有力機関でガイドライン作成の手法として採用されており、他の団体にも普及が進んでいる。本稿では GRADE システムの概要について述べ、わが国での導入について考察を行った。

B. 研究方法

GRADE システムに関する情報は、GRADE Working Group のウェブサイト、文献、及び診療ガイドラインの国際的ネットワークである G-I-N での調査によった。

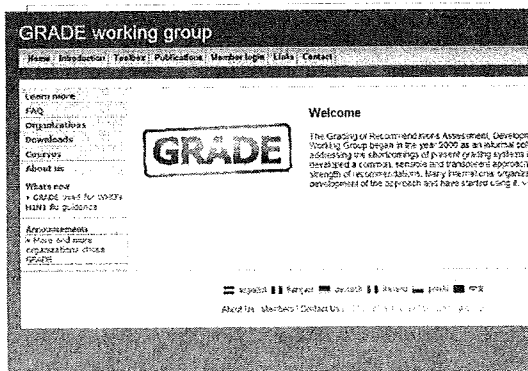
・GRADE Working Group

2000 年よりスタートした国際共同グループで、GRADE システムの開発、普及、ツールの提供を行っている。

(<http://www.gradeworkinggroup.org/index.htm>)

・G-I-N

診療ガイドラインの作成や使用に関わる個人や団体から構成されている。



C. 研究結果

(1) GRADE システムの特徴

GRADE システムは、エビデンスの質及び推奨を分類するための方法で、診療ガイドラインの作成のほか、系統的レビューや医療技術評価にも有用とされている。その特徴として、

- ・ 世界の代表的なガイドライン作成グループが開発している
- ・ エビデンスの質と推奨のレベルが明確に区別されている
- ・ 他の代替的治療法についても明確に評価している
- ・ エビデンスの質のアップグレード、ダウングレード基準が明確かつ包括的である
- ・ エビデンスから推奨の過程の透明性が担保されている

ことが、あげられる。

・ エビデンスの質

従来の考え方では、エビデンスの質は研究デザインにより領域化されており、Randomized Controlled Trial (RCT) を頂点とする階層構造が一般的である。しかし、①分類がシンプルすぎること、②サンプルサイズの小さい、質の高くない RCT も最上級のエビデンスとして位

置づけられる傾向にあること、③ 結果が異なる複数の RCT も最上級のエビデンスとして位置づけられる傾向にあること、④強い説得力があっても、観察研究は最上級のエビデンスとして扱われないこと、等の問題が指摘されていた。GRADE システムではエビデンスの質を、High、Moderate、Low、Very Low の4つに分類し、それぞれ High...今後の研究成果により変更の可能性は低い、Moderate...変更の可能性はある、Low...変更の可能性が高い、Very low...効果が不確実、とされている。

エビデンスの質のレベルについては、RCT は high からスタート、観察研究は low からスタートし以下の事項を勘案しながら決定される。

➤ 判定を下げる要因

- ・ 方法、デザインに問題（限界）がある
- ・ 結果に一貫性に欠く部分がある
- ・ 結果が間接的である
- ・ データが不正確である
- ・ 出版バイアスがある

➤ 判定を上げる要因

- ・ 効果が大きい
- ・ 交絡要因による過小評価の可能性はある
- ・ 量反応関係が見られる

・ 推奨のレベル

GRADE システムでは推奨のレベルを Strong と Weak の二つに分類しており、患者、医療従事者、政策決定者に明確な方向性を示すことができるとしている。

Strong：多くの患者が推奨を希望する。希望しない者はわずかである。

Weak：多くの患者が推奨を希望するが、希望しない者も多数いる。臨床医は患者の価値観や

好みによる選択も尊重する。

推奨レベルの判定に影響するのは、効果と副作用のバランス、エビデンスの質、価値観と選好、コスト（資源の分配）である。

効果と副作用のバランス：効果が副作用に比べて大きいほど強い推奨となり、小さいほど弱い推奨となる可能性が高い。

エビデンスの質：質が高いほど強い推奨となる可能性が高い。

価値観と選好：価値観や選好が多様であるほど、不確定の部分が大きいほど、弱い推奨となる可能性が高い。

コスト（資源の分配）：コスト高であるほど、資源の消費が大きいほど、強い推奨となる可能性は低い。

Quality of evidence	
High quality	⊕⊕⊕⊕ or A
Moderate quality	⊕⊕⊕○ or B
Low quality	⊕⊕○○ or C
Very low quality	⊕○○○ or D

