

図2 推奨度 C の考え方・対応

1. 文献を採用あるいは不採用とする条件の明示

癌種によって情報の多少やエビデンスレベルの高・低、あるいは個々の question に対する情報の多少が生じる。したがって採用・不採用の条件はガイドラインによって異なり、そのため一定の規準を明示しなければならない。

2. 採用文献の批判的吟味の実施

文献内容の要約記載による紹介を必須とした構造化抄録の作成によって、分析にバイアスが加わって誤った結論が導き出されていないかを判断する。また、引用の場合にはガイドラインに添付する。特に構造化抄録の作成者は文献内容の妥当性と質の評価について抄録内容に対するコメントを加えることを条件とする。

3. 採用した文献についての一覧表 (abstract table) の作成

論文コード (著者名, 発表年), 対象患者の概要 (患者数, 年齢, 重症度, 診療施設の内訳など),

患者群の診療内容の内訳, 結果 (outcome など) について randomized controlled trial (RCT), case control study, 相関研究などの研究手法別にわかりやすく配列する。わが国での癌診療ガイドラインにおいては abstract table の作成がほとんどなされていないのが現状である。

4. Clinical question に対する推奨内容のエビデンスレベルの分類と決定

エビデンスレベル決定については, (1) 実験研究の結論は観察研究のそれより, (2) ランダム化比較試験の結論は非ランダム化比較試験のそれより, (3) 分析疫学的研究の結論は専門家個人あるいは専門家組織のそれより真実を反映している可能性が高いとされている⁵⁾。同一あるいは類似したエビデンスが存在する場合には, より質の高いエビデンスを採用することとなるが, たとえ質が若干低くても特記すべき注釈を付して採用することもある。

表4 AGREE instrument の観点と項目

観点	項目	AGREE 評点
対象と目的	1. 目的の具体的な記載	
	2. 取り扱う問題の具体的な記載	
	3. 対象患者の具体的な記載	
利害関係者の参加	4. 専門家代表者の作成への参加	
	5. 患者の価値観や好みへの考慮	
	6. 利用者の明確な定義	
	7. 想定利用者による試行	
作成の厳格さ	8. エビデンスの系統的な検索	
	9. エビデンスの選択基準の記載	
	10. 推奨決定方法の記載	
	11. 健康上の利益・副作用・リスクの考慮	
	12. 推奨とエビデンスの明確な対応関係	
	13. 公表前の外部審査	
	14. 改訂手続きの予定	
明確さと提示の仕方	15. 推奨の具体性	
	16. 患者の状態に応じた選択肢の明示	
	17. 重要な推奨の明確さ	
	18. 利用のツールが用意	
適用可能性	19. 推奨の適用による制度・組織上の考慮	
	20. 推奨適用に伴う費用の考慮	
	21. モニタリング・監査の基準の提示	
編集の独立性	22. ガイドライン編集の独立性	
	23. 作成グループの利益相反の記載	
全体評価		

各項目ごとに1~4点の採点をする。当該ガイドラインの問題点を指摘するとともに、その特徴や長所について全体評価欄に文章をもって評価を行う。

5. 勧告の作成と勧告の強さの決定

エビデンスレベルや結論にばらつきがある場合、また、一定の質の高いエビデンスのある場合にはメタ分析によって勧告を決定する。ただし、十分なエビデンスが存在しないときは、臨床的有用性を示した報告の大きさや有害事象の程度、コスト内容によって勧告を導き出すこともある。その過程を図1に示す。勧告決定に至るエビデンスの程度によって勧告の強さが分類される⁵⁾。AないしD、あるいはEの4~5段階に分けられている。なお、CについてはC₁とC₂に分けて日常臨床の場に誤解を生じさせないように配慮したかたちを

とっているガイドラインがわが国では増えている。その利用法の考え方については図2のような考えを勧めている。

6. 作成・公表した診療ガイドラインについての評価の実施

質の高いガイドラインとして成立しているかを検証するものである。評価項目の内容によっては単純明快な“yes or no”評価が難しいものもあることから、たとえばスコアリングを用いての評価が望ましいように思われる(表4)⁷⁾。そのうえで、段階的に診療ガイドラインの質を向上させていく

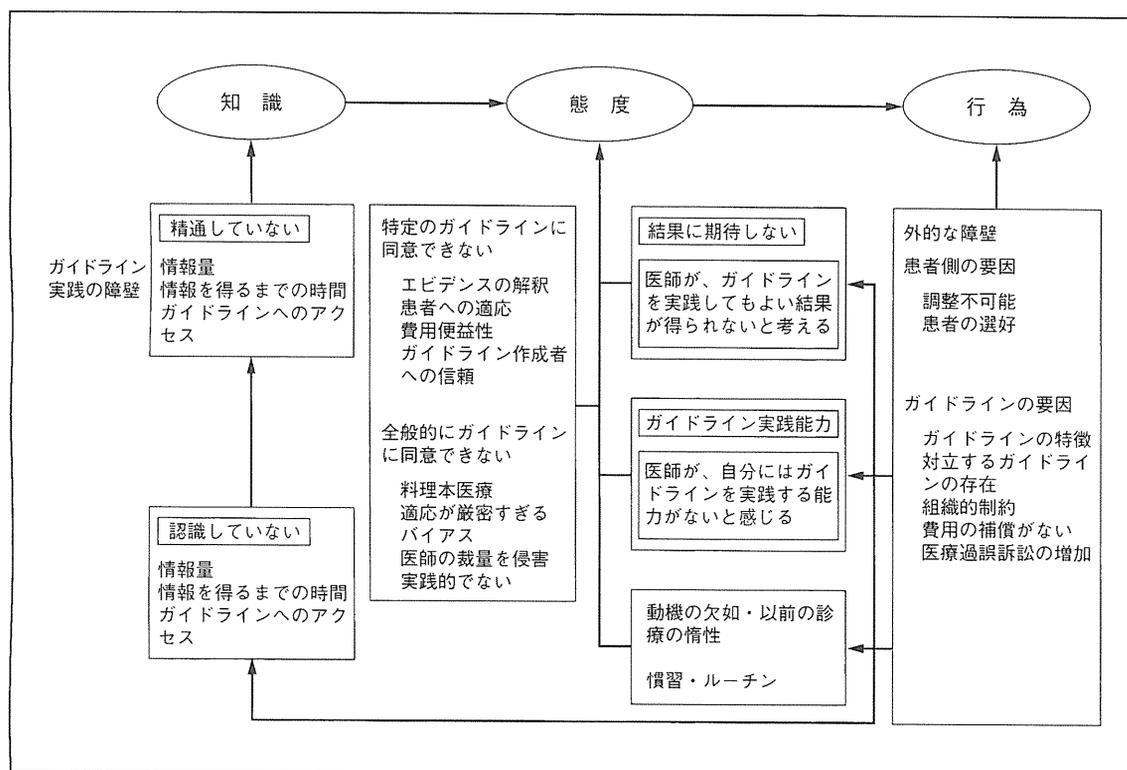


図3 医師のガイドライン実践の障壁
文献13より引用・改変

との意図をもつての評価を行うことがわが国の医療実態に沿っているようにも思われる。しかし、診療内容の globalization が要求されている今日、コンセンサスのある国際的評価法を修飾しないかたちで実施することが正しい在り方なのであろう。

7. 診療ガイドライン作成・公表後の有用性に関する評価の実施（希望的要望項目）

診療行為への影響、患者の健康上でのアウトカム、一定の必要診療と診療経費の関係、費用対効果などについて評価を行うものである（図3）。

医療従事者、特に医師の医療行為の変化の有無については診療機関単位として可能であり、評価は施設内の歴史的対照を設定しての非ランダム化比較試験によって確認することが可能であろう。先に述べたように、診療ガイドラインの存在によって患者の健康上のアウトカムが改善したとする報告は少なくない。医師の裁量下診療が主流であっても有用で適切な医療が行われていたことを反映

しているとも考えられる。一方で、同じ健康上のアウトカムを得るにあたって、総診療費の削減がはかられたとする報告も多い。したがって、診療成果に貢献するほどの役割はないかもしれないが、根拠のない無駄とも判断され得る余分な診療行為を回避できることがその効果として推定し得る。

わが国での医療体制は今日まで国民にとって優しい医療として世界に誇れるものとして理解されてきた。診療ガイドラインの登場によってどのような結果が生じるのであろうか。国際的な視点からも診療ガイドラインのもたらすあらゆるアウトカムは必須な評価項目群と言えよう。