Strength of recommendation	
Strong recommendation for using an intervention	↑↑ or 1
Weak recommendation for using an intervention	↑? or 2
Weak recommendation against using an intervention	↓? or 2
Strong recommendation against using an intervention	↓↓ or 1

参考文献 2 より引用

(2) GRADE システムの導入状況

Gunn Elisabeth VIST らは、2008 年に GRADE システムの導入状況について調査している。それによれば、“GRADE”、“recommendation”、“guideline”の用語を用いて 2004 年以降の状況について National Guideline Clearinghouse に収載されたガイドラインの検索を行ったところ 592 タイトルが検出され、ガイドラインを作成した 280 団体のうち 5 団体者は GRADE システムを用いたことの記載があり、他の 5 団体では類似のシス

テムが用いられていた。しかし 35%の団体で 2004 年以降の新規ガイドライン作成がなく、また 2004 年以降に新たなガイドラインを作成した団体のうち 29%は、GRADE システムを用いていないことが記載されていたとしている。

2010 年の時点で GRADE システム採用団体として 45 の団体が GRADE Working Group のウェブページに紹介されており、主要なガイドライン作成機関、雑誌において GRADE システムの導入が進んでいることが伺われる。最近では、WHO が 2009 年に新型インフルエンザ(H1N1)のガイドラインを作成する際に GRADE システムが用いられた。

我が国の状況について Minds に掲載されている 62 のガイドラインを調査したところ、GRADE システムを導入、あるいは参照しているとの記載があったものは 3 件であった。それぞれ 2006 年、2008 年、2009 年に発行されたもので、わずかながらも増加傾向にある。

D. 考察

(1) GRADE システム導入の問題点と限界

GRADE システムを導入することにより、エビデンスから推奨を作成する手続きが明確となり、透明性が確保されるとされている。しかし、いくつかの限界も指摘されている。例えば、ガイドラインの作成に関わる全員が GRADE システムを熟知していなければならないこと、有害事象等の負の要因がある場合のダウングレード、すなわち望ましい効果と望ましくない作用が同居する場合の判定が容易ではないこと等が挙げられている。

また GRADE システムは、エビデンスの質、推奨のレベルを決定するプロセスを示したも

ので、Clinical Question の設定、Evidence の
收拾等の作業は残る。さらに作業者間の意見調
整が簡素になるわけでもない。

GRADE システム導入の問題点

- ・新システムに対する抵抗感の克服
- ・作業者のトレーニングの問題
- ・作業量の増加
- ・作業者の増加（関わる人数が増える）
- ・作業ステップが増える
- ・導入効果が明確でない

(2) 我が国での導入について

我が国の診療ガイドライン作成における
GRADE システムの導入は、他の欧米諸国に比
べて進んではない。世界の有力団体が相次い
で採用し、GRADE システムが診療ガイドラ
イン作成の標準手法となりつつあり現在、我が国
においても導入を検討する必要があるだろう。

ただし、GRADE システムの導入には、前述
したような問題もあり、トレーニングや普及・
啓発のための体制整備が必要である。

E. 結論

- ・ GRADE システムは、エビデンスの質及び
推奨を分類するための方法で、世界の有力団
体が採用し、診療ガイドライン作成の標準的
手法となりつつある。
- ・ 我が国における GRADE システムの導入
は、他の先進国に比べて遅れており、普及を
進めるべきである。
- ・ 導入について問題点や限界が指摘されて
おり、トレーニングや普及・啓発のための体
制整備が必要である。

F. 研究発表

1. 論文発表
なし
2. 学会発表
なし

G. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

参考文献

1. Guyatt GH, Oxman AD, Vist G, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group. Rating quality of evidence and strength of recommendations GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-926
2. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ; GRADE Working Group. Rating quality of evidence and strength of recommendations: What is "quality of evidence" and why is it important to clinicians? *BMJ*. 2008 May 3;336(7651):995-8
3. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ; GRADE Working Group. Rating quality of evidence and

- strength of recommendations: Going from evidence to recommendations. *BMJ*. 2008 336(7652):1049-51
4. 相原守夫, Gordon Guyatt, Yngve Falck-Ytter. GRADE システムと SoF. *Therapeutic Research*. 2007 28:2335-2354
5. Gunn Elisabeth VIST, et al. Use of GRADE when making recommendations. *ACTA MÉDICA PORTUGUESA* 2009;22:682-683