■ 評価に至るプロセスとその公開への期待 ■

ガイドライン作成過程の項目別に、妥当性や臨床の実践の場における有用性・可能性を客観的に評価する手法に関する報告としては、Appraisal of

Guidelines for Research and Evaluation (AGREE) プロジェクトによる評価法(表4)⁷⁾, Shaneyfelt ら⁸⁾の評価法, Cluzeau ら⁹⁾の評価法などが代表的である。日本癌治療学会は AGREE 評価法を採用し, 作成組織にその結果を伝えている。今後の改訂に重要な参考資料とするように希望したい。

評価を実施する場合において課題となることは, 評価に参加する評価者数(量), 評価参加者の専門性と客観性(質), そして評価の偏りに対する判断と最終決定条件が挙げられる。この過程を結果的に採用している今日のガイドラインとして, 消化器外科領域では食道癌¹⁰⁾, 膵癌¹¹⁾, 胆道癌¹²⁾の診療ガイドラインが挙げられる。このなかで上記に紹介した評価法に準じて評価を行い, 評価法別に比較検討したことを公開報告しているガイドラインは膵癌のそれのみである。しかし, 評価の公表責任者である山口幸二氏も評価結果について, (1) 評価者数が少ないこと, (2) 専門分野別の評価者間の比較を実施し得ていないこと, (3) 評価法別の分析が十分とは言えないこと, (4) 外部評価結果での比較を行っていないこと, などを指摘している。いずれにせよ, 診療ガイドラインの展開法を正確に熟知して公表したことに敬意を表したい。この成果は他領域の癌診療ガイドライン作成および更新の過程に大きな影響を与えていくものと考えられる。

■ ガイドライン作成後に 診療行為は変化したのか ■

臨床の場へのガイドラインの直接的役割は, 臨床の場における膨大な情報量のなかで信頼度の高い情報と, 標準的医療水準とされる客観的情報を提供することにある。

一方で, わが国では使用許可のない薬物の使用などについて強く禁止され得ることが課題とされている。ガイドライン作成にそのような外圧が加わりすぎると古いエビデンスに縛られて新しい診療行為を実行できないという社会的制限を生じかねない。医師の経験不足の領域へ具体的に診療行為を示唆することによって医療の底上げに役立つ

ことも考えられよう。個々の医師の診療行為の評価に利用すべきものではなく, 医療施設(機関)の診療内容に影響を与えるべきものと考えられている。

さて, 診療ガイドライン内容をそのまま当てはめ得る対象条件が多い疾患ほど高いレベルのエビデンスが多く, 勧告(推奨)レベルも高くなる。そのような疾患の診療については, 究極的には教科書的治療の範囲で患者のアウトカムに差を生じないことが推察できる。一方, 病態の偏位が大きく, かつその予後予知の困難な領域では診療行為についての高いレベルのエビデンスは存在しにくく, 医師の判断と能力によって診療結果の良し悪しに委ねることとなる。このような疾病ではたとえ診療ガイドラインが存在しても臨床応用への利用に診療ガイドラインの価値を見出すことができないと考える。その意味では, 担癌状態に外科治療を加えるということとなると病態は変化しやすくなり, その早期診断・治療が要求される。一方で, 治癒の可能性の低い場合には刻々と病態が悪化することなどから診療ガイドラインの内容提示意義は大きいと言える。

さて, 医師によってはエビデンスを重要視しない向きも少なくない。その理由としては, (1) エビデンスの重要性を知らないこと, (2) 知っていても自分の決定した診療行為による患者のアウトカムに差を生じないかむしろ上回ると考えていること, (3) エビデンスそのものに対する疑念を抱いている, すなわちエビデンスを支えた臨床試験内容(研究方法と成績)が実態を反映していないと考えていること, (4) 患者が有する複数以上の病態に対して診療側の配する時間的, 医療経済的制約のなかで診療上の重視度が低いということを意識下あるいは無意識的に判断していること, (5) 診療行為を評価する体制が診療機関内に存在しないこと, (6) 積極的な高齢者治療を当初より否定的に考えているか, 医師が勤務医療機関内にあるそのような風潮の影響を受けていることなどがある。これらについては日常茶飯事的に生じ得ることであろうが, いずれも進歩・前進の視点からは何も生じてこない態度と言える(図4)。これらに

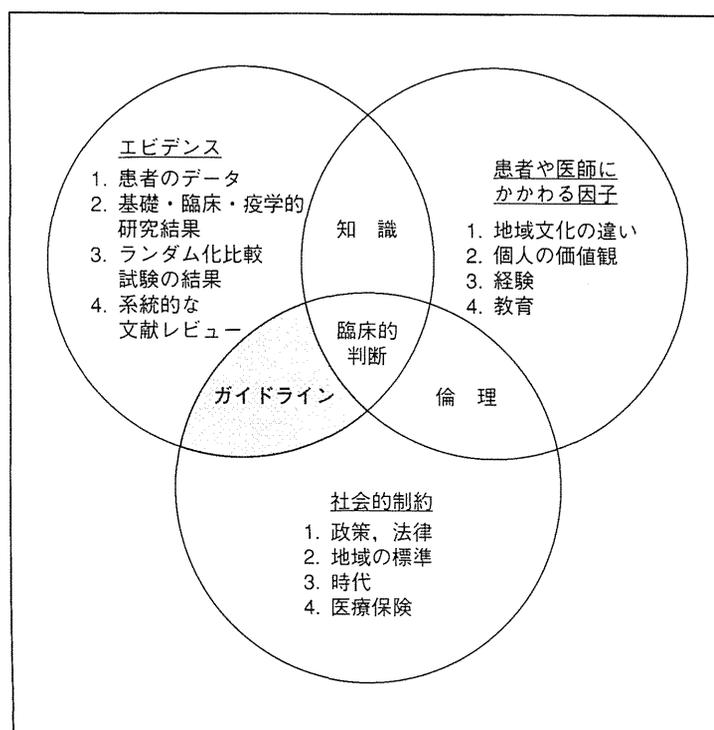


図4 臨床的判断に関する諸因子

については関本¹³⁾が要領よくまとめている。

さて、臨床現場の医師にガイドラインを利用してもらうためには、医療体制のなかで、医療の質の評価の必要対象を「プロセス」と「アウトカム」の2種に分け、医療機関に対する評価を行うことが考えられる。非効率な医療を展開し、合わせて不十分なアウトカムへつないでいることに気付く機会となり得る。しかし一方で、不十分なエビデンスに基づいた推奨内容である場合には、医療内容の押しつけから医療人の積極性を失わせかねないことともなり得る¹⁴⁾。

■ 社会に有用な癌診療ガイドライン利用を目指して ■

診療レベルと施設間格差を1つの根拠にして診療ガイドラインの作成促進のための世論作りはなされた。一方で、生じ得る疑問に対する解答として「ガイドラインは規則でもマニュアルでもない」、あるいは「これさえ守っていればよいというもの

ではない」との説明もなされてきた。しかし、法の現場ではそのような意向とは別に利用されつつある。確かに医療の底上げの目的のために存在していることは明白であるが、大量に出ている診療ガイドライン内容についてそれぞれの記載事項をすべて説明しないなかで不幸な結果を生じると、なぜ記載内容を選択しないかというインフォームド・コンセントがなかったのかという責任追求や、健康アウトカムが不良となった場合のあらゆる診療行為に対する責任追求というかたちで利用されかねない。たとえ、診療結果がその行為と無関係であってもである。診療ガイドラインのネガティブ利用を試みようとする意図をもった集団が発生しやすい社会では、医療への責任回避から高質医療を生じがたい環境が作り出されることになるという「もろ刃の剣」の存在が診療ガイドラインである。上記のほかにも多くの社会的課題が長所・短所として天秤にかけられる内容が含まれている(表5)¹⁵⁾。

さて、臨床に役立つ診療ガイドラインの実践的

表5 癌診療ガイドラインの“the pros and cons”

対象	Pros	Cons
医療体制	最新医療の提供 医療効率の改善 公共的情報の共有 医療費用の抑制	医療費用の増大 ・資源利用の増大 ・資源へ誤解・過剰信頼 民族的特徴の確認不足 新規臨床研究・治験実施の困難化
患者	医療情報入手の簡易化 自主的解釈能力の向上 信頼・確信のもとでの受療 政治・行政への意見提示・介入	個別的医療の制約化 知識不足に由来する誤認 過剰な不信・迷いの助長 政治・行政の偏りの誘因
医療者	診療過程の質の改善・向上 臨床研究内容の質の改善・向上 インフォームド・コンセントへの利用 医学教育への利用	個別化診療に対する制限 不公正な評価 医療訴訟上の有害利用 新規臨床研究・開発の制限

利用がなされるための前提は、(1) そのガイドラインの利用される頻度が高いこと（読みやすい、解釈しやすい、発信源への信頼度が高いことなどが主要因となる）、(2) 診療ガイドラインの利用行為が社会に普及し容認されていることが挙げられる。前者においては、数多い情報のなかで信頼度が高いとされている診療情報の一部を提供しているというものとして、後者においては診療行為決定のプロセスに影響を与えるものの絶対的なものではなく、またアウトカムの決定や改善につながるというものでないこととして認識されていることが背景となっている。そのような条件下であるならば、医療の発展のために一助となるはずである。

■ おわりに

— 癌診療ガイドラインの今後

癌診療ガイドラインはようやくスタートラインに立ち、そして長距離走のスタートを切ったばかりと言える。医療制度の変革のなかで診療の質の向上を目指すための設定としたいが、医療水準を上回る診療を促進させるとの誘導については診療ガイドラインを囲む外的要因（社会経済状況の好転や社会的格差の容認など）に左右される。

癌診療ガイドラインの利用度については、各癌腫別に専門系学会内の会員および評議員などを対

象としたアンケート調査結果として、それぞれの学会の代表者によって学術集会などの場で公表されている。その結果をみると、筆者らの予想を上回って高利用率となっており、若い医師層ほど利用度が高いようである。なお、習熟度別について推し測ることはできていないが、いわゆるベテラン医師においてはその利用頻度は低い。その理由として、診療ガイドラインに対する捉え方に差のあることが主たる背景因子の1つと理解される。

さて、一方でガイドラインでの推奨内容を金科玉条のように上段にかざして勧告内容を“must”あるいは“should”のものとして発言している向きを一部にみることがある。そのような態度は真摯に勉強をしようとする医師に多い傾向にあるが、決して好ましい姿勢ではない。常に“evidence-based”あるいは専門家によるコンセンサスの結果としての勧告であることを念頭に置いての見方をしていただきたい。そして医療、特に外科医療は医療技術の有無あるいはその提供の可・不可が大きくアウトカムに影響させることを強調したい(表5)¹⁶⁾。積極的に新しいエビデンス作りを目指した臨床研究・治験に参加できる個人的・組織的体制を整えることに力を注いでいただきたい。

つぎに、診療行為への不満、クレームあるいは訴訟と診療ガイドラインの関係について触れたい。ガイドラインは医療水準の提示である¹⁷⁾。これまでクレーム内容の判断については、施設あるいは

個人の医療行為上、置かれている条件別要因（たとえば地域や医療施設の大きさなど）で判断されてきた傾向にあった。今後はこれまでの判断とは別に一律的条件のもとで論議されることとなろう。クレームの内容はクレームごとにその要求度のレベルは異なるが、医療水準と同等か、それ以上の行為であるか否かを確認する資料として利用されるであろうし、医療水準を上回っていたとしても、ガイドライン解説文の内容の最高位の診療との比較がなされることとなる¹⁸⁾。この意味でも、医療の集約化、都市集中化、そして高リスク領域への医師の希望者減などを促進する要素になることと思われる。その意味では、国民の深い理解によって誤った方向へと向かないことを願うのみである。

国民が医療に関心を持ち、そして十分に理解が深まることが安定した医療環境を作る。その意味では、癌診療ガイドライン作成に患者（国民）の代表が加わることは望ましい。一部の癌診療ガイドラインについては患者に加わってもらったとの記録が残っているものがあるが、終始一貫しての参加というのは消化器癌領域診療ガイドラインにはみられない。終始参加してもらうためには、医療側の受け入れ体制としての条件がまず必須条件となるが、患者においても参加者（作成委員）として位置付けられるには一定の条件を具備していることが要求される（表6）¹⁹⁾。そのうえで、患者用ガイドラインが作成されていくことについては他言を待たない。そのためには呈示方法や文章表現への工夫が必要となる。現在、魅力的で中味を読みたいと感じさせるデザインや、挿入図をマンガ化するなどの工夫、あるいは勧告に至る多段階のアルゴリズムについてのわかりやすい呈示の仕方などの工夫・研究が進歩しつつある。ガイドラインの普及については、わが国の成績を示さなくてはならない。そのためには、癌診療のデータベースが国民の財産の1つとなり得るとの視点から、国費を導入してのインターネット上でのデータ収集とその公開の在り方を検討すべきであろう。急速な展開がみられる情報研究分野のよい意味で恩恵にあずかりたいところである。それらについての正しい在り方と、未来の医療に役立つとされる

表6 ガイドライン作成にかかわる患者委員の参加への条件

I. 作成委員としての患者代表者への期待像
1. 診療ガイドライン作成への使命感を有する。
2. 高い視点から判断できる能力を有する。
3. 幅広い視覚から患者の視点で考えをまとめる能力を有する。
4. 医学的専門的知識について一定の理解を有する。
5. 患者の抱える問題点を十分に把握し、論理的に要約・発言できる。
II. 患者代表への支援
1. 選択支援のための組織・部署を設ける。
2. Clinical question, topics についてあらかじめ広く要望を集める（患者団体や一般市民などから）。
3. 応募者のデータベースを作成するための公平な広報を行う。
4. 公募制とする。
5. 2名以上の参加とする。
6. 委員にはトレーニングとしての解説の機会を設ける。
7. 理解不可について担当事務局に相談の機会を設ける。
8. グループダイナミクスを学習する機会を設ける。

工学情報の展開について期待と不安を抱きつつも、癌診療の中核をなす消化器外科医の知恵をここにも注いでいただきたいと願っている。

なお、以上の本文、表、図については筆者らの他稿から多くを引用していることをお断りする²⁰⁾。

文 献

- 1) 日本胃癌学会（編）：胃がん治療ガイドラインの解説。一般用、第1版、金原出版、2001
- 2) Canadian Task Force on the Periodic Health Examination : The periodic health examination. Can Med Assoc J 121 : 1193-1254, 1979
- 3) Grimshaw JM, Russell IT : Achieving health gain through clinical guidelines. II : ensuring guidelines change medical practice. Quality in Health Care 3 : 45-52, 1994
- 4) Du Pen SL, Du Pen AR, Polissar N, et al : Implementing guidelines for cancer pain management : results of a randomized controlled clinical trial. J Clin Oncol 17 : 361-370, 1999
- 5) 福井次矢, 丹後俊郎 : 診療ガイドラインの作成の手順, ver4.3. 厚生労働省科学研究費補助金事業「EBMの普及のためのシラバス作成と教育方法およびEBMの有効性評価に関する研究」および「日本におけるEBMのためのデータベース構築および提供利用に関する調査研究」の報告書, 2001
- 6) 日本癌治療学会がん診療ガイドライン委員会 : 日本癌治療学会がん診療ガイドライン作成の手引き, Ver 2, 2004 (<http://www.jsco-cancer-tq.jp>)

- 7) The AGREE Research Trust. <http://www.agreetrust.org> (2007年8月31日アクセス)
- 8) Shaneyfelt TM, Mayo-Smith MF, Rothwan J: Are guideline following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA **281**: 1900-1905, 1999
- 9) Cluzeau FA, Littlejohns P, Grimshaw JM, et al: Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care **11**: 21-28, 1999
- 10) 日本食道学会(編): 食道癌治療ガイドライン. 金原出版, 2007
- 11) 日本膵臓学会膵癌診療ガイドライン作成小委員会(編): 科学的根拠に基づく膵癌診療ガイドライン. 金原出版, 2006
- 12) 胆道癌診療ガイドライン作成出版委員会(編): 胆道癌診療ガイドライン. 医学図書出版, 2007
- 13) 関本美穂: EBM 診療ガイドラインの利用と検証. 診療ガイドラインの実践的利用. 実践の障壁. クリニカルプラクティス **25**: 1020-1024, 2006
- 14) 新保卓郎(訳・解説): 診療ガイドラインはガイドラインに従って作られているか? 医学雑誌に掲載された診療ガイドラインの作成方法の質の検討. JAMA (日本語版) 2000年5月号: 131-137, 2000
- 15) 阿部好文: 診療ガイドラインをどのように利用するか. クリニカルプラクティス **25**: 1010-1013, 2006
- 16) 小川道雄: 癌治療と診療ガイドライン. 外科治療 **96**: 390-393, 2007
- 17) 古川俊治, 北島政樹: 診療ガイドラインと法的「医療水準」. 日消病会誌 **101**: 1-8, 2004
- 18) 河原和夫: 診療ガイドラインの有効性について. 日整会誌 **80**: 720-723, 2006
- 19) 高橋 都, 佐藤(佐久間)りか, 中山健夫: 患者参加型の診療ガイドラインがもたらすもの. クリニカルプラクティス **25**: 1042-1046, 2006
- 20) 平田公一, 古畑智久, 大村東生, 他: 「癌治療ガイドライン」の基本的検証法. 消外 **30**: 1823-1836, 2007
(HIRATA Koichi, et al 札幌医科大学外科腫瘍学・消化器外科学: 〒060-8543 北海道札幌市中央区南1条西16丁目)

◆原稿募集◆

私の工夫—手術・処置・手順

弊誌では「私の工夫—手術・処置・手順」欄への読者の皆様からの投稿を募集しております。

日ごろ実践しておられます手術・処置・手順などに関する工夫を右記の要領でおまとめのうえ、ご寄稿ください。

投稿受付先: ☎113-8719

東京都文京区本郷局私書箱5号
(株)医学書院「臨床外科」編集室宛

- 1) 原稿の内容は、日ごろ実践しておられる手術、処置、検査などに関する独自の工夫点をご紹介いただくものとします。
- 2) 原稿枚数は400字4枚程度とします。レイアウトの関係で、多少変更させていただくことがあります。ほか、図表がある場合は1~2枚。
- 3) 原稿には住所、氏名、E-mail アドレスを明記してください。
- 4) 著作権譲渡同意書ならびに文章データを保存したCD、FDなどを同封してください。
- 5) 原稿の採否は、編集委員会で決定させていただきます。

Quality of guideline development assessed by the Evaluation Committee of the Japan Society of Clinical Oncology

Takuro Shimbo · Tsuguya Fukui · Chikashi Ishioka · Kohji Okamoto · Takahiro Okamoto · Shingo Kameoka · Atsushi Sato · Masakazu Toi · Kunihiko Matsui · Toshihiko Mayumi · Shigetoyo Saji · Masaru Miyazaki · Yuichi Takatsuka · Koichi Hirata

Received: 5 January 2010 / Published online: 24 March 2010
© Japan Society of Clinical Oncology 2010

Abstract

Background The Japan Society of Clinical Oncology started implementing clinical practice guidelines for cancer in 2001. It created a Guideline Committee and has published cancer-related information in collaboration with individual subspecialty cancer societies. The society then established an Evaluation Committee to assess the quality of guidelines.

Methods The quality of development and general characteristics of guidelines were reviewed using the AGREE instrument. The six standardized domain scores and 23-item crude scores were described, and items with a low median score or a wide inter-quartile range were explored. Kappa statistics for inter-rater reproducibility were also described.

T. Shimbo (✉)
Department of Clinical Research and Informatics,
Research Institute, International Medical Center of Japan,
1-21-1 Toyama, Shinjuku-ku, Tokyo, Japan
e-mail: tshimbo@ri.imcj.go.jp

T. Fukui
St. Luke's International Hospital, Tokyo, Japan

C. Ishioka
Department of Clinical Oncology, Institute of Development,
Aging and Cancer, Tohoku University, Sendai, Japan

K. Okamoto
First Department of Surgery, School of Medicine,
University of Occupational and Environmental Health,
Kitakyushu, Japan

T. Okamoto
Department of Endocrine Surgery, School of Medicine,
Tokyo Women's Medical University, Tokyo, Japan

T. Okamoto
Department of Hygiene and Public Health, School of Medicine,
Tokyo Women's Medical University, Tokyo, Japan

S. Kameoka
Second Department of Surgery, Tokyo Women's Medical
University, Tokyo, Japan

A. Sato
Division of Medical Oncology, Department of Internal
Medicine, Showa University School of Medicine, Tokyo, Japan

M. Toi
Department of Breast Surgery, Graduate School of Medicine,
Kyoto University, Kyoto, Japan

K. Matsui
Clinical Education Center, Kumamoto University Hospital,
Kumamoto, Japan

T. Mayumi
Department of Emergency and Critical Care Medicine,
Nagoya University Graduate School of Medicine, Nagoya, Japan

S. Saji
Department of Surgical Oncology, Gifu University Graduate
School of Medicine, Gifu, Japan

M. Miyazaki
Department of General Surgery, Graduate School of Medicine,
Chiba University, Chiba, Japan

Y. Takatsuka
Department of Breast Surgery, Kansai Rosai Hospital,
Amagasaki, Japan

K. Hirata
First Department of Surgery,
Sapporo Medical University, Sapporo, Japan

Results Domains in which the median score was >50 points in 18 guidelines developed between March 2005 and May 2009 included “scope and purpose,” “rigor of development,” and “clarity and presentation.” Domains with a median score < 50 points were “stakeholder involvement,” “applicability,” and “editorial independence.” Scores in all domains except “stakeholder involvement” were higher during the second half of the period than during the first half of the period, although *P* values were 0.10–0.93. Crude scores remained low for items 5, 7, 19, 20, 22, and 23, and the inter-quartile ranges of items 2, 6, 10, and 22 were wide. Kappa statistics ranged from –0.02 to 0.64, and they were especially low for items 3, 5, 7, 18, and 23.

Conclusion Guideline quality has tended to improve during the 10 years since the society started this activity. However, issues remain to be improved through continuous revisions.

Keywords Clinical practice guideline · AGREE instrument · Cancer

Introduction

The Japan Society of Clinical Oncology started implementing clinical practice guidelines (CPGs) for cancer in 2001 in collaboration with allied subspecialty societies. The society has developed summary versions of CPGs and flowcharts, and it has published them on the Internet with structured abstracts of important articles. Around 20 guidelines have been developed by subspecialty societies by November 2009, and 13 of them are presented on the society’s homepage (<http://www.jsco-cpg.jp/>) [1].

The society established a Guideline Committee (GC) for this activity, as well as an Evaluation Committee (EC) to evaluate and ensure the quality of published guidelines. The aims of the present study were to identify issues requiring resolution from a summary of the assessment results generated by the EC.

Methods

Process before publishing the guidelines

The activity of CPG publishing and implementation in the society proceeds as follows. A subcommittee of the GC for a specific cancer writes a draft summary, algorithm, and structured abstract in accordance with the specific subspecialty society, and submits them, or sometimes a complete CPG, to the board of the GC. The board of the GC reviews and sends them to the EC. The EC evaluates them

and reports the result to the chair of the GC and the members of GC subcommittee. If there is no major flaw, a homepage is developed. These tools for implementation of the CPG are then released to the public after the final approval of the GC and the board of the society.

The review in the EC

The EC has ten members, including a chair and four members from outside the society. All members individually review drafts under evaluation before attending a meeting where all members reach a consensus-based final assessment.

The AGREE instrument [2] was used for reviews that focus on the process of CPG development and the general characteristics of the CPGs, but not on the validity of specific statements. The AGREE instrument is a comprehensive tool for evaluation whose validity and reproducibility have been investigated [3, 4]. The EC did not require revision of the content and format of the draft after review, but revisions were expected for a subsequent version. The EC previously presented the appropriate methods for developing evidence-based CPGs to the GC.

Method of review

The present study summarizes the results of the review of the CPGs by the EC.

The AGREE instrument consists of 23 items that assess six domains of the CPG development process: “scope and purpose” (items 1–3), “stakeholder involvement” (items 4–7), “rigor of development” (items 8–14), “clarity and presentation” (items 15–18), “applicability” (items 19–21), and “editorial independence” (items 22–23). For each item, a crude score of 1–4 is assigned based on the reviewers’ certainty of fulfilling the requirements of the items and the quantity of information contained in the CPG. A standardized domain score is calculated for the 6 domains after summing and adjusting the crude scores into a scale from 0 to 100 points. A global assessment could be given, but such global assessments were not recorded for all CPGs. Global quality was described as an aggregated score determined from the summation of all domain scores, although AGREE does not suggest using this strategy for global assessment.

The distributions of the crude scores for the items were determined. Low-score items in which the medians were ≤ 2 and dispersed items, for which the inter-quartile range of the crude score was 1–4, were identified. The dispersed items contained CPGs with both low and high scores, which led to the supposition that they could be easily improved.

Kappa statistics were calculated for each item to determine inter-rater reproducibility [5, 6]. Low kappa values

indicate a trend toward the item scoring differently among raters. When calculating kappa, crude scores of 1 and 2, as well as those of 3 and 4, were combined into one level. The EC used only one representative score based on consensus

for evaluation at meetings and did not use the individual crude scores from which the kappa values were derived.

When members thought that determining a score was difficult, the committee used its own criteria to standardize

Table 1 Guidelines that have been reviewed by the evaluating committee

Type of cancer	Title	Version
Stomach ^a	Japanese Gastric Cancer Association: guidelines for the diagnosis and treatment of carcinoma of the stomach, April 2004 edition	2
Liver ^b	The Japan Society of Hepatology: ^c “clinical practice guidelines for hepatocellular carcinoma:” evidence-based clinical guidelines for the diagnosis and treatment of hepatocellular carcinoma in Japan (the print/web version)	1
GIST ^a	Japanese Gastric Cancer Association, Japan Society of Clinical Oncology, Japanese Study Group on GIST: clinical practice guidelines for gastrointestinal stromal tumors (GIST) in Japan	1
Oral cancer	Japan Society for Oral Tumors: clinical practice guidelines for oral cancer	1
Uterine cervix	The Japan Society of Gynecologic Oncology: treatment guidelines for cervical cancer, 2007 edition	1
Uterine body	The Japan Society of Gynecologic Oncology: treatment guidelines for uterine body cancer, 2006 edition	1
Children’s leukemia	The Japanese Society of Pediatric Hematology: guidelines for the treatment of childhood leukemia/lymphoma, 2007 edition	1
Esophagus ^d	The Japan Esophageal Society: guidelines for the diagnosis and treatment of esophageal cancer	2
Kidney ^d	The Japanese Urological Association: clinical practice guidelines for managing renal carcinoma and the digest edition (web version)	1
Pancreas ^d	Japan Pancreas Society: evidence-based clinical practice guidelines for pancreatic cancer	1
Colon ^d	Japanese Society for Cancer of the Colon and Rectum: guidelines for the treatment of colon cancer, 2005 edition	1
Biliary tract ^d	Japanese Society of Hepato-Biliary-Pancreatic Surgery: clinical practice guidelines for the management of biliary tract and ampullary carcinomas (the print and web digest version)	1
Head and neck	Japan Society for Head and Neck Cancer: clinical practice guidelines for head and neck cancer	1
Breast ^a	The Japanese Breast Cancer Society: evidence-based clinical practice guidelines of the Japanese Breast Cancer Society (5 volumes) and web version 1. Systemic therapy 2. Surgery 3. Radiation therapy 4. Screening and diagnosis 5. Epidemiology and prevention	1
Lung	The Japan Lung Cancer Society: clinical practice guidelines for lung cancer, revised edition	2
Skin ^d	The Japanese Skin Cancer Society: clinical practice guidelines for the management of cutaneous malignancies	1
Ovary ^d	The Japan Society of Gynecologic Oncology: ovarian cancer treatment guidelines, 2004 edition	1
Ovary	The Japan Society of Gynecologic Oncology: ovarian cancer treatment guidelines, 2007 edition	2

Order in table reflects the list in the homepage of the Japan Society of Clinical Oncology (order of Japanese 50 sounds)

^a Presentation was partly funded by the Scientific Study for the Third Term Comprehensive Control Research for Cancer of the Ministry of Health, Labour, and Welfare in 2007

^b Development was funded by the Scientific Study for Supporting Clinical Practice Guidelines of the Ministry of Health, Labour, and Welfare in 2002–2003

^c On October 2009

^d Development and presentation was partly funded by the Scientific Study for the Research on the Medical Safety and Health Technology Assessment of the Ministry of Health, Labour, and Welfare in 2005–2006

Table 2 Domain scores determined using the AGREE instrument for clinical practice guidelines

Domain	Total (<i>n</i> = 18)		The first half, March 2005–March 2007 (<i>n</i> = 10)		The second half, April 2007–May 2009 (<i>n</i> = 8)		<i>P</i> value ^a
	Median	IQR ^b	Median	IQR	Median	IQR	
Scope and purpose	72.2	66.7–100	66.7	55.5–100	83.3	66.7–100	0.38
Stakeholder involvement	41.7	16.7–50.0	43.1	25.0–58.3	41.7	29.2–50.0	0.93
Rigor of development	66.7	38.9–83.3	44.4	16.7–72.2	72.2	61.1–86.1	0.13
Clarity and presentation	75.0	58.3–91.7	70.8	33.3–91.7	83.3	70.8–100	0.18
Applicability	33.3	0–66.6	16.7	0–33.3	50.0	25.0–66.7	0.10
Editorial independence	0	0–50.0	0	0–0	33.3	0–50.0	0.12
Aggregated	56.3	36.5–69.8	48.6	28.6–58.7	65.9	54.8–71.4	0.11

^a Comparison of scores between the first half of the period and the second half of the period was tested using the Wilcoxon rank-sum test

^b Inter-quartile range

the score among its members. Item 13 indicates a requirement for an external review of the CPG. This item was not scored because review by the EC is compatible with this. Item 21 requires the CPG to present key review criteria for monitoring or audit. This item was also omitted from scoring because quality indicators for measuring adherence to CPGs have not been developed.

Results

The EC started reviewing CPGs in March 2005, and 18 of them had been reviewed by May 2009 (Table 1). Table 2 shows the standardized domain scores of these CPGs. The domains with median scores > 50 points during the entire period of review were “scope and purpose,” “rigor of development,” and “clarity and presentation.” The median scores for “stakeholder involvement,” “applicability,” and “editorial independence” were <50 points. All domain scores except “stakeholder involvement” were higher during the second half of the period than during the first half of the period, although the *P* values were 0.10–0.93.

Figure 1 shows the distribution of crude scores for each item in all CPGs. Item numbers with median crude scores ≤ 2.0 were 5 (emphasizing patients’ perspectives), 7 (pre-test before publication), 19 (discussion about potential organizational barriers), 20 (considering cost implications), 22 (editorial independence from funding body), and 23 (records of conflicts of interest). The item numbers with widely distributed crude scores were 2 (description of clinical questions), 6 (target users defined clearly), 10 (presentation of methods for formulating recommendations), and 22 (editorial independence from funding body).

Table 3 shows the inter-rater reproducibility for each item. The kappa statistics were –0.02 to 0.64, and the null hypothesis that the consistency of the results occurred by chance alone could not be rejected for items 3 (target

patients described specifically), 5 (emphasizing patients’ perspectives), 7 (pre-test before publication), 18 (tools for application), and 23 (records of conflicts of interest).

Discussion

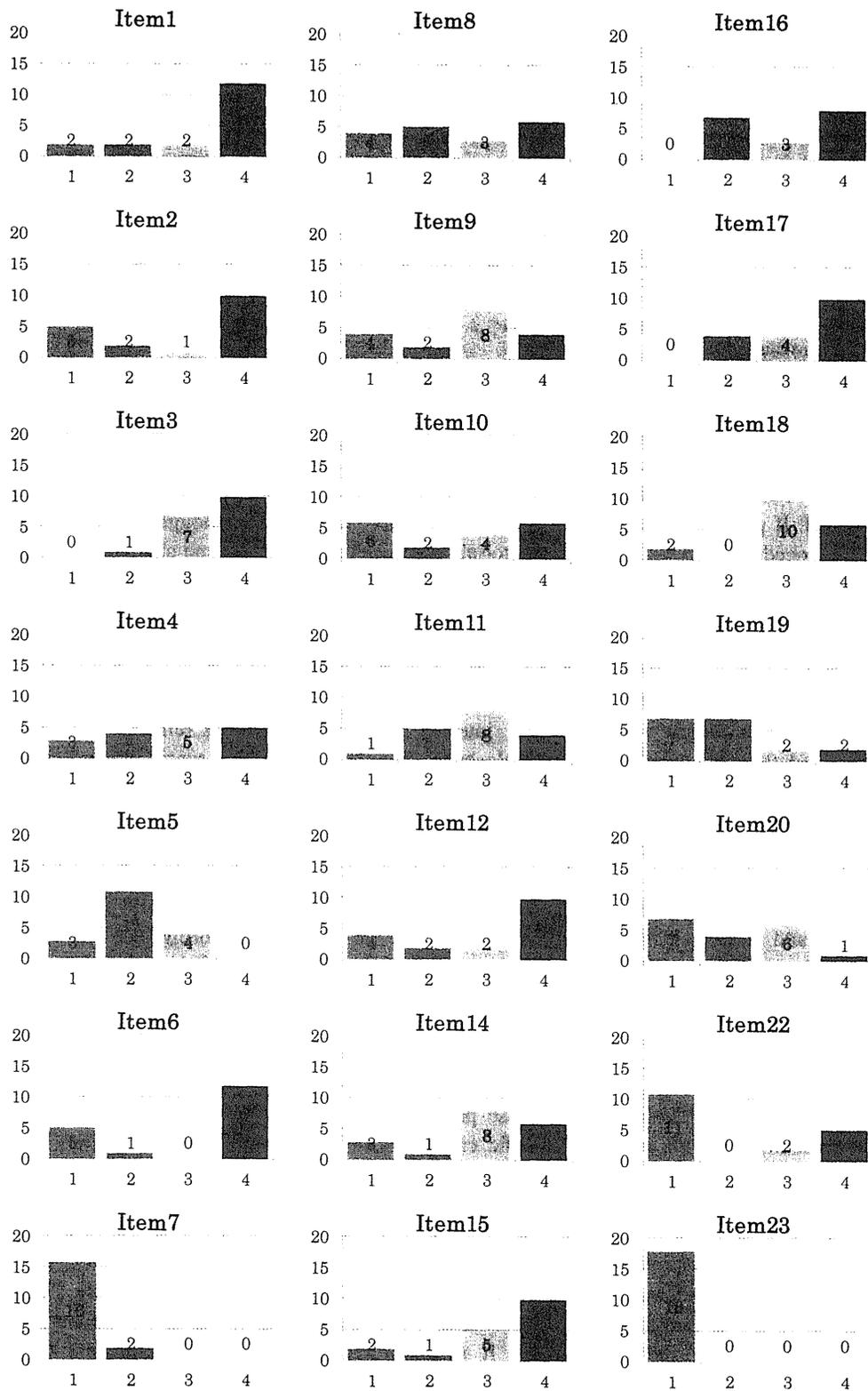
The present report describes the results of continuous evaluation of CPGs assembled by the Japan Society of Clinical Oncology. Changes in standardized domain scores indicated that the methods and organization for developing CPGs have improved slightly, although the differences were not statistically significant and the number of CPGs assessed was small. The domains with median scores > 50 points were “scope and purpose” (items 1–3), “rigor of development” (items 8–14), and “clarity and presentation” (items 15–18). Domains with median scores < 50 points were “stakeholder involvement” (items 4–7), “applicability” (items 19–21), and “editorial independence” (items 22–23). Developers must consider these findings when developing new guidelines or revising those that have been already established. For individual items, low scores were observed in items 5, 7, 19, 20, 22, and 23.

Item 5 emphasizes patients’ perspectives. The values of individual patients with cancer should be considered in clinical decision making. Several guidelines seemed to specifically recommend a single option without providing alternatives. Representatives of patients or paramedical staff should be involved in these processes.

Item 7 addresses the pilot use of the CPG before formal publication. When a pilot is not used to improve the quality of the CPG, early feedback about its validity, implementation, and impact on routine practice after publication should be obtained.

Item 19 addresses potential organizational barriers. Alternatives should be discussed when barriers interfere with CPG implementation.

Fig. 1 Distribution of crude scores for each item. Crude scores of each item were reached by consensus after discussion in a committee meeting and are not simple means or medians of scores supplied by individual members of the Evaluation Committee



Item 20 refers to cost issues. The clinical practice of oncology must be individualized because it is based on patient status and value judgments. In general, the issue of cost is important, especially in preventive medicine and in

the long-term management of prevalent chronic disorders such as hypertension or dyslipidemia. Cost is more urgent in preventive medicine than for oncologists whose patients have cancer.

Table 3 Inter-rater reproducibility of each item

Item	Kappa ^a	P value	Item	Kappa ^a	P value
1	0.23	<0.01	12	0.31	<0.01
2	0.64	<0.01	14	0.49	<0.01
3	0.00	0.49	15	0.15	<0.01
4	0.37	<0.01	16	0.20	<0.01
5	-0.02	0.61	17	0.15	<0.01
6	0.34	<0.01	18	0.05	0.18
7	0.04	0.23	19	0.19	<0.01
8	0.33	<0.01	20	0.28	<0.01
9	0.35	<0.01	22	0.14	0.01
10	0.33	<0.01	23	0.05	0.20
11	0.18	<0.01			

^a Kappa statistics express agreement of several raters above the expected value

Item 22 requires editorial independence from funding bodies. The source of financial support should be documented. If pharmaceutical companies are the source, then the procedure for maintaining editorial independence should also be documented.

Item 23 asks about records of conflicts of interest. None of the CPGs described records for conflicts of interest, although the impact of CPGs on routine practice is substantial. Concern about conflicts of interest is increasing in Japan, where medical journals have not managed this issue as foreign journals have. The Japan Society of Clinical Oncology and the Japan Society of Medical Oncology have developed the “Clinical Oncology Research Conflict of Interest Policy (ver. 1)” [7, 8]. According to this policy, all members of the society must report their status regarding conflict of interest when they report and publish in the society, and these reports are centrally reviewed. This procedure must be followed when CPGs are developed, and records about conflicts of interest should be explicit.

The distribution of crude scores was wide for items 2, 6, 10, and 22, for which the same item scored low and high in several CPGs. Improving these points might not be difficult, although guideline-specific conditions might be involved. The involvement of experts specialized in the field of guidelines will be useful. Item 2 requires clear descriptions of clinical questions. When “Clinical Question” is first described for each CPG topic, it may help focus readers to understand the content more easily. This format of clinical question is preferable. Item 6 asks for a clear definition of the target users. It is important to define that clearly when developing and using CPGs. Item 10 addresses an explicit document that describes the methods of formulating recommendations; however, many CPGs did not provide this information. The impact of an assessment of benefits and harms after a systematic review on formulating a recommendation should be addressed. If disagreement about a recommendation

arises, the methods used to reach consensus should be described.

Although the EC has reviewed a dozen CPGs, this report has some challenging issues as limitations. First, the inter-rater reproducibility of several items of the AGREE instrument was poor. Previous studies have identified good validity and reproducibility [3, 4], but we found that reproducibility was not easily achieved in our setting. Although AGREE is a good method of evaluation, the scoring remains subjective. We did not directly use the crude scores of individual members to reach the final assessment. Nevertheless, low reproducibility means that judgment by a member using the AGREE items is not a simple matter. Among low-score items, the score of items 5, 7, and 23 might be influenced by a difficult evaluation. Consensus will be achieved if the committee has criteria for scoring that maintain the original concept of the AGREE items.

Second, common scoring methods are not applicable to all CPGs, because solid evidence is not available in some fields of cancer. Although all CPGs of the society are related to cancer, each type of cancer has specific characteristics. AGREE itself does not recommend establishing a threshold to differentiate CPGs of “good” or “bad” quality.

The activity of CPG development is continuous, and CPGs of the subspecialty societies and the published material of the society (<http://www.jsco-cpg.jp/>) will be revised sequentially. These guidelines have also been published on the homepages of the subspecialty societies and of the Medical Information Network Distribution Service (MINDS), thus bringing the CPGs closer not only to medical professionals but also to patients. The activities of publishing and implementing CPGs within the society over the first decade seem to have begun well. Efforts to improve quality must be maintained, and users, including patients, should be able to easily understand the contents.

Acknowledgments The authors thank Ms. Misao Oda, who manages the affairs of the Evaluation Committee and maintains detailed records. We also acknowledge the assistance of Ms. Kaoruko Nakazawa, who helped with drafting the manuscript. This report was partly funded by the “Study for Development and Revision of Clinical Practice Guidelines for Cancer and Maintenance of Publishing,” which is supported by the Ministry of Health, Labour, and Welfare, and by the fund for Scientific Studies for Cancer Clinical Research.

Conflict of interest statement M. Toi received honoraria from GlaxoSmithKline. The other authors have no conflict of interest.

References

1. Japan Society of Clinical Oncology (2007) Clinical practice guidelines for cancer. <http://www.jSCO-cpg.jp/>, accessed on November 10, 2009
2. AGREE collaboration (2004) <http://www.agreecollaboration.org/>, accessed on July 6, 2009
3. AGREE collaboration (2003) Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* 12:18–23
4. MacDermid JC, Brooks D, Solway S et al (2005) Reliability and validity of the AGREE instrument used by physical therapists in assessment of clinical practice guidelines. *BMC Health Serv Res* 5:18
5. Fleiss JL (1981) *Statistical methods for rates and proportion*, 2nd edn. Wiley, New York, pp 212–236
6. Kappa-interrater agreement (2007) STATA base reference manual, vol 2, I-P, Release 10. STATA Press, Texas, pp 83–96
7. Joint Formulation Committee (Japan Society of Clinical Oncology/ Japan Society of Medical Oncology) (2009) Clinical Oncology Research Conflict of Interest Policy (ver.1). <http://jsco.umin.ac.jp/outline/COI.html>
8. Joint Formulation Committee (Japan Society of Clinical Oncology/ Japan Society of Medical Oncology) (2009) Bylaws to the Clinical Oncology Research Conflict of Interest Policy (ver.1). <http://jsco.umin.ac.jp/outline/saisoku4.html#saisoku4>

総 説

診療ガイドラインの活用のための方策と評価法

真弓 俊彦¹⁾ 吉田 雅博²⁾ 平田 公一³⁾
高田 忠敬⁴⁾

要旨：現在、種々の診療ガイドラインが作成されているが、ガイドラインの重要な目的である診療内容の標準化と患者の予後改善達成のため、ガイドライン流布のための種々の方策は不可欠である。また、ガイドラインに対する評価が必要で、現在までは AGREE などのガイドライン作成法に対する評価が主であった。今後は、ガイドラインの有用性に基づく評価が必要と考えられ、そのためには症例登録などによってガイドライン内容と臨床効果が合致するか否かの検証が重要で、そのためのシステムの設置と整備が必要となる。また、遵守率なども分析のための因子として必須で、ガイドラインの有用性の評価のためには、臨床指標をあらかじめ提示することが好ましい。

索引用語：診療ガイドライン、臨床指標、プロスペクティブスタディ（前向き研究）、急性腎炎、急性胆道炎

はじめに

本邦においても厚生労働省研究班や学術団体によって多数の診療ガイドラインが作成され、日本消化器病学会でも精力的に診療ガイドラインを作成した^{1)~6)}。そこで、本稿では、ガイドライン作成後に必要となる方策とガイドラインの評価そして本邦での診療ガイドラインの将来像について概説する。

1 診療ガイドラインの現状と問題点

診療ガイドラインは診療内容の標準化、患者の activity of daily living (ADL) や予後の改善のために、限りある資源を有効に活用し、効率の良い医療を行うことを目的として作成されることが多い。時に、その領域の専門家から、「ガイドラインは無用」との意見もみられるが、そもそもガイ

ドライン作成の目的は診療内容のばらつきを是正し、その疾患を専門としない一般臨床医の診療内容の底上げを図ることを目的としており、診療ガイドラインを参考に診療にあたるのは一般臨床医である。専門家にとってはガイドラインは不要ともいえ、最先端の診療行為はガイドラインから逸脱した行為である場合もあり得る。

一方、ガイドライン作成過程で問題点は明確となる。たとえば、急性胆管炎、胆嚢炎のガイドライン作成過程で、それらの診断基準や重症度判定基準が存在しないことが明らかになった⁷⁾。同様に、スタンダードと思われてきた診断法や治療法でも、無作為化比較対照試験 (RCT) などの質の高い研究が行われておらず、十分なエビデンスに基づいていないことが判明し、今後のリサーチ

1) 名古屋大学大学院医学研究科救急・集中治療医学 2) 国際医療福祉大学化学療法研究所附属病院人工透析・一般外科 3) 札幌医科大学第一外科 4) 帝京大学外科

Evaluations and strategies for acceleration of practical use of practical guidelines
Toshihiko MAYUMI¹⁾, Masahiro YOSHIDA²⁾, Koichi HIRATA³⁾ and Tadahiro TAKADA⁴⁾

1) Department of Emergency and Critical Care Medicine, Nagoya University Graduate School of Medicine. 2) Department of Hemodialysis and Surgery, International University of Health and Welfare, Clinical Research Center, Kaken Hospital, 3) Department of Surgical Oncology and Gastroenterological Surgery, Sapporo Medical University Graduate School of Medicine, 4) Department of Surgery, Teikyo University School of Medicine
Corresponding author : 真弓 俊彦 (mtoshi@med.nagoya-u.ac.jp)

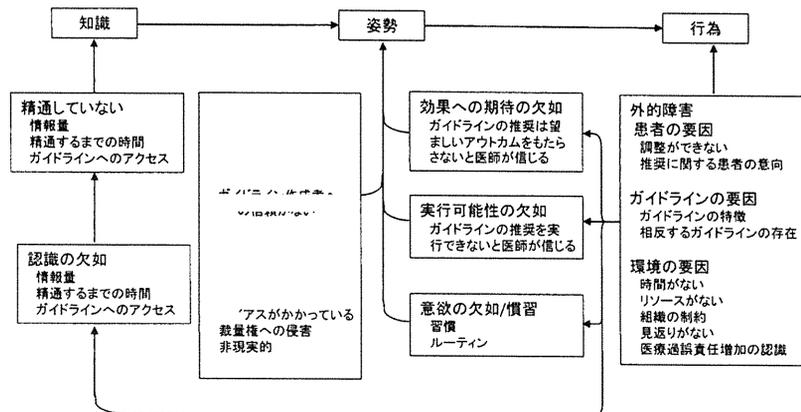


Figure 1. 医師がガイドラインに沿った診療行為を行うにあたってのバリアー(文献11より引用).

課題を定めることには意義がある。

また、近年、司法の場で診療ガイドラインが引用されることも少なくなく、ガイドライン作成者は日本の医療の現状も加味してガイドラインを作成することが必要である。一方、ガイドライン使用者は、「すべての症例でガイドライン内容を遵守すればよく、遵守しないことが悪い」ということではないことを十分理解する必要がある。臨床では、患者の状態、患者・家族の意向、施設の状態などを加味し診療内容を決定すべきである。ガイドライン内容から逸脱する場合には患者、家族にその理由を説明し、インフォームドコンセントを得ることが不可欠である。

II ガイドライン作成後の方策

ガイドライン作成の目的は、診療の標準化や患者のADLや予後の改善であり、ガイドラインの完成はまだその過程に過ぎない。ガイドラインを流布させ、使用により、診療内容が変化し、患者のADLや予後の向上することが必要となる。

胃潰瘍や急性膵炎のガイドラインが作成されているが、実際には、その領域の専門家でも十分周知されていないかったり、アルゴリズムや推奨文しかみておらず、ガイドライン本文を読まれることは少ないことが報告されている^{8)~10)}。また、ガイドラインの普及以外にもガイドライン内容の受容、実行にともなう効果への期待度、モチベーション、内的外的障害など、医師がガイドラインに沿った診療行為を行うにあたってのバリアーが提起さ

れている (Figure 1)¹¹⁾。ガイドラインに沿った診療が行われ、患者の予後を改善するためには、その流布のための種々の方策が必要である (Table 1)。

1. 使用しやすいガイドラインの作成

まず第一に、ガイドラインそのものが、使用者が利用しやすいものでなくてはならない。利用しやすいように、診療上の疑問を網羅したクリニカルクエスチョン方式とする、診療方針が一目でわかるアルゴリズムを掲載する、目次や索引、クリニカルクエスチョン一覧を作成する、ガイドライン本文や推奨文にもエビデンスレベルを記載するなどのことが求められている。また、後述のような臨床指標 (clinical indicator) を掲げ、いくつかの診療行為をまとめて提示し、それらの施行の有無の評価を行いやすくすることも必要である。

2. ガイドラインの普及

第二は普及の促進である。どんなに質が高く使い勝手がよいガイドラインでも、その存在が知られていないと、使用してもらえない。ガイドライン内容の概略を解りやすく表現し、携帯可能な小冊子の作成、重要事項を列挙したカードの作成配布、さらには、学会ホームページなどで紹介し、ガイドライン簡略版や小冊子、カードなどをいつでも閲覧かつダウンロードもできる、などの方策が必要である。可能であれば、専用のホームページに、必要な情報が集約されているとなお望ましい。また、学会などでシンポジウムや公開討論会

Table 1. 日本におけるガイドライン作成後の方策

<p>1. 利用しやすいガイドラインの作成</p> <ul style="list-style-type: none"> ・目次や索引を作成する ・ガイドライン本文や推奨文にもエビデンスレベルを記載する ・診療上の疑問を網羅したクリニカルクエスチョン方式とする ・診療方針が一目でわかるアルゴリズムを掲載する ・臨床指標を掲げ、bundleとして、施行の有無を明確化する <p>2. ガイドラインの普及の促進</p> <ul style="list-style-type: none"> ・学会などでのシンポジウムや講演を開催する ・携帯可能な小冊子を作成する ・重要事項を列挙したポスター、チェックリスト、カードを作成配布する ・市民への啓蒙のために市民向けのガイドラインを作成する ・学会ホームページなどに掲載する ・ガイドライン専用のホームページを作成し、必要な情報を集約する ・ホームページなどから上記資料をダウンロード可能とする ・診療ガイドラインを集計している Minds などのホームページに掲載してもらう ・臨床例との整合性を feedback し、新知見を盛り込み定期的な改訂を行う <p>3. ガイドラインの遵守</p> <ul style="list-style-type: none"> ・症例登録を行う ・臨床指標を掲げ、bundleの施行の有無を明確化する ・上記の施設間比較を可能とする ・症例登録が可能なスタッフの配置などの整備を行う ・上記を可能とする保険診療上の優遇処置などを整備する ・予後改善や政策に反映するために市民へ啓蒙する
--

などを重ねることもよい。また、ある基準以上の診療ガイドラインを集計し公開している日本医療評価機能の医療情報サービス (Minds) のようなホームページ (<http://minds.jcqh.or.jp/>) に掲載してもらうことも必要であろう。

3. ガイドライン内容の遵守の推進

第三は、ガイドライン内容の遵守の推進である。前向きに症例登録を行い、ある時期までにある診断や治療などを行うことを具体的に示した臨床指標 (clinical indicator) を提示し、その施行の有無を登録することによって、ガイドラインの遵守を促進する。これによって後述のガイドライン評価も可能となる。

模範となる事例としては、Surviving Sepsis Campaign のホームページがあげられる¹²⁾。このcampaign は European Society of Intensive Care Medicine (ESICM), International Sepsis Forum (ISF), Society of Critical Care Medicine (SCCM)

を中心に多数の国の学会が参加し、5年間で敗血症による死亡率を25%減少させることを目的に2002年から開始された。そのホームページにはガイドラインを始め、ガイドライン内容に沿った11の臨床指標を各々の時間内に実施するように集めた sepsis bundle、それらの実施の有無をチェックするためのカードなどが掲げてあり、いつでも閲覧、ダウンロード可能となっている。また、重症敗血症症例を前向きに登録する sepsis registry も行っており、症例登録時には、sepsis bundle の施行の有無を入力することによって施設ごとの遵守率を feedback している。また、一般市民向けのページも設け、啓蒙活動も行っている。

これらを利用し、sepsis campaign の内容を医師、看護師に教育することによって敗血症患者の入院および、28日死亡率が低下したことが報告されている¹³⁾。

4. ガイドライン改訂

一方、ガイドライン作成後にも新たな知見が報告され、また、医療内容や環境も変化する。そのため、ガイドラインにも寿命があるといわれ¹⁴⁾、一般に4年程度で改訂することが望ましいとされている。日本の多くのガイドラインでは、改訂に際して新しい知見は取り入れられるものの、それ以外の意見についてはあまり反映されていない。臨床現場からのフィードバックも活かしたガイドラインの定期的な改訂は、普及、継続、患者の予後改善のためにも必要である。

III ガイドラインの評価

1. ガイドラインの評価法

ガイドラインが一般臨床医に実際に使用されるか否かは、そのガイドラインの質によっており、ガイドラインが信頼に値する作成法によって作成されていることが不可欠である¹⁵⁾¹⁶⁾。作成母体が学会など権威のあるものであっても、信頼に値する作成方法でない場合には、使用してもらえない。

従来、ガイドラインの評価法についてはいくつか提唱があるが、当該領域の専門家でなくても評価が可能との考え方に基いて、ガイドラインの作成法で評価する場合がほとんどであった¹⁵⁾¹⁷⁾¹⁸⁾。そのうち、最も使用されているのはAGREEの評価法である (Table 2)¹⁹⁾²⁰⁾。本邦での診療ガイドラインも当初は、専門家のコンセンサスなどに基づくものなど、その質が十分とはいえなかったが、日本癌治療学会がん診療ガイドライン評価委員会で行われている日本で作成されてきた種々の癌の診療ガイドラインの評価では、作成方法のレベルは近年徐々に向上してきていることが報告されている²¹⁾。

しかしながら、ガイドラインの評価は、作成方法による評価だけではなく、本来の目的である、ガイドラインによる診療内容の変化、患者のADLや予後の変化を評価し、その有用性も評価されるべきである。これらの評価にはガイドラインが発表され数年を要することや、実際の評価が煩雑であることもあり、多数作成された本邦のガイドラインでも、実際のガイドラインの活用度やその有

効性に関してはあまり関心が払われてこなかった。ガイドラインを評価する場合、ガイドラインの作成方法の評価とガイドラインの有効性の評価とでは時間軸が異なり、両者は分けて考える必要がある²²⁾。

2. 症例登録によるガイドラインの検証

ガイドラインが十分なエビデンスに基づいたものであれば臨床指標の遵守率などで評価することもできるが、ガイドラインが限られたエビデンスから作成されたものであれば、まず、実際の臨床例とガイドラインとの整合性の検証を行う必要がある。そのためには、アンケート調査や症例登録などによってガイドラインを検証することを優先しなくてはならない。

ガイドラインの普及率やガイドラインによる診療内容、患者のADLや予後の変化の評価は、アンケート調査や症例集積でもある程度可能であるが、さらに好ましいものは、前向きな症例登録である。該当症例の概要や臨床指標の実施の有無を登録させることによって、毎回ガイドライン内容を確認させ、臨床指標の実施を促すことができる。また、アンケート調査よりもより質の高い、ガイドラインと臨床例での整合性の評価、ガイドライン内容の実施の有無、さらに、その実施の有無と患者のADL、予後の評価、ひいてはガイドラインの評価が可能となる。

しかしながら、症例登録の実施にあたって解決すべき課題は少なくない。まず、症例登録システムの構築が必要であるが、これには大学病院医療情報ネットワーク研究センター (UMINセンター) のインターネット医学研究データセンター (INDICE: <http://indice.umin.ac.jp/>) のシステムの利用は安価で信頼でき、すでに平成22年4月現在93万5千例が登録されており、その利用が推奨される。次に、個々の症例の登録が煩雑であることが最大の問題である。そのためには、登録内容を簡潔とし、また、入力を容易とするとともに、臨床医だけではなくclinical clerkやclinical nurseなどを配置し、症例登録を容易に行えるシステム体制の整備も必要である。ガイドライン普及による効率化や予後改善効果を考慮すれば、保

Table 2. AGREE での評価項目 (文献 19, 20 より引用)

	項目
1	ガイドライン全体の目的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的問題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。
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。
5	患者の価値観や好みが十分に考慮されている。
6	ガイドラインの利用者が明確に定義されている。
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。
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。
9	エビデンスの選択基準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。
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。
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。
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。
14	ガイドラインの改訂手続きが予定されている。
15	推奨が具体的であり、曖昧でない。
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。
17	どれが重要な推奨か容易に見分けられる。
18	利用のためのツールが用意されている。
19	推奨の適用にあたって予想される制度・組織上の障害が論じられている。
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。
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。
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。
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。
全体評価	あなたはこれらのガイドラインを診療に用いることを推奨しますか？

*各項目を1(全く当てはまらない)～4(強く当てはまる)で評価する。

険点数などの優遇処置も妥当と思われる。

一般診療では前向き登録は必ずしも必要ないかもしれないが、先の急性胆管炎、胆嚢炎のガイドラインは質の高い知見が少なくコンセンサスによって形成された診断基準や重症度判定であったため、それらの検証のために、現在、「CLASS Tokyo Study」としてインターネットを利用して世界から急性胆管炎を登録する研究を開始している(<http://class.umin.jp/>)。

3. 臨床指標、Bundleによるガイドラインの評価

多数あるガイドライン内容すべての実施の有無やそれらによる影響を評価することは困難なため、ガイドラインに「臨床指標」を掲げ、その実施の有無で評価することが提案されている。この臨床指標の提示は、ガイドライン作成での評価項目として前述のAGREEでの項目21にも掲げられている¹⁹⁾²⁰⁾。ガイドラインに臨床指標を設定すると、ガイドラインの遵守率が向上するとされて